

LANGUAGE AND LINGUISTICS MONOGRAPH SERIES 50

《語言暨語言學》專刊系列之五十

**Breaking Down the Barriers:
Interdisciplinary Studies in
Chinese Linguistics and Beyond**

綜古述今 鉤深取極

Volume 2

Edited by

Guangshun Cao, Hilary Chappell, Redouane Djamouri and
Thekla Wiebusch

曹廣順 曹茜蕾 羅 端 魏婷蘭／編輯

Institute of Linguistics, Academia Sinica, Taipei, Taiwan

中央研究院 語言學研究所

2013

LANGUAGE AND LINGUISTICS MONOGRAPH SERIES

EDITORIAL BOARD

<i>Editor in Chief</i>	Chih-Chen Jane Tang (Academia Sinica)
<i>Associate Editors</i>	Henry Y. Chang (Academia Sinica)
	Thomas Hun-tak Lee (The Chinese University of Hong Kong)
	Pei-chuan Wei (Academia Sinica)

Editorial Board Members

- Mark C. Baker (Rutgers University)
- Robert A. Blust (University of Hawai'i, Mānoa)
- Keh-jiann Chen (Academia Sinica)
- Gennaro Chierchia (Harvard University)
- Guglielmo Cinque (University of Venice)
- Bernard Comrie (Max Planck Institute for Evolutionary Anthropology)
- Hiroya Fujisaki (University of Tokyo)
- Martin Haspelmath (Max Planck Institute for Evolutionary Anthropology)
- C.-T. James Huang (Harvard University)
- Larry M. Hyman (University of California, Berkeley)
- Christine Lamarre (Institut National des Langues et Civilisations Orientales)
- Richard K. Larson (Stony Brook University)
- Paul Jen-kuei Li (Academia Sinica)
- Chinfu Lien (National Tsing Hua University)
- Yen-hwei Lin (Michigan State University)
- Ian Maddieson (University of California, Berkeley)
- Alec Marantz (New York University)
- James A. Matisoff (University of California, Berkeley)
- Hsiu-fang Yang (National Taiwan University)

Editorial Assistants

Joyce Chun-yu Kuo (Academia Sinica)
Greta Ya-Ying Chuang (Academia Sinica)
Olivia Yating Weng (Academia Sinica)

Table of Contents

目 錄

Volume 1

Preface.....	i
Photographs of Alain Peyraube.....	ix
Tribute..... William S-Y. Wang (王士元).....	xvii
Tribute..... Guangshun Cao (曹廣順).....	xix
Biography of Alain Peyraube.....	xxi
List of publications of Alain Peyraube.....	xxix

Cognition

The Evolutionary Origin of Language Reconsidered Charles N. Li (李 訥).....	1
Language Learning and the Brain: An Evolutionary Perspective William S-Y. Wang (王士元).....	21
On Moving in Conceptual Space: Readers' Responses to a Parallel Text in Chinese and English Gisela Bruche-Schulz.....	49
Encoding Motion Events in Taiwan Sign Language and Mandarin Chinese: Some Typological Implications James H-Y. Tai (戴浩一) and Shiou-fen Su (蘇秀芬).....	79
Psycholinguistique comparative du traitement des phrases: référence spéciale au chinois Michèle Kail.....	99

Modern Chinese

Une des formes d'expression de la subjectivité en chinois mandarin:

L'auxiliaire modal 要 *yao*

Viviane Alleton (艾樂桐)..... 129

Réalisation syntaxique ou formation lexicale? Le statut grammatical des
adjectifs formés par reduplication en chinois contemporain

Zhitang Yang-Drocourt (楊志棠)..... 145

說“還是”——祈願情態的浮現

方 梅..... 171

Reflexive Nominal Compounds

Ekkehard König 187

Case, Tense and Clauses

Yen-Hui Audrey Li (李豔惠)..... 205

漢語句法的雙層式分析法：功能與形式的會聚

曹逢甫..... 237

Verbal Reduplication and Verbal Classifiers in Chinese

Marie-Claude Paris (白梅麗)..... 257

Decoding the Multifunctional Usages of *gen* in Chinese

Yun-hsiang Chuang (莊雲翔) and Elizabeth Zeitoun (齊莉莎) 279

Are A-not-A Questions Alternative Questions or Wh-Questions?

Liejiong Xu (徐烈炯)..... 305

Lexicon

詞義演變三例

蔣紹愚..... 321

試論與漢字相關的「言語錯誤」在漢語詞彙歷史演變中的作用 朱慶之.....	345
---	-----

Frequency Effects in Chinese Morphology: Diachronic Evidence from a Synchronic Corpus Chao-Jan Chen (陳超然)	371
---	-----

Ancient and Classical Chinese

Improper Subjects: Towards a Philosophical Grammar of Existence Predicates in Classical Chinese Christoph Harbsmeier (何莫邪).....	383
---	-----

The Causative in Old and Middle Chinese Tsu-Lin Mei (梅祖麟).....	403
---	-----

A Syntactic Analysis of Modal <i>bì</i> 必: Auxiliary Verb or Adverb? Barbara Meisterernst (梅思德)	425
--	-----

Medieval Chinese

說“定知”：古漢語中的一種長距離副詞移位 李 明.....	451
----------------------------------	-----

中古以來表患病類詞語的搭配組合 方一新、王雲路.....	469
---------------------------------	-----

Contact and Change in the History of the Chinese Language Hsiao-jung Yu (遇笑容).....	485
---	-----

Early Mandarin Chinese

元白話語言性質再研究 曹廣順.....	503
------------------------	-----

也說“漢兒言語” 江藍生.....	517
----------------------	-----

Volume 2

Diachrony

- “完結短語”及漢語結果補語的語法化和完成體標記的演變過程
沈陽、玄珮..... 553
- Syntactic Change in Chinese and the Argument-Adjunct Asymmetry
Redouane Djamouri (羅端), Waltraud Paul (包華莉) and
John Whitman (魏德中)..... 577
- Short Passives in Modern and Classical Chinese
Shengli Feng (馮勝利)..... 595
- Chinese Resultative Verb Compounds: Lexicalization and
Grammaticalization
Chaofen Sun (孫朝奮)..... 625
- 近、現代漢語裡“給+VP”的形成
張伯江..... 651
- 歷時與共時：起始類介詞的語意發展和多義現象
熊慧如..... 665
- 試說“連X+都VP”構式的語法化
楊永龍..... 687

Classification

- Peut-on compter les langues?
Gérard Lenclud 707
- On Principles and Practices of Language Classification
Katia Chirkova (齊卡佳) 715
- The Classification of Sinitic Languages: What is “Chinese”?
Victor H. Mair (梅維恒) 735

Language Contact, Typology and Chinese Dialects

On Isomorphism and Formulas of Equivalence in Language Contact

Bernd Heine 755

Pan-Sinitic Object Marking: Morphology and Syntax

Hilary Chappell (曹茜蕾) 785

The Development of VOC to VCO in Mai-Cantonese of Sanya in Hainan

Andy C. Chin (錢志安) and Benjamin K. Tsou (鄒嘉彥) 817

南方民族語言比較句語序的演變和變異

吳福祥 831

Classifiers and Some Typological Considerations

Dan Xu (徐 丹) and Jingqi Fu (傅京起) 865

When Lexicalization Meets Grammaticalization: The Development of 'wang+path' Adverbials in Northern Chinese

Christine Lamarre (柯理思) 887

再論“們”的來源

李 藍 911

Shape Classifiers in Earlier Southern Min Texts

Chinfa Lien (連金發) and Liching Livy Chiu (邱力璟) 933

粵語“先”、“添”虛實兩用的跨域投射解釋

劉丹青 951

Cantonese *sin* 先 and the Question of Microvariation and Macrovariation

Rint Sybesma (司馬翎) 971

湖南南部處置和被動同標記現象的再思考

伍雲姬 995

Script

- La marque du redoublement graphique en chinois archaïque
Chrystelle Maréchal (麥里筱) 1011
- Le développement des livres de rimes en dictionnaires: *Qièyùn* et
Kānmiù bǔquē Qièyùn
Françoise Bottéro (蒲芳莎) 1031
- L'étymologie graphique et l'héritage des concepts fondamentaux de la
culture chinoise
Shun-chiu Yau (游順釗) 1051
- Graphic Designs: A New Approach to Palaeographic Analysis
Ken-ichi Takashima (高嶋謙一) 1059
- Palaeography Meets Linguistics: Analysis of 𠂔 and 𠂔 in the
Oracle-Bone Inscriptions
Anne O. Yue (余靄芹) 1079
- Radicals as Ontologies: Concept Derivation and Knowledge
Representation of Four-Hoofed Mammals as Semantic Symbols
Chu-Ren Huang (黃居仁), Ya-Jun Yang (楊雅君) and
Sheng-Yi Chen (陳聖怡) 1117

Diverse

- The Last Text of the Last Pazih Speaker
Paul Jen-kuei Li (李壬癸) 1135
- Some Issues in the Study of Chinese Poetic Prosody
Hongming Zhang (張洪明) and Chenqing Song (宋晨清) 1149

“完結短語”及漢語結果補語的語法化 和完成體標記的演變過程

沈 陽¹ 玄 玥²

北京大學¹

北京語言大學²

“完結短語”假設是在生成語法理論框架內提出的關於漢語虛化結果補語的理論，認為虛化結果補語在動詞短語內部投射為主要謂語動詞的內部體“完結短語”。依據這個假設理論，可以解釋漢語完成體標記“了”從結果補語到虛化結果補語、最終發展為體標記的語法化過程，並在句法結構上支持動詞後存在兩個不同的“了”。

關鍵詞：“完結短語”假設，結果補語，語法化，完成體標記

1. “完結短語”假設的理論背景

漢語結果補語中有一個小類，粘合式述補結構即動結式中的補語 R，其自身的謂詞詞彙意義很不明顯，語義指向述語，主要表示動作事件“完成、結束”的意義，與前項動詞 V 有比較廣泛的搭配，具有很强的能產性，本文稱為“虛化結果補語”。¹ 該類補語在現代漢語中基本上是一個封閉的類，都是由實義動詞虛

¹ 這類虛化結果補語的術語，在有些研究中稱為“動相補語”，這源自於趙元任 (Chao 1968) 的術語“phase complement”，及呂叔湘 (1980) 的翻譯。“相 (phase)”多數情況下指情狀的某一特定階段 (Comrie 1976)；鄭良偉 (1988)、郭銳 (1997)、吳福祥 (1998) 等都採用這種說法。陳平 (1988) 所說的“時相 (phase)”主要指句子的體貌。而語法學研究中還有研究將“aspect”，通常譯為“體（體態、體貌）”，也稱為“動相”，如呂叔湘 (1942) 的《中國文法要略》，這裡指一個動作過程中的各個階段。可見有關“相”的術語用法並不統一。而且從字面上理解，“動相”也表示動作的“相（階段）”，這裡應該不只包含動作的完成結束狀態，還包含動作過程的其他階段，如開始、進行等。本文認為“動相補語”這一術語雖然有一定的使用基礎，但是容易造成誤解，所以建議不採用它。至於張洪年 (1972)、趙元任 (1980)、梅祖麟 (1994)、吳福祥 (2001-2002) 等研究中都使用“狀態補語”這個術語，這與本文前面介紹的表示結果狀態的帶“得”組合式述補結構的稱謂也有衝突，所以也不建議採用。其他如“唯補詞”、“指動補語”、“時相詞”等術語也都有一定使用局限。因此本文建議採用體現這類補語語法演變過程的術語“虛化結果補語”，顯示

化而來的，例如“完、好、掉、住、成、了(liǎo)、著(zháo)、過(guò)”等。例如：

- (1) a. 我早就擦完了客廳。 b. 小朋友們捂好了眼睛。
 c. 我徹底忘掉了那個約會。 d. 你盯住那個可疑的人。
 e. 他好不容易聽成了演唱會。 f. 你把雞蛋吃了(liǎo)了再上學。
 g. 他可惹著(zháo)了小麗了。 h. 昨天放過(guò)了那部片子了。

這些補語動詞都可以（至少曾經可以）作實義結果補語，表示動結式某論元的結果狀態，而不是動結式整個事件的完成結束。學者們的研究都表明虛化結果補語是由實義結果補語語法化而來的。虛化結果補語與實義結果補語不同，具有後附性和粘著性、語義指向動詞、不能指派論元、不是動結式的語義重心；虛化結果補語又不完全是漢語的體標記，與典型的完成體標記“了”相比，它與前項動詞具有選擇限制性、可以與否定詞“沒”共現、可以與完成體標記共現、可以形成可能補語、可以做句子焦點等。

以往對虛化結果補語的研究多是從結果補語、詞類劃分、體貌系統和語法化等角度出發，生成語法理論框架內對結果補語的研究幾乎都沒提到虛化的結果補語。玄玥(2008)提出“完結短語”假設，以期解決漢語虛化結果補語的句法結構問題。根據她的研究，漢語虛化結果補語是一種內部體，即表達謂語動詞的“完結”功能，在句中將投射“完結短語(Telic Phrase)”，簡稱 TelP，它的功能是給予主要謂語動詞一個終結點，是緊鄰動詞的功能短語，在 VP 之內，低於 VP 內的其他功能性成分“體短語”。漢語動結式的動詞短語 VP² 的句法結構可描述如下：

- (2) [_{VP} [_{V'} V [_{AspP} [_{Asp'} Asp [_{TelP} [_{Tel'} Tel [_{VP} [_{V'} V PP³]]]]]]]]]

其本來是一種結果補語，在發展的過程中虛化了。

² 句法結構中有兩個不同的 VP。雙層 VP 結構中上層的輕動詞短語 vP（以及本文下文中的“體短語 AspP”、“完結短語 TelP”）是功能性的成分，下層的 V 才是在句中起到謂語作用的主要動詞。在行文中，句法結構中下層的 VP 被稱為“主要動詞短語 VP”（相應的有“主要動詞 V”），而相對於 IP、CP 等句子的功能層的詞彙層 VP 直接稱為“動詞短語 VP”，或無標記地稱為 VP。

³ 我們認為主要動詞 V 的補足語(Complement)是介詞短語 PP 或其他非論元成分。漢語中有多种補語形式，除了粘合式動結式的謂詞性補語外，還有其他表示動作的結果性質的補語，也起到補充、說明、對動作結點限定的作用，如帶“得”的結果補語從句（或稱狀態補語）（如“他笑得岔氣了”、“他打得手腫了”），介詞結構補語（如“走到天安門”、“寫在黑板上”），時量短語補語（如“他哭了三小時”、“他喝了三天粥”），動量短語補語（如“他來過幾次”、“他走了幾趟”）等。這些補語形式在本文理論假設中置於主要謂語動詞的補足語位置，篇幅所限不能詳

先要說明一下 (2) 中體短語 AspP 的位置。與謂語動詞的時態相關的時間有三個，即“事件時間 (event time)”（動作發生的時間），“參照時間 (reference time)”和“敘述時間 (speech time)”，涉及事件時間和參照時間的是“體 (Aspect)”範疇，而涉及事件時間和句子敘述時間的是“時 (Tense)”範疇。現在的研究多傾向於將二者分開，在句子中形成爲不同的投射，以表明是謂語所涉及的不同體貌性質。所以與“時”相比，“體”是更接近於謂語動詞時間性的功能成分，句法位置上也更接近於謂語動詞。以往研究多將“體”處理爲句子層面的功能性成分，是 TP 裡的一個投射，句法位置高於 VP，但這種處理並無理論上或經驗上的證據。生成語法的“小句理論”⁴ 認爲體短語是動詞短語 VP 內部的一種體貌成分，位於主要動詞短語 VP 之後，本文也將其處理爲動詞短語 VP 內部的功能性投射。功能性短語在句法上都高於詞彙性短語，而“完結短語 TelP”是主要動詞 V 更基本的“體貌”特徵，與動詞關係更緊密，所以體短語的句法位置就是：在 VP 之內，完結短語 TelP 之上，輕動詞短語 vP 之下。不過無論是否同意體短語 AspP 在動詞短語內部，它比完結短語投射更加功能化、在完結短語 TelP 上方的事實是不可否認的。也就是說，如果某些理論或某些研究不支持將體短語置於動詞短語 VP 內部，而是如外部體說法置於動詞短語 VP 之外，也不會影響體短語和動詞內部體“完結短語”的基本層次結構。

本文的目的是：在“完結短語”假設的理論背景下討論漢語完成體標記“了”的演變過程，力圖在生成語法理論框架內對結果補語的語法化現象進行解釋。

2. 完成體標記“了”的歷時發展線索

本節主要引述前人研究，簡要說明完成體標記“了”的發展演變情況。

王力 (1958) 指出：“動詞詞尾‘了’和‘著’的產生，是近代漢語語法史上

細討論。由於這個位置不涉及動結式結構中主要成分，所以暫不闡釋這個位置的理論意義對本文的研究沒有太多影響。

⁴ 小句理論是 Hoekstra (1988)、Sybesma (1992, 1999) 等提出的關於結果補語的理論。基本觀點是漢語結果補語結構中的述語 V 表達一個非狀態性的動作行爲，該動作行爲內部沒有作用的範圍和到達的終點；V 可帶有一個由簡單主謂結構構成的小句性補足語 (V 的補足語還可以是名詞短語，形成動賓結構)，這個結果補語小句則爲主句謂語動詞提供作用的範圍和到達的終點。同時補語小句含有非時態 (tense) (即不是 TP) 的功能性成分，漢語補語小句的功能性成分是體標記“了”。小句謂語是動結式的補語 R。小句的主語有兩個來源：當主要動詞述語爲及物動詞時，小句主語是動結式的賓語，即句子表層的賓語；當述語爲不及物動詞時，小句主語是動結式的主語，即句子的表層主語。

劃時代的大事”。對完成體標記“了”已有很多研究成果，如王力(1958)、太田辰夫(1958)、趙金銘(1979)、梅祖麟(1981, 1988)、潘允中(1982)、曹廣順(1995, 2000)、吳福祥(1996, 1998, 2003)、廖名春(1990)、柳士鎮(1992)、楊秀芳(1992)、劉堅等(1992)、孫錫信(1992)、孫朝奮(1997)、馮春田(1988)、蔣冀騁、吳福祥(1997)、志村良治(1984)、張頴(2000)、石鎰(2000)、蔣紹愚(1994, 1999)等。

根據上述研究，“了”最初是表示“了結”意義的動詞，最早見於晉宋時期（曹廣順 1995）。這種處於“V+O+了”位置的“了”不是動態助詞，只是連動式的一個謂語。例如：

- (3) a. 公留我了矣，明府不能止。（《三國志·蜀書·楊洪傳》）
- b. 益郭耆舊傳令送，想催驅寫取了，慎不可過淹留。（王獻之《雜帖》）
- c. 左來右去，二主各了。若還悔者，罰毯一張供獻。（《吐魯番出土文書·前涼升平十一年王念賣駝券》）

到了晚唐五代時候，連動式已經轉化為動結式，不及物動詞也可進入第二個謂語的位置、並帶有賓語。可以作為補語的不及物動詞“了”的位置也出現了變化，出現了“V+了+O”的格式，“了”也從處於賓語後變到緊貼在動詞後，研究者一般把這種“了”看作詞尾或動態助詞。據劉堅等(1992)、孫錫信(1992)等的研究，“了”字的虛化是從中晚唐開始的，在當時的韻文中能看到幾例。例如：

- (4) a. 見了師兄便入來。（《變文集·難陀出家緣起》）
- b. 唱喏走入，拜了起居，再拜走出。（《變文集·唐太宗入冥記》）
- c. 看經須解義須修行。若依了義教，即入涅槃城。（《祖堂集》）

到了宋代，動態助詞“了”大量出現；到了南宋，“此期的禪宗語錄，已經很難找到不用‘動+了+賓’格式的了”（劉堅等 1992）。尤其“了”出現在動結式謂語之後，更體現了它的體標記性質。例如：

- (5) a. 長大了擇時聘與人，六親九族皆歡美。（《變文》P686）
- b. 恰則心頭托托地，放下了日多縈系。（毛滂《惜分飛》）
- c. 如今都教壞了學生，個個不肯讀書。（《朱子語類》，卷一）

漢語史的研究主要集中在體標記“了”的產生年代和形成原因兩方面。動態助詞⁵“了”的產生時代的確定是與動態助詞“了”的判斷標準聯繫在一起的，這方面有很多爭論，如梅祖麟(1994)把“動作動詞”後面的“了”看成“狀態補語”(即虛化結果補語，見附注1)，把“成就動詞”後面的“了”視作“完成貌詞尾”；蔣紹愚(2001)把出現在持續動詞和非持續動詞後面的“了”都看成動相補語(即虛化結果補語，見附注1)，前者表示完結，後者表示完成，語義都指向動作；吳福祥(1996)把用於“瞬間動詞”後作為判定動態助詞“了”的標準之一，其後(1998)又把持續動詞和非持續動詞後面的“了”分別處理作“結果補語”和“動相補語”。雖然各家所用術語不同，但在判定動態助詞“了”上有一個基本共識，即“了”用於持續動詞和非持續動詞後面時虛化程度不同，後者的虛化程度高於前者。不過各家對兩類格式中“了”的虛化程度的認定和歸類上仍然有分歧。關於動態助詞“了”的形成原因，梅祖麟(1981)認為是詞彙興替的結果，代替了“畢、已、訖、竟”等表示完成的動詞，而且“了”位於動詞之後、賓語之前是由於受到同類格式補語前移的影響；曹廣順(1986)、梅祖麟(1994)認為動態助詞“了”的形成是受到“卻”的影響，只是前者認為“卻”是唐代表示完成體的助詞，而後者認為它是狀態補語。吳福祥(1998)認為是“了”先在“動+了”格式中虛化為動相補語，形態化後帶上賓語。但蔣紹愚(1994, 2005)仍認為前移說正確，即“V+O+了”中“了”是完成動詞，有的則已經發展為動相補語(如“辭父娘了”)；這種格式裡的“了”語法化的程度進一步加深，對動詞的依附性加強，就可以改變語法位置，“前移”到動詞後面，成為“Vt+了+O”。也是受到同類結構的影響，如“卻、取、得、將”等，可在動賓間，所以“了”類化作用也可以進入這種語法槽。

其實如何判斷歷時材料中“了”的性質，即體標記“了”產生的年代，與本文對完成體標記“了”的語法化分析並無直接關係，本文更注重探討在生成語法理論框架內，如何解釋漢語體標記的形成和發展。所以只要在漢語史研究找到了這樣共識性的結論，即“了”的語法化過程是“實義動詞做謂語——實義結果補語——虛化結果補語——完成體標記”，就可以印證本文的理論假設。至於體標記形成原因和發展軌跡，本文也只是在理論層面結合句法結構中的“完結短語”假設來探討。

⁵ “動態助詞”是從詞類角度的名稱，與本文的“體標記”所指相同。為保持以往研究的原貌，此處也用這個術語。

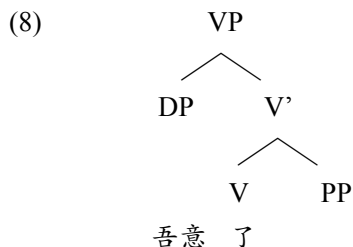
3. 完成體標記“了”的形成過程——從完結短語 TelP 到體短語 AspP

本節主要結合前面對虛化結果補語的研究，用完結短語假設為基礎的生成語法理論來解釋完成體標記“了”的語法化歷程。

最初“了(liǎo)”作為實義動詞，有“瞭解”的意思，如(6)；也表示“完成、結束、了結”的意思，可獨立作謂語動詞，如(7)。因為“了”前還可以有副詞作狀語進行修飾，證明“了”是謂語；獨立作謂語時，就是句子的唯一主要動詞，結構也比較簡單，如(8)：

(6) 恐此三事，當待賢能然後了耳。(《三國志·魏書九·夏侯玄傳》)

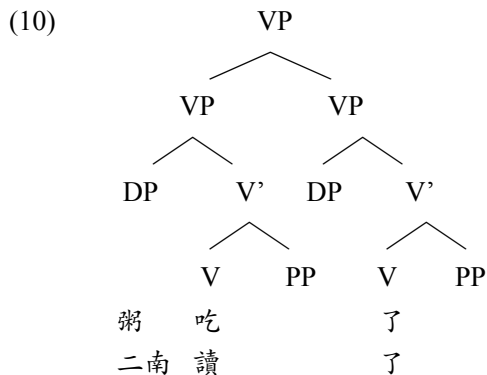
(7) 吾常疑汝於文偉優劣未別也，而今而後，吾意了矣。(《三國志·蜀書十四·費禕傳》)



作為動詞的“了”還會出現在連動式中，表示前一個動詞短語完成、結束，如(9)。此時結構是兩個並列的謂語，其句法結構如(10)：

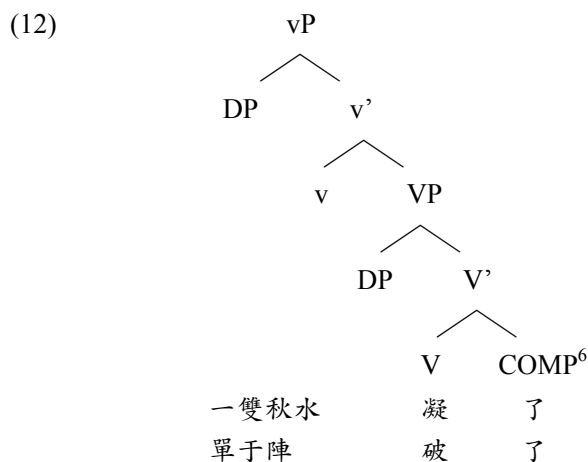
(9) a. 師住庵時，有一僧吃粥了，便辭師。(《祖堂集》)

b. 今公讀二南了，還能不正牆面而立否？(《朱子語類》卷二十五)



當“了”直接在動詞後作補語，先仍是保留動詞意義，表示某事物的完結，在此結構中就是表示動結式賓語所表示的事物完成了，這是實義結果補語，如(11)。其句法結構為(12)：

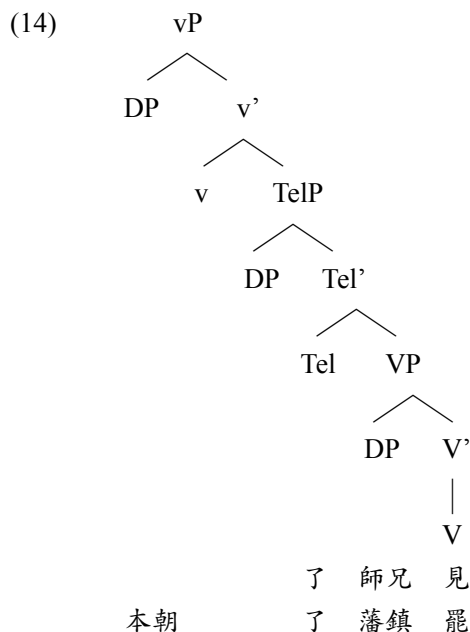
- (11) a. 凝了一雙秋水。(白居易《如夢令》)
b. 將軍破了單于陣，更把兵書仔細看。(沈傳師《寄大府兄侍史》)



當“了”在補語位置上，表示整個事件的完成時，是本文所說的虛化結果補語，如(13)。此時“了”已虛化，因為不是“見師兄+師兄了”，“了”指的是“見師兄”這個事件完成了。此時的“了”就是本文所說的虛化結果補語，不能指派論元，是主要動詞的內部體，位於其上。“了(liǎo)”語法化為虛化結果補語的初期，漢語中還沒有發展出完成體標記，所以 vP 和 VP 之間只有“完結短語 TelP”功能層，其句法結構如(14)：

- (13) a. 見了師兄便入來。(《變文集·難陀出家緣起》)
b. 自本朝罷了藩鎮，(《朱子語類》卷二十五)

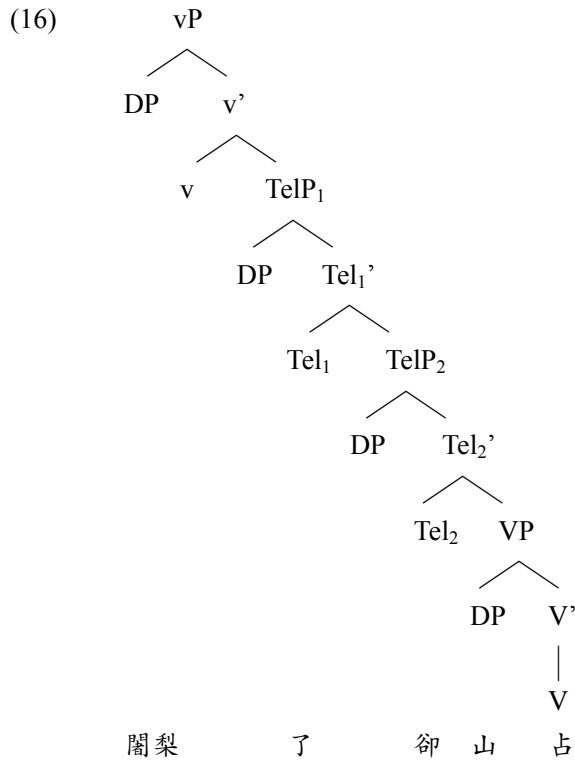
⁶ 實義結果補語的句法位置有很多爭論，篇幅所限不能詳細討論。本文暫且將其置於述語動詞的補足語位置，以區別於虛化結果補語。



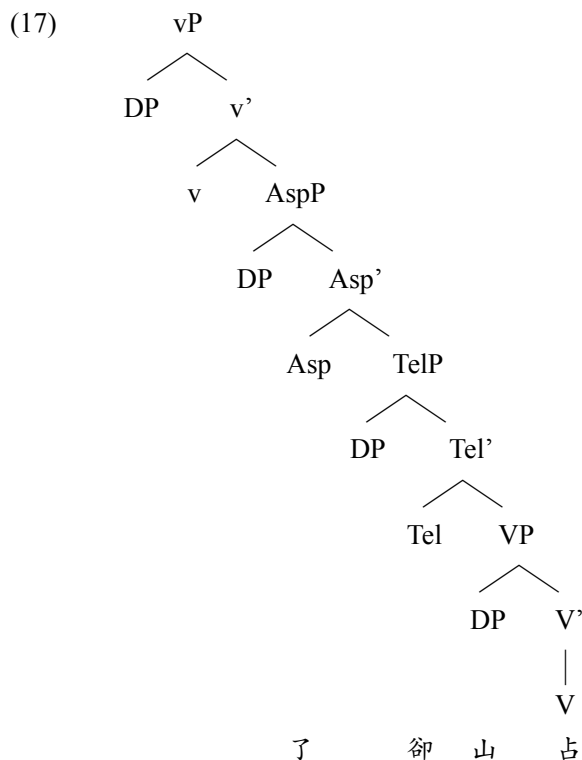
在長期頻繁使用中，“了”成為典型的表完結的詞語，但語義也漸漸弱化，相應地完結功能漸漸弱化。因單獨出現“了”做完結短語中心語的句子已不足以表達主句謂語動詞的完結點，所以又加入新的完結短語，“了”出現在帶有虛化結果補語的動結式謂語之後，如 (15)。

(15) 與摩則大唐國內山總被閻梨占卻了也。(《祖堂集》卷六)

上例中補語“卻”也是近代漢語中經常表示完結的詞語，主要謂語動詞“占”後先有一個表示完成的“卻”，接著是個表示完成的“了”，更可證明“了”由補語逐漸虛化為體標記的過程。表示完成、完結的補語“卻”也是比較容易虛化的，許多研究表明“卻”在漢語史上也作完成體標記。(15)中的“卻”是虛化的結果補語，所以也可以認為“了”已經從一種較虛化的完結短語變成體短語，而自身也可看作是完成體標記。其句法結構如 (16)：



由於“了(liǎo)”表示“完成”、“完結”的意義，與完結短語的基本意義趨同，與其他表示完結的動詞“畢、已、訖、竟、卻”等競爭而勝出，成為最經常使用的表示事件完成的完結短語中心語。在長期頻繁使用中，“了(liǎo)”漸漸語法化，成為完結短語中心語的典型。但是由於經常使用，它的完結功能也在高頻使用中漸漸弱化。



單獨出現“了(le)”做完結短語中心語的句子，已不足以表達主句動詞的完結點，所以又加入新的完結短語，而以“了”作為中心語的完結短語也會保留在句中。最終長期作為完結短語中心語的“了(liǎo)”演變成了體貌標記，成為完成體必備的成分，並伴隨語音弱化，變成“了(le)”。這樣處於上層的完結短語₁(TelP₁)演變為體短語(AspP)，其句法結構也就成了(17)。

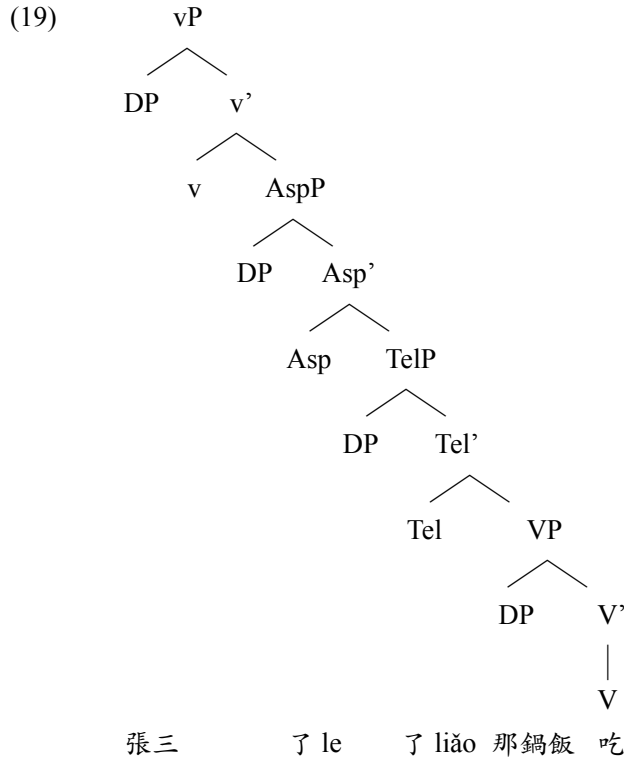
由並列結構轉向主從結構，符合語法化的軌跡，英語助動詞由句子的連接詞發展而來、漢語的體標記由連動式到動補式到動態助詞都體現了這個人類語言共同的發展軌跡。

4. 動詞後有兩個不同的性質的“了(le)”

上文對完成體標記“了(liǎo)”的語法化歷程的解釋符合人類語言體標記的語法化方式。語法化的過程是從詞彙層向功能層轉變、由位置較低的功能層向較高的功能層轉變的過程。在從V到T的語法化發展中是有中間狀態的，比如現代

英語中“dare（敢）”和“need（需要）”都有動詞和情態詞兩種用法，可以看出情態詞由動詞演變而來的痕跡。漢語中也有類似情況。“了(le)”在普通話中雖然已經是典型的完成體標記了，但是北方方言中還保留動詞“了(liǎo)”。這個動詞幾乎很少單獨做謂語，而主要做補語，如(18)。此時句子經常使用的“完結短語”變爲了“體短語”，其句法結構爲(19)：

(18) 張三吃了(liǎo)了(le)那鍋飯。



Travis (2010) 也指出兩種語素共現說明他們是不同短語的中心語 (the co-occurrence of these two morphemes suggests that they are generated in different heads)。而漢語中還能出現“了(liǎo)”與“了(le)”共現，說明動詞性的“了(liǎo)”和體標記“了(le)”已經是不同的語素，需要分別投射不同的短語，這應該也是完成體標記“了”語法化的痕跡。

與之相關的是關於動詞後“了”的語法意義，這也是漢語研究中一直有較大爭議的問題。有的學者認爲動詞後“了”表示動作的完成，是完成體的標記，如

趙元任 (Chao 1968)、王力 (1944)、高名凱、石安石 (1979)、呂叔湘主編 (1980)、盧英順 (1994)、吳福祥 (1998) 等。有的學者認為“了”表示實現，如劉勛寧 (1988)、劉月華 (1988)、竟成 (1993)、Sybesma (1997)、張家驊 (2004)、潘文國 (2003) 等。也有的學者認為它表示動作的開始，是“起始體 (inchoativity 或 inceptive aspect)”，如 Rohsenow (1978)、《語言與語言學詞典》(1984)、戴耀晶 (1990, 1997) 等；還有“完成-延續”體標記 (金立鑫 1998)、表示“實現-延續” (房玉清 1992)、表示“達成” (史有為 2003) 等各種提法。也有學者不拘泥於名稱之爭，認為“完成”不等於動作結束，“完成”和“實現”是一致的，如王還 (1990) 等。以上對動詞後“了”的語法性質的定義都基於“了”的語法意義是均質統一的，即只有一個動詞後“了”。但也有研究對“了”進行語法環境的區分，有人甚至區分出謂語動詞後的幾種“了”。如朱德熙 (1982) 提出動詞後的“了”是“表示動作的完成”，而形容詞後的“了”是“表示已經實現了的事”。馬希文 (1983) 區分具有補語性質及動詞後的體標記“了”。木村英樹 (1983) 提出“結果補語性詞尾”這一個語法範疇來區別完成體標記的“了”。房玉清 (1992) 提出動詞後“了”在結束性動詞 (包括述補式動詞或短語) 後表示“完成”，用在持續性動詞後則表示“實現”。李小凡 (2000) 也提出動詞後“了”所表示的完成體有兩個次範疇，一個是“完畢體”，當它用在表示動作的謂詞後面，表示動作“完畢”，基本都是單音節動詞；另一個是“生成體”，當它用在形容詞和動結式、動趨式之後，表示“動作變化在某一時刻已經生成了某種結果”。其他研究如范方蓮 (1963)、于根元 (1983)、戚曉杰 (1995) 等對動詞後的“了”有兩個意義也有模糊的意識。左思民 (1997)、金立鑫 (2002) 等也主張把動詞後的“了”進行細化分析。上面這些研究雖然都注意到動詞後“了”的兩種不同情況，但只是將它們解釋為不同的義項，並主要從語義上來區分。

本文也主張對動詞後“了”做異質分析，但認為這一點可歸結為句法結構的差異。前面的結構分析證明，動詞後應該存在兩個同形的“了”，一個是表示完結，是謂語核心語義的一部分，一個是虛化結果補語“了 (liǎo)”進一步虛化的形式。例如“你吃了晚飯再走”中的“了”就是如此。但由於動詞後表示完結的“了 (le)”和表示完成體的“了 (le)”同形，很難區分開，因此只能通過否定形式進行證明。朱德熙 (1982) 指出，“和‘VP 了’相對應的否定形式是‘沒(有) VP’，‘沒’和‘了’通常是不能在同一個句法單位中‘共現 (co-occur)’的”。宋玉柱 (1980) 也曾明確指出漢語裡沒有“沒去了 (le) 幾次”這種結構，只有讀“liǎo”才可以。這些觀點得到學界普遍承認。例如：

(20) *我沒吃了 (le) 那塊餅。

但是普通話中也確實存在“沒（沒有）+謂語+了 (le)”的結構，如陳剛 (1981, 1985) 和毛敬修 (1985) 等就舉出很多這樣的例句，如 (21)；而且他們所舉的表示損失義的句子中“了”都可以讀“(le)”，如 (22)。雖然他們也指出這些例句中的“了”都可以替換為“掉”或“上”，而且有的例句也可以讀“(liǎo)”，但是讀作“(le)”也是完全可以接受的。例如：

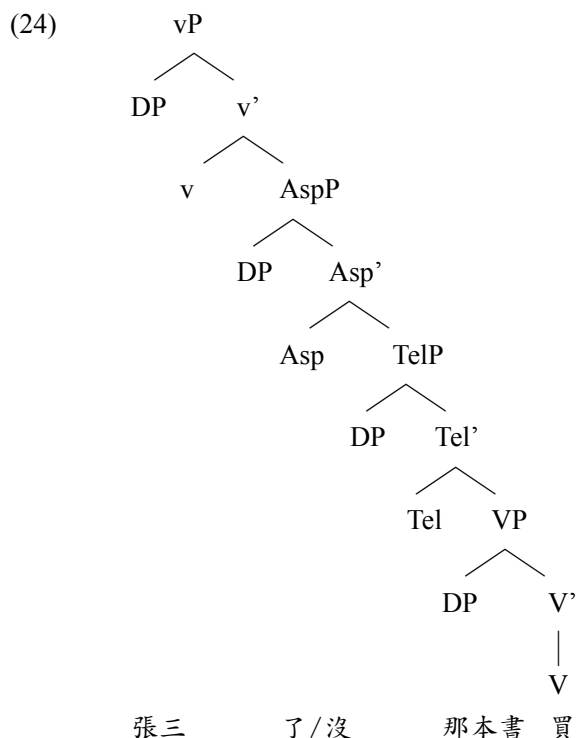
- (21) a. 虎妞並沒有懷了孕。
b. 他又活了，但地風沒治服了他。
- (22) a. 他丟了官與錢財，但是還沒丟失了自信與希望。（《倫生·十九》）
b. 項羽後悔當初沒殺了劉邦。
c. 他可沒爛了心肝肺。
d. 這回總算沒賠了本。
e. 我可沒得罪了誰。
f. 說真格的，二爺可沒虧待了你。
g. 並沒有為了這事掉了腦袋。
h. 沒忘了教育他的孩子。
i. 祥子自己可並沒隨便忘了這件事。
j. 在家沒住了幾天，我又到外邊去了兩個月。
k. 我家自管蠶花不好，可並沒害了誰。（矛盾《春蠶》）
l. 但還沒有磨了三升，就向她叫。（柔石《為奴隸的母親》）
m. 但他並沒有放棄了他的專業。（徐遲《哥德巴赫猜想》）
n. 肉切得好不好的，總算切出來了，沒耽誤了事。

呂叔湘主編 (1980) 的《現代漢語八百詞》也指出漢語有一類動詞如“忘、丟、關、喝、吃、咽、吞、潑、酒、扔、放、塗、抹、擦、碰、砸、摔、磕、撞、踩、傷、殺、宰、切、沖、賣、還、毀”後面的“了_i”表示動作有了結果，相當於補語“掉”。還指出“用‘沒’的否定句裡的動詞後面一般不能再有‘了_i’，但與‘掉’相似的‘了_i’可以有”，如下面 (23) 中“扔了”的“了 (le)”就不是動詞義很明顯的“了 (liǎo)”。比較：

(23) 幸虧沒扔了它，今天又用上了。

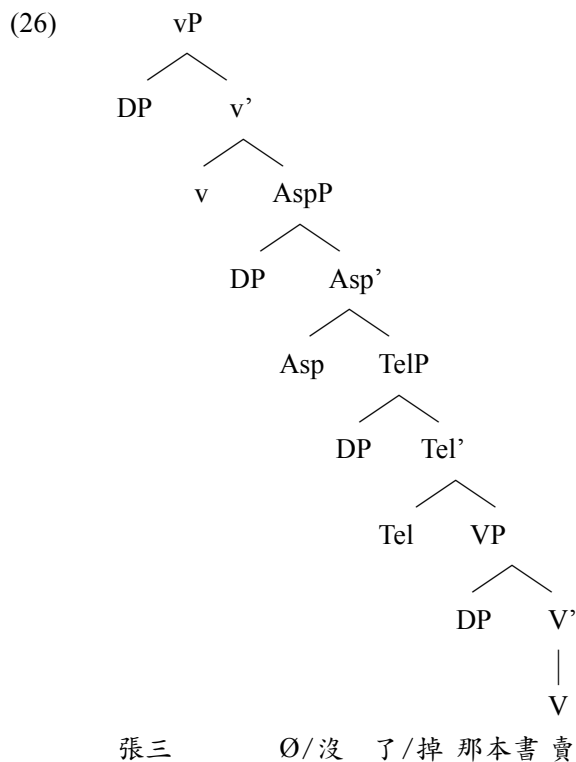
所以我們認為之所以會造成前述各家對漢語中“了”的語法意義的爭議，其根本原因就在於動詞後有同形的兩個“了(le)”：其中一個確實是完成體標記。而另一個是具有完結動作功能的結點標記，只是動作的完結點，並不標誌動作的完成實現。能與完成體否定形式“沒”共現的是後一種表示完結的“了(le)”。表示完結的“了(le)”是同樣可以表示完結的動詞“了(liǎo)”的虛化形式，在語音上更加弱化，但是在句法位置上還是“完結短語 TelP”的中心語；表示完結的“了(le)”也是完成體標記“了(le)”的前身，因為它在語音的進一步虛化，使得它在使用中表意更加不清晰，也促使新的完結短語的使用出現，並最終變成同形的體標記。

按照這種分析，漢語中有的句子中的“了”是完成體標記，是體短語中心語，其句法結構如(24)所示；因此可以形成(25)這樣的句子。例如：



- (25) a. 張三買了那本書。 b. 張三沒買那本書。
c. *張三沒買了(le)那本書。

而漢語中還有一些句子中“了”是謂語動詞的完結點，是“完結短語”中心語，其句法結構如 (26) 所示；因此才可以形成 (27) 這樣的句子。例如：



- (27) a. 張三賣了那本書。
b. 張三沒賣了/掉那本書。

現代漢語中兩個“了 (le)”的存在也是在共時層面證明了我們對完成體標記“了 (le)”的語法化過程的分析。

5. 餘論

本文不僅是在生成語法理論框架內分析漢語完成體標記的語法化過程，同時也驗證了“完結短語”理論假設的一個優點。其他動結式的研究往往忽略漢語的體貌，也不考慮完成體標記與動結式的補語的淵源關係。“小句理論”雖然將體

標記“了”放在了補語小句中，力圖找出補語與體標記的關係，但是作為功能成分的體標記在主句動詞之下也不符合生成語法的一般原則；而且在生成過程中兩次與謂語動詞有關的移動並不一致，理論上顯得不夠統一。⁷ 本文的理論假設可以順理成章地解釋補語與體標記的關係：虛化結果補語作為主句動詞的完結功能成分，投射在動詞的上方，而真正的體具有的功能性更強，位置也更高。從句法功能上看，完結短語和體短語都是給動詞一個終結點，補語所形成的完結短語就是一種內部體，從內部體向真正的體貌標記發生語法化的轉變是普遍的語法化途徑；從句法位置上看，二者都是在 *vP* 和 *VP* 之間，十分接近，存在語法化的可能性，體短語比完結短語句法位置略高正表明它的語法化程度高。而且，共時平面的句子結構如此，歷時平面的句法結構演變也如此。二者同為功能性成分，但強弱有別，漢語的虛化結果補語最終演變為句法位置臨近的體標記具有很强的理論可行性。

⁷ 小句理論解釋的生成過程如下：a [IP[*vP* 張三 [*v'*[*VP*[*v'* 打 [*AspP* 了 [*SC* 李四 傷]]]]]] → b [IP[*vP* 張三 [*v'*[*VP*[*v'* 打 [*AspP* 傷 *i*+了 [*SC* 李四 *t_i]]]]]]] → c [IP[*vP* 張三 [*v'*[*VP*[*v'* 打+(傷 *i*+了)_{*j*} [*AspP* *t_j* [*SC* 李四 *t_i]]]]]]]。兩次移動的性質相同，都是核心詞移動 (head movement)；目的也相同，都是句中所有與謂語有關的特徵或成分要形成一個鏈，連接在一起。但是兩次移動的方式卻不同，b 中補語上升到體標記的左側，形成“傷了”；而 c 中這個組合“傷了”移到主句動詞的“打”的右側。這種為了表層線性化的語序而設定的不同的合併策略就有特設之嫌。這個生成過程的分析是小句理論的一個明顯不足之處。**

引用文獻

- Cao, Guangshun (曹廣順). 1986. Zutangji zhong de 'di (di)', 'que (le)', 'zhe' 《祖堂集》中的“底(地)”、“卻(了)”、“著” [The particles *di (di)*, *que (le)*, *zhe* in *Zutangji*]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1986.3: 192-202.
- Cao, Guangshun (曹廣順). 1995. *Jindai Hanyu Zhuci* 近代漢語助詞 [*Function Words in Early Mandarin Chinese*]. Beijing: Language and Culture Press.
- Cao, Guangshun (曹廣順). 2000. Shi lun Hanyu dongtai zhuci de xingcheng guocheng 試論漢語動態助詞的形成過程 [Auxiliary words in Pre-Modern Chinese]. *Hanyushi Yanjiu Jikan* 漢語史研究集刊 [*Studies on the History of Chinese Language*] 2:74-89. Chengdu: Bashu Shushe.
- Chao, Yuen Ren (趙元任). 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chao, Yuen Ren (趙元任). 1980. *Zhongguohua de Wenfa* 中國話的文法 [*A Grammar of Spoken Chinese*], translated by Pang-Hsin Ting. Hong Kong: The Chinese University Press.
- Chen, Gang (陳剛). 1981. Tan “mei dong le bin/bu” shi 談“沒動了賓/補”式 [On the construction of *mei* verb object/complement]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1981.1:45.
- Chen, Gang (陳剛). 1985. Guanyu “mei V le₁” shi 關於“沒 V 了₁”式 [On the construction of *mei*+V+*le*₁]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1985.5:329-333.
- Chen, Ping (陳平). 1988. Lun Xiandai Hanyu shijian xitong de sanyuan jieou 論現代漢語時間系統的三元結構 [On the tripartite structure of the temporal system of Modern Chinese]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1988.6:401.
- Cheng, Robert L. (鄭良偉). 1988. Shiti, dongliang he dongci chongdie 時體、動量和動詞重疊. *Shijie Hanyu Jiaoxue* 世界漢語教學 [*Chinese Teaching in the World*] 1988.2:73-80.
- Cheung, Hung-nin Samuel (張洪年). 1972. *Xianggang Yueyu Yufa de Yanjiu* 香港粵語語法的研究 [*A Grammar of Cantonese as Spoken in Hong Kong*]. Hong Kong: The Chinese University Press.
- Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Dai, Yaojing (戴耀晶). 1990. *Lun Xiandai Hanyu de Ti* 論現代漢語的體 [On *Aspect in Modern Chinese*]. Shanghai: Fudan University dissertation.

- Dai, Yaojing (戴耀晶). 1997. *Xiandai Hanyu Shi Ti Xitong Yanjiu* 現代漢語時體系統研究 [Research on Tense and Aspect Systems of Modern Chinese]. Hangzhou: Zhejiang Education Press.
- Fan, Fanglian (范方蓮). 1963. Cunzaiju 存在句 [Existential sentences]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1963.5:386-395.
- Fang, Yuqing (房玉清). 1992. Dongtai zhuci *le, zhe, guo* de yuyi tezheng ji qi yongfa bijiao 動態助詞“了、著、過”的語義特徵及其用法比較 [A comparison of the use and functions of the dynamic auxiliaries *le, zhe, guo*]. *Hanyu Xuexi* 漢語學習 [Chinese Language Learning] 1992.1:14-20.
- Feng, Chuntian (馮春田). 1988. *Jindai Hanyu Yufa Yanjiu* 近代漢語語法研究 [Research on Early Modern Chinese Grammar]. Ji'nan: Shandong Education Press.
- Gao, Mingkai (高名凱), and Anshi Shi (石安石). 1979. *Yuyanxue Gailun* 語言學概論 [Introduction to Linguistics]. Beijing: Zhonghua Book Company.
- Guo, Rui (郭銳). 1997. Guocheng he fei guocheng: Hanyu weicixing chengfen de liang zhong waizai shijian leixing 過程和非過程：漢語謂詞性成分的兩種外在時間類型 [Process and non-process: two external time types of Chinese predicate constituent]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1997.3:162-175.
- Hoekstra, Teun. 1988. Small clause results. *Lingua* 74.2-3:101-139.
- Jiang, Jicheng (蔣冀騁), and Fuxiang Wu (吳福祥). 1997. *Jindai Hanyu Gangyao* 近代漢語綱要 [Outline of Early Modern Chinese Grammar]. Changsha: Hunan Education Press.
- Jiang, Shaoyu (蔣紹愚). 1994. *Jindai Hanyu Yanjiu Gaikuang* 近代漢語研究概況 [A Survey of Modern Chinese Studies]. Beijing: Peking University Press.
- Jiang, Shaoyu (蔣紹愚). 1999. Hanyu dongjieshi chansheng de shidai 漢語動結式產生的時代 [Time of formation of Chinese VCR]. *Guoxue Yanjiu* 國學研究 [Studies in Sinology] 6:327-348. Beijing: Peking University Press.
- Jiang, Shaoyu (蔣紹愚). 2001. Shishuo Xinyu, Qimin Yaoshu, Luoyan Qielan Ji, Xianyu Jing, Baiyu Jing zhong de 'yi', 'jing', 'qi', 'bi' 《世說新語》、《齊民要術》、《洛陽伽藍記》、《賢愚經》、《百喻經》中的“已”、“竟”、“訖”、“畢” [Yi, jing, qi and bi in the Shishuo Xinyu, Qimin Yaoshu, Luoyan Qielan Ji, Xianyu Jing and the Baiyu Jing]. *Yuyan Yanjiu* 語言研究 [Studies in Language and Linguistics] 2001.1:73-78.
- Jin, Lixin (金立鑫). 1998. Shi lun 'le' de shi ti tezheng 試論“了”的時體特徵 [On the tense and aspect features of *le*]. *Yuyan Jiaoxue yu Yanjiu* 語言教學與研究 [Language Teaching and Linguistic Studies] 1998.1:105-120.
- Jin, Lixin (金立鑫). 2002. Ciwei 'le' de shi ti yiyi ji qi jufa tiaojian 詞尾“了”的時體意義及其句法條件 [The aspectual significance of the suffix *le* and its syntactic

- conditions]. *Shijie Hanyu Jiaoxue* 世界漢語教學 [*Chinese Teaching in the World*] 2002.1:34-43.
- Jing, Cheng (竟成). 1993. Guanyu dongtai zhuci 'le' de yufa yiyi wenti 關於動態助詞“了”的語法意義問題 [On the grammatical meaning of the dynamic auxiliary *le*]. *Yuwen Yanjiu* 語文研究 [*Linguistic Researches*] 1993.1:52-57.
- Kimura, Hideki (木村英樹). 1983. Guanyu buyuxing ciwei 'zhe' he 'le' 關於補語性詞尾“著”和“了” [On the complementary verb *zhe* and *le*]. *Yuwen Yanjiu* 語文研究 [*Linguistic Researches*] 1983.2:24-32.
- Li, Xiaofan (李小凡). 2000. Xiandai Hanyu ciwei 'le' de yufa yiyi zai tantao 現代漢語詞尾“了”的語法意義再探討 [Discussion of the grammatical meaning of the Modern Chinese verb *le*]. *Yufa Yanjiu he Tansuo* 語法研究和探索 [*Grammar Studies and Exploration*] 10:203-209. Beijing: The Commercial Press.
- Liao, Mingchun (廖名春). 1990. Tulufan chutu wenshu yuci guankui 吐魯番出土文書語詞管窺 [A study of the Turpan unearthed documents]. *Gu Hanyu Yanjiu* 古漢語研究 [*Research in Ancient Chinese Language*] 1990.1:21-27.
- Liu, Jian (劉堅), Lansheng Jiang (江藍生), Weiguo Bai (白維國), and Guangshun Cao (曹廣順). 1992. *Jindai Hanyu Xuci Yanjiu* 近代漢語虛詞研究 [*A Study on the Function Words of Modern Chinese Language*]. Beijing: Language and Culture Press.
- Liu, Shizhen (柳士鎮). 1992. *Wei Jin Nanbeichao Lishi Yufa* 魏晉南北朝歷史語法 [*The Historical Grammar in the Wei, Jin, and Southern and Northern Dynasties*]. Nanjing: Nanjing University Press.
- Liu, Xunning (劉勛寧). 1988. Xiandai Hanyu ciwei 'le' de yufa yiyi 現代漢語詞尾“了”的語法意義 [The grammatical meaning of *le* in Modern Chinese]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1988.5:321-330.
- Liu, Yuehua (劉月華). 1988. Dongtai zhuci 'guo₂', 'guo₁', 'le₁' yongfa bijiao 動態助詞“過₂”、“過₁”、“了₁”用法比較 [A comparison of the use of the dynamic auxiliaries *guo₂*, *guo₁*, *le₁*]. *Yuwen Yanjiu* 語文研究 [*Linguistic Researches*] 1988.1:6-16.
- Lü, Shuxiang (呂叔湘). 1942. *Zhongguo Wenfa Yaolue* 中國文法要略 [*An Outline of Chinese Grammar*]. Beijing: The Commercial Press.
- Lü, Shuxiang (呂叔湘). (ed.) 1980. *Xiandai Hanyu Baibai Ci* 現代漢語八百詞 [*Eight Hundred Words in Modern Chinese*]. Beijing: The Commercial Press.
- Lu, Yingshun (盧英順). 1994. Guanyu 'le₁' shiyong qingkuang de kaocha 關於“了₁”使用情況的考察 [An investigation on the usage of *le₁*]. *Journal of Anhui Normal University (Philosophy & Social Sciences)* 1994.2:194-200.
- Ma, Xiwen (馬希文). 1983. Guanyu dongci 'le' de ruohua xingshi /lou/ 關於動詞“了”的弱化形式 /lou/ [On the weakened form *lou* of the verb *le*]. *Zhongguo*

- Yuyan Xuebao* 中國語言學報 [*Chinese Linguistics: The Journal of Linguistic Society of China*] 1:1-14.
- Mao, Jingxiu (毛敬修). 1985. Guanyu 'V(C)le' zhong de 'le' 關於“V(C)了”中的“了” [On the *le* in *V(C)le*]. *Tianjin Normal University Journal* 1985.1:78-86.
- Mei, Tsu-Lin (梅祖麟). 1981. Xiandai Hanyu wanchengmao jushi he ciwei de lai yuan 現代漢語完成貌句式和詞尾的來源 [The origin of the perfective aspect construction and the perfective suffix in Modern Chinese]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 1981.1:65-77.
- Mei, Tsu-Lin (梅祖麟). 1988. Hanyu fangyan li xuci 'zhe' zi sanzong yongfa de lai yuan 漢語方言裡虛詞“著”字三種用法的來源 [The source of three usages of the functional word *zhe* in Chinese dialects]. *Zhongguo Yuyan Xuebao* 中國語言學報 [*Chinese Linguistics: The Journal of Linguistic Society of China*] 3:193-216.
- Mei, Tsu-Lin (梅祖麟). 1994. Tangdai, Songdai gongtongyu de yufa he xiandai fangyan de yufa 唐代、宋代共同語的語法和現代方言的語法 [The grammar of the common language in Tang and Song Dynasties and modern dialect grammar]. *Chinese Languages and Linguistics*, Vol. 2: *Historical Linguistics*, ed. by Paul Jen-kuei Li, Chu-Ren Huang & Chih-Chen Jane Tang, 61-97. Taipei: Institute of History and Philology, Academia Sinica.
- Ota, Tatsuo (太田辰夫). 1987. *Zhongguoyu Lishi Wenfa* 中國語歷史文法 [*A Historical Grammar of Modern Chinese*], translated by Shaoyu Jiang & Changhua Xu. Beijing: Peking University Press.
- Pan, Wenguo (潘文國). 2003. Cong 'le' de Ying Han kan Hanyu de shi ti wenti 從“了”的英漢看漢語的時體問題 [Tense and Aspect in Chinese: a case study of the word *le* and its English translations]. *Journal of East China Normal University (Philosophy and Social Sciences)* 2003.4:62-69.
- Pan, Yunzhong (潘允中). 1982. *Hanyu Yufashi Gaiyao* 漢語語法史概要 [*A Summary of Chinese Grammar History*]. Zhengzhou: Zhongzhou Ancient Books Publishing House.
- Qi, Xiaojie (戚曉杰). 1995. Tan Hanyu de wanchengti ji qi biao xian xingshi 談漢語的完成體及其表現形式 [On Chinese perfective and its form]. *Dongfang Luntan* 東方論壇 [*The Orient Forum*] 1995.3:87-91.
- Rohsenow, John. 1978. Perfect *le*: aspect and relative tense in Mandarin Chinese. *Proceedings of Symposium on Chinese Linguistics*, ed. by Robert L. Cheng, Yingche Li & Ting-chi Tang, 267-291. Taipei: Student Book.
- Shi, Qin (石鏡). 2000. Qian tan zhuci 'le' yufahua guocheng zhong de jige wenti 淺談助詞“了”語法化過程中的幾個問題 [Several questions about the grammaticalization of the auxiliary *le*]. *Hanyushi Yanjiu Jikan* 漢語史研究集刊 [*Studies on the History of Chinese Language*] 2:90-105. Chengdu: Bashu Shushe.

- Shi, Youwei (史有爲). 2003. Hanyu fangyan “dacheng” lei qingmao de leixingxue kaocha 漢語方言“達成”類情貌的類型學考察 [A typological investigation on the pan-aspect of accomplishment in Chinese dialects]. *Yuyan Yanjiu* 語言研究 [Studies in Language and Linguistics] 2003.3:91-103.
- Shimura, Ryoji (志村良治). 1984. *Zhongguo Zhongshi Yufashi Yanjiu* 中國中世語法史研究 [Studies on the History of Middle Chinese Grammar], translated by Lansheng Jiang & Weiguo Bai. Beijing: Zhonghua Book Company.
- Song, Yuzhu (宋玉柱). 1980. Hanyu li you “mei qu le jici” zhe zhong jiegou ma? 漢語裡有“沒去了幾次”這種結構嗎? [Is there *mei qu le jici* in Chinese?]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1980.6:478.
- Sun, Chaofen (孫朝奮). 1997. Zai lun zhuci ‘zhe’ de yongfa ji qi lai yuan 再論助詞“著”的用法及其來源 [The usage and source of the auxiliary *zhe*]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1997.2:139-146.
- Sun, Xixin (孫錫信). 1992. *Hanyu Lishi Yufa Yaolue* 漢語歷史語法要略 [A Summary of Chinese Historical Grammar]. Shanghai: Fudan University Press.
- Sybesma, Rint. 1992. *Causatives and Accomplishments: The Case of Chinese ba*. Leiden: Leiden University dissertation.
- Sybesma, Rint. 1997. Why Chinese verb-*le* is a resultative predicate. *Journal of East Asian Linguistics* 6.3:215-261.
- Sybesma, Rint. 1999. *The Mandarin VP*. Dordrecht: Kluwer.
- Sybesma, Rint (司馬翎), and Yang Shen (沈陽). 2006. Jieguo buyu xiaoju fenxi he xiaoju de neibu jiegou 結果補語小句分析和小句的內部結構 [Small clause results and the internal structure of the small clause]. *Journal of Huazhong University of Science and Technology (Social Science Edition)* 2006.4:40-46.
- Travis, Lisa. 2010. *Inner Aspect: The Articulation of VP*. Dordrecht: Springer.
- Wang, Huan (王還). 1990. Zai tan Xiandai Hanyu ciwei ‘le’ de yufa yi yi 再談現代漢語詞尾“了”的語法意義 [The grammatical meaning of the Modern Chinese verb *le*]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1990.3:180-181.
- Wang, Li (王力). 1944. *Zhongguo Yufa Lilun* 中國語法理論 [The Theory of Chinese Grammar]. Beijing: The Commercial Press.
- Wang, Li (王力). 1958[1980]. *Hanyu Shi Gao* 漢語史稿 [An Outline of Chinese Language History]. Beijing: Zhonghua Book Company.
- Wu, Fuxiang (吳福祥). 1996. *Dunhuang Bianwen Yufa Yanjiu* 敦煌變文語法研究 [Studies on the Grammar of the Dunhuang Bianwen]. Changsha: Yuelu Shushe.
- Wu, Fuxiang (吳福祥). 1998. Chong tan ‘dong+le+bin’ geshi de lai yuan he wanchengti zhuci ‘le’ de chansheng 重談“動+了+賓”格式的來源和完成體助詞“了”的產生 [The source of the construction *V+le+O* and the formation of perfective

- auxiliary *le*]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1998.6:452-462.
- Wu, Fuxiang (吳福祥). 2001-2002. Nanfang fangyan jige zhuangtai buyu biaoji de lai yuan 南方方言幾個狀態補語標記的來源 [Origin of some markers of stative complements in the southern Chinese dialects]. *Fangyan* 方言 [*Dialect*] 2001.4: 344-354; 2002.1:24-34.
- Wu, Fuxiang (吳福祥). 2003. Guanyu yufahua de danxiangxing wenti 關於語法化的單向性問題 [On the unidirectionality of grammaticalization]. *Dangdai Yuyanxue* 當代語言學 [*Contemporary Linguistics*] 2003.4:307-322.
- Xuan, Yue (玄玥). 2008. *Wanjie Duanyu Jiashe he Hanyu Xuhua Jieguo Buyu Yanjiu: Jian Lun Hanyu Jieguo Buyu, Ti Biaoji he Quxiang Buyu de Jufa Wenti* 完結短語假設和漢語虛化結果補語研究：兼論漢語結果補語、體標記和趨向補語的句法問題 [On the Telic Phrase Hypothesis and Chinese Weak Resultative Complements: With the Syntax of Resultative Complements, Aspect Markers and Directional Complements in Chinese]. Beijing: Peking University dissertation.
- Yang, Hsiu-fang (楊秀芳). 1992. Cong lishi yufa de guandian lun Minnanyu 'zhe' ji chixumao 從歷史語法的觀點論閩南語“著”及持續貌 [*Zhe* and the continuing Aspect of Southern Min from the perspective of historical grammar]. *Chinese Studies* 10.1:349-394.
- Yu, Genyuan (于根元). 1983. Guanyu dongci houfu 'zhe' de shiyong 關於動詞後附“著”的使用 [On the usage of the verb *zhe*]. *Yufa Yanjiu he Tansuo* 語法研究和探索 [*Grammar Studies and Exploration*] 1:106-119. Beijing: Peking University Press.
- Zhang, Cheng (張楨). 2000. Wei Jin Nanbeichao shiqi 'zhuo' zi de yongfa 魏晉南北朝時期“著”字的用法 [The usages of “*zhuo*” in Chinese in the period of Wei and Jin Kingdoms and Southern and Northern Dynasties]. *Zhongwen Xuekan* 中文學刊 [*The Chinese Academic Journal*] 2:121-138.
- Zhang, Jiahua (張家驊). 2004. Tongguo Han E duibi kan 'le₁' de changti yiyi 通過漢俄對比看“了₁”的常體意義 [The context-independent perfective meaning of the Chinese *le₁*: a contrastive study with its Russian counterpart]. *Dangdai Yuyanxue* 當代語言學 [*Contemporary Linguistics*] 2004.2:97-109.
- Zhao, Jinming (趙金銘). 1979. Dunhuang Bianwen zhong suojian de 'le' he 'zhe' 敦煌變文中所見的“了”和“著” [*Le* and *zhe* in the Dunhuang Bianwen]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1979.1:65-69.
- Zhu, Dexi (朱德熙). 1982. *Yufa Jiangyi* 語法講義 [*Lectures on Grammar*]. Beijing: The Commercial Press.

Zuo, Simin (左思民). 1997. *Xiandai Hanyu Ti de Zai Renshi* 現代漢語體的再認識 [Further Recognition on Aspect of Contemporary Chinese]. Shanghai: Shanghai Normal University dissertation.

The Grammaticalization of Chinese Resultative Complement as Telic Phrase and the Development of the Perfective Aspect Marker

Yang Shen¹ and Yue Xuan²

Peking University¹

Beijing Language and Culture University²

Telic Phrase Hypothesis is a theory of Chinese weak resultative complements in Generative Grammar. The weak resultative complement projects a *Telic Phrase*, an inner aspect of the main verb in syntax. The grammaticalization of *le* as a perfective aspect marker from a resultative complement to a weak resultative complement can be explained according to this hypothesis, which also supports the fact that there are two different verb *le* in syntax.

Key words: Telic Phrase Hypothesis, resultative complements, grammaticalization, perfective aspect markers

Syntactic Change in Chinese and the Argument-Adjunct Asymmetry

Redouane Djamouri (羅 端)¹

Waltraud Paul (包華莉)¹

John Whitman (魏德中)²

CNRS-EHESS-INALCO¹

Cornell University²

The word order in Chinese has always been SVO, from the earliest attested documents (14th-11th c. BC) up to Modern Mandarin. Examined carefully, the observed SOV cases in pre-Archaic Chinese turn out to either involve focalization of the object or object pronouns in the context of negation. Importantly, both structures observe head-complement order, i.e. a pattern consistent with VO. This removes any coherent basis for the claim that pre-Archaic Chinese was a SOV language. Against this background of stable VO order, important changes can, however, be observed for the distribution of adjunct phrases, from both pre- and postverbal position in pre-Archaic Chinese to exclusively preverbal position in Modern Mandarin, reflecting changes in the format of the *v*P.

Key words: word order SVO vs. SOV, argument vs. adjunct, cleft construction, pre-Archaic Chinese, Shang inscriptions, Modern Mandarin

1. Introduction

This paper argues for what historical linguists have long called the *uniformitarian principle*: the postulate that the same principles of analysis that apply to synchronic grammars also apply to earlier stages of a language. To apply distinct methodologies is simply incompatible with the fact that each (past or present) synchronic stage represents a stable system which can be acquired by a learner and hence must be consistent with the universal constraints observed for language in general (for an extensive discussion of this issue, cf. Hale 1998, 2007).

Unfortunately, many diachronic studies on Chinese take the linear sequence at face value and do not offer a structural analysis. However, as pointed out by e.g. Hale (2007: 5) “It is not possible in any meaningful sense to know what ‘changed’ between Stage I and Stage II of some ‘language’ without knowing what Stage I and Stage II were, as synchronic systems.”

A good case at hand is the issue of word order in Chinese. According to Li & Thompson (1974:208) pre-Archaic Chinese¹ (< 12th c. BC) was an SOV language, which changed to SVO between the 10th and the 3rd c. BCE, before starting to shifting back to SOV, a change purported to be still incomplete in Modern Mandarin.

Besides the fact that Li & Thompson did not take into account the available data for pre-Archaic Chinese, their incorrect (but still influential) statement concerning major word order changes attains superficial plausibility only because they (as well as subsequent linguists taking up their claim) at no point provide an analysis of the alleged SOV cases.

Examined carefully, all of the observed SOV cases in pre-Archaic Chinese turn out to either involve focalization of the object or object pronouns in the context of negation. Importantly, the relevant focus pattern in pre-Archaic Chinese was restricted to a type of cleft construction, where the focused constituent follows an item that functions as a matrix copular predicate. Needless to say this pattern instantiates VO order. Likewise, under an analysis where the object pronoun occupies the specifier of a functional projection the examples illustrating an at first sight preverbal object position also show a head-complement structure. This removes any coherent basis for the claim that Chinese was predominantly SOV before the 11th c. BC.

While Chinese has always been SVO, from the earliest textual sources, the Shang inscriptions (14th-11th c. BC), up to Modern Mandarin, there have been important changes in the distribution of adjunct phrases, from both pre- and postverbal position in pre-Archaic Chinese to exclusively preverbal position in Modern Mandarin, reflecting changes in the format of the *v*P (cf. Djamouri & Paul 1997, 2009).

2. VO word order in the Shang inscriptions (14th-11th c. BC)

Formal studies on Chinese historical syntax rarely include the earliest sources from the pre-Archaic Chinese period, i.e. the Shang inscriptions (14th-11th c. BC), although the latter constitute the very basis for examining the subsequent development of Chinese

¹ Current Western terminology for the periods of older Chinese is confusing for the neophyte. Karlgren (1923) uses the term “Archaic Chinese” to refer to the language of the early and middle Zhou period (11th c. - 221 BC), as it is reconstructed on the basis of (i) the rhymes in the *Shi Jing* (*Book of Odes*) (roughly 800-600 BC), (ii) the phonetic series revealed by the Chinese script, and (iii) information available from Middle Chinese (6th and early 7th c. AD). More recently, in historical phonology, the term “Old Chinese” has replaced the term “Archaic Chinese” (cf. Baxter 1992, Sagart 1999:4). Following the periodicization established by Peyraube (1988), “Pre-Archaic Chinese” in this paper refers to the language of the Shang bone inscriptions (14th-11th c. BC), which antedates Old Chinese and Middle Chinese. Note that Chinese is thus a language whose syntax is recoverable at an earlier stage than its phonology.

syntax. Of the 26,000 complete sentences in the Shang corpus, 94% have SVO order, and only 6% SOV (cf. Chen Mengjia 1956, Djamouri 1988, Shen Pei 1992 a.o.).

2.1 The postverbal position as canonical position for arguments

The argument(s) subcategorized for by a verb occupy the postverbal position. This holds both for argument NPs ((1)-(4)) and PPs ((5)-(9)). Accordingly, both the direct and the indirect object follow the verb in the double object construction (cf. (3), (4), (6)).

- (1) 王伐舌方
wáng fá [_{NP} gōng fāng]² (Heji 6223)
king fight Gong tribe
'The king will fight the Gong tribe.'
- (2) ...王𪔐麋...
wáng jǐng mí (Heji 10361)
king trap elk
'The king will trap elks.'
- (3) 帝受我年
dì shòu [_{IO} wǒ] [_{DO} nián] (Heji 09731 recto)
Di give us harvest
'[The ancestor] Di will give us a harvest.'
- (4) 𪔐祖乙三宰
yòu zǔyǐ sān lǎo (Heji 01610)
present Zuyi three penned:sheep
'One will present to Zuyi three penned sheep (as sacrifice).'
- (5) 王往于田
wáng wǎng [_{PP} yú tián]
king go to field
'The king will go to the fields.'
- (6) 𪔐于祖乙一牛
yòu [_{PP} yú zǔyǐ] [_{NP} yī niú] (Heji 06945)
present to Zuyi one ox
'One will present to Zuyi an ox (as sacrifice).'

² The following abbreviations are used in glossing examples: CL classifier; DUR durative aspect; EXP experiential aspect; NEG negation; PART sentence-final particle; PERF perfective aspect; PL plural (e.g. 3PL = 3rd person plural); SG singular; SUB subordinator.

- (7) 子商亡斷在囿 (Heji 02940)
 zǐ shāng wáng duàn [pp zài huò]
 prince Shang NEG end in misfortune
 ‘The prince Shang will not end in misfortune.’
- (8) 方允其來于沚 (Heji 6728)
 fāng yǔn qí lái [pp yú zhǐ]
 Fang effectively FUT come to Zhi
 ‘Fang will effectively come to Zhi.’
- (9) 我乎往于西 (Heji 10050)
 wǒ hū wǎng [pp yú xī]
 IPR order go to west
 ‘We will order to go west.’

The rich corpus of data from Pre-Archaic Chinese clearly invalidates Li & Thompson’s (1974:208) claim that ‘pre-Archaic Chinese’ (< 12th c. BC) was an SOV language, which changed to SVO between the 10th and the 3rd c. BC. As for the subsequent alleged “gradual” shift “back” to SOV, a change purported to be not completed yet, i.e. 2000 years after it started, it does not bear further scrutiny, either. It is simply not correct to view Modern Mandarin as still in the process of “becoming” “more and more” SOV (cf. §4 below). Chinese is and has always been VO, and there is thus no basis whatsoever for a “cyclic change” ‘OV > VO > OV’.

Concerning the (surface) SOV cases in Pre-Archaic Chinese, they can be divided into two classes, i.e. focalization of the object, on the one hand, and object pronouns in the context of negation, on the other. Crucially, both turn out to involve head-complement configurations consistent with VO. For reasons of space, we will limit ourselves to focalization of the object. (For a detailed discussion of the structure ‘Neg pronoun V’, cf. Djamouri 2000).

2.2 Focalization structures

Complete sets of predictions in the Shang inscriptions such as (10)-(11) permit us to identify superficial OV structures as clear cases of focalization. (10) presents a prediction in the form of a simple assertion displaying VO order. Against this background, two alternatives, (11a-b), are proposed. In these alternatives, ‘follow someone (in order to fight Xia Wei)’ presents the presupposition, and the object of the verb *bǐ* ‘follow’ the focus:

- (10) 王比望乘伐下危 (Heji 6476)
wáng bǐ wáng chéng fā xià wēi
king follow Wang Cheng fight Xia Wei
‘The king will follow Wang Cheng to fight Xia Wei.’
- (11) a. 王勿唯望乘比 (Heji 6476)
wáng wù wéi [_{NP} wáng chéng] bǐ
king NEG be Wang Cheng follow
‘It must not be Wang Cheng that the king will follow.’
- b. 王亟望乘比 (Heji 6476)
wáng huì [_{NP} wáng chéng] bǐ
king must:be Wang Cheng follow
‘It must be Wang Cheng that the king will follow.’

All of the attested examples where an argument NP or PP occupies a (surface) preverbal position involve focalization. Importantly, the relevant focus pattern in pre-Archaic Chinese was restricted to a type of cleft construction, as in modern Mandarin *shi...de* clefts (cf. Paul & Whitman 2008). The cleft structure is clear in (11a-b), where the focused constituent *Wáng Chéng* follows the negated matrix copula *wù wéi* ‘NEG be’ in (11a) and the modal matrix copula *huì* ‘must be’ in (11b). On the cleft analysis, the focused constituent is postverbal, because to the right of the copula: it occupies the specifier position of the projection selected as complement by the copula. Accordingly, this construction illustrates head-complement, not complement-head order. The same facts are exemplified in the discourses in (12)-(15), where the same matrix copula elements *huì* and (*wù*) *wéi* are attested.

- (12) a. 王亟易白嫫比 (Heji 6460 r.)
wáng [_{VP} huì [[_{NP} yáng bó shǐ]_i [_{VP} bǐ t_i]]]
king must:be Yang lord Shi follow
‘It must be Shi, lord of Yang, that the king will follow.’
- b. 王勿唯易白嫫比 (Heji 6460 r.)
wáng [_{NegP} wù [_{VP} wéi [[_{NP} yáng bó shǐ]_i [_{VP} bǐ t_i]]]]
king NEG be Yang lord Shi follow
‘It must not be Shi, lord of Yang, that the king will follow.’
- (13) a. 王勿唯龍方伐 (Heji 6476)
wáng [_{NegP} wù [_{VP} wéi [[_{NP} lóng fāng]_i [_{VP} fā t_i]]]]
king NEG be Long tribe fight
‘It must not be the Long tribe that the king will fight.’

- b. 王蚩龍方伐 (Heji 6476)
 wáng [vP huì [[NP lóng fāng]_i [vP fǎ t_i]]]
 king must:be Long tribe fight
 ‘It must be the Long tribe that the king will fight.’

- (14) 蚩羊侑于母丙 (Heji 2523)
 [vP huì [[NP yáng] [vP yòu [PP yú mǔ bǐng]]]]
 must:be sheep offer to ancestress Bing
 ‘It must be a sheep that one will offer to Ancestress Bing.’

- (15) 唯祖乙侑仁 (Heji 1573)
 [vP wéi [[NP zǔyǐ] [vP yòu pǒ]]]
 be Zuyi offer po.sacrifice
 ‘It is to Zuyi that one will offer a *po* sacrifice.’

To summarize, the preceding discussion has shown the importance of a precise syntactic analysis of the synchronic stage at hand. The surface ‘O V’ sequence in focalization structures turns out to involve head-complement order in accordance with the main word order ‘VO’. The fact that argument PPs pattern with argument NPs and occur in the postverbal position (cf. §2.1) further corroborates the head-initial property of clause structure in pre-Archaic Chinese.

3. Distribution of adjunct phrases

Unlike arguments, adjuncts (both PPs and NPs) in pre-Archaic Chinese can appear in three positions: preceding the subject, between the subject and the verb or postverbally (after the object when present) (cf. §3.1-§3.3. below). Note here that it is not sufficient to correlate the position of PPs *per se* with VO vs. OV word order (‘PP V’ with OV and ‘V PP’ with VO, cf. Dryer 2003:48-49). The argumental vs. non argumental status of the PPs needs to be taken into account as illustrated in (16). The argument PP *yú shāng* ‘in(to) Shang’ subcategorized for by the verb *rù* ‘enter’ must occupy the postverbal position, whereas the adjunct PP *yú qī yuè* ‘in the seventh month’ precedes the verb.

- (16) 王于七月入于商 (Heji 7780 r.)
 wáng [vP [PP yú qī-yuè] [vP rù [PP yú shāng]]]
 king in seven-month enter in Shang
 ‘The king in the seventh month will enter the Shang city.’

Non-phrasal adverbs such as *yì* ‘also’, *yǐn* ‘indeed’ are confined to the preverbal position below the subject and excluded from postverbal position:

- (17) 五月癸巳雨乙巳亦雨 (Heji 20943)
 [wǔ yuè guǐsì] yǔ yǐsì yì [v_P yǔ]
 five month *guisi* rain *yisi* also rain
 ‘On the day *guisi* of the fifth month, it rained; on the day *yisi*, it also rained.’
- (18) 出伐于黃尹亦出于蔑 (Heji 00970)
 yòu fá [p_P yú Huángyǐn] yì [v_P yòu [p_P yú Miè]]
 offer victim to Huangyin also offer to Mie
 ‘We will offer victims (as sacrifice) to Huanyin, and also to Mie.’
- (19) 壬辰允不雨風 (Heji 12921 v.)
 rénchén yǔn [NegP bù [v_P yǔ]] fēng
renchen indeed NEG rain blow
 ‘On the Renchen day, indeed it did not rain, but the wind blew.’

This property is consistent with VO languages, and equally holds for English and Modern Mandarin (where in general the verb does not raise to T°).

3.1 S V (O) [adjunct PP/NP]

Let us next examine sentences with adjunct phrases in postverbal position. This is a feature in which pre-Archaic Chinese patterns more strongly with typical head-initial languages than modern Mandarin, since in modern Mandarin adjunct phrases must precede the verb. Accordingly, the equivalents of (20)-(24) in Modern Mandarin would be unacceptable.

- (20) 乎多犬网鹿于麓 (Heji 10976 r.)
 hū duō quǎn [v_P wǎng lù [p_P yú nóng]]
 order numerous dog.officer net deer at Nong
 ‘Call upon the many dog-officers to net deer at Nong.’
- (21) 乞令吳以多馬亞省在南 (Heji 564 r.)
 qì ling wú yǐ duō mǎyǎ [v_P xǐng [p_P zài nán]]
 Qi order Wu lead numerous military.officer inspect at south
 ‘Officer Qi will order Wu to lead the numerous military officers to carry out an inspection in the south.’
- (22) 其品祠于王出 (Heji 23713)
 qí [v_P pǐn cí [p_P yú [T_P wáng chū]]]
 FUT *pin*.sacrifice *ci*.sacrifice at king go.out
 ‘One will perform a *pin* and a *ci* sacrifice when the king goes out.’

Both adjunct PPs headed by *yú* and *zài* can occur to the right of the verb (and object, when present). (21) is noteworthy insofar as it neatly illustrates pervasive head-complement order, where each embedding verb takes its propositional complement to its right. (23)-(25) illustrate temporal adjunct NPs in postverbal position:

- (23) 王入今月 (Heji 20038)
wáng [vP rù [NP jīn yuè]]
king enter present month
‘The king will enter (the city) this month.’
- (24) a. 其雨丁 (Heji 33943)
qí [vP yǔ [NP dīng]] yǔn [vP yǔ [NP dīng]]
FUT rain *dīng* indeed rain *dīng*
‘It will rain on the day Ding.’ ‘Indeed, it rained on the day Ding.’
- (25) 出于河來辛酉 (Tun 1119)
yòu yú hé [lái xīn-yǒu]
present to He next *xinyou*.day
‘[We will] present a sacrifice to the divinity He on the next *xinyou* day.’

3.2 S [adjunct PP/NP] V (O)

In contrast to the postverbal position where only one adjunct is permitted, multiple adjuncts are attested in the preverbal position to the right of the subject:

- (26) 王在十二月在襄卜 (Heji 24237)
wáng [[PP zài shí’èr yuè] [[PP zài xiāng] [vP bǔ]]]
king at 12 month at Xiang divine
‘The king in the twelfth month at the place Xiang made the divination.’

As illustrated in (27)-(28), NP and PP adjuncts show the same distribution:

- (27) 王今丁巳出 (Heji 07942)
wáng [NP jīn dīngsì] chū
king actual *dingsi* go.out
‘The king on this *dingsi* day goes out.’
- (28) 王自余入 (Heji 3458)
wáng [PP zì yú] rù
king from Yu enter
‘The king will enter from Yu.’

3.3 [Adjunct PP/NP] S V (O)

Finally, adjunct phrases can also occupy the sentence-initial position to the left of the subject:

- (29) 于辛巳王圍召方 (Heji 33023)
 [PP yú xīnsì] wáng wéi shào fāng
 at *xinsi* king surround Shao tribe
 ‘On the *xinsi* day, the king will surround the Shao tribe.’
- (30) 今六月王入于商 (Heji 7775)
 [NP jīn liù yuè] wáng rù yú shāng
 actual six month king enter in Shang
 ‘This sixth month, the king will enter the Shang city.’
- (31) 在𤯂王其先邁捍 (Ying 593)
 [PP zài nǚ] wáng qí xiān gòu hàn
 at Nü king FUT advance meet opposition
 ‘At Nü, the king will advance and meet an armed opposition.’

3.4 Focalization of adjuncts

The structure for the focalization of adjuncts is the same as that for arguments, i.e. it involves a cleft structure with a matrix copular predicate selecting a complement, whose specifier hosts the focalized adjunct.

- (32) 王勿佳今日往 (Heji 07351)
 wáng [NegP wù [vP wéi [[NP jīn rì] [vP wǎng]]]]
 king NEG be actual day go
 ‘It must not be today that the king will go.’
- (33) 唯于辛巳其雨 (Heji 20912)
 [TP [vP wéi [[PP yú xīnsì] [vP qí yǔ]]]]
 be at *xinsi* FUT rain
 ‘It is on the day *xinsi* that it will rain.’

In surface order terms, a focalized adjunct again is postverbal, i.e. it follows the copula. It cannot be confused with an “ordinary” preverbal adjunct (as illustrated in §3.2), given the obligatory presence of the copula when clefting adjuncts.³

³ In contrast to adjunct PPs, argument PPs can be focalized without an overt copula (cf. (ii)). *Yòu*

3.5 Pre- and postverbal distribution of adjunct phrases in Classical Chinese

The possibility of adjunct phrases in three positions (to the left or the right of the subject as well as following the verb) naturally raises the question whether there are any syntactic, semantic or pragmatic constraints governing the distribution of these adjunct phrases. While this problem has not been discussed in the literature on pre-Archaic Chinese, there have been studies devoted to possible semantic motivations governing the distribution of adjunct PPs in later stages such as Classical Chinese. Let us therefore briefly look at Classical Chinese and see whether we can gain any insight from it for the situation in pre-Archaic Chinese.

As can be seen in the following example from *Mengzi*, adjunct PPs in Classical Chinese can occur both in the pre- and postverbal position:

- (34) ... 故以羊易之 (Mengzi, Liang hui wang, 4th-3rd c. BC)
 ... gù [PP yǐ yáng] yì zhī
 therefore with sheep replace 3SG
 ‘... therefore [I] replace it [i.e. the ox] with a sheep.’
- (35) 我非愛其財而易之以羊也。 (ibid.)
 wǒ fēi ài qí cái ér yì zhī [PP yǐ yáng] yě
 1SG NEG cherish 3SG value CONJ replace 3SG with sheep PART
 ‘It is not that I attach a great importance to its value [i.e. the value of the ox] and therefore replaced it with a sheep.’

There seems to exist no consensus about possible semantico-pragmatic differences between the preverbal and the postverbal positions for adjunct PPs. While Lu Guoyao (1982) claims that the PP *yǐ yáng* carries emphasis in both positions, Liu Jingnong

sui ‘present an immolation’ constitutes the presupposition in (ii), and *yú Fùdīng* ‘to Fuding’ the focus, thus contrasting with *yú zǔyǐ* ‘to Zuyi’ in the first prediction (i) (displaying VO order):

- (i) 王侑歲于祖乙 (Heji 32113)
 wáng yòu suì [PP yú zǔyǐ]
 king present immolation to Zuyi
 ‘The king will present an immolation sacrifice to Zuyi.’
- (ii) 于父丁侑歲 (ibid.)
 [PP yú fùdīng] yòu suì
 to Fuding present immolation
 ‘It is to Fuding that [the king] will present an immolation.’

This corroborates the necessity to distinguish between argument PPs and adjunct PPs.

(1998) suggests that in ‘[*yǐ* NP] V O’ the VP is emphasized, whereas in ‘V O [*yǐ* NP]’ the PP is emphasized.

Concentrating on the syntactic aspect of this adjunct distribution, Huang (2006) proposes to derive postverbal adjunct phrases in Classical Chinese via VP fronting over the uniformly preverbal adjunct phrases to a non-specified landing site XP above *v*P. (This same VP fronting mechanism must be optional, since as we have seen adjunct phrases may surface in preverbal position in Classical Chinese).

Let us briefly examine how Huang’s suggestion can account for the following related facts.

First, non-branching adverbs such as *yǐ* ‘already’, *yǔn* ‘indeed’ were confined to the preverbal position below the subject from pre-Archaic Chinese (cf. (36)) through to Classical Chinese,

- (36) 允雨丁 (Heji 33943) (=24b)
yǔn [_{VP} yǔ [dīng]]
indeed rain ding.day
‘Indeed, it rained on the day Ding.’

Second, in pre-Archaic Chinese, preverbal and postverbal adjunct phrases could co-occur in the same sentence:

- (37) 其衣，翌日其征于室 (Heji 30373)
qí yī [yì rì] qí yán zūn [yú shì]
FUT sacrifice next day FUT continue offer at temple
‘After having performed a *yì* sacrifice, the next day one will continue to make offerings in the temple hall.’
- (38) 晡允雨自西 (Heji 20965)
[zè] yǔn yǔ [zì xī]
evening effectively rain from west
‘In the evening effectively it rained from the west.’
- (39) 晡亦出虹自北飲于河 (Heji 10405 v.)
[zè] [_{VP} yì [_{VP} yǒu chū hóng [_{PP} zì běi]] [_{VP} yǔn [_{PP} yú hé]]
evening also have come.out rainbow from north drink in river
‘In the evening there was also a rainbow coming out from the north and drinking in the river.’

On a VP fronting analysis, the most obvious way to derive the correct linear sequence in (36) is to position non-branching adverbs such as *yǔn* ‘indeed’ in a

projection above νP , perhaps as heads, along the lines of Cinque (1999). VP then fronts around the phrasal adverb (here *ding* ‘Ding day’) but adjoins below the non-branching adverb.

(37)-(39) might eventually be accommodated by generating the first adjunct phrase in a higher projection above νP (e.g. AuxP in (37)) or in a position outside TP ((38)-(39)) and by fronting the VP over the unique νP internal adjunct phrase, thus obtaining one preverbal and one postverbal adjunct phrase.

The problem posed by the existence of multiple adjuncts in preverbal position (cf. (26)) and the concomitant lack of multiple adjuncts in postverbal position, however, remains, Huang’s proposal predicting a strictly symmetric distribution for both positions. Importantly, this state of affairs did not only hold in pre-Archaic Chinese, but multiple adjuncts to the right of the VP remained impossible in the subsequent stages where adjuncts could still occur in postverbal position (until 2nd c. AD).

Accordingly, we suggest the Shang and Classical Chinese facts can be handled by allowing the verb to select exactly one VP shell (cf. Larson 1988) in these periods of Chinese. This results in the following structure for sentences such as (37)-(39):

$$(40) \text{ AdvP } [_{\nu P} \text{ V } [_{\nu P} \text{ t}_V \text{ AdvP}]]$$

3.6 Intermediate summary

In pre-Archaic Chinese, adjunct phrases can appear in three positions, to the left or the right of the subject and postverbally (i.e. after the object when present). While the constraints governing the distribution of adjuncts remain to be elucidated, it is evident that preverbal adjunct position cannot be likened to focus, since focalization of adjuncts requires a cleft structure with an overt matrix copular predicate (cf. §3.4 above).

The fact that only ‘V O adjunct’ is attested, to the exclusion of ‘V adjunct O’, argues against a V-raising analysis as in French, where a sentence such as (41) is derived by raising the verb to the functional category Infl/Tense, the adverb *souvent* ‘often’ taken to indicate the left margin of νP .

$$(41) \begin{array}{ccccccc} [_{IP/TP} \text{ Jean } [_{\text{Infl/Tense}} \text{ embrasse } [_{\nu P} \text{ souvent } [_{\nu P} \text{ t}_{\text{embrasse}} \text{ Marie}]]]] \\ \text{Jean} & & \text{kiss} & & \text{often} & & \text{Marie} \\ \text{‘John often kisses Mary.’} \end{array}$$

Furthermore, while multiple adjuncts in preverbal position are attested, no such data can be observed for the postverbal position, i.e. ‘[V O] adjunct adjunct’ is excluded. This fact cannot be captured by Huang’s (2006) VP fronting analysis. Importantly,

multiple adjuncts to the right of the VP remained impossible in the subsequent stages where adjuncts could still occur in postverbal position (until 2nd c. AD).

All these observations lead us to adopt a VP-shell structure à la Larson (1988) for pre-Archaic Chinese, where the postverbal adjunct is a complement of the verb and thus within the VP: [_{VP} V [_{VP} O [_{V'} t_V adjunct XP]]]. The possibility of exactly one branching adjunct to the right of the verb indicates that pre-Archaic Chinese allowed selection of just one such shell. The change resulting in the disappearance of postverbal branching adjuncts can then be formulated as loss of the VP shell structure.

4. VO word order in Modern Mandarin: Ban against adjuncts in postverbal position

A small sample of data suffices to invalidate Li & Thompson's (1974) claim that Modern Mandarin is in the process of changing into an SOV language, a process allegedly initiated more than two thousand years ago. (For studies arguing against the alleged OV status of Modern Mandarin, cf. *inter alia* Light 1979, Huang 1978, Mulder & Sybesma 1992, Whitman & Paul 2005).

Only arguments subcategorized for by the verb and “quasi” arguments depending on the verb's aktionsart, i.e. quantifier phrases indicating duration or frequency (cf. (45)-(46)) are admitted in postverbal position (cf. Huang 1982, Paul 1988).

- (42) 她打掃房子。
 tā dǎsǎo fǎngzi
 3SG sweep room
 ‘She’s cleaning the room.’

Accordingly, in the case of double object verbs, both the direct object and the indirect object follow the verb:

- (43) 他送了孩子很多錢。
 tā sòng-le [_{NP} háizi] [_{NP} hěn duō qián]
 3SG offer-PERF child very much money
 ‘He gave the child a lot of money (as a present).’
- (44) 我賣了一輛汽車給他。
 wǒ mài-le [_{NP} yí-liàng qìchē] [_{PP} gěi tā]
 1SG sell-PERF 1-CL car to 3SG
 ‘I sold him a car.’

- (45) 他等了半個小時。
 tā děng-le [QP bàn-ge xiǎoshí]
 3SG wait -PERF half-CL hour
 ‘He waited for half an hour.’
- (46) 他已經來了五次了。
 tā yǐjīng lái-le [QP wǔ-cì] le
 3SG already come-PERF 5-time PART
 ‘He has already come five times.’

By contrast, adjuncts (phrasal and non-phrasal) are totally excluded from the postverbal position (unlike in English) and have to precede the verb.

- (47) 他也/每天/常常來 (*也/*每天/*常常)。
 tā yě / měi-tiān / chángcháng lái (*yě /*měi-tiān /*chángcháng)
 3SG also / every-day / often come also/ every-day / often
 ‘He also comes every day/often.’
- (48) (在家裡/白天) 他 (在家裡/白天) 休息 (*在家裡/*白天)。
 (zài jiāli / báitiān) tā (zài jiāli / báitiān) xiūxi (*zài jiā-li /*báitiān)
 at home / daytime he at home / daytime rest at home / daytime
 ‘(At home/during daytime) he rests (at home/during daytime).’
- (49) 我給他當翻譯 (*給他)。
 wǒ gěi tā dāng fānyì (*gěi tā)
 1SG for 3SG act interpreter for 3SG
 ‘I serve as an interpreter for him.’

The ban against adjuncts in postverbal position illustrated in the data from Modern Mandarin (and observable from approximately the 3rd c. AD onwards) indicates major changes in the format of the vP, against the backdrop of constant VO word order. Consequently, more than the simple loss of a feature (optionally) triggering VP fronting over adjuncts (cf. Huang 2006) must be involved here. Instead, it is rather the loss of the VP shell structure as reflected in the impossibility for the verb to merge with a non-argument which is one of the factors at stake.

5. Conclusion

Using the issue of word order change as an illustration, we have demonstrated that in order to obtain meaningful results, the same analytical tools must be applied in

synchronic and diachronic studies alike.

Implementing this methodological principle, we have shown in detail that the changes observed in Chinese did not concern basic word order, but the distribution of adjunct phrases. These differences in the positional constraints on adjuncts in turn reflect a major change within the *v*P. While in pre-Archaic and Archaic Chinese (up to the 2nd c. AD), an adjunct could be complement of the verb (in the spirit of Larson 1988), in Modern Mandarin, adjuncts are precluded from the postverbal position and must precede the verb. Accordingly, only (quasi) arguments can be merged with the verb in Modern Mandarin.⁴ We have suggested to capture this change by the loss of the VP shell structure.

The incorrect assumptions about major word order changes in Chinese à la Li & Thompson (1974), which are still influential today (cf. a.o. Newmeyer 1998:242) are partly due to their not going back to the earliest available data. However, it is in the first place the concept of change as a panchronic, pangenerational event which leads to this faulty hypothesis, as is clearly illustrated by Li & Thompson's (1974:206) statement that "The shift [to SOV; DPW] is obviously incomplete, since Modern Mandarin Chinese still permits SVO word order in certain constructions". Only within this kind of conception is it possible to present individual changes as mere incremental steps of a "macro" change (cf. Hale 1998, 2007 for detailed discussion of these pervasive misconceptions).

⁴ Bennett (1981) and Jepson (1991) also noted changes resulting in a "reduction" of postverbal material; however, they did not realize that this was due to the exclusion of adjuncts in postverbal position.

References

- Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin & New York: Mouton de Gruyter.
- Bennett, David C. 1981. The Li & Thompson findings on word order and word order change in Mandarin: an alternative analysis. *York Papers in Linguistics* 8:183-191.
- Chen, Mengjia. 1956. *Yinxu Buci Zongshu* [*A General Study of the Yin Site Oracle Inscriptions*]. Beijing: Science Press.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford & New York: Oxford University Press.
- Djamouri, Redouane. 1988. *Etude des formes syntaxiques dans les écrits oraculaires gravés sur os et écaille de tortue (Chine 14^e-11^e av. J.-C.)*. Paris: Ecole des Hautes Etudes en Sciences Sociales.
- Djamouri, Redouane. 2000. Preverbal position of the pronominal object in Archaic Chinese. Paper presented at the 9th International Conference on Chinese Linguistics, June 26-28, 2000. Singapore: The National University of Singapore.
- Djamouri, Redouane. 2001. Markers of predication in Shang bone inscriptions. *Sinitic Grammar: Synchronic and Diachronic Perspectives*, ed. by Hilary Chappell, 143-171. Oxford & New York: Oxford University Press.
- Djamouri, Redouane. 2005. Chinois ancien et constructions verbales en série. <http://www.typologie.cnrs.fr/spip.php?article107>.
- Djamouri, Redouane, and Waltraud Paul. 1997. Les syntagmes prépositionnels en *yu* et *zai* en chinois archaïque. *Cahiers de Linguistique – Asie Orientale* 26.2:221-248. Electronic version available at: <http://crlao.ehess.fr/document.php?id=58>.
- Djamouri, Redouane, and Waltraud Paul. 2009. Verb-to-preposition reanalysis in Chinese. *Historical Syntax and Linguistic Theory*, ed. by Paola Crisma & Giuseppe Longobardi, 194-211. Oxford & New York: Oxford University Press.
- Dryer, Matthew. 2003. Word order in Sino-Tibetan languages from a typological and geographical perspective. *The Sino-Tibetan Languages*, ed. by Graham Thurgood & Randy J. LaPolla, 43-55. London: Routledge.
- Ernst, Thomas B. 2002. *The Syntax of Adjuncts*. Cambridge & New York: Cambridge University Press.
- Hale, Mark. 1998. Diachronic syntax. *Syntax* 1.1:1-18.
- Hale, Mark. 2007. *Historical Linguistics: Theory and Method*. Oxford: Blackwell.
- Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge: MIT dissertation.
- Huang, C.-T. James. 2006. The macro-history of Chinese and the theory of change. Paper presented at the Chinese Linguistics Workshop, December 1-2, 2006. Chicago: University of Chicago.

- Huang, Shuanfan. 1978. Historical change of prepositions and emergence of SOV order. *Journal of Chinese Linguistics* 6.2:212-242.
- Jepson, Jill. 1991. Typological models and diachronic change in Chinese. *Linguistics* 29.1:5-32.
- Karlgren, Bernhard. 1923. *Analytic Dictionary of Chinese and Sino-Japanese*. Paris: Geuthner.
- Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19.3: 335-391.
- Li, Charles N., and Sandra A. Thompson. 1974. An explanation of word order change SVO > SOV. *Foundations of Language* 12.2:201-214.
- Light, Timothy. 1979. Word order and word order change in Mandarin Chinese. *Journal of Chinese Linguistics* 7.2:149-180.
- Liu, Jingnong. 1998. *Hanyu Wenyan Yufa [The Grammar of Classical Chinese]*. Beijing: Zhonghua Book Company.
- Lu, Guoyao. 1982. Mengzi 'yi yang yi zhi', 'yi zhi yi yang' liang zhong jiegou leixing de duibi yanjiu [A comparative analysis of two similar structures, 'yi yang yi zhi' and 'yi zhi yi yang']. *Xianqin Hanyu Yanjiu [Research on Pre-Qin Chinese]*, ed. by Xiangqing Cheng, 272-290. Ji'nan: Shandong Education Press. [Reprinted in *Lu Guoyao Zi Xuan Ji [Selected Works by Lu Guoyao]*, 1-18. Zhengzhou: Henan Education Press, 1994.]
- Mulder, René, and Rint Sybesma. 1992. Chinese is a VO language. *Natural Language and Linguistic Theory* 10.3:439-476.
- Newmeyer, Frederick J. 1998. *Language Form and Language Function*. Cambridge: MIT Press.
- Paul, Waltraud. 1988. *The Syntax of Verb-Object Phrases in Chinese: Constraints and Reanalysis*. Paris: Editions Langues Croisées. Electronic version available at: <http://crlao.ehess.fr/document.php?id=177>.
- Paul, Waltraud, and John Whitman. 2008. 'Shi...de' focus clefts in Mandarin Chinese. *The Linguistic Review* 25.3-4:413-451.
- Peyraube, Alain. 1988. *Syntaxe diachronique du chinois. Evolution des constructions datives du 14e siècle av. J.-C. au 18e siècle*. Paris: Collège de France, Institut des Hautes Études Chinoises.
- Sagart, Laurent. 1999. *The Roots of Old Chinese*. Amsterdam & Philadelphia: John Benjamins.
- Shen, Pei. 1992. *Yinxu Jiagu Buci Yuxu Yanjiu [Research on Word Order in the Oracle Inscriptions of the Yin Site]*. Taipei: Wenjin Press.
- Whitman, John, and Waltraud Paul. 2005. Reanalysis and conservancy of structure in Chinese. *Grammaticalization and Parametric Change*, ed. by Montserrat Batllori,

Redouane Djamouri, Waltraud Paul, and John Whitman

Maria-Lluisa Hernanz, Carme Picallo & Francesc Roca, 82-94. Oxford & New York: Oxford University Press.

Short Passives in Modern and Classical Chinese

Shengli Feng (馮勝利)

*The Chinese University of Hong Kong
Beijing Language and Culture University*

While the long passive *bei*-sentences in Mandarin Chinese involve an A'-movement, the short passives, as argued in the paper, can be divided into two types of constructions: one is viewed as a lexical compound; while the other is viewed as a complex verb. From both synchronic and diachronic perspectives, it is further argued that if the chronology of emergence of the [*bei* VP] forms is taken into account, three types of operations ranging from morphology (the [*bei*-V] compound), to morphosyntax (the [[*bei* V] NP] generated by the verb complex via Verb Incorporation) and to phrasal syntax (the [*bei* V __] derived by NP-movement) are all observed in history and maintained in Mandarin Chinese. It is then concluded that while synchronic analyses should refer to historical facts, diachronic studies can just as effectively cast light on synchronic research.

Key words: long passive, short passive, A'-movement, Verb-incorporation

1. Introduction

In formal syntactic theory, passive constructions have been taken as fundamental syntactic processes involving NP-movement. According to Chomsky (1995), movements never occur unless the interactions of some principles or subtheories require it. Passivization, a case of NP-movement, is triggered by passive morphology and forced by case theory. But such kind of syntactic analysis is found difficult to apply to languages like Chinese, since there is simply no such passive morphology in Chinese as in English. For example:

- (1) a. 在戰爭中張三被敵人炸掉了一條腿。
Zài zhànzhēng zhōng Zhāngsān bèi dírén zhà-diào le yìtiáo tuǐ.
at war inside Zhangsan BEI enemy blown-off ASP one.CL leg
'Zhangsan's one leg was blown off by the enemy in the war.'
- b. 張三_i被李四打了他_i好幾次。
Zhāngsān_i bèi Lǐsì dǎ le tā_i hǎo-jǐ cì.
Zhangsan_i BEI Lisi hit ASP him_i good-many times.
'Zhangsan was hit by Lisi many times.'

bei is traditionally called a passive marker and therefore the pattern of a full passive sentence in Modern Chinese is like the following.

- (2) [NP₁ *bei* NP₂ V NP₃/pronoun]

NP₁, the subject of a passive, is the logical object, and NP₂ is the logical subject. Since a pronoun or an NP in the object position needs Case according to the Case-filter, the verbs in (1) must have the ability to assign Case to the object. Therefore the passive marker *bei* does not function like *-en* in English, which absorbs the accusative case and the subject theta role from the verb.

Furthermore, although a pronoun is allowed to appear in the object position as seen in (1b), the sentence becomes ungrammatical if the agent NP is absent:

- (3) *張三_i被__打了他_i好幾次。
 *Zhāngsān_i bèi ____ dǎ le tā_i hǎo-jǐ cì.
 Zhangsan_i BEI ____ hit ASP him_i good-many times
 Intended meaning: ‘Zhangsan was hit many times.’

However, the absence of the agent NP would not affect the grammaticality of (1a):

- (4) 在戰爭中張三被炸掉了一條腿。
 Zài zhànzhēng zhōng Zhāngsān bèi ____ zhà-diào le yìtiáo tuǐ.
 in war inside Zhangsan BEI ____ blown-off ASP one.CL leg
 ‘One of Zhangsan’s leg was blown off in the war.’

Why is the agent NP so crucial in licensing the object pronoun and why are object NPs not affected by the presence or absence of the agent? *Bei* sentences without the agent NP have been called *Short Passive* (Ting 1995, Huang 1999), hence the question is why pronouns should make any difference from NPs in the short passive construction.

In this paper, I argue that the synchronic problems may very well be resolved from a diachronic perspective. First, according to Feng (1998), there were three types of *bei* passive structures that developed one after the other in Chinese history. The [*bei*-V] compound passive is attested in Archaic Chinese (400 BC). The [*bei* VP] phrasal passive is attested during the Eastern Han (100 AD) and the [*bei* IP] full (long) passive is attested by the Late Han (200 AD) and further developed by the Six-dynasties (420-589 AD). These three stages of the developments of *bei* passives are historically related but structurally independent, and all of the three forms are kept in Modern Chinese.

This paper is organized as follows. Section 2 provides background from previous studies on short passives in Chinese. Section 3 supports the compound-analysis proposed

by McCawley (1992) from a historical perspective. Section 4 describes the origin of the three types of *bei* passive constructions in Chinese history. Section 5 argues for a Verb Incorporation (VI) analysis for the phrasal compounds, and §6 is a summary of this study.

2. Previous studies on short passive construction

The short passive has always been a problem in Chinese syntax. For example (taken from McCawley 1992).

- (5) *我被____從身上偷了手錶。
 *Wǒ bèi ____ cóng shēnshàng tōu le shǒubiǎo.
 I BEI ____ from body steal ASP watch
 ‘I had my watch stolen off my wrist.’

When the agent NP is absent, no prepositional phrases are allowed to be adjacent to the *bei*. To this problem, Cheng (1986) proposes a reanalysis rule, cited in (6) which seems to work very well for (5).

- (6) $[_{VP} P + e + V] \rightarrow [_{V} P V]$ (Cheng 1986)

It states that in the domain of VP, a preposition (i.e., *bei*) which is followed by an empty NP is reanalyzed as part of a compound verb. In Cheng’s system the reanalysis is obligatory if the agent NP is empty. Cheng’s (6) would rule out sentences like (5) by designating that *bei* is not adjacent to the verb and thus cannot be compounded with the verb. It seems to be a perfect solution to the problems caused by (5). There are, unfortunately, grammatical sentences in which the *bei* is not followed by a verb when the agent NP is absent, as pointed out by McCawley (1992) and Ting (1995):

- (7) a. 張三被秘密地逮捕了。
 Zhāngsān bèi mìmì.de dàibǔ le.
 Zhangsan BEI secretly arrest ASP
 ‘Zhangsan was arrested secretly.’
 b. 張三被殘忍地殺死了。
 Zhāngsān bèi cánrěn.de shā-sǐ le.
 Zhangsan BEI cruelly kill-die ASP
 ‘Zhangsan was killed cruelly.’

Obviously, when the reanalysis rule successfully rules out the ungrammatical ones in (5), it also rules out the grammatical ones in (7). It should be noted, however, that the reanalysis rule given by Cheng has captured an important property of the short passive sentences, as we will see in §3.

There is another problem brought up by short passives to the A'-movement analysis proposed by Feng (1990, 1995) and supported in Chiu (1995), Ting (1995, 1996) and Cheng, Huang, Li & Tang (1996). The A'-movement analysis was originally motivated by long passive sentences like the following (Feng 1990):

- (8) a. 張三_i被人把他_i打死了。
 Zhāngsān_i bèi rén bǎ tā_i dǎsǐ le.
 Zhangsan_i BEI someone BA him_i kill ASP
 'Zhangsan was killed by someone.'
- b. 教室_i被老師派李四找人掃乾淨_{t_i}了。
 Jiàoshì_i bèi lǎoshī pài Lǐsì zhǎo rén sǎo-gānjìng t_i le.
 classroom_i BEI teacher send Lisi find men sweep-clean t_i ASP
 '*The classroom was sent Lisi to find someone to sweep by the teacher.'
 'The classroom was cleaned by someone who the teacher sent Lisi to find.'

Feng (1990) argues that the *bei*-sentences in (8) involve an empty operator movement, taking the following structural analysis:

- (9) 張三被我打死了。
 Zhāngsān bèi [_{S'} OP_i [_S wǒ dǎ-le e_i]]

The operator *OP* is an EC and is semantically empty, therefore, [_{e_i}] is in effect a free variable, assigned no value by its operator in the embedded clause. As a result, it violates the principle barring vacuous operators. To satisfy the requirements that each LF variable be assigned a value by an antecedent that A-binds it, [_{e_i}] must be bound by an NP in the matrix subject position as in (10a):

- (10) a. 張三被我打了。
 Zhāngsān_i [bèi [OP_i [wǒ dǎ le e_i]]]
 Zhangsan BEI I hit ASP
 'Zhangsan was hit by me.'
- b. *Zhāngsān_i bèi [OP_j [wǒ dǎ-le e_j]]

Under this analysis, sentence (10b) can be readily ruled out, because [_{e_j}] is bound

only by OP_j , and is in effect a free variable, which must be ruled out. The advantage of this analysis is that it explains not only the grammaticality of long passives in Modern Chinese, but also those in Classical Chinese¹ (see Wei 1994), as in (11).

- (11) a. 墓上人…被石酒氣充入鼻中。(Suoshenji 搜神記 19, ca. 363 AD)
 Mù shàng rén... bèi Shí jiǔ qì chōng rù bí zhōng
 tomb up man... BEI Shi alcohol smell penetrate enter nose inside
 ‘The men on the tomb...got penetrated by Shi’s alcoholic smell into their nose.’
- b. (金丹)若被諸物犯之…(Baopuzi 抱朴子, Neipian 內篇, ca. 300 AD)
 (Jīndān) ruò bèi zhūwù fàn zhī...
 (Jindan) if BEI everything attack him...
 ‘If Jindan was attacked (him) by everything...’

A retained object (11a) and especially a resumptive pronoun (11b) were also found to appear in the object position of long passive sentences in Classical Chinese which is exactly what the A'-movement analysis predicts.

However, the A'-movement analysis suffers from the challenge of short passive sentences. First, why are the following sentences ungrammatical?

- (12) *張三_i被打了他_i了。
 *Zhāngsān_i bèi dǎ le tā_i le.
 Zhangsan BEI hit ASP him ASP
 Intended meaning: ‘Zhangsan was hit (him).’

There seems to be no reason for the A'-movement analysis to rule out the ungrammatical sentences, because the absence of the agent NP should not affect the structure for the Null Operator to move, if short passives are derived from the long form via deletion of the agent NP. In order to account for these sentences, Feng (1990), following Cheng (1986), proposes a reanalysis rule similar to (6) and claims that the $_{VP}[\dots bei X \dots]$ is reanalyzed as a compound. However, the reanalysis account, as seen before, is seriously challenged by the following examples (Ting 1995):

¹ The term ‘Classical Chinese’ is used here as a cover term referring to the literary language in Archaic and Ancient Chinese. ‘Archaic Chinese’ refers to the language used before the Qin Dynasty (220-206 BC) and ‘Ancient Chinese’ refers to the language from the Han Dynasty (206 BC-220 AD) to the Tang Dynasty (618-907 AD).

- (13) a. *張三被在公園打死了。
 *Zhāngsān bèi zài gōngyuán dǎ-sǐ le.
 Zhangsan BEI at park hit-die ASP
 Intended meaning: ‘Zhangsan was hit to death in a park.’
- b. 張三被秘密地打死了。
 Zhāngsān bèi mìmì.de dǎ-sǐ le.
 Zhangsan BEI secretly hit-die ASP
 ‘Zhangsan was hit to death secretly.’

The question is: if *bei* must be reanalyzed with the next element, why can it be so in (13b) but not in (13a)? Obviously, the reanalysis approach cannot account for the difference between (13a) and (13b).

In order to explain why there are two entirely different syntactic derivations operating in *bei* sentences, Ting (1995) proposes that there are actually two separate BEIs. BEI-1 is a two place predicate, taking an infinitive clause as its internal argument and an NP as its external argument, hence giving rise to an A'-movement. But BEI-2 can only take a VP as its internal argument, hence an A-movement is involved.

The Two-BEI analysis avoids the problem caused by the Reanalysis Hypothesis (Cheng 1986, Feng 1990), because what must be said about the short passives like (13b) in terms of a reanalysis can now be attributed directly to the syntactic properties of the BEI-2. Since BEI-2 only takes a bare VP, only VP adverbials (13b) but not sentential adverbials (13a) are allowed (See Ting 1996 for detailed analyses).

While distinguishing two BEIs is really a correct solution, as can be seen in §5, the NP-movement analysis for short passives is not without problems as it stands. First, as Huang (1999) points out, the *bei* passives (both long and short) may contain a subject-oriented adverb like *gùyì* ‘intentionally’, which the *bei* passive in English does not allow. For example (taken from Huang 1999):

- (14) a. 張三故意被打了。
 Zhāngsān gùyì bèi dǎ le.
 Zhangsan intentionally BEI hit ASP
 ‘Zhangsan intentionally got hit.’
- b. 張三故意被李四打了。
 Zhāngsān gùyì bèi Lǐsì dǎ le.
 Zhangsan intentionally BEI Lisi hit ASP
 ‘Zhangsan intentionally got hit by Lisi.’
- c. *Zhangsan was hit intentionally.’

This suggests that the subject is base-generated in place and receives an independent thematic role from *bei*. If this is so, the Theme object cannot undergo an NP-movement to the subject position, because the subject position is not dethematized.

Second and most importantly, it is difficult for an NP-movement analysis to account for the grammatical sentence properly, as seen in (4) and more in (15).

- (15) a. 張三被咬掉了一隻手。
 Zhāngsān bèi yǎo-diào le yìzhī shǒu.
 Zhangsan BEI bite-off ASP one hand
 ‘Zhangsan had one of his hands bitten off.’
- b. 房頂被吹掉了一個角。
 Fángdǐng bèi chuī-diào le yíge jiǎo.
 roof BEI blow-off ASP one corner
 ‘A corner of the roof was blown off.’
- c. 愚人被他打頭。(Baiyujing 百喻經, ca. 500 AD)
 Yúrén bèi tā dǎ tóu.
 simpleton BEI he hit head
 ‘The simpleton was hit (his head) by someone.’
- d. 近者被截髮。(Luoyang Qielanji 洛陽伽藍記, ca. 500 AD)
 Jìn zhě bèi jié fà.
 close one BEI cut hair
 ‘The intimate ones had their hair cut off.’
- e. 龜被生揭其甲。(Lingbiao Luyi 嶺表錄異, ca. 850 AD)
 Guī bèi shēng jiē qí jiǎ.
 turtle BEI alive take-off its shell
 ‘The turtle’s shell was taken off while it was still alive.’

In these sentences, a retained object can still appear in the object position. The above grammatical sentences therefore argue against an A-movement analysis as given in the standard theory that NP-movement is forced by the lack of Case.² In fact, for whatever motivations that one can come up with to make the object NP (or part of it) to

² Note that the retained object can probably be said to have an inherent Case not necessarily assigned by the verb. If this is so, why are the same sentences ungrammatical in English?

*The roof was blown off a corner.

*Zhangsan was bombed off one leg in the war.

*Zhangsan was bitten off one hand.

As seen below, the difference between Chinese and English can be explained by the theory developed here.

move (or stay), the movement process must be optional and not obligatory, which obviously leads to a situation in which an NP-movement is not called for.³ In what follows I argue that an appropriate analysis for the short passives should not only account for synchronic examples, but also accommodate diachronic facts. This is particularly so in the study of *bei* passives because, as we will see later, the short passives were not only independent of the long passives, but also developed during several historical stages. Most importantly, each stage of their developments created an independent structure used in Modern Chinese. Therefore, without a historical perspective, it is hard to see the independence of the short form from the long one and difficult to capture the variety of short passives established in history which have endured until the present day.

3. [*bei*-V] as compounds

Although current syntactic analyses have adopted the A'-movement analysis for long passive *bei* sentences, there is no agreement upon the short passive sentences. The problems involved in various types of short passives can be summarized in examples given in (16):

- (16) a. 張三被人秘密地逮捕了。
 Zhāngsān bèi rén mìmi.de dàibǔ le.
 Zhangsan BEI people secretly arrest ASP
 'Zhangsan was secretly arrested by someone.'
- a'. 張三被____秘密地逮捕了。
 Zhāngsān bèi _____ mìmi.de dàibǔ le.
 Zhangsan BEI _____ secretly arrest ASP
 'Zhangsan was secretly arrested.'
- b. 張三_i被人秘密地把他_i殺害了。
 Zhāngsān_i bèi rén mìmi.de bǎ tā_i shāhài le.
 Zhangsan BEI people secretly BA him murder ASP
 'Zhangsan was secretly murdered (him) by someone.'
- b'. *張三_i被____秘密地把他_i殺害了。
 *Zhāngsān_i bèi _____ mìmi.de bǎ tā_i shāhài le.
 Zhangsan BEI _____ secretly BA him murder ASP
 Intended meaning: 'Zhangsan was secretly murdered (him).'

³ Huang (1999) proposes an 'NP-movement-plus-control analysis' for the [NP *bei* V__] short passives and, as seen in below, diachronic facts also argue for the NP-movement in this type of *bei* sentences.

- c. 張三_i被人把他_i打死了。
 Zhāngsān_i bèi rén bǎ tā_i dǎ-sǐ le.
 Zhangsan BEI people BA him hit-die ASP
 ‘Zhangsan was killed (him) by someone.’
- c'. *張三_i被把他_i打死了。
 *Zhāngsān_i bèi bǎ tā_i dǎ-sǐ le.
 Zhangsan BEI BA him hit-die ASP
 Intended meaning: ‘Zhangsan was killed (him).’
- c''. *張三_i被打死了他_i。
 *Zhāngsān_i bèi dǎ-sǐ le tā_i.
 Zhangsan BEI hit-die ASP him
 Intended meaning: ‘Zhangsan was killed (him).’
- d. 那隻狗被人打掉了一條腿。
 Nèizhī gǒu bèi rén dǎ-diào le yìtiáo tuǐ.
 that dog BEI one strike-off ASP one.CL leg
 ‘That dog had his one leg struck off by someone.’
- d'. 那隻狗被____打掉了一條腿。
 Nèizhī gǒu bèi ____ dǎ-diào le yìtiáo tuǐ.
 that dog BEI strike-off ASP one.CL leg
 ‘That dog has his one leg struck off.’
- d''. 那隻狗被____活活地打掉了一條腿。
 Nèizhī gǒu bèi ____ huóhuó.de dǎ-diào le yìtiáo tuǐ.
 that dog BEI alive strike-off ASP one.CL leg
 ‘That dog had his one leg struck off alive.’

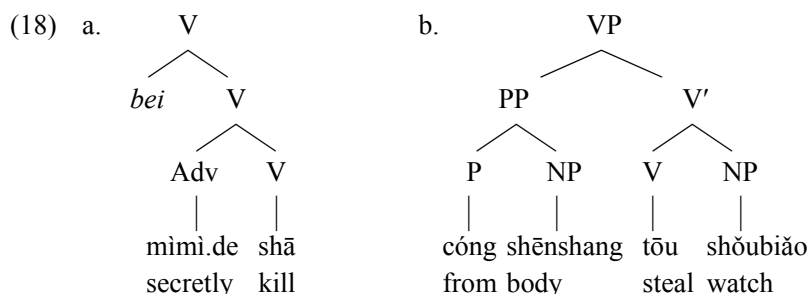
The problems are: first, if [*bei* V] cannot take a resumptive pronoun (16c''), why can it take an NP (16d')? Second, why do the [*bei* ADV V] (16d'') and the [*bei* V] (16d') behave alike in taking an object NP, and why does the [*bei* ADV V] in (16b') and the [*bei* V] in (16c'') act alike in rejecting a resumptive pronoun? All in all, why should the presence or absence of the agent NP make any difference to the acceptability of a *bei* sentence with a pronoun? As seen below, all these properties would naturally follow if the [*bei* V] and [*bei* ADV V] are analyzed as compounds.

McCawley (1992) argues that in short passives of Modern Chinese, *bei* is not an independent word but a part of a compound verb. First, he observes that the omission of the agent NP is not possible when anything intervenes between the *bei* and the V (see Li 1990:156), as seen in (17) (taken from McCawley 1992).

- (17) a. 張三被人在公園打死了。
 Zhāngsān bèi rén zài gōngyuán dǎ-sǐ le.
 Zhangsan BEI man ZAI park hit-die ASP
 ‘Zhangsan was killed by someone in a park.’
- a'. *張三被____在公園打死了。
 *Zhāngsān bèi ____ zài gōngyuán dǎ-sǐ le.
 Zhangsan BEI at park hit-die ASP
 Intended meaning: ‘Zhangsan was killed in a park.’
- b. 我被他從身上偷了手錶。
 Wǒ bèi tā cóng shēnshang tōu le shǒubiǎo.
 I BEI him from body steal ASP watch
 ‘I had my watch stolen off my wrist by him.’
- b'. *我被____從身上偷了手錶。
 *Wǒ bèi ____ cóng shēnshang tōu le shǒubiǎo.
 I BEI from body steal ASP watch
 Intended meaning: ‘I had my watch stolen off my wrist.’
- c. 那些人被李四都請來了。
 Nèixiē rén bèi Lǐsì dōu qǐnglái le.
 those people BEI Lisi all invite ASP
 ‘Those people were all invited by Lisi.’
- c'. *那些人被都請來了。
 *Nèixiē rén bèi dōu qǐnglái le.
 those people BEI all invite ASP
 Intended meaning: ‘Those people were all invited.’
- d. 我被八戒怎麼領導起來了？
 Wǒ bèi Bājiè zěnmē lǐngdǎo qǐlái le?!
 I BEI Bajie why lead start ASP
 ‘Why I am starting to be led by Bajie?’
- d'. *我被____怎麼領導起來了？
 *Wǒ bèi ____ zěnmē lǐngdǎo qǐlái le?!
 I BEI why lead start ASP
 Intended meaning: ‘Why I am starting to be led?’

McCawley notes that the adverbial quantifier *dōu* 都 can appear before the V if the agent NP is present, but cannot do so if that NP is omitted. When the agent NP is absent, however, the following sentences are still grammatical even if there are interventions between the *bei* and the V, as seen in (7). McCawley argues that the items that can

separate *bei* from the verb are verb-modifiers, not V'-modifiers, S-modifiers, or Verb-complements. According to McCawley, “the combination of modifier and modified constituent is of the same category as the modified constituent.” If this is so, it follows that the combination of a V-modifier with a V is still a V (and not a V'). Thus, *mìmì.de dǎsǐ* in (7a) is a V and so is *mìmì.de shā* in (7b), given the assumption that *mìmì.de* and *cánrěn.de* are V modifiers, rather than V' modifiers. If both *mìmì.de dǎsǐ* and *cánrěn.de shā* are verbs, they are eligible for compounding with *bei*, whereas expressions of the other types do not combine with a V into any syntactic unit at all, let alone a V, and are thus not eligible for compounding with *bei*. This is shown in (18).



The [*bei*-ADV-V] in (18a) is a compound, because the *bei* and the [ADV-V] belong to the same syntactic category, while the combination of *bei* with a P in (18b) would not result in any syntactic unit at all. This is why sentences given in (17a'-b') and (17c'-d') are all ungrammatical.

I will adopt McCawley's hypothesis and provide further evidence for his analysis. Note that, given McCawley's assumption, sentences such as *Zhāngsān bèi mìmì.de dàibǔ le* 'Zhangsan was arrested secretly' is in fact an [NP V] structure. No resumptive pronoun is therefore allowed to appear in the object position coindexed with the subject, because of an obvious violation of the Principle B of the Binding theory. This is exactly the case as we have seen before:

- (19) a. 李四_i被警察偷偷地打了他_i一頓。
 Lǐsì_i bèi jǐngchá tōutōu.de dǎ le tā_i yídùn.
 Lisi BEI policeman secretly hit ASP him once
 'Lisi was hit secretly by a policeman.'
- b. *李四_i被____偷偷地打了他_i一頓。
 *Lǐsì_i bèi ____ tōutōu.de dǎ le tā_i yídùn.
 Lisi BEI secretly hit ASP him once
 Intended meaning: 'Lisi was secretly hit (him).'

- c. *李四_i被____打了他_i一頓。
 *Lǐsì_i bèi ____ dǎ le tā_i yídùn.
 Lisi BEI hit ASP him once
 Intended meaning: ‘Lisi was hit (him).’

Why is (19b) not grammatical while (19a) is? The different grammaticality status between (19a) and (19b) shows that the two structures [*bei* NP ADV V] and [*bei* ADV V] must be different. On the other hand, why are (19b) and (19c) both ungrammatical? The same grammaticality status between (19b) and (19c) shows that they must have identical structures. This will lead naturally to a conclusion that the [*bei*-ADV-V] must be a single compound verb like [*bei* V].

4. History of short passives

The history of *bei* passives in Classical Chinese also confirms the hypothesis that the [*bei* V] and [*bei*-ADV-V] are compounds. First, the [*bei* V] short passives originated from the [*bei* NP] structure (‘*bei*’ means ‘to receive’ or ‘to get’), for example:

- (20) a. 天下被其利。(*Mozi* 墨子, *Shangxian* 尙賢, ca. 468-376 BC)
 Tiānxià bèi(pī) qí lì.
 world receive its benefit
 ‘The world got benefit from it.’
 b. 身被三累。(*Hanfeizi* 韓非子 2.3.52, ca. 281-233 BC)
 shēn bèi sān lěi.
 body receive three trouble
 ‘You got three troubles.’

Through a reanalysis, i.e., [*bei* N] → [*bei* V], the N in [*bei* NP] changed into a V, giving rise to the [*bei* V] structure during the Warring States (475-221 BC) (See Peyraube 1996, Wang 1989, Wei 1994), as seen in (21).

- | | | | |
|---------|-----------|-------------------------------------|----|
| (21) a. | bèi lù | ‘got killed’, ‘got a death penalty’ | 被戮 |
| b. | bèi shāng | ‘got wounded’, ‘got a wound’ | 被傷 |
| c. | bèi rǔ | ‘got insulted’, ‘got an insult’ | 被辱 |
| d. | bèi qīn | ‘got invaded’, ‘got an invasion’ | 被侵 |
| e. | bèi xíng | ‘got executed’, ‘got an execution’ | 被刑 |
| f. | bèi fēn | ‘got disrupted’, ‘got a disruption’ | 被分 |

Although there is no argument against the category change from N to V in *[bei NP]* sentences, one fact seems not to have received enough attention in the literature, namely only disyllabic *[bei X⁰]* forms are possible for an interpretation as either a *[bei N]* or a *[bei V]*, as seen in (21). (21) shows that all recognizable *[bei V]* forms prior to the Han Dynasty (221 BC-206 AD) are overwhelmingly disyllabic forms, which I will call the Archaic forms. Of course there were trisyllabic *[bei σσ]* forms before Han, however, these forms like (22) were more likely a *[bei NP]* rather than a *[bei V]*, as argued by Wang (1980:425).⁴

(22) Pre-Qin (before 221 BC)

厚者入刑罰，薄者被毀醜。(Mozi 墨子, *Gui yi* 貴義, ca. 468-376 BC)

Hòu zhě rù xíngfá, bó zhě bèi huǐchǒu.

thick one enter punishment thin one get condemnation

‘The kind ones get into punishment and the mean ones get condemnation.’

(22) shows that the trisyllabic *[bei σσ]* is different from the disyllabic *[bei σ]* structure (21). The disyllabic complements of *bei* before Han are more likely to be NP and therefore are unlikely to be reinterpreted as a verb. However the change from *[bei NP]* to *[bei VP]* needs a proper environment in which a reanalysis of N as V could take place, i.e., an environment that could enable the next generation to have trigger experiences for generating a new grammar (see Lightfoot 1991). Obviously, only the disyllabic *[bei σ]* forms can serve this purpose, as indicated in (21). The contrast between (21) and (22) shows clearly that the transition from *[bei NP]* to *[bei VP]* proceeded from a disyllabic environment. That is:

(23) $N \rightarrow V / [bei __]_{\sigma\sigma}$

The necessary disyllabic environment for this change therefore leads naturally to a conclusion that the new *[bei V]* structure was born as a Prosodic Word, given the theory of Prosodic Morphology of Classical Chinese developed in Feng (1997).

⁴ One may argue that *bèi xínglù* 被刑戮 ‘was executed’ is also found in *Guǎnzǐ* 管子 (Pre-Qin?). If *bèi xínglù* is a *[bei V]* structure in *Shǐjì* 史記, it should also be analyzed as a *[bei V]* in *Guǎnzǐ*. As a result, the *[bei σσ]* forms that are analyzed as *[bei V]* must date from before the Han. However, it is well known that *Guǎnzǐ* was not written by a single person at one period of time. Although the phrase *bèi xínglù* did appear in *Guǎnzǐ*, according to previous studies (see Luo 1981, Xie 1983, and many others), the Section in which *bèi xínglù* occurs (i.e., *Chéngmǎshù* 乘馬數) was written by Han Scholars during the time of Emperor Wu and Zhao—exactly the same time that the *Shǐjì* was written. As a result, *Chéngmǎshù* cannot be considered as a Pre-Qin text and *bèi xínglù* is therefore not a Pre-Qin language.

Furthermore, the fact that no adverbs can appear inside the [*bei* V] PrWds in Pre-Qin (221 BC) indicates that the [*bei* V] forms may have been lexicalized as a compound. This is shown in (24).

- (24) a. 被圍于趙 (*Zhanguoce* 戰國策, *Qice* 齊策, ca. 300 BC)
 bèi wéi yú Zhào
 ‘got surrounded by Zhao’
 b. *被必圍于趙
 bèi bì wéi yú Zhào (=not attested)
 Intended meaning: ‘got certainly surrounded by Zhao’
 c. *被未圍于趙
 bèi wèi wéi yú Zhào (=not attested)
 Intended meaning: ‘got not surrounded by Zhao’

If the Archaic short passives were born as PrWds and lexicalized as compounds, then it is not surprising that the modern short passives also inherit this property.⁵

Third, as seen before, a resumptive pronoun in Ancient Chinese is allowed to appear after the embedded verb co-occurring with an agent NP. For example:

- (25) a. (李子敖) 被鳴鶴吞之。 (*Soushenji* 搜神記, ca. 363 AD)
 (Lǐ Zǐáo) bèi míng hè tún zhī.
 (Li Zi’ao) get chirping crane swallow him.
 ‘Li Zi’ao got swallowed (him) by a chirping crane.’

⁵ Note that, even if the [*bei* -V] forms in Archaic Chinese are compounds, it does not mean that after the Han the two elements in the compound cannot form a phrase. Hence, synchronically (at any particular time after the Qin) there may be both compounds and phrases that share the same surface form. However, the compounds are inherited from Archaic Chinese, and their counterpart phrases are created anew. As a result, it would be extremely difficult to identify whether a given form such as *bèi hài* 被害 ‘be harmed’ in Han is an (old) compound or a (new) phrase according to their non-analyzed surface form. In fact, it may be a compound in one environment (without an object) but a phrase in another (with an adverb). This situation is similar to that of Modern Chinese: it is difficult to decide whether *guān-xīn* ‘concern heart, to concern’ is a compound or a phrase without the kind of syntactic information available to the researcher (for example, *guan-xin* is a compound when it takes an object, but it is a phrase in *guān shén.me xīn* ‘what on earth are you concerning about’). The point I am trying to make here is this: the [*bei*-V] compounds and the [*bei* VP] phrases can be distinguished clearly by their diachronic developments. Hence diachronic studies can reveal distinct syntactic structures that emerged one after another historically, which synchronized and overlapped later on, making it difficult (if not impossible) to distinguish them in some cases.

- b. (天女) 被池主見之。(*Soushenji* 搜神記, ca. 363 AD)
 (Tiān nǚ) bèi chí-zhǔ jiàn zhī.
 (Heaven woman) BEI pool-owner see them.
 ‘The Heaven fairies were seen (them) by the pool-owner.’
- c. (金丹) 若被諸物犯之。(*Baopuzi* 抱朴子, *Neipian* 內篇, ca. 300 AD)
 (Jīndān) ruò bèi zhūwù fàn zhī...
 (Jindan) if BEI everything attack it...
 ‘If Jindan was attacked (it) by everything...’

There is no sentence found with a structure of $[NP_i \text{ bei } ______ V \text{ pro}_i]$. Given the historical origin of the $[bei-V]$ PrWd and compound, and given our analysis above, it is clear that the $[bei V]$ sequence must be a verb compound, resulting in the ungrammatical structure: $*[NP_i [bei-V]_V \text{ Pro}_i]$. The fact is that there were many sentences formed with a $[NP_i \text{ bei NP V pro}_i]$ structure, but hardly any $[NP_i \text{ bei V pro}_i]$ cases in the language.⁶ The non-existent $*[NP_i \text{ bei V pro}_i]$ sentences strongly indicate that the *bei* and the

⁶ There is only a single exception among all the grammatical sentences in my data:

- (i) 過在什麼處，即被打之？(*Zutangji* 祖堂集, 4. *Shitou Heshang* 四 • 石頭和尚, ca. 10th century)

Guò zài shén.me chù, jí bèi dǎ zhī.
 wrong at what place then *bei* hit it
 ‘What was wrong, then (he) was hit?’

However, the *zhī* in this case may be used as a place holder, rather than a resumptive pronoun. The reason is as follows: First, we cannot find *zhī* appearing in $*[bèi dǎ zhī]$ ‘was hit *it’, but only in a four-syllable environment: $[(j)bèi] (dǎzhī)]$. Hence it is possible that *zhī* was used as a place holder for filling up the four-syllable pattern. Second, the function of *zhī* in (i) parallels the function of *suǒ* (an object clitic which is also licensed in the context of a null operator, see Chiu 1995, Ting 1996) in the same environment:

- (ii) a. 恐被所算 (*Sanguo Yanyi* 58 三國演義 • 第58回)
 kǒng bèi suǒ suàn
 afraid got SUO trick
 ‘afraid to be tricked’
- b. 必被所侵 (*Sanguo Yanyi* 94 三國演義 • 第94回)
 bì bèi suǒ qīn
 definitely get SUO catch
 ‘(will) definitely be caught’

Both *zhī* 之 and *suǒ* 所 are disallowed by the null operator movement in (i-ii), but both *zhī* and *suǒ* did appear there. However, all of the environments in which they appear are four-syllable environments, and we could not find examples of $*[bèi suǒ V]$ or $*[bèi V zhī]$. This indicates that it is the four-syllable environment (i.e., the prosody) that forces *zhī* and *suǒ* to appear, making them function as a place holder.

following V must be a single predicate. This undoubtedly supports the analysis that [*bei* V] is a compound, not only in Modern Chinese, but also in Classical Chinese.

Fourth, contrary to the case of resumptive pronoun, an NP is very commonly seen to occur after the [*bei*-V] compound, as seen in (26) below.

- (26) a. 行至雪山南畔，當被劫奪國策及諸敕信。(Dunhuang Bianwen 敦煌變文, 8-10th c.)
 (Wáng Duānzhāng) xíng zhì xuě shān nán pàn,...
 (Wang Duanzhang) travel arrive snow mountain south side
 dāng bèi jiéduó guócè jí zhū chì xìn.
 immediately get seize credential and every imperial letter.
 ‘When Wang Duanzhang traveled to the south of the snow mountain,...
 he got his credentials and every imperial letters seized.’
- b. 被損落一板齒。(Dunhuang Bianwen 敦煌變文, 8-10th century)
 (Qín Gùrán) bèi sǔn-luò yì bǎnchǐ.
 (Qin Guyan) et harm-fall one incisor
 ‘Qin Guyan got his incisor knocked out.’
- c. 任奸終被變形儀。(Dunhuang Bianwen 敦煌變文, 8-10th century)
 Rèn jiān zhōng bèi biàn xíng-yí.
 indulge evil finally get change appearance-bearing
 ‘If one indulges evil, he will finally get his appearance and bearing changed.’

The non-existent [NP_i [*bei*-V]_V zhi_i] and the commonly occurring [NP *bei*-V NP] follow exactly what the theory predicts. That is, the [*bei*-V] is a verb compound hence a pronoun is not allowed by Principle B of the Binding Theory, and since the NP would not be affected by the binding requirement, it is legitimate to appear in the object position.

Fifth, if the [NP *bei* V NP] sentences are grammatical in Classical Chinese, the [NP *bei* ADV V NP] structure should also turn out to be grammatical. This is because if [*bei*-V] is a compound and thus does not reject the NP object, the [*bei* ADV V] would act similarly in allowing an NP object because according to McCawley’s analysis we adopt here, the [*bei*-ADV-V] is a verb compound as well. This, as seen in (27), is exactly the case.

- (27) a. 龜被生揭其甲。(Lingbiao Luyi 嶺表錄異, ca. 850 AD)
 Guī bèi shēng jiē qí jiǎ.
 turtle BEI lively take-off its shell
 ‘The turtle’s shell was taken off while it was still alive.’

- b. (燕子) 橫被強奪窟。 (*Dunhuang Bianwen* 敦煌變文, 8-10th century)
 (Yànzǐ) héng bèi qiáng duó kū.
 (swallow) flagrantly get forcibly seize nest
 ‘The swallow got his nest seized flagrantly and forcibly.’

Undoubtedly, *shēng* ‘alive’ and *qiáng* ‘forcibly’ were V-modifiers in Ancient Chinese, hence the [*bèi-shēng/qiáng*-V] is a legitimate verb compound. The theory predicts that if the [*bei*-ADV-V] sequence is a legitimate compound, it can only take an NP, rather than a resumptive pronoun, as its object. The existent [*bei*-ADV-V NP] and the non-existent *[NP_i *bei*-ADV-V pro_i] indicate, once again that the [*bei*-ADV-V] in Classical Chinese must be a verb compound.⁷

5. Verb incorporation

So far we have seen that [*bei*-V] and [*bei*-ADV-V] are compounds in both Modern and Archaic Chinese. While the compound analysis has solved the V-adverbial and the resumptive pronoun problems, there are still questions difficult to answer.

First, if the [*bei*-ADV-V] is on a par with the [*bei*-V] compound such as *bèi-bǔ* ‘got arrested’ (28) in Modern Chinese, and *bèi-wéi* ‘got surrounded’ (24c) in Archaic, an immediate question is why the [*bei*-V] compounds cannot take an adverb to form a [*bei*-ADV-V] compound. For example:

- (28) *張三被秘密地捕了。
 *Zhāngsān bèi mìmì.de bǔ le.
 Zhangsan BEI secretly arrest ASP
 Intended meaning: ‘Zhangsan got arrested secretly.’
- (24) c. *被未圍于趙
 bèi wèi wéi yú Zhào (=not attested)
 Intended meaning: ‘got not surrounded by Zhao’

Bèi-bǔ in Modern Chinese and *bèi-wéi* in Archaic Chinese are compounds hence no intervention of adverbs is allowed between the *bei* and the verb. If the [*bei*-ADV-V]

⁷ There is a part-whole relation between the subject and the retained object and this relation can be captured by assuming that the subject of the *bei* passive is an experiencer base-generated on the Spec of IP, and the embedded VP is a predicate of the *bei* (see Huang 1999). Thus, the event/action denoted by the VP must be something that has ‘happened to X’ where X is the experiencer. This thematic structure will give rise to a part-whole relation observed in examples like (27). I would like to thank the anonymous reviewer for pointing out this question to me.

forms are also compounds, what is the difference between the [*bei*-ADV-V] compounds which allow an adverb inside the [*bei* V] and the [*bei*-V] compounds which do not?

Second, as the following examples in (29) show, the lexical [*bei*-V] compounds (both Modern and Archaic) cannot take an object, but an NP can easily appear in the object position of [*bei* VP] short passives, as seen in (4) for Modern Chinese and (15) for Archaic Chinese.

- (29) a. *他們家被捕了一個人。
 *Tā.men jiā bèi bǔ le yí.ge rén.
 they family BEI arrest ASP one.CL person
 Intended meaning: ‘Their family had one person arrested.’
- b. *他們家被警察捕了一個人。
 *Tā.men jiā bèi jǐngchá bǔ le yí.ge rén.
 they family BEI policeman arrest ASP one person
 Intended meaning: ‘Their family had one person arrested by a policeman.’
- c. *萬乘之國被圍城 (in Archaic Chinese)
 Wàn shèng zhī guó bèi wéi chéng. (=not attested)
 ten.thousand chariot’s country get surround city.
 Intended meaning: ‘A state with ten thousand chariots got surrounded
 *their city.’

Both the Modern and the Archaic [*bei*-V] compounds reject an object NP and both the Modern and the Ancient [*bei*-ADV-V] compounds accommodate an object NP⁸. Why is that so?

Third, [*bei*-V] compounds are (almost) all disyllabic forms and they are undoubtedly lexical items, while the [*bei*-ADV-V] forms can hardly be considered, and have never been used as individual lexical items. Of course sizes and usages cannot be taken as an argument against a syntactic analysis of compounding. However, together with their different behaviors with respect to the intervention of adverbs and the ability to take objects, we have good reasons to conclude that the [*bei*-V] passives are different from the [*bei*-ADV-V] ones. The former is a morphological compound, while the latter is

⁸ Aside from Archaic Chinese, in Modern and Classical Chinese there are some [*bei* V] forms which can also take an NP object, as shown in (15). According to the present analysis, the [*bei* V NP] forms have a D-structure like [[*bei* [V NP]]], rather than [[*bei*-V] NP], even though after VI, the [*bei* [V NP]] becomes [[*bei*-V] NP], as we will see later. This is so because otherwise there would be no distinction between the grammatical [*bei* [V NP]], as in (15), and the ungrammatical [[*bei*-V] *NP], as in [[*bèi-bǔ*] **rén*] 被捕人 ‘got arrested *person’ in Modern Chinese and [[*bèi wéi*] **chéng*] 被圍城 ‘got surrounded *city’ in Archaic Chinese.

more like a phrase-word (Zwicky 1990:206) which involves an utterance larger than a word. Just as Huang (1999) characterizes, it is “combining *bei* with a verb phrase to form a ‘passive verb phrase’”. As a result, even though the [*bei*-ADV-V] forms should better be analyzed as a compound as argued by McCawley and supported by historical facts given above, they seem to be a different type of compounds from the [*bei*-V] forms, given that “compound” has been used as a cover term for a collection of related, but by no means identical phenomena in the literature.

Considering the lexical property of [*bei*-V] compounds and the phrasal characteristics of the [*bei*-ADV-V] passives, I suggest that the former be seen as morphological compounds formed by (prosodic) morphological rules, whereas the latter are phrasal compounds derived by a syntactic operation of Verb Incorporation (Feng 2001). There are good reasons to consider this hypothesis to be highly plausible, as seen below.

First, Verb Incorporation is a process of both morphological and syntactic structure: syntactic in that its distribution and consequences for the structure are determined by syntactic principles; morphological in that the resulting [V+V] structure acts like a compound. Obviously, those properties are what we want the [*bei*-ADV-V] to have — a syntactically determined phrasal compound.

Second, by Verb Incorporation, the [*bei*-ADV-V] compound would have a biclausal underlying structure and a combined surface structure. This is also what we expect. If [*bei*-ADV-V] is derived from a source containing two verbs (*bei* and the embedded V), the adverb is naturally expected to appear before the embedded verb. When VI takes place, the verb could first form a compound with the adverb and then combine with *bei*, yielding a larger verb form. As a result, the compound properties and the phrasal characteristics can both be captured in the resulting complex predicate.

Third, historical facts provide solid evidence for the possibility of VI in the [*bei*-ADV-V] structure. As seen below, the [*bei*-ADV-V] has a [*bei* VP] origin developed independently in the history of Chinese. As argued by Feng (1998), there were three types of *bei* passive structures clearly attested in the history of Chinese:

I. [*bei* V] compound passive (ca. 300 BC)

- (30) I. 萬乘之國被圍于趙。(Zhanguoce 戰國策, Qice 齊策, ca. 300 BC)
 Wàn shèng zhī guó bèi wéi yú Zhào.
 ten thousand chariot 's country get surround by Zhao
 ‘A state with ten thousand chariots got surrounded by Zhao.’

II. [bei VP] phrasal passive (ca. 100 AD)

- II.a. 橫被暴誅 (*Wuyuechunqiu* 吳越春秋, *Helü Neizhuan* 闔閭內傳, ca. 50 AD)
 héng bèi bào zhū
 unexpectedly get violently kill
 ‘got violently killed unexpectedly.’
- II.b. 紛遭外禍，累害之也。 (*Lunheng* 論衡, *Leihai* 累害, ca. 100 AD)
 Fēnzāo wài huò, lěi hài zhī yě...
 suffer external disaster accumulatively harm them PRT
 生動之累，咸被累害。
 shēng-dòng zhī lèi, xián bèi lěi hài.
 live-move ’s kind all get accumulatively harm
 ‘One suffers from disasters caused by external factors (because) the situation accumulatively harmed them...all kinds of living and moving creatures will get accumulatively harmed.’
- II.c. 被棺殮 (*Lunheng* 論衡, *Duizuo* 對作, ca. 100 AD)
 bèi guān liàn
 get coffin bury
 ‘got coffin-buried’
- II.d. 多被陰害 (*Taipingjing* 太平經, ca. 100 AD)
 duō bèi yīn hài
 frequently get invisibly harm
 ‘frequently got invisibly harmed.’
- II.e. 即被怒責 (*Weishu* 魏書, *Guangpingwang zhuan* 廣平王傳, ca. 550 AD)
 jí bèi nù zé
 then get angrily scold
 ‘immediately got angrily scolded.’
- II.f. 嘗被急召 (*Nanshi* 南史, *Renfang zhuan* 任昉傳, ca. 650 AD)
 cháng bèi jí zhāo
 ever get urgently call
 ‘once got urgently called in’

III. [*bei* IP] long passive (ca. 200 AD)⁹

III.a. 臣被尚書召問。(Cai Yong 蔡邕, *Beishoushibiao* 被收時表 132-192 AD)

Chén bèi Shàngshū zhāo-wèn.

I get Minister call-ask

‘I, the subject, got questioned by the Minister.’

III.b. 被明公逼，且喜且懼。(Dongguan Hanji 東觀漢記, *Min Gong zhuan* 明公傳, ca. 200 AD)

(Wú) bèi Míng-Gōng bī, qiě xǐ qiě jù.

(I) get Honorable-duke summon also glad also fear

‘I was summoned by (your) Honorable-Duke, I felt both happiness and fear.’

III.c. 被火氣劫故也。(Shanghanlun 傷寒論, ca. 200 AD)

Bèi huǒqì jié gù yě.

get fire attack reason PRT

‘Because (the body) was attacked by internal heat.’

As seen in §4, the [*bei*-V] forms in Archaic Chinese were compounds and the compound passives are kept in Modern Chinese as shown in (16). On the other hand, the examples given in (30-II) must not be analyzed as a compound: an adverb appears before the V, indicating that the V can take a modifier of its own and hence the [ADV V] is not part of a compound but an independent constituent (see Feng 1998) for more arguments). This type of example is extremely important because if the [ADV V] is not part of a compound, it must be analyzed as a VP embedded under *bei*. If so, it follows that the [*bei* VP] structure was developed before the long passives in a chronological order of [*bei* VP] → [*bei* IP]. This further entails that the [*bei* VP] structure was established independently of the long passives in history. This discovery confirms Ting’s hypothesis that there are two *bei* structures in Modern Chinese: one is a long passive ([*bei* IP]) and the other is a short ([*bei* VP]). Given the historical evidence, the questions why there are two BEIs in Modern Chinese and how they come about are no longer a mystery: the two *bei* structures are inherited from Classical Chinese. In fact, the [*bei* VP] short passive is not only independent of the [*bei* IP] long passive, but also older than the long form in the history. This provides solid evidence for the argument

⁹ *Shanghanlun* 傷寒論 was completed by 205 AD, and Cai Yong 蔡邕 (132 AD-192 AD) was the last person who was involved in compiling the *Dongguan Hanji* 東觀漢記. Given this, examples in (30III) can be taken as the earliest full passives, dating from around 200 AD.

that the short passives ([*bei* VP]) are not derived from deletion of the agent NP in long passives, since when the [*bei* VP] structure was established, the [*bei* IP] structure had not even appeared. The diachronic fact shows that not only the short did not come out of the long, but the other way around might be the case, that is, the long passives were developed from the [*bei* VP] short passives by introducing the agent NP into the embedded clause, creating an IP when the lexical strength of the matrix verb *bei* was weakened from a transitive (in [*bei* VP]) to an intransitive (or may be an auxiliary) verb (in [*bei* IP]).

According to previous studies (Peyraube 1989, 1996, Tang 1987, Wei 1994, and many others), a relative chronological order of the developments of all *bei* passives can be summarized as follows:

Table 1: Relative chronology of *bei* passives in Classical Chinese

Chronology	Origin	Step-I	Step-II	Step-III
400 BC	<i>bei</i> NP			
300 BC	<i>bei</i> NP	<i>bei</i> N/V		
200 BC	<i>bei</i> NP	<i>bei</i> V		
100 AD	<i>bei</i> NP	<i>bei</i> V	<i>bei</i> ADV V	
200 AD	<i>bei</i> NP	<i>bei</i> V	<i>bei</i> (ADV) V	<i>bei</i> NP V
400 AD	<i>bei</i> NP	<i>bei</i> V	<i>bei</i> (ADV) V (NP)	<i>bei</i> NP V (NP)
...				
Modern		<i>bei</i> V	<i>bei</i> (ADV) V (NP)	<i>bei</i> NP V (NP)

Table 1 shows that the *bei* sentences originated from a [*bei* NP] structure and further developed a [*bei* N/V] alternation. As seen before, only in disyllabic environments can the [*bei* N/V] alternation be observed and the N be reanalyzed as V, giving rise to the [*bei*-V] compound passive. The [*bei*-V] compound passive was further developed into a [*bei* [ADV V]_{VP}] phrasal passive around the Eastern Han dynasty (ca. 100 AD) and finally gave rise to the full passive [*bei* [NP V]] structure by the late Han (ca. 200 AD). What is important to note here is the fact that every step of the development of the *bei* sentence was realized by the creation of a new structure: first the [*bei*-V] compound passive, then the [*bei* VP] phrasal passive, and finally the [*bei* IP] full passive. Each structure is historically related but syntactically independent: related in the sense that the latter form would not have come into being without the former one; independent in the sense that the former was not destroyed by the latter one. Instead they both survived and developed into their respective present versions. They had coexisted throughout the history. This is why in Modern Chinese there are still [*bei*-V] compound passives and why the [*bei* VP] and [*bei* IP] behave differently syntactically. Of course, this does not

mean that they would not influence each other. In fact, as seen in Table 1, the appearance of [*bei* V NP] cases was found only after the birth of the [*bei* IP] structure, suggesting that the [*bei* V NP] was further developed under the influence of [*bei* IP].

Given the historical evidence above, we now have a solid ground to propose that the [*bei* VP] structure is independent of the [*bei*-V] compound structure and of the [*bei* IP] full passive structure. If this is true, it is only natural to expect an independent syntactic operation involving only the [*bei* VP] structure. As we will see below, the Verb Incorporation is permitted only in the [*bei* VP] structure.

First, the [*bei*-V] compound would not involve the VI, simply because they are formed in morphology. Second, the [*bei* IP] cannot be derived by VI either, because as argued by Li Yafei (1990), the lower verb in an IP complement cannot be incorporated into the matrix verb, as shown in (31)

- (31) $_{VP1}[V_1 \text{ IP}[I \text{ }_{VP2}[NP \text{ }_v[V_2 \text{ NP}]]]]$
 bèi zhūwù fàn zhī (被諸物犯之)
 get everything attack it
 ‘got attacked by everything.’

By Minimality Condition (Rizzi 1990), V_2 must first move to I, and then to V_1 , in order to form a complex [V_1+V_2] compound. However, such a movement results in what Li Yafei has called a sandwiched chain: [V, I, V], which is not allowed by Binding Condition C in the theory developed in Y. Li (1990). Therefore, the embedded verb in [*bei* IP] is not incorporable with *bei*. The situation, however, is completely different in the [*bei* VP] structure:

- (32) $_{VP1}[V_1 \text{ }_{VP2}[V_2 \text{ NP}]]$
 bèi qǔ dǎn (被取膽)
 get take gall

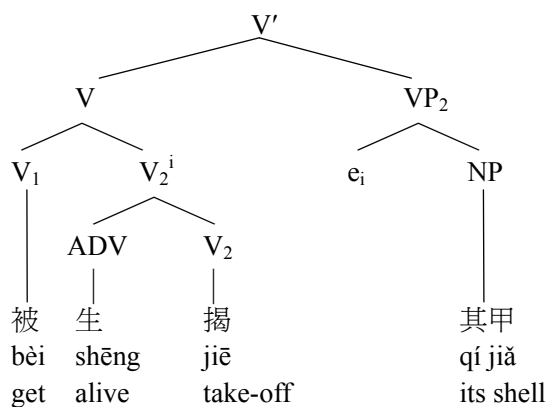
The V_2 can be minimally moved into V_1 forming a [V, V] chain, and no violation of any kind would be caused. Hence VI is permissible only in the [*bei* VP] structure, given the following surface structure:

- (33) $_v[_v[V_1-V_{2i}] \text{ }_{VP2}[e_i \text{ NP}]]$
 bèi-qǔ_i t_i dǎn (被取膽)
 get-take gall

The advantages of the VI-analysis can also be seen from several aspects. First, morphology and modal verbs have been universally considered as triggers for Verb

Incorporation. The Verb Incorporation in [*bei* VP] may also be subject to a morphological factor: the monosyllabic *bei* cannot stand alone as an independent verb after the grammaticalization in the historical development of the *bei* passive construction. Huang (1999) distinguishes two stages of grammaticalization of *bei* in history: originally *bei* was a transitive verb with two arguments. Later it has weakened to the status of an intransitive and possibly even to the status of an auxiliary or a bound morpheme. The morphological status of being an auxiliary or a bound morpheme provides good reasons for *bei* to be considered as a trigger for the VI.¹⁰ If this is so, it follows that VI would not happen until *bei* was weakened. It is commonly agreed that *bei* was weakened after the establishment of the full passive [*bei* IP] (by the Eastern Han) and it is true that only after the appearance of the [*bei* IP] long form, could one find the [*bei* V NP] short forms, as seen in Table 1. This strongly favors a VI analysis for the [*bei* V NP] sentences. If *bei* needs support morphologically and if *bei* takes only a bare VP complement in [*bei* [V NP]_{VP}], VI is forced in this situation. That is, VI must also take place in [NP *bei* [ADV V NP]_{VP}], so that it first triggers a [ADV-V] compound and then the [ADV-V] is incorporated into *bei*, forming a [*bei*-ADV-V] complex, as shown in (34).

(34)



¹⁰ VI may also be triggered by the oldest [*bei*-V] compounds, in that *bei* always tends to form a compound with a monosyllabic verb and the tendency may finally become a general requirement when *bei* loses its lexical strength and is weakened to an auxiliary. Thus VI must take place in the environment where *bei* takes only a bare VP. This is possible for the following reasons: if there were no [*bei* V] compounds, there would have been no [*bei* Adv V] forms in Mandarin, given the fact that no short passives, except the *bei* forms, are allowed in Mandarin, or in other Chinese dialects, because other passive verbs, such as *ho* 與 (Taiwanese) or *ràng* 讓 (Mandarin), are not developed from a compound.

Second, the asymmetry between [*bei*-V NP] and [*bei*-V **pro*] can also be captured along this line of reasoning. Since the object NP in [[*bei* V] NP] can still be assigned a case after Verb Incorporation (see Baker 1988, Y. Li 1992), [[*bei* V] NP] sentences are/were grammatical throughout Chinese history. An object pronoun is not allowed by VI, however, since incorporation changes government relations by allowing the (matrix) verb to govern something which it otherwise would not have governed (Baker 1988: 101). For example, in Kinyarwanda, P Incorporation changes government relations (taken from Baker 1988:496):

- (35) a. Abaana ba-ra-shyir-a ibitabo kuri bo.
 children SP-PRES-put-ASP books on them
 ‘The children are putting books on themselves.’
 b. Abaana ba-r-ii-shyir-a-ho ibitabo.
 children SP-PRES-REFL-put-ASP-on books
 ‘The children are putting books on themselves.’

In (35a), the [NP, PP] is a lexical pronoun and may be coreferential with the matrix subject (the PP counts as ‘Complete Functional Complex’ and hence binding domains in Kinyarwanda). However, when the P is incorporated into the verb as in (35b), morphological reflexivization must apply. This is because, as argued by Baker (1988), the binding domain of the locative NP has been expanded by Incorporation. That is, the Incorporation causes the object of P to be externally governed by the matrix verb. Thus its binding domain must include that matrix verb and the matrix subject. What is important here is the fact that when the government relation is changed after the Incorporation process, the governing category for binding relations is also forced to extend to the entire matrix clause (see Baker 1988:101 for parallel cases of NI examples of possessor stranding).

Parallel to the P Incorporation in Kinyarwanda, when the embedded V is incorporated into the matrix verb *bei* in Chinese, an object pronoun is also disallowed to be coreferential with the subject. This is because the object position of the embedded verb is externally governed by the *bei* after the VI, thus the governing category for the embedded object pronoun is not the embedded VP, but the entire matrix clause.¹¹

¹¹ One may wonder how to account for the following ungrammatical sentence in a VI analysis:

(i) *教室被派約翰找人去掃乾淨了。

* Jiàoshì_i bèi pài Yuēhàn zhǎo rén qù sǎo-gānjìng t_i le
 classroom BEI send John find men to sweep-clean ASP

Intended meaning: ‘*The classroom was sent John to find people to clean.’

Note that without the agent the embedded clause is a VP, rather than an IP. The [NP_i bèi_i [VP

Another argument in favor of the Verb Incorporation for [*bei* VP] is that this analysis now has the virtues of both the compound property (as discussed in §4) and the phrasal characteristics (as discussed in §5) of the [*bei* ADV V NP] short passives. It also has the virtues of both the non-lexical and non-phrasal properties of the forms. Recall that an important feature for morphological (verb) compound is the prosodic constraint of disyllabicity. This constraint is obviously not followed by the verb complex. Now the distinction between the lexical compounds and the complex verb compounds can be made clearly: lexical compounds are formed in (Prosodic) Morphology and must obey prosodic morphological rules, while complex verb compounds are derived in syntax via VI from natural phrases. No length constraint is required for syntactic phrases, hence complex verbs are not constrained by prosodic morphological rules. As a result, lexical compounds always coincide with PrWds, while complex verb compounds involve an utterance larger than a word in size. This is why in both Classical and Modern Chinese lexicon, there are no lexical [*bei*-V] compounds that are formed into a size longer than a PrWd.¹²

6. Summary

In this paper I have argued from both synchronic and diachronic perspectives that short passives in Chinese can be divided into two types of constructions: one is viewed as a lexical compound which supports Ting's analysis (1996); while the other is viewed as a complex verb, which accords with McCawley's argument (1992). Furthermore, if the chronology of emergence of the [*bei* VP] forms is taken into account, three types of operations ranging from morphology (the [*bei*-V] compound), to morphosyntax (the [[*bei* V] NP] generated by verb complex via Verb Incorporation) and to phrasal syntax (the [*bei* V__] derived by NP-movement) can all be observed in history. The varieties of *bei* forms are not only distinguished chronologically but also structurally, and all of

pài ..**t_i*...]] must be ruled out because VI, as seen above, requires a bare VP to be the complement of *bei*, hence there is neither [Spec, CP] nor [Spec, IP] in the embedded clause. However, the *t_i* in (i) is left by a long-distance movement; hence it must be an A'-trace. An A'-movement needs a position like [Spec, CP] or [Spec, IP] in order for the deeply embedded element to be moved out or to be coindexed with its antecedent. Since the bare VP in VI does not have a position of the sort, (i) must be ruled out.

¹² This holds without exception in Classical Chinese. In Modern Chinese, however, there seems to be an exception in *bèi-yāpò jiējí* 'be-oppressed class'. However, *bèi-yāpò* can occur in *bèi lǎobǎn yāpò* 'be oppressed by the boss', indicating that *bèi yāpò* is different from lexical compounds such as *bèi-pò* 'be forced' that cannot be separated at all. As a result, *bèi yāpò* is not a lexical compound.

them are kept in Modern Mandarin Chinese. It follows that a syntactic analysis must not only take care of the synchronic facts but also their histories. The Modern *bei* sentences are typical examples of this sort: all four structures, the [*bei*-V], the [*bei* V __], the [*bei* V NP] and the [*bei* IP] are handed down from Classical Chinese.

The implication of the present analysis is that while synchronic analyses should refer to historical facts, diachronic studies can equally well cast light on synchronic research. Clearly the syntactic structure of Classical long passives are captured by the A'-movement analysis developed from Mandarin Chinese (Feng 1990) and the [*bei*-V] compound analysis for Archaic Chinese is significant in understanding the synchronic behaviors of [*bei* V] short passives in Modern Chinese. From the synchronic analysis, we draw insights that allow us to account for long passives in Classical Chinese, and based on this diachronic study we were able to discover a variety of short passives in Modern Chinese.

References

- Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Cheng, Lisa L.-S. 1986. *Clause Structure in Mandarin Chinese*. Toronto: University of Toronto MA thesis.
- Cheng, Lisa L.-S., C.-T. James Huang, Y.-H. Audrey Li, and Chih-Chen Jane Tang. 1999. *Hoo, hoo, hoo: syntax of the causative, dative, and passive constructions in Taiwanese*. *Contemporary Studies on the Min Dialects*, ed. by Pang-Hsin Ting, 146-203. Berkeley: Project on Linguistic Analysis, University of California.
- Chiu, Bonnie. 1995. An object clitic projection in Mandarin Chinese. *Journal of East Asian Linguistics* 4.2:77-117.
- Chomsky, Noam. 1988. *Lectures on Government and Binding: The Pisa Lectures* (5th edition). Dordrecht: Foris.
- Chomsky, Norm. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Chou, Fa-kao. 1962. *Zhongguo Gudai Yufa [Classical Chinese Grammar]*. Taipei: Institute of History and Philology, Academia Sinica.
- Feng, Shengli. 1990. *The Passive Construction in Chinese*. Manuscript. Philadelphia: University of Pennsylvania.
- Feng, Shengli. 1995. GB-theory and the passive construction in Chinese. *Studies in Chinese Linguistics* 1:1-28.
- Feng, Shengli. 1997. Prosodic structure and compound words in Classical Chinese. *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon*

- in Modern and Ancient Chinese*, ed. by Jerome L. Packard, 197-260. Berlin & New York: Mouton de Gruyter.
- Feng, Shengli. 1998. Prosodically motivated passive *bei* constructions in Classical Chinese. *SKY (Suomen kielitieteellinen yhdistys) Yearbook 1998*, 41-48. Helsinki: The Linguistic Association of Finland.
- Feng, Shengli. 2001. Lun ci de duoweixing [The multidimensional properties of word]. *Dangdai Yuyanxue [Contemporary Chinese Linguistics]* 2001.3:161-174.
- Goodall, Grant. 1991. Mandarin Chinese passives and the nature of case absorption. Paper presented at the third North American Conference on Chinese Linguistics (NACCL-3). Ithaca: Cornell University.
- Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge: MIT dissertation.
- Huang, C.-T. James. 1999. Chinese passive in comparative perspective. *Tsing Hua Journal of Chinese Studies*, New Series 29.4:423-509.
- Kayne, Richard S. 1975. *French Syntax: The Transformational Cycle*. Cambridge: MIT Press.
- Kroch, Anthony S. 1989. Asymmetries in long-distance extraction in a tree-adjoining grammar. *Alternative Conceptions of Phrase Structure*, ed. by Mark R. Baltin & Anthony S. Kroch, 66-98. Chicago: University of Chicago Press.
- Li, Grace Hui-Ju. 1994. *Bei: An Affected Head – The bei Construction Revisited*. Taipei: National Taiwan Normal University MA thesis.
- Li, Yafei. 1990. *Conditions on X⁰-movement*. Cambridge: MIT dissertation.
- Li, Y.-H. Audrey. 1985. *Abstract Case in Chinese*. Los Angeles: University of Southern California dissertation.
- Li, Y.-H. Audrey. 1990. *Order and Constituency in Mandarin Chinese*. Dordrecht & Boston: Kluwer.
- Li, Y.-H. Audrey. 1992. Indefinite *wh* in Mandarin Chinese. *Journal of East Asian Linguistics* 1.2:125-155.
- Liberman, Mark, and Alan Prince. 1977. On stress and linguistic rhythm. *Linguistics Inquiry* 8.2:249-336.
- Lightfoot, David. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge: MIT Press.
- Luo, Genze. 1981. *Guanzi Tanyuan [An Inquiry of Guanzi's Origin]*. Taipei: Lejin Books.
- McCarthy, John, and Alan Prince. 1993. *Prosodic Morphology I*. Manuscript. Amherst: University of Massachusetts; New Brunswick: Rutgers University.
- McCawley, James D. 1992. Justifying part-of-speech assignments in Mandarin Chinese. *Journal of Chinese Linguistics* 20.2:211-246.
- Norman, Jerry. 1988. *Chinese*. Cambridge & New York: Cambridge University Press.

- Pierre, Pica. 1987. On the nature of the reflexivization cycle. Manuscript. Cambridge: MIT.
- Peyraube, Alain. 1989. History of the passive constructions in Chinese until the 10th century. *Journal of Chinese Linguistics* 17.2:335-372.
- Peyraube, Alain. 1996. Recent issues in Chinese historical syntax. *New Horizons in Chinese Linguistics*, ed. by C.-T. James Huang & Y.-H. Audrey Li, 161-213. Dordrecht & Boston: Kluwer.
- Reynolds, Robert. 1995. "Passive" morphology in Mandarin. Paper presented at the 7th North American Conference on Chinese Linguistics (NACCL-7). Madison: University of Wisconsin-Madison.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge: MIT Press.
- Tang, Yuming. 1987. Han Wei Liuchao beidongshi luelun [A brief survey of the passive construction in Chinese of the period from 206 BC to 589 AD]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 1987.3:216-222.
- Ting, Jen. 1995. *A Non-Uniform Analysis of the Passive Construction in Mandarin Chinese*. New York: University of Rochester dissertation.
- Ting, Jen. 1996. A non-uniform analysis of the passive construction in Mandarin Chinese. Paper presented at the 8th North American Conference on Chinese Linguistics (NACCL-8). Urbana-Champaign: University of Illinois.
- Wang, Li. 1989. *Hanyu Shi Gao* [*An Outline of Chinese Language History*]. Beijing: Zhonghua Book Company.
- Wei, Pei-chuan. 1994. Gu Hanyu beidongshi de fazhan yu yanbian jizhi [The mechanism of the development and changes of passive constructions in Classical Chinese]. *Chinese Languages and Linguistics*, Vol. 2: *Historical Linguistics*, ed. by Paul Jen-kuei Li, Chu-Ren Huang & Chih-Chen Jane Tang, 293-319. Taipei: Institute of History and Philology, Academia Sinica.
- Xie, Yunfei. 1983. *Guanzi Xilun* [*Analysis of Guanzi*]. Taipei: Student Book.
- Zwicky, Arnold M. 1990. Syntactic words and morphological words, simple and composite. *Yearbook of Morphology* 3 (1990), ed. by Geert Booij & Jaap van Marle, 201-216. Dordrecht: Foris.

Chinese Resultative Verb Compounds: Lexicalization and Grammaticalization

Chaofen Sun (孫朝奮)

Stanford University

This paper is an historical study of the formation of the Chinese resultative verb compounds (RVCs) that signal a resultant state of a non-agent with a V1V2 predicate. Metaphorization and metonymization, understood within the theoretical framework of Brinton & Traugott (2005), are proposed to have played a most important role in the formation of the RVC in Middle Chinese. Many scholars noted (Wang 1958, Ota 1958, and Mei 1991) that the Modern Chinese RVCs as a linguistic form did not exist in Old Chinese. Mei (1991) further claimed that the neutralization of the second verb in a matching transitive verb-verb sequence gave rise to the present-date RVC with a transitive-intransitive format. This paper, however, demonstrates that, in light of the polysemous structure of many verbs and Levin & Hovav's theory on argument realization (2005), the ambitransitive nature of the verbs in these verbal sequences ultimately was responsible for the unification of the two verbs as a process of metonymization exemplified by hyperbole and its spread as a process of metaphorization. Moreover, it is also argued that, on the basis of several formal distinctions, Chinese RVCs can be separated into two different categories: those resulting from lexicalization and those resulting from grammaticalization.

Key words: resultative verb compounds, grammaticalization, lexicalization, ambitransitive

1. Introduction

Over the last two millennia, hardly any other change in Chinese history has impacted Chinese syntax more notably than the formation of the Chinese resultative verbs (hereafter RVC) in Middle Chinese¹ as it directly fed into, among other things, the emergence of the BA construction, the potential construction, etc. (Ota 1958, Mei 1991,

¹ In this paper, Old Chinese is conveniently used to refer to the language of historical texts from the 5th century BCE to the 2nd century CE, Middle Chinese from the 3rd to 10th century CE, Early Modern Chinese from the 11th century to 19th century, and Modern Chinese from the 20th century to the present day.

Sun 1996, and Liang 2006). As compared to the Old Chinese grammar, it drastically changed the landscape of Chinese syntax in subsequent times. Thus, a good understanding of the formation of the Chinese RVC is essential in the study of the history of Chinese syntax.

In this paper, a Modern Chinese RVC construct can be taken to be a unified verb compound composed of two verbal morphemes such as *da-po* ‘hit-break’, *kan-wan* ‘see-finish’, *he-zui* ‘drink-drunk’, and *tui-chu* ‘retreat-exit’ as shown in Example (1), with the second verb (hereafter V2) indicating the result of the action signified by the initial verb (hereafter V1). This kind of compound forms an independent lexical word, as the verbal suffix *le* signaling perfective aspect cannot be inserted between the two unified components, as shown in example (1e) (Li & Thompson 1981).

- (1) a. 他打破了杯子
 ta da-po-le beizi
 3rd hit-break-ASP cup
 ‘He broke the cup.’
- b. 他看完了這本書
 ta kan-wan-le zhe-ben shu
 3rd see-finish-ASP DEM-CL book
 ‘He finished reading the book.’
- c. 他喝醉了酒
 ta he-zui-le jiu
 3rd drink-drunk-ASP liquor
 ‘He got drunk.’
- d. 他退出了比賽
 ta tui-chu-le bisai
 3rd retreat-exit-ASP competition
 ‘He withdrew from the competition.’
- e. *打了破、 *看了完、 *喝了醉、 *退了出
 da-le-po kan-le-wan he-le-zui tui-le-chu
 hit-ASP-break see-ASP-finish drink-ASP-drunk retreat-ASP-exit

Functionally speaking, a Chinese RVC serves to predicate a non-agent argument signaling a resultant state. In terms of the semantic case roles of its arguments, the RVCs are not totally consistent. In (1a) the object *beizi* ‘cup’ occurring after an RVC is the patient of V1 and V2 of the RVC, *da* ‘hit’ and *po* ‘break’. Accordingly, the resultant state is a predicate of the patient, i.e., *the broken cup* resulting from the action of

‘hitting’. In (1b), the object *shu* ‘book’ is the theme of V1 *kan* ‘see’ and *wan* ‘finish’. In other words, the resultant state ‘finish reading a book’ is a predicate of the theme. But in (1c) the postverbal object² *jiu* ‘liquor’ is the theme of the V1 ‘drink’ but not the V2 *zui* ‘drunk’, which is a predicate of the preverbal subject experiencer, *ta* ‘3rd person pronoun’, not a predicate of its theme object *jiu* of V1 *he*. It will be shown that throughout history, the Chinese RVCs have not been limited to predicating a syntactic object exclusively, but rather a non-agent element of any given RVC.

Moreover, traditionally examples such as *tui-chu* ‘retreat-exit’ in (1d) have an intransitive V2 indicating a resultant state of a theme, or an experiencer, and are treated as a type of RVC (Li & Thompson 1981).

In a Chinese RVC construct there are two essential constituents, V1 and V2. Typically, the former is a transitive verb (hereafter Vt) and the latter an intransitive verb (hereafter Vi) associated with a certain number of subcategorized arguments. Most V2s in (1), as intransitive verbs (Vi), can function on their own as free morphemes, such as the intransitive verb *si* ‘die’ in (2b).

- (2) a. 他打死了隻蒼蠅
 ta da-si-le zhi cangying
 3rd hit-die-ASP CL fly
 ‘He hit a fly (to death).’
 b. 蒼蠅死了
 cangying si-le
 fly die-ASP
 ‘The fly died.’

Note that an intransitive verb such as *si* ‘die’ usually takes an experiencer subject such as the one in (2b). But in an RVC, the object *cangying* ‘fly’ is the patient/object of V1 but a logical experiencer/subject of V2 *si* in (2a). In this paper, this fact is taken to imply that the V2 in a Chinese RVC, as a resultative component of a lexical item with a V1+V2 formation, typically does not participate in the argument selection, such as the sequence *da-si* in (2a), in which the *fly* is the patient object of V1 *da*. However, the intransitive V2 functions to predicate a resultant state, or extent of change, of a non-agent argument of the V1.

² Note that in this paper, the grammatical terms subject and object are used very loosely for ease of elucidation following the tradition adopted by most Chinese linguists. Typically, a subject refers to an agentive NP in a preverbal position, and an “object” an NP in patient/theme in a postverbal position.

Li & Thompson (1981) also define the Modern Chinese RVCs from other verb compounds by an infixal potential marker, *-de/bu-* ‘positive/negative potential’, such as the pairs in (3), in which the potential marker and V2 together indicate the extent of an event, for example, beat to a possible extent of dying, or not dying.

- (3) a. 打得/不死
da-de/bu-si
‘can/cannot be beaten to death’
- b. 打得/不破
da-de/bu-po
‘can/cannot be broken’
- c. 看得/不完
kan-de/bu-wan
‘can/cannot finish reading’
- d. 喝得/不醉
he-de/bu-zui
‘can/cannot drink to be drunk’

Those in (4) that cannot co-occur with the potential markers *de/bu* are treated as coordinate verb compounds by Li & Thompson (1981). Whereas the two morphemes in (4a) and (4b) can be treated as semantically parallel, it is, nevertheless, unclear how the sequences in (4c) and (4d) are semantically different from the resultative verb compounds in (3) as the same resultative relationship can be easily construed from these sequences, *kuo-da* and *gai-liang* ‘expand to a resultant state that is big’ and ‘change to a resultant state that is good’.

- | | | |
|--------|--|---------------------------|
| (4) a. | 批評
pi-ping
criticize-comment
‘to criticize’ | *批得/不評
pi-de/bu ping |
| b. | 加強
jia-qiang
add-strong
‘to strengthen’ | *加得/不強
jia-de/bu qiang |
| c. | 擴大
kuo-da
expand-big
‘to expand’ | *擴得/不大
kuo-de/bu-da |

- d. 改良 *改得/不良
gai-liang gai-de/bu-liang
change-good
‘to improve’

It then follows that RVCs as a semantic group may not be completely homogeneous and there can be RVCs of different kinds. Similar observations have been made by Liang (2006). In the following sections, I will show how this issue can be fully accounted for within a framework of lexicalization/grammaticalization proposed by Brinton & Traugott (2005).

Historically, many scholars (Wang 1958, Ota 1958, and Mei 1991) noted that this kind of Modern Chinese RVC as a linguistic category did not exist in Old Chinese. Section 2 deals with insights and insufficiencies in the literature in order to fully account for the formation of the Chinese RVCs. Section 3 proposes a new analysis of the Old and Middle Chinese verbs in terms of transitivity with some insights from Levin & Havov’s work (2005). Section 4 discusses how regularization motivated the emergence of the Chinese RVC as a process characteristic of metonymization. Section 5 summarizes the historical developments of Chinese RVCs within a lexicalization/grammaticalization framework (Brinton & Traugott 2005).

2. Previous studies on the formation of Chinese resultative verb compounds

It has been observed by many (Wang 1958, Ota 1958, Mei 1991, and Xu 2006) that in Old Chinese there were no RVCs, i.e., lexical items composed in the Vt+Vi format. All the verbal sequences in Old Chinese were parallel in nature. For example, even though in (5a) the two verbs in sequence can be construed to have a resultative relationship, they are, nevertheless, both transitive verbs, meaning 激 *ji* ‘to excite’ and 怒 *nu* ‘to irritate’ respectively, systematically different from the Modern Chinese RVCs which have a Vt+Vi format. Similarly, the two verbs in (5b) should also be analyzed as a sequence in a Vt+Vt format. Although V2 *zou* ‘to run’ looks like an intransitive verb, it was used as a transitive verb with a causative meaning ‘to make someone run’.

Vt+Vt:

- (5) a. 乃激怒張儀 (*Shiji* 史記: 蘇秦列傳)
nai **ji** **nu** Zhang Yi
then **excite** **irritate** Name
‘Then made Zhang Yi become angry.’

- b. 陳餘擊走常山王張耳 (Shiji 史記: 張丞相列傳)
 Chen Yu **ji** **zou** Changshan wang Zhang Er
 Name **strike run** Title king Name
 ‘Chen Yu strike (and cause) King Changshan Zhang Er to run.’

In other words, the constraint for two Old Chinese verbs to occur in this type of parallel construct was that the two verbs must share a common argument structure. The examples in (6) are two cases with a parallel construct encompassing two intransitive verbs, Vi+Vi.

Vi+Vi:

- (6) a. 聞陳王戰敗 (Shiji 史記: 蘇秦列傳)
 wen Chen wang **zhan bai**
 hear Name king **battle lose**
 ‘(Someone) hears that King Chen fought and lost.’

Vi+Vi:

- b. 恐帝長大後見怨 (Hanshu 漢書: 王莽傳)
 kong di **zhang da** hou jian yuan
 afraid emperor **grow big** after see resent
 ‘(Wang Yu) was afraid that after the emperor (Ping Di) would have grown up, he would face (the emperor’s) resentment.’

It is generally believed (Wang 1958, Ota 1958, Li 1984, Mei 1991, and Xu 2006) that the typical Modern RVC pattern, Vt+Vi, did not surface until Middle Chinese, which Ota (1958) and Mei (1991) treated as a case of neutralization in transitivity. Note that among the Middle Chinese data, there are cases in which the V2s must be treated as an intransitive verb such as the two cases of *si* ‘die’ in (7). Note that on the one hand, the Vi *si* in (7a) is followed by an objective noun phrase, in spite of the fact that it is an intransitive verb. On the other hand, as Old Chinese is a VO-type language, the argument in objective case typically is expected to occur after a verb. There is no such objective NP after the so-called transitive verb *shao* in (7b).

Vt+Vi+O:

- (7) a. 乃打死之 (You Ming Lu 劉義慶: 幽明錄)
 nai **da si** zhi
 then **hit die** 3rd
 ‘... then, beat it (the dog) to dead.’

when V2 was an intransitive verb of displacement such as the one in (10). What actually might have happened was that the kind of V2s that could occur in this construction expanded from Old Chinese verbs of displacement to commonly include intransitive verbs of other kinds in Middle Chinese.

Vt+NP+Vi

- (10) 持劍盾步走 (Shiji 史記: 項羽)
chi **jian** **dun** **bu** **zou**
hold sword shield step **run**
 ‘(They) held their sword and shield running.’

Nonetheless, no full account would directly follow from the four reasons as a whole. If the Old Chinese V1+V2 sequence must be parallel, i.e., Vt+Vt or Vi+Vi, this constraint would have required that a transitive verb, or an intransitive verb, must be followed by a V of the same kind. Logically, this matching constraint should have prevented the emergence of a Vt+Vi sequence of any kind, as it would have blatantly violated the very constraint. My hypothesis is that the neutralization may follow more closely from the ambitransitive nature of many of the Chinese verbs.

3. Historical ambitransitivity of many Chinese verbs

In this section, pertaining to the issue of neutralization in transitivity, it will be demonstrated that Middle Chinese was an essentially monosyllabic language without any overt transitivity markers and many verbs could easily be used either as transitive verbs or as intransitive verbs. Furthermore, I will show that at a certain time many verbs could be used predominantly one way but, at another time, this changed. Ambitransitivity, or the ability to function as either a transitive or an intransitive verb, therefore, may have led to ambiguity and been a more important contributing factor to the formation of the Chinese RVCs.

3.1 Ambitransitivity

The verb ‘to kill’ is generally considered to be one of the more prototypical transitive verbs, implying a volitional initiator of an event. Ota (1958) and Mei (1991) noted that whereas the parallel VV sequence in (11a), with *sha* ‘to kill’ as its V2, is a transitive pattern, (11b), with *si* ‘to die’ as its V2, is an intransitive pattern. Typically, the transitive pattern has an NP following V2 functioning as the common object of V1 and V2. In contrast, there is no NP following V2 in an intransitive pattern.

Vt+Vt:

- (11) a. 岸崩，盡壓殺臥者 (Shiji 史記: 外戚世家)
 an beng, jing **ya sha** wo zhe
 bank collapse all **press kill** lie one
 ‘The collapsed bank (bury) killed all the sleepers.’

Vi+Vi:

- b. 百餘人炭崩盡壓死 (Lunheng 論衡: 吉驗)
 bai yu ren tan beng jin **ya si**
 hundred more person mine collapse all **press die**
 ‘More than 100 people (bury) died in the collapse of the mine.’

However, if the transitivity properties of *sha* ‘kill’ and *si* ‘die’ are unambiguously so, *ya* ‘press’ must be capable of functioning either as a transitive or an intransitive verb. Otherwise, the Old Chinese matching constraint for a VV sequence would have prevented it from co-occurring with both transitive *sha* ‘kill’ and intransitive *si* ‘die’. The matching condition for the Old Chinese V1+V2 sequences dictated that either the Vt+Vt format or the Vi+Vi format is acceptable, but not *Vt+Vi, nor Vi+Vt. It appears that *ya* ‘press’ and *sha* ‘kill’ in (11a) both have an instrument as its subject and a patient as its object. In (11b), *ya* ‘press’ and *si* ‘die’ share an experiencer as the subject without an overt object. Then, if *ya* ‘press’ can be used either transitively or intransitively, one would wonder if transitivity for any given verb can be unequivocally determined.

As Liang (2006) and Xu (2006) noted, even though *sha* is used predominantly as a transitive verb and *si* predominantly as an intransitive one, in (12a) *sha* and *si*, with an experiencer subject, are predicates parallel to each other, functioning intransitively as a middle-voice verb meaning something close to ‘to die’. In contrast, the verb *si* is used transitively as a causative verb, with an agent and a patient in (12b), functioning somewhat like a transitive verb ‘kill’.

- (12) a. 秦孝王死，公孫鞅殺 (Zhangguo Zonghengjia Shu 戰國縱橫家書)
 Qin Xiao wang **si**, Gongsun Yang **sha**
 Name Name king **die** Name **kill**
 ‘King Xiao of Qin died, and Gongsun Yang died (too).’
- b. 崔子之徒以戈斫公而死之 (Hanfeizi 韓非子)
 Chui-zi zhi tu yi ge zhuo gong er **si zhi**
 Name REL disciple with ax smash lord and **die 3rd**
 ‘The disciples of Chuizi smashed at the lord with an ax and caused him to die.’

3.2 Changes in its dominant uses in terms of transitivity

It then follows that there might be no clear-cut distinction between the Old Chinese verbs ‘to kill’ and ‘to die’ completely without reference to a given context, as they can be used synonymously either as a transitive verb or as an intransitive verb in different contexts. Such a phenomenon persisted throughout Chinese history. As a matter of fact, *sha* was used more and more with an intransitive meaning and eventually ceased to occur in a RVC as a V2 in Early Modern Chinese. The examples in (13) are some cases of *sha* used in *Zhuzi Yulei* (13th century CE).

- (13) a. 婦殺夫 (Zhuzi Yulei 朱子語類 3)
 fu sha fu
 wife kill husband
 ‘The wife killed the husband.’
- b. 打殺那要向便門底心 (Zhuzi Yulei 朱子語類 16)
 da-sha na yao xiang bian men de xin
 hit-extinguish DEM want toward convenient door REL heart
 ‘Give up on the idea of wanting the door of convenience.’
- c. 便自殺身 (Zhuzi Yulei 朱子語類 5)
 bian zi sha shen
 then self kill body
 ‘Then, kill oneself.’
- d. 草木春生秋殺 (Zhuzi Yulei 朱子語類 15)
 cao mu chun sheng qiu sha
 grass wood spring life autumn die
 ‘The plants grow in the spring and die in the autumn.’

The verb *sha* is used as a transitive verb in (13a-c) and an intransitive verb in (13d). The increase of intransitive usage is shown by the table in Table 1. Furthermore, it is interesting to note that as a V2 in (13b) *sha* is used metaphorically. With an agent/theme case roles for its argument, there is not an obviously affected patient that would die resulting from the action of ‘to kill’. Furthermore, its intransitive use in (13d) simply means something like ‘die/wither’ that would only have a theme as its subject.

Between the 10th century CE text *Dunhuang* and the 13th century CE text *Zhuzi Yulei*, the ratios between the transitive pattern (suo) *sha* (+V2) N and the intransitive pattern (V1+) *sha* decreased from 3.8:1 to 1.4:1. This may suggest that *sha*’s intransitive usage became more common over the 300 years. But the key is that the verb that was originally more closely associated with the transitive ‘kill’ could be used more and more

intransitively over time. With regard to its function as a V2, in Early Modern Chinese the intransitive use of *sha* was completely replaced by *si* ‘die’ as is in *da-si* hit-die ‘kill’ of (2) (I will return to discuss the significance of this change in §4).

Table 1: The increase of intransitive occurrences of *sha* ‘to kill’ over time

	(suo) <i>sha</i> (+V2) N ³	(V1+) <i>sha</i>	Total
<i>Dunhuang Bianwen</i> 敦煌變文	68	18	86
Ratio	3.8	1	
<i>Zhuzi Yulei</i> 朱子語類 (1-21)	30	21	51
Ratio	1.4	1	

Ambitransitivity is a prevalent feature for a multitude of Chinese verbs. In relation to the history of the RVC, *po* 破 ‘break’ is another commonly used V2 in Modern Chinese with a converging ratio between transitive and intransitive usages over time in history. Recall that one of the reasons given by Mei (1991) for the neutralization of the V2 in transitivity was the emergence of a new serial-verb construction, V1+NP+V2. It was also noted in §2 that such a sequence is not completely new in Middle Chinese and, in Old Chinese, a verb of displacement following a transitive verb already took this pattern. What happened was that verbs, other than the displacement type, started to appear in this position. Mei (1991) noted that there is, for this type of V2, an alternative pattern V1+V2+NP, in which V2 should be treated as an intransitive verb just like its counterpart in the V1+NP+V2, such as *po* ‘break’ in the pair of examples in (14).

- (14) a. 當打汝口破 (You Ming Lu 幽明錄)
dang da ru kou **po**
should hit your mouth **break**
‘(Someone) should break your mouth.’
- b. 以梨打破頭 (Bai Yu Jing 百喻經)
yi li da **po** tou
with pear hit **break** head
‘(Someone) break open the head with a pear.’

³ *Suo* 所 in historical Chinese texts is an optionally used preverbal pro-form signaling a syntactic object of the following transitive verb. So any instance of a verb co-occurring with it is considered a transitive use and, thus, categorized under the transitive pattern.

However, historically *po* ‘break’, like *sha* ‘kill’ and *si* ‘die’, could commonly function as either a transitive verb with a following NP (15a) or an intransitive verb without an NP (15b-c). As a transitive verb, it may have an agent and patient thematically such as the one in (15a). As an intransitive verb, it can be construed as a middle voice verb, in other words informally it is syntactically active but semantically passive resembling English verbs ‘open’ and ‘cook’ in the sentences such as ‘The door opens easily’ or ‘The casserole cooked in the oven’.

- (15) a. 破塔壞寺 (*Dunhuang Foshuo Amituo Jing* 2 敦煌佛說阿彌陀經講經文二)
 po ta huai si
 break pagoda bad temple
 ‘break down pagodas and ruin temples’
- b. 宅舍破壞無投處 (*Dunhuang Mulian jiu mu* 敦煌大目乾連冥間救母)
 zai-she po huai wu tou chu
 residence-hut break bad Neg send place
 ‘All houses were destroyed and (we) have no place to go.’
- c. 刀剗骨肉片片破 (*Dunhuang Mulian jiu mu* 敦煌大目乾連冥間救母)
 dao gua gu-rou pian-pian po
 knife cut bone-flesh CL-CL break
 ‘A knife cut the flesh off the bone, (and) every piece is cut.’

After counting its occurrences in the *Shiji* and *Dunhuang* texts, (from the 1st century BCE and the 10th century CE, respectively), we found that the transitive use of *po* ‘break’ was much more common in Old Chinese (6:1), but that its intransitive use became nearly just as common in Middle Chinese (2.2:1).

Table 2: The decrease of the transitive occurrences of 破 *po* ‘to break’ over time

	(suo) 破 (+V2) N	(V1+) 破	<i>Total</i>
<i>Shiji: biographies</i> 史記列傳	278	46	324
<i>Ratio</i>	6	1	
<i>Dunhuang Bianwen</i> 敦煌變文	60	27	87
<i>Ratio</i>	2.2	1	

In this light, possibly those occurring in the V1+NP+V2 pattern signify *po* ‘to break’ being treated more as an intransitive verb and those in the V1+V2+NP pattern

were modeled after a transitive verb by the Middle Chinese native speakers. Furthermore, the upshot of the above discussion shows that the forging of the Chinese RVC cannot be fully explained simply in terms of neutralization in transitivity as many of them are ambitransitive.

The examples in (16) are a case illustrating the ambitransitive nature of the V1 in the Old Chinese's V1+V2 sequence. As a V1, *e* 餓 could mean 'to starve' transitively, or intransitively. In (16a), the two verbs, *e* 'to starve' and *sha* 'to kill' are both transitive verbs with a common agent and a common patient. In (16b) both *e* 'to starve' and *si* 'to die' are intransitive verbs with a common experiencer and a locative *Shaqiugong*.

- (16) a. 餓殺其子 (Lunheng 論衡: 齊世家)
 e *sha* *qi* *zi*
 starve kill 3rd son
 'starved and killed his son'
- b. 餓死沙丘宮 (Shiji 史記: 趙世家)
 e *si* Sha-qiu-gong
 starve die Name
 'starve to death in Shaqiugong'

In the remainder of this paper, explanations within a grammaticalization framework will be given to illustrate how ambitransitivity might have contributed to the formation of Chinese RVCs throughout history due to the potential ambiguity it generates in many contexts.

3.3 Event structures

In studying argument realization, Levin & Hovav (2005) do not look at the event structure of transitive and intransitive verbs in a mutually exclusive manner. On the one hand, an event, with verbs like *sweep* and *wipe* as the predicate, is simple as these verbs only take surface contact and do not entail any result state. On the other hand, an event, with causative change-of-state verbs like *break* and *dry* as the predicate, is treated as complex. In other words, the event structure of a proposition, in which there is an NP following a verb, such as 'sweep the floor', still can be represented by a simple event structure similar to that with an intransitive verb. The VP 殺人 in (17a) *sha ren* 'to kill people' is hyperbole entailing a situation with a simple event structure in which nobody would die and is, thus, similar to 'sweep the floor' without implying a causative change of state. Thus, the notion of transitivity alone cannot fully explain how the Vt+Vi pattern emerged out of the Old Chinese Vt+Vt sequence under the matching condition.

In the next section, we will discuss some factors motivating the creation of the Vt+Vi RVC.

4. Hyperbole, metonymization, metaphorization, and the emergence of RVCs

In the last section, it was noted that ambitransitivity, exemplified by verbs such as *sha* 殺 ‘to kill’, *si* 死 ‘to die’, and *po* 破 ‘to break’, is prevalent. Recall that in Old Chinese there were only Vt+Vt and Vi+Vi sequences that are parallel in nature. The three verbs in the examples given so far are all common occurrences of V2 in an RVC sequence. In this section, it will be shown how the polysemous structure of these verbs allowed them to forge Chinese RVCs as processes of metonymization and metaphorization. As shown by Hopper & Traugott (2003), metonymic and metaphorical inference are complementary, not exclusive, processes. The development of the Chinese RVC is highly context-dependent involving both processes, be it lexicalization or grammaticalization.

Metonymization refers to a process from which a new category, or meaning, arises in context (Brinton & Traugott 2005). It is a kind of semanticization of a new meaning initially induced from a specific context. The historical development of Vt+Vi RVCs is such a case. For instance, *sha* ‘to kill’, as a V2, originally occurred mostly in the Vt+Vt sequence. However, a new meaning of *sha* in an otherwise transitive pattern Vt+Vt might have arisen, motivated by hyperbole and subsequent regularization, in a context similar to the one exemplified by the late Old Chinese poem in (17a). As a transitive verb, V1 *chou* ‘to worry’, with a causer/experiencer structure is a less prototypical transitive verb than V2 *sha* ‘to kill’ that normally has an agent/patient argument structure. However, in this case, *baiyang* ‘white willow’, the causer of *chou* is inanimate. Suppose that the polysemous structure of *sha* allows either an agent or an inanimate causer as its subject. Then, the verbal sequence, *chou* and *sha* 愁殺, satisfies the Old Chinese matching condition as both verbs share two common arguments, even though it does not entail the typical meaning of ‘to kill’ as a non-animate tree is not likely to be a volitional murderer in its normal sense. Obviously a tree that ‘kills’ a person is used as a case of hyperbole highlighting the extent of a dire situation in which a dreadful sadness would affect a person. The significance of this usage is at least twofold. First, whereas the earliest cases of RVCs may not violate the matching condition for Vt+Vt sequences a non-matching Vt+Vi RVC might arise out of the transitive V2. That is, the Vt+Vi resultative meaning is contextually induced, or pragmatically inferred, from this kind of sequence that satisfied the Old Chinese matching condition. Second, the transitive V2, without entailing a loss of life, is a Vt with a simple event structure resembling an intransitive

verb. Thus the derivation of the Vt+Vi went through a transitional stage that is marked by a V2 predicate with a simple event structure.

- (17) a. 白楊多悲風，瀟瀟愁殺人 (*Gushi shijiu shou* 古詩十九首)
 bai yang duo bei feng, xiao-xiao **chou sha** ren
 white willow much sad wind gentle **worry kill** person
 ‘White willow with much sorrow, gently worried people to death.’
- b. 打殺那要向便門底心 (*Zhuzi Yulei* 朱子語類 16)
 da-sha na yao xiang bian men de xin
 hit extinguish DEM want toward convenient door REL heart
 ‘Give up on the idea of wanting the door of convenience.’
- c. 可殺頭痛 (*Zutangji* 祖堂集 14)
 ke sha tou tong
 can kill head hurt
 ‘It is a killing headache.’
- d. 愁死人 (present-day Standard Chinese)
 chou si ren
 worry die people
 ‘worry to death’

Levin (1999:225-243) and Levin & Hovav (2005) noted that there are two kinds of transitive verbs: change-of-state verbs with a complex event structure (18c) and verbs with a simple event structure the does not entail any result state (18a-b). It is hypothesized that the sequence *sha ren* in (17a) a simple event structure resembling the English ‘sweep the floor’ as the metaphorical use here signals an event that only touches a person without causing someone to die.

- (18) a. Pat swept the floor.
 [x ACT <*sweep*> y]
- b. Pat left.
 [x ACT <*leave*>]
- c. Pat broke the vase.
 [[x ACT <*manner*>] CAUSE [BECOME [y <*BROKEN*>]]]

Example (13b), repeated as (17b), is an additional example of this in Early Modern Chinese, in which the sentence indicates a mental activity. The complex structure for (17b) as an RVC can be [[x ACT <打 do>] CAUSE [BECOME [y <殺那...心 NO

MORE>]]]. It then follows that intransitive verbs in simple event structures can easily occur in this slot. The example in (17d) is given to show that, in present-day standard Chinese, *sha* ‘to kill’ is replaced by the intransitive *si* ‘to die’ as a V2 of the RVC. But when it is used as an RVC, *sha ren* indicates an exaggerated resultant state of the situation in (17a-b) somewhat comparable to the English ‘sweep the floor clean’ in which a causative change-of-state is involved. The example in (17c) was taken from Liang (2006) who claims that when *sha* ‘to kill’ is used to indicate an extent (or resultant state) it can even be placed in front of the verb signifying the initiative event, that it modifies in *sha tou tong* ‘killing headache’. Here this phenomenon is taken to reveal *sha*’s ability to indicate a static meaning that has nothing to do with real killing.

Still another feature of the early RVC sequence such as those in (17a-b) is that both V1 and V2 have an experiencer as their subject. ‘The willow tree’ is not an animate subject but a figurative causer. However, the person who is worried, as is signaled by the verb *chou*, is typically a non-volitional experiencer. Note that in the pair of Old Chinese examples cited in (11), *ya-sha* ‘to press and kill’, has a subject *an* ‘bank’ that is in an instrumental case and *ya-si* ‘to press and die’ has a subject *bai yu ren* ‘more than a hundred people’ that is in an experiencer case. Neither is representative of a high transitivity case that should have a deliberate and volitional agent. It is therefore claimed here that the Vt+Vi RVC was derived out of the context of low transitivity. Furthermore, the fundamental function of the RVC is to indicate the resultant state of a non-agent. That is, V2 in this sequence does not imply a change of state. It indicates a change of state only in the V1+V2 sequence, as V2 expresses a state of affairs that bears a cause-and-effect relationship with V1.

Example (19) is an additional piece of supporting evidence from the Modern Shanghai dialect for the intransitive analysis of *sha* ‘kill’ in (17a-b). The cognate of Old Chinese Vt *sha* in the Modern Shanghai dialect is the V2 *sa?* in (19). It is interestingly ambiguous in the Modern Shanghai dialect as it can imply either a complex event structure, or a simple event structure, given a specific context. For instance, it is a Vt+Vt sequence, if the hearer and the speaker are enemies and if the speaker is violent enough to murder the hearer. However, it is a Vt+Vi sequence if the hearer and the speaker are, for instance, lovers and if what is implied is an event that is exaggerated without implying any killing but a resultant state following the V1 *dā* ‘to hit’.

- (19) 打殺儂 (Shanghainese)
 dā sa? nong
 hit kill/die 2nd
 ‘kill you’ or ‘beat you up’

The derivation of the Chinese Vt+Vi RVC can also be understood as a process of metaphorization that assumes “the outcome of pragmatic extensions of meaning that become widely accepted and eventually semanticized in the context of speaker-hearer interaction.” (Brinton & Traugott 2005:106). In this paper the spread of the Vt+Vi RVC originally triggered by instances exemplified by those in (11) and (17a-b) is then fittingly treated as a process of metaphorization. Note that it was observed in (14) that *po* ‘to break’ is ambitransitive in Middle Chinese. Whereas the string in (20a) is the Old Chinese pattern for an intransitive V2 following a transitive V1, the string in (20b) is the pattern for a Vt+Vt in which the two verbs share a common subject and a common object. Now we can say that the *po* as V2 in the string (20b) can then be reanalyzed as a Vi for two good reasons. First, metonymically speaking, it was commonly used either as an intransitive verb or as a transitive verb in Old Chinese. Second, metaphorically speaking, by using the Vt+Vt sequence to convey hyperbole, where V2 has a simple event structure comparable to that of an intransitive verb, the Vt+Vi format systematically emerged making the Old Chinese verb-matching requirements irrelevant. A metaphorical extension is made to enable a Vi, or an adjective, to occur in the V2 position as analogous to the Vt that has a simple event structure.

- (20) a. 當打汝口破 (You Ming Lu 幽明錄)
dang da ru kou po
should hit your mouth break
‘(Someone) should break your mouth.’
- b. 以梨打破頭 (Bai Yu Jing 百喻經)
yi li da po tou
with pear hit break head
‘(Someone) break open the head with a pear.’

More support for the current hypothesis comes from the so-called verbs of completion as a V2. Mei (1981) and Cao (1986) claim that the present-day Chinese perfective marker *-le* 了 originally was a verb of completion commonly occurring as a V2. These verbs of completion⁴ as the V2 such as *chi* 訖 in V1+N+*chi* ‘to complete’ (21a), V1+*chi* (21b), and V1+*que* (21c), did not appear until Middle Chinese when the Vt+Vi sequence was formed.

⁴ Other verbs of completion include *jing* 竟, *yi* 已, and *bi* 畢. This group of verbs was substituted by *le* 了 and became lost in Early Modern Chinese.

- (21) a. 讀之訖 (Shiji 史記: 太史公自序)
 du zhi qi
 read 3rd complete
 ‘to finish reading it’
- b. 聽訖雲 (Shishuo Xinyu 世說新語: 賞譽)
 ting qi yun
 listen finish say
 ‘to finish listening to the discussion’
- c. 迎門卻行。高祖大驚 (Shiji 史記: 高祖本紀)
 ying men que xing. gaozu da jing
 welcome door retreat walk Name great surprise
 ‘Welcome at the door walking backward, Gaozu was greatly startled.’
- d. 設卻百日齋，渾家忘卻你 (Wang Fanzhi 王梵志詩)
 she que bai ri zhai, hunjia wang que ni
 set PERF hundred day feast relatives forget PERF 2nd
 ‘(Your relatives) will have forgotten you when organizing the 100th celebrating day (of your death?).’

Examples (21a-b) show that the syntactic distributions of the verbs of completion such as *qi* are consistent with the V2s under investigation such as *po*. As put forth by Cao (1986), the first Chinese perfective aspect marker is *que* in (21d), which was used originally as a directional verb meaning ‘to retreat, to go away, etc.’ as is the case in (21c). A directional verb like *que* was rarely used as a transitive verb in Middle Chinese (Sun 1999). Therefore, the *V-que* sequence in front of a noun in (21d) arguably was also made possible by the formation of the *Vt+Vi* sequence and resulted from a process of metaphorization. Its aspectual meanings were substituted by grammatical suffixes such as *-le* 了 in Early Modern Chinese (Cao 1986, and Sun 1996).

5. A framework of lexicalization and grammaticalization

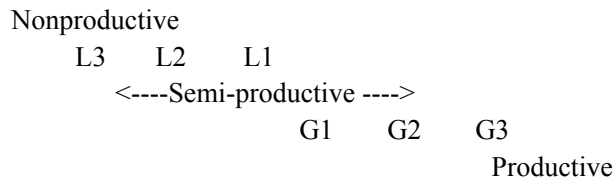
This section sums up the formation of the Chinese RVC presented so far within a theoretical framework of grammaticalization. Data will be given to show that Chinese RVCs belong to two kinds of processes, lexicalization and grammaticalization. Typologically speaking, the Chinese resultative indicates a change of state of a non-agent, going beyond the notion of a syntactic object or a semantic patient.

Within Brinton & Traugott’s framework (2005:96), lexicalization is defined as “the change whereby in certain linguistic contexts speakers use a syntactic construction

or word formation as a new contentful form with formal and semantic properties that are not completely derivable or predictable from the constituents of the construction or the word formation pattern. Over time there may be further loss of internal constituency and the item may become more lexical.” On the other hand, grammaticalization is defined as “the change whereby in certain linguistic contexts speakers use parts of a construction with a grammatical function. Over time the resulting grammatical item may become more grammatical by acquiring more grammatical functions and expanding its host-classes” (Brinton & Traugott 2005:99).

Therefore, both grammaticalization and lexicalization are contextually induced and may involve similar mechanisms of change such as metonymization, metaphorization or other features of change. They are nevertheless two distinct processes. Lexicalization typically involves contentful and referential meanings. Grammaticalization is often characterized by semantic bleaching, decategorization, and linguistic meanings. Chinese RVCs, as a process of unification (forming a V1+V2 compound word), include both lexicalization and grammaticalization. The schema in (22) is proposed by Brinton & Traugott (2005) to account for the variations and similarities of the two processes. It concerns adopting items into an inventory (traditionally known as a lexicon) and modifying items within the structure of the inventory with respect to their function, productivity, fusion, coalescence, compositionality, and degree of contentfulness. The arrows in (22) signal that the changes are synchronically bidirectional but diachronically unidirectional.

(22) Diachronic change along clines of lexicality and grammaticality



Chinese RVCs roughly fall into three categories, L2, L1 and G1, with L standing for lexicalization and G for grammaticalization. L2 represents a category of lexicalization that is less productive, or compositionally more idiosyncratic, than those in L1. Members of this type of RVCs take account of those in (4), 批評 *pi-ping* criticize-comment ‘to criticize’, 改良 *gai-liang* change-good ‘to improve’, 加强 *jiang-qiang* add-strong ‘to strengthen’, 擴大 *kuo-da* expand-big ‘to expand’, etc., that are fused together, semi-idiosyncratic and contentful. Formally, the infixal potential markers *de* and *bu*, noted in (4) and repeated in (23), cannot be inserted between the two components.

(24) a.	打得破	打不破
	da-de-po	da-bu-po
	hit-INF-break	hit-NINF-break
	‘breakable’	‘unbreakable’
b.	退得出	退不出
	tui-de-chu	tui-bu-chu
	retreat-INF-exit	retreat-NINF-exit
	‘can be withdrawn’	‘cannot be withdrawn’

- | | | |
|----|-----------------|--------------------|
| c. | 打得死 | 打不死 |
| | da-de-si | da-bu-si |
| | hit-INF-die | hit-NINF-die |
| | 'can be killed' | 'cannot be killed' |
| d. | 喝得醉 | 喝不醉 |
| | he-de-zui | he-bu-zui |
| | drink-INF-drunk | drink-NINF-drunk |
| | 'can get drunk' | 'cannot get drunk' |

Furthermore, some of the V2 are actually quite productive in forming an RVC with another verb. For example, 死 *si* 'to die' as a V2 is quite productive, as in 罵死 *ma-si* scold-die 'scold to death', 說死 *shuo-si* say-die 'say too definitely', 看死 *kan-si* look-die 'have a fixed view', etc.

G1 in (24) refers to the RVCs with a closed class of V2s including 完 *-wan* 'finish', 好 *hao* 'good', 到 *-dao* 'arrive', etc. that have practically lost all of their contentful meanings, and function to indicate an aspectual meaning that Li & Thompson (1981) define as a phase marker, such as *-hao* in (25), which has lost its contentful meaning 'good' but indicates a resultant state of 'to be finished'.

- (25) 我吃好飯了
 wo **chi hao** fan le
 1st **eat good** rice CRS⁵
 'I have eaten.'

In terms of productivity, G1 and L1 RVCs are similarly not as predictable as the non-RVCs V1-G2 sequences such as those with the experiential aspect marker *-guo* in (26b). The host-class (here referring to the V1s in sequence) that can co-occur with a G2 marker, such as *-guo* in (26b), is much larger than the V1s that can co-occur with any given V2s of the L1 and G1 types. Whereas a G1 type is essentially a periphrastic V1+V2 construct, V1-G2, such as *-guo* in (26b) and G3, such as *-le* in (26c), are morphologically bound, functioning as verbal suffixes.

⁵ CRS stands for currently relevant state and PERF for perfective aspect (See Li & Thompson 1981). Other abbreviations used in this article include G for grammaticalization, L for lexicalization, V for verb, Vt for transitive verb, Vi for intransitive verb, EXP for experiential aspect, ASP for aspect, RVC for resultative verb compound, INF for infix, NINF for negative infix, etc.

- (26) a. *他也紅完
 ta ye **hong-wan**
 3rd also red-finish
- b. 他也紅過
 ta ye **hong-guo**
 3rd also red-EXP
 ‘He once was quite popular.’
- c. 他也紅了
 ta ye **hong-le**
 3rd also red-PERF
 ‘He has been quite popular.’

Therefore, the sequences of V1-G2, or V1-G3 such as 過 *-guo* and 了 *-le* in (26c) are not RVCs. Although there are not many G2 or G3 markers in Chinese, and those that exist historically derive from verbs of completion similar to those V2s in G1, a detailed study of them falls outside the scope of the current investigation. They simply are given to illustrate the differences between them and the G1 RVC type.

6. Summary

To summarize it is hypothesized that the present-day RVC sequence with a Vt+Vi structure ultimately grew out of the Middle Chinese parallel Vt+Vt sequence. The change of the V2 in this sequence from Vt to Vi was originally triggered by a processes similar to hyperbole such as *chou-sha ren* grieve-kill people (17a) as *sha* ‘to kill’ acquires a new function to signal a new meaning without involving a change of state such as ‘to die’. Furthermore, the spread of the RVCs in the language took hundreds of years to develop, leading to cases of lexicalization and grammaticalization.

Many Chinese verbs throughout history have had dual functions like *sha* ‘to kill/to die’ that could be used transitively or intransitively. It is then further hypothesized that the *sha ren* ‘kill person’ in the specific context signified by example (17a) is semantically analogous to a possible intransitive meaning that does not involve any death but a certain state of affairs with a simple temporal structure. These early RVCs were mostly cases of lexicalization with perhaps the exception of *que* 卻 in (21d) that has a grammatical meaning related to time. Table 3 illustrates the historical development correlating to the V2’s ability to participate in selecting an argument. The V2 in L1 RVCs still functions like a predicate, even though it is already combined with V1 to form a single predicate. However, unlike the Middle Chinese V1+V2 sequences in which V1 and V2 are a two

contentful verbs functioning as two predicates in selecting two sets of shared arguments, the V1s and V2s in the L1 RVCs are more closely unified into a single predicate. Some of them, such as the V2 *zui* in *he-zui jiu* drink-drunk wine ‘to get drunk with wine’ show signs of losing the ability to independently select a complete set of arguments. For example, the theme *jiu* ‘wine’ immediately following it is not its object but the object of V1-*he* ‘to drink’. Second, the V2s in L1 are either intransitive verbs or adjectives, correlating with its basic function to indicate a resultant state of a non-agent brought about by V1, thus different from its original parallel Vt+Vt sequence. The V2s, verbs of completion such as *hao* 好 ‘good’, in the G1 RVCs are no longer a contentful predicate, and thus do not select any argument, but are morphologically more integrated and semantically bleached verbs signaling, instead, a temporal meaning of the unified compound. Some of these V2s are totally grammaticalized into verbal suffixes expressing different temporal meanings with a different degree of high productivity.

Table 3: The historical development correlating to the V2’s ability to participate in selecting an argument

	L2	L1	G1	G2/G3
Meaning	contentful	contentful	grammatical	grammatical
V2	fused with V1	predicate of result	periphrastic	bound morpheme

In Middle Chinese the spread of the Vt+Vi sequence in forming the present-day RVCs as a special type of construction is also a process of metaphorization, as cases such as *chou-sha ren* grieve-kill people (17a) might have served as a model for other occurrences to emulate. Finally, in such a lexicalization/grammaticalization model, on one end of the continuum, there is the L2 type with items that are lexically fused together such as *pi-ping* criticize-comment ‘to criticize’ that is not analyzable, or unpredictable, without allowing the potential markers *de/bu* to co-occur with them. On the other end of the RVCs is the G1 type, with the V2 in a periphrastic sequence functioning as a grammaticalized phase marker of the V1+V2 predicate. L1, somewhat similar to G1 in terms of productivity and predictability, is the prototypical Vt+Vi RVC with the V2 as a sign entailing a simple event structure indicating the effect (resultant) state caused by V1. The most grammaticalized G2 and G3 morphemes, though derived from G1 V2s, are not part of an RVC but verbal suffixes.

References

- Brinton, Laurel J., and Elizabeth Closs Traugott. 2005. *Lexicalization and Language Change*. Cambridge & New York: Cambridge University Press.
- Cao, Guangshun. 1986. Zutangji zhong de 'di (di)', 'que (le)', 'zhe' [The particles *di* (*di*), *que* (*le*), *zhe* in *Zutangji*]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 1986.3:192-202.
- Gong, Hwang-chenng. 2000. Cong Han-Zangyu de bijiao kan Shanggu Hanyu de citou wenti [Prefixes in Old Chinese from the perspective of Sino-Tibetan comparative studies]. *Language and Linguistics* 1.2:39-62.
- Hopper, Paul J., and Elizabeth Closs Traugott. 2003. *Grammaticalization* (2nd edition). Cambridge & New York: Cambridge University Press.
- Levin, Beth. 1999. Objecthood: an event structure perspective. *Chicago Linguistic Society (CLS)* 35.1:275-289. Chicago: Chicago Linguistic Society.
- Levin, Beth, and Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge & New York: Cambridge University Press.
- Li, Ping. 1984. Shishuo Xinyu he Baiyujing zhong de dongbu jiegou [The verbal complement construction in *Shishuo Xinyu* and *Baiyujing*]. *Yuyanxue Luncong* [*Essays on Linguistics*] 14:129-157. Beijing: The Commercial Press.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Liang, Yinfeng. 2006. *Hanyu Dongbu Jiegou de Chansheng yu Yanbian* [*The Formation and Developments of the Chinese Verbal Complements*]. Shanghai: Xuelin Press.
- Mei, Tsu-Lin. 1981. Xiandai Hanyu wanchengmao jushi he ciwei de lai yuan [The origin of the perfective aspect construction and the perfective suffix in Modern Chinese]. *Yuyan Yanjiu* [*Studies in Language and Linguistics*] 1981.1:65-77.
- Mei, Tsu-Lin. 1988. Neibu gouni Hanyu san li [Three Chinese cases of internal reconstruction]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 1988.3:169-181.
- Mei, Tsu-Lin. 1991. Cong Handai de 'dong, sha', 'dong, si' lai kan dongbu jiegou de fazhan [The development of the verbal complement from the perspective the V-sha and V-si in the Han Dynasty]. *Yuyanxue Luncong* [*Essays on Linguistics*] 16:112-136. Beijing: The Commercial Press.
- Ota, Tatsuo. 1958. *Chūgokugo Rekishi Bunpō* [*A Historical Grammar of Modern Chinese*]. Tokyo: Konan Shoin.
- Sun, Chaofen. 1996. *Word-Order Change and Grammaticalization in the History of Chinese*. Stanford: Stanford University Press.

- Sun, Chaofen. 1999. The origin of the Chinese verbal suffixes. *Studies on Chinese Historical Syntax and Morphology*, ed. by Alain Peyraube & Chaofen Sun, 183-202. Paris: EHESS, CRLAO.
- Xu, Dan. 2006. *Typological Change in Chinese Syntax*. Oxford & New York: Oxford University Press.
- Wang, Li. 1958. *Hanyu Shi Gao* [*An Outline of Chinese Language History*]. Beijing: Science Press.
- Yakhontov, Sergey E. 1960[1986]. Shanggu Hanyu de fufuyin shengmu [The initial consonant clusters in Old Chinese], translated by Feisheng Ye, Chongye Chen & Jianqiao Yang. *Hanyushi Lunji* [*Research on History of Chinese Language*], ed. by Zuofan Tang & Shuangbao Hu, 42-55. Beijing: Peking University Press.

近、現代漢語裡“給+VP”的形成

張伯江

中國社會科學院

本文研究“拿了三吊讓老王給拿出去”（《小額》）這樣例句中“給”字用法的來歷。這種用法是否跟表示給予的動詞和引介受惠者 (benefactive) 或與事 (dative) 的介詞有關，學界意見不一，大多認為表示“影響義”；迄今所見的世界其他語言“給予”義動詞的演化報告中，未見有這種直接作用於動詞的現象。本文考察“給+VP”產生初期的一些材料，用實例說明，介詞“給”後面省略受惠者和與事的現象在早期材料中大量存在，省略的原因跟人稱代詞的弱化傾向有關。這種省略現象與“影響義‘給+VP’”同時存在，過渡痕跡明顯。本文用有說服力的過渡性的例子論證省略受惠者/與事的“給+VP”是如何演化為影響義的“給+VP”的，從理論上說就是論證普通語言學的 GIVE > BENEFACTIVE > DATIVE 發展鏈條在漢語裡為什麼進一步延伸出了更新的一個環節——影響義動詞標記的。本文的結論還在更多事實中顯現出進一步的解釋力。首先可以合理解釋這種“給+VP”格式為什麼常常與“把/被”句伴生。其次可以解釋“給+VP”中 VP 可以是一元謂詞（即不及物動詞和形容詞）的現象，當代漢語單個一元謂詞現象的新發展，完全是靠“給”的作用而實現影響義的。

關鍵詞：受惠者，與事，省略，影響

1. 問題

“給”的主要身分，一個是動詞，表示給予，一個是介詞，引介受惠者 (benefactive) 或者與事 (dative) 成分。這兩種“給”都是要求帶名詞性賓語的，後者從前者虛化而來，句法和語義脈絡也是清楚的。近、現代北方口語裡，還有一種“給”字後面不帶名詞賓語，直接挨著動詞的用法，如“拿了三吊讓老王給拿出去”（《小額》）。這種現象近來受到不少學者的重視，如齊滄揚 (1995)、張誼生 (2001)、王彥杰 (2001)、李煒 (2004)、石毓智 (2004)、洪波 (2004)、李宇明、陳前瑞 (2005)、柯航 (2005)、劉永耕 (2005)、楊霽楚 (2008)、顏力濤 (2008)

等都會論及。經過這些學者的努力，這個現象在現代漢語裡的表現已經揭示得相當清楚了。對這個“給”字的語法意義，比較一致的看法是認為有“加強動詞的影響力”的作用，本文暫且依從這種說法，並在下文簡稱之為“影響義‘給’”。但是，關於這個影響義“給”的性質和作用，“給+VP”格式究竟是怎麼形成的，都還沒有令人完全信服的結論。儘管多數學者相信這種格式裡的“給”是動詞“給”語法化的結果，但究竟是通過省略介詞賓語的途徑而來，還是動詞“給”的直接虛化，這是筆者認為值得探究的，因為迄今所見的世界其他語言“給予”義動詞的演化報告中，還未見有這種直接與動詞組合的現象，其中的發展機制是什麼，有什麼理論蘊含，是本文重點關心的。

有相當多的學者相信這種直接加在動詞前的“給”是由引介受惠者/與事的介詞“給”進一步虛化而來的，具體地說，就是認為“給+VP”中間省略了一個受惠/與事成分。這種解釋用在現代漢語的共時系統中有一定的困難，楊霽楚(2008)和顏力濤(2008)對此作了有說服力的討論。楊文提出的論證是：首先，有些句子不能補上受惠/與事成分（把廠長給感動了→*把廠長給他感動了）；其次，有受惠/與事成分出現的，刪除該成分與原句意思並不相符（她早把她姥姥那圍裙給我預備好了。刪除受惠者，優勢理解不是“給我”）。顏文則嚴格依據“省略處填補的詞語只有一種可能”的原則，通過對大量現代文學作品實例的逐一檢驗，證明“省略受惠/與事”的說法並不成立。

不過，這是不是表明歷史上“給+VP”的形成也就一定不與省略相關呢？爲了弄清這個事實，我們考察了“給+VP”產生初期的一些材料：大致成書於1860-1880年間的《兒女英雄傳》，印行於1898年的《談論新篇》，刊行於1908年的《小額》和1910年的《京語會話》。已有幾位學者的考察表明，影響義“給+VP”是從這幾種作品中開始出現的。

2. 事實

2.1 考察這些早期北京口語材料，一個突出的現象就是，省略受惠者和與事的現象大量存在。我們先看兩個實例：

- (1) 他當時也沒說甚麼，趕人家把房子蓋得了，給他加幾兩銀子的房錢，他不答應，叫人家總得給Ø加十兩銀子的房租，人家不肯給他加，他說若不依著他那個數兒辦，就叫人家搬家，上別處做買賣去。（《談論新篇》第五十六章）

- (2) 直過了一年多，他們才查出來是別人給 Ø洩漏的，這才把我們舍弟洗出來。（《談論新篇》第五十七章）

爲什麼斷定這兩個例子是省略受惠者的呢？例(1)“給加十兩銀子”後面緊接著有“人家不肯給他加”的說法；例(2)“別人給洩漏的”這句話，是照應著前邊不遠處的上文“他們就不免疑惑我給他們洩漏的”而言的。都有顯形的受惠者出現，都是明顯的證據。

由是我們可以判斷出來，下面這些例子，都屬於省略受惠者的情況：

- (3) 早有本地長班預先給 Ø找下公館，沿河接見。（《兒女英雄傳》第二回）
- (4) 別的我不知道，內囊兒舅母都給 Ø張羅齊了，外妝公婆都給 Ø辦妥了。（同上，第二十六回）
- (5) 您給 Ø哄哄孩子我給 Ø打點飯去。（《小額》）
- (6) 一五一十的合盤托出求他給 Ø想個主意。（同上）
- (7) 將來財政部裡頭求您給 Ø謀一個位置罷。（《京語會話·敘親》）
- (8) 咱們還不竟逛市場，你看著什麼好吃，爺爺我給 Ø買什麼，你看好不好哇？（《京語會話·勸善》）
- (9) 托我給 Ø請一位先生，不知道您意中有合宜的人沒有？（《談論新篇》第四十四章）
- (10) 您可以給 Ø雇一個人送了去麼？（同上，第四十六章）

2.2 再看省略與事的情況。“與事”這個成分，向來有與“受惠者”界限不清的情況，很多語言都有這兩種角色共用一個標記詞的情況，漢語便是如此。這裡我們分別出的“受惠者”和“與事”主要是從意義角度看的：偏向於“幫助”語義的，確定爲受惠者；偏向於“針對”語義的，處理爲與事。嚴格地說，我們處理爲與事的，都應該看成廣義的受惠者，因爲從形式上說，“給 N”只有到了主要動詞的後頭，才可以清楚地與受惠者區分開，才便於處理爲與事。不過，漢語事實上沒有動詞後面“給 N”省略名詞的現象，而且，這樣的現象也不與本文關心的影響義“給+VP”現象相關。

下面是一些我們認爲接近於“與事”意義的實例：

- (11) 自然也該照著外省那怯禮兒，說定了親，婆婆家先給Ø送匹紅綢子掛紅，那叫“紅定在先”。（《兒女英雄傳》，第二十六回）
- (12) 先頭啦，聽見要給Ø磕頭賠不是，伊太太倒很願意。（《小額》）
- (13) 倒帶著挨打的，到他門上給Ø賠了個不是。（同上）
- (14) 趕是時候兒，叫他做得了給Ø送到這店裡來。（《談論新篇》第十六章）

爲什麼介詞“給”後面的受惠/與事成分會發生省略？我們認爲這跟人稱代詞的弱化傾向有關。我們發現，被省略掉的介詞賓語，大多是人稱代詞，而人稱代詞由於其上下文已知性高，信息量弱小，在很多語言裡都有失去獨立的成詞能力、弱化爲附綴形式的現象；漢語沒有“一致關係”等語法範疇，也就沒有“代詞附綴化”的句法要求，在弱化到一定程度時就完全省略掉了。

2.3 接下來，看看以上兩種情況之外的，即所謂“影響義‘給+VP’”的實例。我們所考察的清末四部作品裡這種類型的用例都不少，總的來看，與“把”字句伴生現象比較多見；但是仔細說來，有兩方面的特點：十九世紀末的兩部作品《兒女英雄傳》和《談論新篇》裡，影響義“給+VP”並不總與“把”字句伴生，且未見“被（叫/讓）”字句；二十世紀初的兩部作品《小額》和《京語會話》裡，與“叫/讓”句伴生的多了起來，且不與“把/被（叫/讓）”字句伴生的例子明顯減少。以下是實例：

- (15) 甚至如新買的馬桶，新打的夜壺，都給預備在床底下。（《兒女英雄傳》第三十九回）
- (16) 最奇不過的是這老頭兒家裡竟會有書，案頭還給擺了幾套書。（同上，第三十九回）
- (17) 烏大人接過去，又給收拾了收拾，便叫安公子戴上。（同上，第四十回）
- (18) 又托他家的門館先生管待程相公，又囑咐把酒先給收在倉裡，閒來自己去收。（同上，第三十九回）
- (19) 就這麼給哄下來了，這不是碰了個大釘子麼？（《談論新篇》第五十六章）
- (20) 你在外頭耽擱了這麼幾天，差一點兒把你們大掌櫃的給急死。（《談論新篇》第四十三章）

- (21) 兩個人早把青皮連給拉了走啦。(《小額》)
- (22) 王媽你可給我瞧著點兒狗，上回我就讓他給咬了一下子。(同上)
- (23) 那一溜兒都叫他給得罪透啦。(同上)
- (24) 有一點兒家當兒都讓他舅舅給花光啦。(同上)
- (25) 您要肯辦，就有那個事，昨天我有一個朋友，因為在家裡打牌，教偵緝隊給抓了去啦。(《京語會話》)

3. “給+V”的形成

我們要探討的是既非省略受惠者亦非省略與事的影響義“給+VP”現象的來歷。以往研究中對這個問題沒有一致的意見。洪波(2004)強調“給”字這個用法不與介詞省略賓語相關，但承認是“給”在清代中葉以後的“進一步語法化”，沒有講出變化的過程和動因；齊滬揚(1995)、張誼生(2001)對此倒是有明確的說法，他們認為這裡有一個“給”字賓語從後面移動到“給”字之前的過程。但正如柯航(2005)所批評的，“把”語法化為介詞遠遠早於“給”，介詞“給”的賓語移到“把”字賓語說不過去，移位的證據和動因講不清楚。柯文相信“給+VP”是由省略與事形成的，但她過於強調了“把”字在這個過程中的作用，而我們上面已經揭示，影響義“給+VP”的產生並不必然依賴於“把”字句。

的確，在我們考察的清代語言材料中，省略受惠者、省略與事以及影響義“給+VP”現象，三者是同時存在的。也就是說，我們無法從不同年代的語言材料中觀察誰先誰後的規律。但是我們仍然可以從學理上推斷三者之間的因果關係。在語法化的研究上，我們常常遇到這樣的困難：幾個顯然相關的現象無法從文獻角度判斷時間上的先後，有的比較易於解釋，有的難以解釋。那麼我們可以有把握地說，應該是容易解釋的發生在先，不容易解釋的發生在後。

需要強調的是，我們談到的這種辦法，是基於確信容易解釋的和不易解釋的現象之間一定是相關的，不意味著所有無法解釋的現象都可以這樣獲得解釋。具體而言，就是我們覺得，省略受惠者/與事的“給+VP”現象跟影響義的“給+VP”現象，在早期文獻中，有可以觀察到的聯繫。看以下例子：

- (26) 好容易出來幾個善人給說合完了。(《小額》)
- (27) 過來好幾個街房才給勸開。(同上)
- (28) 勸完了婆婆，又到西廂房勸媳婦兒，好容易才都給勸住。(同上)

首先，這幾個例子從句義看，可以認為是與前面討論過的“省略受惠者/與事”同類的，同時也可以看出“說合”、“勸開”、“勸住”這幾個動詞由於有了“給”的存在，其中“影響義”的增強。這都是從意義上說的。再從形式上說，“給”字後面都還可以比較自然地補上受惠者成分：

(26') 好容易出來幾個善人給[他們]說合完了。

(27') 過來好幾個街房才給[他們]勸開。

(28') 勸完了婆婆，又到西廂房勸媳婦兒，好容易才都給[他們]勸住。

但是，更自然的說法是變換成“把”字句：

(26'') 好容易出來幾個善人把他們給說合完了。

(27'') 過來好幾個街房才把他們給勸開。

(28'') 勸完了婆婆，又到西廂房勸媳婦兒，好容易才把他們都給勸住。

這樣的例證，可以顯現出“省略受惠者/與事”的“給+VP”與“影響義”的“給+VP”在產生之初是有聯繫的。

事實上，早期的“給+VP”實例，當不與“把/被”句伴生時，常常可以有“省略”和“影響”兩種不同的理解。龔千炎(1994)稱前者為“介詞”，後者為“助詞”，有趣的是，對本文例(4)“內囊兒舅母都給張羅齊了，外妝公婆都給辦妥了”這兩個明顯是同類並列的小句，龔著把“內囊兒舅母都給張羅齊了”歸入“助詞”類，而把“外妝公婆都給辦妥了”歸入“介詞”類。這也說明，面對每一個例子孤立地判斷時，語法學家也有游移不定的時候。

這種“省略”和“影響”具有明顯聯繫的事實，使我們有信心推斷影響義的“給+VP”是由省略受惠者/與事現象發展而來。我們認為易於解釋的變化發生在不易解釋的變化之先，這是用的“邏輯先後”的觀點，同時也是對“歷史先後”的一種構擬。沈家煊(2008)指出，“邏輯先後”關注的是共時系統中語法形式形成的“抽象機制”，這種基於推導的假設有時是跟歷史事實相吻合的，有時未必完全吻合，但是不因跟歷史不合而被否定。本文構擬的歷史過程，主要是依靠邏輯推導，同時也不與史料相衝突。如洪波(2004)發現《紅樓夢》裡就有了“給”省略受惠者的現象，而“影響義”的用法是到了《兒女英雄傳》裡才出現的；本文上面展示的事實也說明，普通的影響義用法並不晚於“給”與“把/被”句（尤其是與被動形式）伴生的現象出現，而伴生現象是後來更加普遍的。這都支持本文構擬的變化過程。

4. 解釋

誠如許多學者指出的，影響義的“給+VP”中，“給”字後面不再能補出受惠者等語義成分，也就是說，這時的“給”已經完全喪失了引介功能，只能解釋為是作用於後面的動詞的。這種作用，就是所謂的“影響義”。

我們關心的問題是，這“影響義”是怎麼來的？

實際上，這裡有兩個問題需要解釋：一是來源於介詞的“給”，在句法上怎麼可以直接加在動詞前邊；二是為什麼它會有加強動詞影響語義的作用。

Heine & Kuteva (2002) 羅列了迄今所知“給予”意義動詞的幾種語法化模式：

GIVE > (1) BENEFACTIVE

GIVE > (2) CAUSATIVE

GIVE > (3) CONCERN

GIVE > (4) DATIVE

GIVE > (5) PURPOSE

沒有一種是 GIVE 後面直接跟動詞組合的。其中與本文有關的是“GIVE > BENEFACTIVE”和“GIVE > DATIVE”兩項。作者特意說明，從給予義動詞發展為受惠者標記詞，再發展為與事標記詞，即 GIVE > BENEFACTIVE > DATIVE，是一個常見的語法化鏈條。作者指出，有些語言裡，從受惠者標記發展為與事標記的條件是，句中的主要動詞為“告訴”等言說動詞，或者“賣”等傳遞動詞的時候，受惠者角色就容易獲得與事的意義。這與本文上面舉例中的情況完全一致：例(11)-(14)中的主要動詞分別是言說動詞“賠不是”和傳遞動詞“送”，因此獲得了與事意義。用這個視點看，我們重點觀察的例(26)-(28)，其中的主要動詞也是與言說有關的“說合”、“勸住”等，可以說，“給”從引介受惠者的介詞發展為影響義的動詞前加成分，也是經由引介與事語義這個過渡站而實現的。

為什麼普通語言學的 GIVE > BENEFACTIVE > DATIVE 發展鏈條到了漢語裡又進一步延伸出了更新的一個環節——影響義動詞標記了呢？從形式上說，首先是因為在這個環節的第一次語法化階段，即動詞“給”發展成介詞“給”的時候，就出現了“給”的賓語省略的情況。省略既已發生，便導致了“V+給”格式的成形，並隨著語義的進一步虛化日益固定下來。再從語義方面說。我們認為，當一個動詞短語前面加上一個受惠者的時候，就等於為主要動詞所表示的行為確定了一個明確的服務目標；當主要動詞是傳遞意義（含言說傳遞）的動詞時，“給+

(N)”的目標意義就更顯豁。動詞的目標意義越明確，動作行為對目標的影響力就越強。“給”字後的成分經歷了這樣幾個角色的變化：

給_動＋接受者＞

給_介＋(受惠者)＋主要動詞＞

給_介＋(與事)＋主要動詞＞

給＋主要動詞

在這幾個階段裡，“給”從其所帶的賓語越來越多地獲得“影響”語義，並在最後一個階段完全實現為自身的語法意義。

這樣，我們就從形式和語義兩個角度解釋了影響義的“給＋VP”格式的形成以及其中“影響義”的來歷。下面我們可以進一步看看這個說法的解釋力。

首先，可以合理解釋這種“給＋VP”格式為什麼常常與“把/被”句伴生。在許多學者的論著中，都把這種影響義的“給＋VP”格式看成“把”字句和“被”字句的次類，也有許多學者發現這種“給”的用法並不局限於“把/被”句，但出現在“把/被”句裡也確實是最多見的。我們知道，漢語“把/被”句謂語的一個最重要的語義特點，就是動作影響力的強化（張伯江 2009），因此，這種“給”用在“把/被”句的主要動詞前是十分恰當的。

其次，可以解釋“給＋VP”中 VP 可以是一元謂詞（即不及物動詞和形容詞）的現象。洪波（2004）和楊霽楚（2008）都曾指出這種現象，例如：

- (29) 我一下子給傻了，腦子裡一片空白。（引自洪文）
- (30) 猛可裡叫了他一聲，他一下子給愣住了。（同上）
- (31) 有一艘船，限載 100 人，超過的話就會沉掉，現在上面剛上了 99 人，就給沉了。（引自楊文）
- (32) 有一次下課去找她，正好宋欣在她旁邊坐著，說她：“你看，你給胖成什麼樣子啦！”（同上）
- (33) 開始我記不清了，只記得我在上網，突然螢幕給黑了，然後所有一切都黑了。（同上）

其實，這種現象在晚清作品中就已經出現了，例如：

- (34) 勸說嘍半天，好容易才把他的火兒給平下去啦，您聽著好些個人嚷嚷，就是為這個。（《京語會話》）
- (35) 這麼看起來，火輪船火輪車那兩樣兒買賣真能把地方兒給興旺起來，所以是萬不可沒有的。（《談論新篇》第七十二章）

“平”和“興旺”都是形容詞，詞彙語義都是與句式對謂語的“強影響性”要求所不符的，可以說，正是“給”的“強化影響力”作用使得它們可以正常出現在句中。如果說早期用例“平下去”、“興旺起來”都還是動補式的，有一定的動態意義的話，那麼當代例子如(29)(31)(33)則完全是單個的“一元謂詞”，完全是靠“給”的作用而實現影響義的。這可以說是“給+V”這個結構在當代漢語裡的進一步發展。弱及物性謂詞在句子裡獲得強影響義，靠的是句式以及其中關鍵性語法成分的作用，這個現象在漢語裡並不是孤立的，“把/被”句就是通過這樣的作用容許非及物性謂詞充當主要謂語的（張伯江 2009）。

5. 結語

本文是關於“給”字一個特殊用法的微觀考察。當我們追究這個特殊用法的句法-語義來歷的時候，涉及了漢語受患者標記、與事標記、“把”字句和“被（叫/讓）”句等一系列關乎語法系統的因素的互動關係和歷史演變問題。貝羅貝(1986)研究漢語雙賓語結構的發展歷史時，就看到了詞彙替換和結構發展之間的複雜關係。他指出：“研究漢語語法發展的歷史過程，不能忽視漢語在各個時期的語法描述細節。只有當語言的每個歷史時期的不少語法問題具有詳細的描述以後，我們才可能提出語法變化真正原因的假設，以及有關這方面的一些預測。”本文就是在這種細節描寫的方向上做的一點努力。

引用文獻

- Gong, Qianyan (龔千炎). 1994. *Ernü Yingxiong Zhuan Xuci Lihui* 兒女英雄傳虛詞例匯 [Data Compilation on Function Words in Legend of Heroes]. Beijing: Language and Culture Press.
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge & New York: Cambridge University Press.
- Hong, Bo (洪波). 2004. 'Gei' zi de yufahua “給”字的語法化 [Grammaticalization of *gei*]. *Nankai Linguistics* 2004.2:138-145.
- Huang, Zanhui (黃瓚輝). 2001. Jieci 'gei' 'wei' 'ti' yongfa buyi 介詞“給”“爲”“替”用法補議 [A supplement to the usage of the prepositions *gei*, *wei* and *ti*]. *TCSOL Studies* 2001.1:49-54.
- Jiang, Shaoyu (蔣紹愚). 2002. 'Gei' ziju, 'jiao' ziju biao beidong de lai yuan: jian tan yufahua, leitui he gongneng de kuozhan “給”字句、“教”字句表被動的來源：兼談語法化、類推和功能的擴展 [The origin of passive sentence with *gei* and *jiao*: on grammaticalization, analogy and expansion of grammatical function]. *Yuyanxue Luncong* 語言學論叢 [Essays on Linguistics] 26:159-177. Beijing: The Commercial Press.
- Ke, Hang (柯航). 2005. 'Ba...gei VP' *Jushi de Lishi Kaocha* “把...給 VP”句式的歷時考察 [A Diachronic Investigation on the “ba...gei VP” Construction in Chinese]. Wuhan: Central China Normal University MA thesis.
- Li, Wei (李煒). 2004. Jiaqiang chuzhi/beidong yushi de zhuci 'gei' 加強處置/被動語勢的助詞“給” [The auxiliary word *gei* in reinforcing the disposition/passive voice]. *Yuyan Jiaoxue yu Yanjiu* 語言教學與研究 [Language Teaching and Linguistic Studies] 2004.1:55-61.
- Li, Yuming (李宇明), and Qianrui Chen (陳前瑞). 2005. Beijinghua 'gei' zi beidongju de diwei ji qi lishi fazhan 北京話“給”字被動句的地位及其歷史發展 [The status and development of the passive *gei* construction in Beijing dialect]. *Fangyan* 方言 [Dialect] 2005.4:289-297.
- Liu, Yonggeng (劉永耕). 2005. Dongci 'gei' yufahua guocheng de yisu chuancheng ji xiangguan wenti 動詞“給”語法化過程的義素傳承及相關問題 [The sememe succession and relevant problems in the grammaticalization of the verb *gei*]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 2005.2:130-138.
- Lü, Shuxiang (呂叔湘). 1948[1984]. *Ba zi yongfa de yanjiu* 把字用法的研究 [On the uses of the pre-transitive *ba*]. *Hanyu Yufa Lunwenji* 漢語語法論文集 [Studies in Chinese Grammar], 176-199. Beijing: The Commercial Press.
- Lü, Shuxiang (呂叔湘) (ed.) 1980. *Xiandai Hanyu Babai Ci* 現代漢語八百詞 [Eight Hundred Words in Modern Chinese]. Beijing: The Commercial Press.

- Peyraube, Alain (貝羅貝). 1986. Shuang binyu jieyou cong Handai zhi Tangdai de lishi fazhan 雙賓語結構從漢代至唐代的歷史發展 [Double-object constructions from Han to Tang]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1986.3:204-216.
- Qi, Huyang (齊滬揚). 1995. Youguan jiecì ‘gei’ de zhipai chengfen de shenglue de wenti 有關介詞“給”的支配成分的省略的問題 [On the ellipsis of governed elements of Preposition *gei*]. *Journal of Shanghai Normal University* 1995.4:83-89.
- Shen, Jiaxuan (沈家煊). 2008. ‘Luoji xianhou’ he ‘lishi xianhou’ “邏輯先後”和“歷史先後” [The ‘logical sequence’ and ‘diachronic sequence’]. *Weiguoyu* 外國語 [*Journal of Foreign Languages*] 2008.5:91-92.
- Shi, Dingxu. 1997. Issues on Chinese passives. *Journal of Chinese Linguistics* 25.1:41-70.
- Shi, Yuzhi (石毓智). 2004. Jian biao beidong he chuzhi de ‘gei’ de yufahua 兼表被動和處置的“給”的語法化 [*Gěi*: Its double functions as passive and disposal markers]. *Shijie Hanyu Jiaoxue* 世界漢語教學 [*Chinese Teaching in the World*] 2004.3:15-26.
- Tang, Sze-Wing (鄧思穎). 2003. *Hanyu Fangyan Yufa de Canshu Lilun* 漢語方言語法的參數理論 [*A Parametric Theory of Chinese Dialectal Grammar*]. Beijing: Peking University Press.
- Wang, Yanjie (王彥杰). 2001. ‘Ba...gei V’ jushi zhong zhuci ‘gei’ de shiyong tiaojian he biaoda gongneng “把...給 V” 句式中的助詞“給”的使用條件和表達功能 [The conditions required and the expressive functions of *gei* in the sentence pattern of ‘ba...gei V’]. *Yuyan Jiaoxue yu Yanjiu* 語言教學與研究 [*Language Teaching and Linguistic Studies*] 2001.2:64-70.
- Wen, Suolin (溫鎖林), and Qun Fan (范群). 2006. Xiandai Hanyu kouyu zhong ziran jiaodian biaoji ‘gei’ 現代漢語口語中自然焦點標記詞“給” [*Gei* as a natural-focus marker in contemporary spoken Chinese]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2006.1:19-25.
- Yan, Litao (顏力濤). 2008. Fuhe ‘ba’ ziju yu fuhe beidongju zhong ‘gei’ hou binyu de shenglue wenti ji qi youyin 複合把字句與複合被動句中“給”後賓語的省略問題及其誘因 [The ellipsis of the object of *gei* in the compound *ba* construction and compound passive construction and its reason]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2008.6:535-543.
- Yang, Jichu (楊霽楚). 2008. *Xiandai Hanyu ‘Gei+VP’ Jiegou Yanjiu* 現代漢語“給+VP”結構研究 [*A Study on ‘Gei+VP’ in Mandarin Chinese*]. Beijing: Peking University MA thesis.

- Yang, Xiao (楊嘯). 2003. *Qingdai 'Gei' Ziju Kaocha: Jian Lun Biao Beidong 'Gei' Ziju de Qiyuan* 清代“給”字句考察：兼論表被動“給”字句的起源 [Investigation on 'Gei' Sentence of Qing Dynasty: Involvement on Origin of Passive Sentence with 'Gei']. Shanghai: Shanghai Normal University MA thesis.
- Zhang, Bojiang (張伯江). 2009. *Cong Shi Shou Guanxi Dao Jushi Yuyi* 從施受關係到句式語義 [Grammatical Relations and Grammatical Constructions]. Beijing: The Commercial Press.
- Zhang, Yisheng (張誼生). 2001. Zhuci 'gei' ji qi xiangguan de jushi 助詞“給”及其相關的句式 [On the auxiliary word *gei* (給) and sentence types]. *Hanyu Xuebao* 漢語學報 [Chinese Linguistics] 3:24-33.

On the Development of *Gei*+VP in Modern Chinese

Bojiang Zhang

Chinese Academy of Social Sciences

It is still under heated discussion whether the affective ‘*gei*’ in such a sentence as ‘*na le sandiao ranglaowang gei nachuqule*’ is related to verbs of giving and the preposition marking benefactive or dative. Though most scholars take it as affective, the grammaticalization of the ‘giving’ verb in Chinese is unique in the world languages. Chinese data show that the omission of the preposition ‘*gei*’ was common in the early modern Chinese, which is thought to be related to the information weakness of personal pronoun. Since the omission co-occurs with the affective ‘*gei*’ in *gei*+VP with clear transition tendency, it is argued that the affective *gei* was the development of the omission of benefactive or dative. Theoretically, the assumption explains why the grammaticalization chain of GIVE > BENEFACTIVE > DATIVE evolved into the marking of verbal affectiveness. What’s more, it can explain other linguistic facts such as why the *gei*+VP is embedded in ‘*bei/ba*’ construction, and why the VP in *gei*+VP can be one-argument predicate.

Key words: benefactive, dative, omission, affective

歷時與共時： 起始類介詞的語意發展和多義現象*

熊慧如

國立台南大學

現代漢語常見的起始類介詞 (ablative prepositions) 有「於」、「自」、「從」、「由」、「自從」、「打」和「打從」，其中除了「打」和「打從」遲至中古漢語晚期至近代漢語前期之間才首次出現用例之外，其餘皆為古老介詞，在上古時期就已經存在。此類介詞在發展過程中歷經語意的演變，不僅引介動作行為發生的處所出發點與時間開始點，亦能表達其他與空間、時間或抽象概念相關的語意，充分表現出介詞一詞多義的特徵。起始類介詞的多義性現象反映出語言使用者的心理認知活動，藉由「起點-路徑-終點」空間位移基本圖式以及隱喻的機制，不僅使得同一詞彙能夠表達動作行為發生的不同空間位置，並且進一步將空間概念域的用法轉移到時間或其他抽象的概念域。詞彙一旦在演變過程中發展成為多義詞之後，很容易在「相因生義」的類推作用下取得其他詞彙的義項，而彼此成為同義詞。而當二個詞彙成為同義詞之後，又會互相影響取得彼此的其他義項，繼而發展出更多語意。介詞的多義和同義之間關係密切且互相影響，既是造成詞彙演變的因，也是詞彙演變之後的果。

關鍵詞：起始類介詞，語意演變，多義詞，同義詞，空間位移基本圖示，隱喻，相因生義

1. 緒論

現代漢語常見的起始類介詞 (ablative prepositions) 有七個：「於」、「自」、「從」、「由」、「自從」、「打」和「打從」，其主要功能為引介動作行為發生的處所出發點或時間開始點。此類介詞的歷史非常久遠，最早可追溯至上古前期¹（西

* 本文部分研究內容接受國科會專題計畫之經費補助（計畫編號：NSC 96-2411-H-024-004）。

¹ 學者對漢語發展史分期的標準不盡相同，名稱使用也略有出入。有學者將 Modern Chinese 稱為現代漢語，也有學者稱之為近代漢語，而當代 (contemporary) 使用的漢語才稱為現代漢語。本文採

元前第 14 至第 11 世紀)，例如「於」(于)²、「自」、「從」(从)和「由」(繇)，在甲骨文中即已使用，其他介詞則在較晚的時期陸續出現：雙音介詞「自從」於春秋戰國時期出現（西元前 770 至 221 年）；「打」於中古晚期（西元第 6 世紀至 1250 年）發現用例；「打從」則遲至宋代（約西元第 13 世紀）才在文獻中出現 (Hsiung 2007, 2008)。

除了表達處所或時間的起始點之外，大部分的起始類介詞也都能表達其他與空間、時間或抽象概念相關的語意。Taylor (1995) 認為不論哪種語言，介詞都是所有詞類當中最為多義 (polysemous) 的一類。雖然一詞多義看似不同語意之間的一種共時關係，然而實際上一個詞的不同義項之間存在著歷史上的淵源。隨著語言的演變，詞彙在不同的時期發展出新語意，但舊語意不一定消亡，甚至可能一直存在，新舊語意共存的情況下就使得一個單詞能夠表達二個以上的語意。

當詞彙一旦產生不同的語意而形成多義詞之後，若其中的某個語意和其他詞彙的語意相同時，則與該詞成為同義詞（或近義詞）。王力 (1991) 指出：「所謂同義，是說這個詞的某一意義和那個詞的某一個意義相同，不是說這個詞的所有意義和那個詞的所有意義都相同」。除了詞義相同才能形成同義詞之外，「時間」也是重要的因素之一。有些詞可能在上古時期互為同義詞，到了中古時期或近代卻不同義；相反的，也有些詞彙之間原來並非同義詞，而是隨著時間的發展才產生相同的語意。歷時與共時的作用不斷影響詞彙，促使詞彙產生多項語意，往多義詞的方向發展；同時，也促使不同詞彙之間因具有相同語意而形成同義詞，而同義詞的形成，又誘發新語意的產生。因此，詞彙的多義與同義之間，關係緊密，互為因果。

本文將從起始類介詞的語意發展，觀察同義詞的形成以及介詞的多義現象，並且試著對介詞的多義現象提出一些解釋。首先，我們從歷時的面向詳細考察起始類介詞從甲骨文時期開始，橫跨上古時期、中古時期、近代漢語時期，一直到現代漢語的語意演變。其次，我們將從共時的面向，以認知語言學與詞彙語意學

用貝羅貝先生 (Peyraube 1988) 所提出的分期標準與名稱：上古前期 (Pre-archaic Chinese)：西元前第 14 世紀至西元前第 11 世紀；上古早期 (Early Archaic Chinese)：西元前第 11 世紀至西元前第 6 世紀；上古晚期 (Late Archaic Chinese)：西元前第 5 世紀至西元前第 3 世紀；中古前期 (Pre-medieval Chinese)：西元前 206 年至西元 220 年；中古早期 (Early Medieval Chinese)：西元第 3 世紀至西元第 6 世紀；中古晚期 (Late Medieval Chinese)：西元第 6 世紀至 1250 年；近代前期 (Pre-modern Chinese)：西元第 13 世紀至西元第 14 世紀；近代漢語 (Modern Chinese)：西元第 15 世紀至西元第 19 世紀；現代漢語 (Contemporary Chinese)：西元第 20 世紀至今。

² 「于」和「於」為上古時期的二個變體，但二者之間的用法並無不同（張楨 2002），可當作是時空中的區別（魏培泉 1993）。「從」和「从」以及「由」和「繇」的情形也是一樣的。因此，為了行文的便利，本文皆以「於」、「從」、「由」來表示，但在引用例句時則依照原文的用字。

的方法來分析多義詞形成的可能因素。

2. 語意發展

起始類介詞在發展過程中，歷經舊語意衰落、新語意興起，甚至有些語意消失又出現的情形。因此這類介詞的語意龐雜，除了「自從」之外，所有介詞都能同時表達與空間及時間範疇有關的語意，有些還能表達其他概念範疇的語意，例如對象、原因等。爲了能夠清楚呈現同一時期語意發展的脈絡並進行比較，本文針對起始類介詞的空間和時間語意進行觀察，包括：處所起點、處所定點、處所經由和處所方向/終點、時間起點、時間定點和時間終點等七種語意，藉此考據出每個介詞在不同時期的語意變化情形。

2.1 「於」的語意變化

「於」在上古時期爲高頻介詞，也是起始類介詞中語意最豐富的一個。尤其在中古前期和中古早期語意蓬勃發展，可以引介所有空間和時間範疇的七個語意。然而從中古晚期開始，有些語意逐漸衰落，到了現代漢語，只剩下處所起點、處所定點和時間定點三個用法。由於「於」屬於文言和書面語體的詞彙，因此大約從中古早期以後，受到白話語體逐漸流行的影響，頻率逐漸降低，到了現代漢語已不是常用介詞。以下例句爲「於」曾經出現過或目前仍能表達的語意，所有語意的變化情形列於〈表 1〉。

(1) 命長之囚，出 <u>於</u> 牢中（論衡·偶會）	表處所起點
(2) 辛酉王田 <u>于</u> 雞泉（合集 37848）	表處所定點
(3) 遇桓 <u>於</u> 岸上過（世說新語·任誕）	表處所經由
(4) 太公乃遷康公 <u>於</u> 海上（史記·田敬仲完世家）	表處所終點
(5) 始 <u>於</u> 今日（中本起經）	表時間起點
(6) 雖 <u>於</u> 時有盛名而行不由本者（三國志·魏書十二）	表時間定點
(7) 不常厥邑， <u>于</u> 今五邦（尚書·盤庚）	表時間終點

〈表 1〉「於」的語意發展

	P-A ³	E-A	L-A	P-ME	E-ME	L-ME	P-MO	MO	C
空間語意									
處所起點		+	+	+	+	+	+	+	+
處所定點	+	+	+	+	+	+	+	+	+
處所經由				+	+	+			
處所方向/終點	+	+	+	+	+	+	+	+	
時間語意									
時間起點			+	+	+				
時間定點	+	+	+	+	+	+		+	+
時間終點		+	+	+	+	+			

2.2 「自」的語意變化

「自」是上古時期最重要的起始類介詞，雖然甲骨文中只有處所起點和時間起點二個用法，但一直沿用到現代漢語都不曾間斷，是最主要的語意。從上古早期至中古晚期是語意發展最蓬勃的階段，但有些語意存在的時間並不長，例如表示處所方向/終點的用法；也有些用法只出現零星用例，例如表示時間定點的用法。和「於」一樣，「自」亦偏向書面語體，因此從中古時期以後使用頻率明顯下降，在現代漢語已不是常用介詞。

- | | |
|---------------------------------------|-------|
| (8) 有朋 <u>自</u> 遠方來，不亦樂乎（論語·學而） | 表處所起點 |
| (9) 夫人 <u>自</u> 帷中再拜（史記·孔子世家） | 表處所定點 |
| (10) 必有 <u>自</u> 海道至者（朱子語類卷一百一） | 表處所經由 |
| (11) 達 <u>自</u> 荊州（高僧傳卷七） | 表處所終點 |
| (12) <u>自</u> 古仙人出入乘五色雲（敦煌變文集·太子成道變文） | 表時間起點 |
| (13) <u>自</u> 虞、夏時，貢賦備矣（史記·夏本紀） | 表時間定點 |

³ P-A：上古前期，E-A：上古早期，L-A：上古晚期，P-ME：中古前期，E-ME：中古早期，L-ME：中古晚期，P-MO：近代漢語前期，MO：近代漢語，C：現代漢語。

〈表 2〉「自」的語意發展

	P-A	E-A	L-A	P-ME	E-ME	L-ME	P-MO	MO	C
空間語意									
處所起點	+	+	+	+	+	+	+	+	+
處所定點		+	+	+	+				
處所經由		+	+	+	+	+			+
處所方向/終點					+	+			
時間語意									
時間起點	+	+	+	+	+	+	+	+	+
時間定點		+		+		+			
時間終點									

2.3 「從」的語意變化

「從」不僅是在甲骨文即已出現的古老介詞，也是在現代漢語中使用頻率最高的起始類介詞。在上古早期曾經失去介詞的用法，但在上古晚期重新出現之後，即發展出四項語意：處所起點、處所定點、處所經由和時間起點，並且沿用至今。「從」在演變過程中也發展出其他語意，但是維持的時間不長，用例也不多。以時間定點為例，在近代漢語時期只在《朴通事諺解》出現一例，《老乞大諺解》和《金瓶梅》各出現二例，而同時期的《紅樓夢》則不見此用法。

- | | |
|-------------------------------------|----------|
| (14) 蝗蟲 <u>從</u> 東方來（史記·秦始皇本紀） | 表處所起點 |
| (15) 夫 <u>從</u> 坐上語侵之（史記·魏其武安侯列傳） | 表處所定點 |
| (16) 大道人 <u>從</u> 何徑來（中本起經） | 表處所經由 |
| (17) 劉姥姥感謝不盡，仍 <u>從</u> 後門去了（紅樓夢 6） | 表處所方向/終點 |
| (18) 你 <u>從</u> 幾時離了王京（老乞大諺解） | 表時間定點 |

〈表3〉「從」的語意發展

	P-A	E-A	L-A	P-ME	E-ME	L-ME	P-MO	MO	C
空間語意									
處所起點	+		+	+	+	+	+	+	+
處所定點			+	+	+	+	+	+	+
處所經由	+		+	+	+	+		+	+
處所方向/終點						+	+	+	
時間語意									
時間起點	+		+	+	+	+	+	+	+
時間定點							+	+	
時間終點									

2.4 「由」的語意變化

和其他古老的起始類介詞「於」、「自」和「從」相較起來，「由」的語意功能單純許多，總共只發展出三項與空間及時間相關的語意：處所起點、處所經由和時間起點，其中以處所經由為主要用法。這三項語意在上古晚期就已發展完成，直至後期都沒有新的變化。

- (19) 他日由鄒之任（孟子·告子下）表處所起點
 (20) 而鄧艾自陰平由景谷道傍入（三國志·蜀書十四）表處所經由
 (21) 由湯至於武丁，聖賢之君六七作（孟子·公孫丑上）表時間起點

〈表4〉「由」的語意發展

	P-A	E-A	L-A	P-ME	E-ME	L-ME	P-MO	MO	C
空間語意									
處所起點		+	+	+	+	+		+	+
處所定點									
處所經由		+	+	+	+	+		+	+
處所方向/終點									
時間語意									
時間起點	+		+	+	+	+		+	+
時間定點									
時間終點									

2.5 「自從」的語意變化

陳昌來 (2002) 認為「自從」的出現代表漢語真正的雙音節介詞開始產生，具有重要的意義。然而此介詞並非如陳文所言是在《敦煌變文集》中第一次出現，事實上，「自從」在《戰國策》中即已出現一個用例（參見例句 (22)），在下一期的《史記》中又有一例（參見例句 (23)）。因此，「自從」初次出現的時期應在西元前第三世紀末至西元前第二世紀初之間 (Hsiung 2007)。

「自從」從出現一直到現代漢語之間，只發展出引介時間起點的用法，是起始類介詞中語意功能最單一的，也是現代漢語中此用法頻率最高的介詞。

- (22) 自從先君文王以至不穀之身（戰國策・楚策） 表時間起點
 (23) 自從窮蟬以至帝舜，皆微為庶人（史記・五帝本紀） 表時間起點

〈表 5〉「自從」的語意發展

	P-A	E-A	L-A	P-ME	E-ME	L-ME	P-MO	MO	C
空間語意									
處所起點									
處所定點									
處所經由									
處所方向/終點									
時間語意									
時間起點			+	+	+	+	+	+	+
時間定點									
時間終點									

2.6 「打」的語意變化

「打」屬於口語性較強的介詞，主要使用於中國北方的方言（呂叔湘 1980），但在南方的某些方言中亦不乏用例，例如江西的石城方言（曾毅平 2000）和海南的屯昌方言（錢燮香 2000）。「打」的介詞用法在中古早期以前的文獻中尚未見到用例，應是在宋代（西元 960 至 1279 年）才出現。而在近代漢語前期之前只能引介處所經由，到了近代漢語時期才發展出其他語意，包括處所起點、時間起點和處所終點。

在現代漢語中，「打」主要引介處所起點，雖然這個用法出現的時間晚於處所經由，但用例反而較多。時間起點的用法罕見，在《紅樓夢》中仍不見用例，我們同意白維國 (1999) 的看法，這個語意的形成應不會早於 18 世紀。而處所終點的用法只在近代漢語短暫出現，在現代漢語已不復見。

- (24) 調領四路兵馬，打清河縣起身（金瓶梅 99） 表處所起點
 (25) 是打哪條路去好（宋代卷・崔待詔生死冤家） 表處所經由
 (26) 正打你行過去（元雜劇・諸葛亮博望） 表處所方向/終點
 (27) 主兒打毛團子似的掇弄到這麼大（兒女英雄傳 40） 表時間起點

〈表 6〉「打」的語意發展

	P-A	E-A	L-A	P-ME	E-ME	L-ME	P-MO	MO	C
空間語意									
處所起點								+	+
處所定點									
處所經由						+	+	+	+
處所方向/終點								+	
時間語意									
時間起點								+	+
時間定點									
時間終點									

2.7 「打從」的語意變化

「打從」和「打」一樣，都是屬於口語語體的方言詞彙，二者的發展過程也非常類似，皆以處所經由為原始語意，再逐漸發展出其他用法。在現代漢語中，二個介詞的語意也無二致，皆可引介處所起點、處所經由和時間起點。

- (28) 我竟打從心底羨慕起他們了（現代漢語） 表處所起點
 (29) 打從胡州過（張協狀元） 表處所經由
 (30) 如今打從那裏查去（紅樓夢 111） 表處所終點
 (31) 打從我們出生的那一剎那到目前為止（現代漢語） 表時間起點

〈表 7〉「打從」的語意發展

	P-MO	MO	C
空間語意			
處所起點		+	+
處所定點			
處所經由	+	+	+
處所方向/終點		+	
時間語意			
時間起點		+	+
時間定點			
時間終點			

3. 起始類介詞的多義現象

Lakoff (1987) 認為，多義詞的不同詞義之間，某些詞義可能比其他詞義更具代表性，而且彼此之間透過各種方法被緊密的聯繫在一起。起始類介詞的每個義項之間也存在著某種關連性，而使這些不同詞義產生關連的原因，可能來自於語言使用者的認知機制，也可能是詞彙與詞彙之間互相影響所形成的結果。本節將探討起始類介詞在發展過程中產生的多義現象，試著從認知心理和詞彙語意二種角度解釋一詞多義形成的原因。

3.1 空間認知圖示

Lakoff (1987) 與 Brugman (1988) 曾針對英語介詞 *over* 做過詳盡的分析，列舉這個介詞所有的空間意義與引伸的抽象意義，他們認為介詞在句子裡承擔了全部與空間有關的語意，包括與空間位移有關的起點、路徑、終點……等。但也有語言學家並不贊同這樣的看法 (Evans & Tyler 2005)，他們認為句子表現出來的動態語意並非僅由介詞所賦予，而是由介詞與運動動詞 (motion verb) 以及對世界的知識共同整合的結果，因為英語的空間介詞，例如 *over*、*to* 和 *through*，本身並不具有移動性 (movement)，既然不具移動性，就不能產生路徑、方向或位移等語意。我們同意句子的動態意義應該由介詞詞組、動詞詞組以及語境共同產生，但漢語介詞是否和英語介詞一樣，亦不具動態語意呢？我們以「從」和「在」為例來探討這個問題：

- (32) a. 從台北看天下
b. 在台北看天下

以上二個句子所表達的意思並無差別，介詞「從」和介詞「在」都引介動作行為「看天下」發生的處所「台北」。然而，從介詞的選擇可以看出二個句子所存在的細微差異：(32a)由「從」引導出動作行為發生的處所，使得動作「看」產生視覺空間的延伸，「台北」與「天下」是視線範圍的二個端點，視線從「台北」發出，落在遠端的「天下」，句子隱含虛擬的空間位移。而(32b)由「在」引導動作行為發生的處所，表示動作行為「看」發生在一個定點，想像某個人坐在台北的客廳裡看報紙上的國際新聞，主事者所處的位置與動作行為發生的位置為同一個定點。因此二個句子相較之下，前者具有動態的色彩，後者則只有靜態的語意。儘管動詞「看」並非運動動詞，但是(32a)仍隱含空間的位移，這顯示動態語意來自空間介詞「從」而非來自動詞「看」。

漢語介詞的運動性與介詞的來源有關。許多語言學家已證實大部分的漢語介詞是經由不同的機制從動詞虛化而來的(Peyraube 1994, 1999, Sun 1996, 吳福祥 2003 等)。當動詞轉變成介詞之後，句法功能改變，但二者之間的語意仍有緊密的關係。以介詞「於」和介詞「從」為例，語言學家大多同意這二個介詞是由動詞虛化而來。

郭錫良(1997)與梅祖麟(2004)對介詞「於」的來源做過考察，並且證實是從動詞「於」虛化而來。「於」的動詞用法可在甲骨文中找到用例：

- (33) a. 今日于亳，亡災？(合集 36567)
b. 王往于田，亡災？(合集 557)

(33a)的「于」做動詞用，有「往」、「到」之意；(33b)的「于」有二種可能的分析，第一是將「往于田」視為連動式「V1 + V2 O」；第二是將句子分析為動詞後接介賓詞組的結構，亦即「V + P + O」。梅祖麟認為動詞「於」在甲骨文歷經連動結構的重新分析(reanalysis)，使得「V1 + V2 O」變成「V1 V2 O」再變成「V + P + O」，因而促使「於」從動詞向介詞虛化。

介詞「從」的來源也與動詞有關。從字形來看，「從」(「从」)表示一人跟隨另一人，《說文》釋義為：「从，相聽也」，以及「從，隨行也」。在甲骨文可找到「從」做為「隨行」、「相從」的動詞用法，因此可歸為運動類動詞(陳年福 2001)，例如：

- (34) 王从望乘伐下危（合集 6476）
 (35) 庚午卜，月辛來从斗（乙 174）

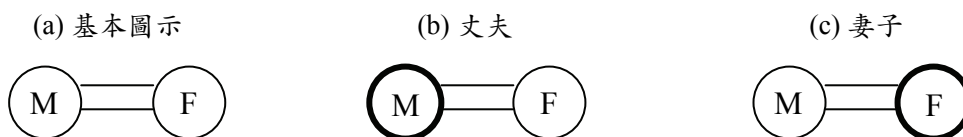
「從」在甲骨文中除了做為動詞之外，也具有介詞的用法。根據馬貝加 (2002) 的研究，「從」由原本與人相隨發展到與空間相隨，便完成了動詞向介詞的轉化，如以下三例（馬貝加 2002:23）：

- (36) 蔡人從之伐戴（左傳・隱公十年）
 (37) 其後余從狄君以田渭濱（左傳・僖公二十四年）
 (38) 子從阪道，劉子從尹道伐尹（左傳・昭公二十年）

「從」在例句 (36) 和 (37) 中仍作動詞，表示人與人相隨，其位置與介詞的句法位置相同，為介詞的形成提供必要條件。例句 (38) 不再是人與人相隨，而是人與空間相隨，已經轉變為介詞用法，引介動作行為經由的處所。由此判斷介詞「從」的來源也與連動式有關，演變的過程為「V1 + O + V2」>「V1 O + V2」>「P + O + V」。

由此不難發現，「於」和「從」由運動動詞虛化為介詞之後，仍保存著運動的特性。物體一旦發生運動，即造成空間的位移，包括起點、路徑和終點。這項運動位移的特性與起始類介詞的多義現象之間有密切的關連性。

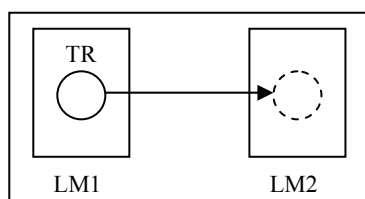
根據 Langacker (1987, 2003) 的認知語法 (Cognitive Grammar)，一個語意結構代表一個敘述 (predication)，而每一個敘述都是由認知域來描述，這個認知域可以是概念結構或是感知的經驗，例如空間、時間，或者其他更抽象的概念。每一個認知域都可以用一個基本圖式 (base) 來表示，這個基本圖式包含了一個表達所要引導出來的意義。而當這個基本圖式中的某個次結構被標明出來時，也就是一個表達所要強調的部分或是最明顯的部分。〈圖 1〉中的三個圖表示一個基本圖式下所包含的不同次結構：



〈圖 1〉夫妻基本圖式 (Langacker 2003:252)

〈圖 1〉表達的是一男 (M) 一女 (F) 所組成的婚姻關係，(a)、(b)、(c) 有類似的圖式，差別在於以圓圈表示的次結構是否以粗線條來標明，以粗線條標明的次結構也就是一個表達中所要凸顯 (profile) 的部分。(a) 為基本圖式，二個圓圈沒有粗細之分，而 (b) 圖和 (c) 圖中的粗體圓圈分別落在 M 和 F。這表示 (b) 圖和 (c) 圖所強調的次結構不一樣：(b) 強調的是婚姻關係中的男性，亦即丈夫，而 (c) 強調的是婚姻關係中的女性，亦即妻子。因此，當一個表達的基本結構中有不同的部分被凸顯時，就會使得該表達具有不同的語意。

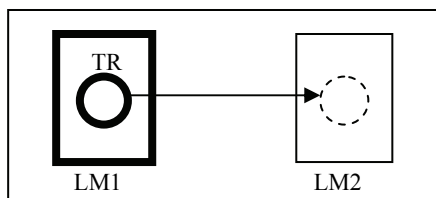
起始類介詞的多義現象也具有同樣的認知基礎，介詞的運動特性產生一個完整的「起點-路徑-終點」空間位移基本圖式，如〈圖 2〉所示：



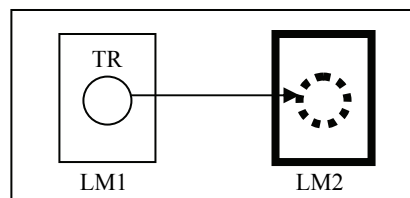
〈圖 2〉起始類介詞的空間位移基本圖式

圖中包含一個 TR，二個 LM 以及一個箭頭，TR (trajector) 指的是在位移圖式中會移動的物體，以實線的圓形表示，而移動之後的 TR 則以虛線的圓形表示。LM1 (landmark 1) 和 LM2 (landmark 2) 為標示 TR 所在位置的界標，以長方形表示；箭頭則表示 TR 移動的路徑。

引介處所起點的圖式可由〈圖 3〉表示，當空間位移基本圖式中的 LM1 被凸顯的時候（以粗線條標示），表示 TR 是以 LM1 為起點移動到 LM2，因此「起點」的語意在「起點-路徑-終點」空間位移基本圖式中被引導出來。相反的，當 LM2 做為被凸顯的部分時，則表示 TR 從 LM1 移動之後將要到達或是已經到達的位置，在「起點-路徑-終點」空間位移基本圖式中引導出「方向」或「終點」的語意，如〈圖 4〉所示。

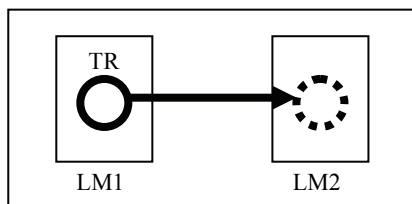


〈圖 3〉起點



〈圖 4〉終點

當「起點-路徑-終點」空間位移基本圖式中被凸顯的部分為箭頭的時候，如〈圖 5〉所示，強調的是 TR 從 LM1 運動到 LM2 之間的路徑，因此引介動作行為經由之處所這個語意就被引導出來。

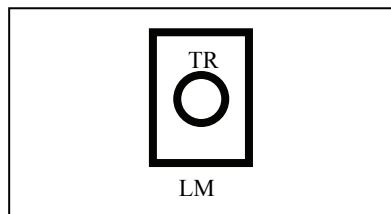


〈圖 5〉路徑

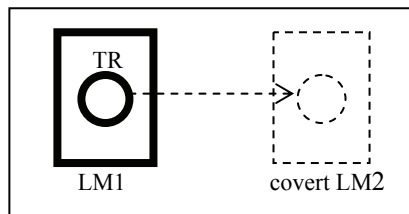
起始類介詞除了表達所自、所經、所到的語意之外，也有可能引介「所在」的用法，亦即表示動作行為發生的定點。在「起點-路徑-終點」空間位移基本圖式中，只包含一個 LM，且 TR 座落在由 LM 標明的位置上。這種空間關係可以進一步分為二種，以例句 (39) 說明：

(39) 長史從門外下車，步入尚書（世說新語·容止）

例句 (39) 由二個獨立的句子組成，包含二個事件，而且二個事件為一連續性的動作。第一個句子「長史從門外下車」沒有提及主事者從何處來，要往何處去，只描述動作（下車）發生的處所，TR 與 LM 之間的空間關係如〈圖 6〉所示。第二個句子「步入尚書」表示動作發生的處所終點，雖然是獨立的句子，但主事者從下車直到步入尚書才完成整個動作。整個句子看來，「從」的介詞詞組引介的是動作發生的處所定點（即所在），但是由於第二個句子隱含處所終點的語意，使得「從」引導的詞組產生表達處所起點（即所自）的語意，因此第二個句子的處所（「尚書」）可以看成是隱性的 (covert) LM2，如〈圖 7〉中以虛線標示的部分。



〈圖 6〉處所定點



〈圖 7〉處所定點 (含隱性 LM2)

當動作行為與看視類動詞有關時，起始類介詞詞組所引介的處所經常具有和〈圖 7〉一樣的認知圖示，例如：

(40) 諸將皆從壁上觀（史記·項羽本紀）

雖然觀看的動作是發生在一個定點（「壁上」），如認知圖式〈圖 6〉所示。但視覺的延伸和視覺的落點造成一個虛擬空間位移的路徑與終點，亦即〈圖 7〉中隱性的箭號和 LM2。因此語言使用者會選擇使用「從」而不是「在」，確實是認知運作的結果。

一詞多義的現象是詞彙發展的產物，所有的詞義不一定在詞彙產生之初就會同時顯現出來，但是語言使用者的認知機制提供詞彙本身具有多種詞義的環境，有些詞義受到凸顯之後，漸為社會接受而使用頻率增高，進而促使潛在語意變成顯性語意。我們可以從句子的歧義性獲得證實：當一個句子出現二種以上的詮釋時，是由於不同的語言使用者凸顯了認知圖式中不同部分之緣故，如下列甲骨文中「從」的例子：

(41) 从向歸（殷契粹編 1067）

(42) 王往于田从東（林 2.22.11）

在上列二個句子當中的介詞「從」，有的學者將之解釋為引導動作行為發生的處所起點；也有學者將之視為動作行為經由的處所。事實上，在缺乏明確語境的情況下，這二種解釋都是合理的，因為「所自」和「所經」皆為「從」的可能語意，當說話者凸顯的部分不同，就使得一個表達產生不同的語意。

3.2 隱喻

起始類介詞除了引介具體的空間概念，也引介抽象的時間概念。空間經驗是人類存在這個世界上最基本的一種經驗，比其他經驗更為中心或典型，許多抽象概念間接地或直接地透過空間概念來理解或表達。其中透過隱喻 (metaphor) 的方式將空間範疇映射到時間範疇已被大量討論並且獲得證實 (Lakoff 1987, Heine et al. 1991, Haspelmath 1997, Talmy 2000, Traugott 2002, etc.)，因此，隱喻造成二個不同概念域之間的轉換也是促使起始類介詞一詞多義的重要機制之一。

以「於」為例，在甲骨文中，此介詞在空間範疇可引介處所定點和處所終點，而在時間範疇也可引介時間定點，雖然在這個時期尚無引介時間終點的用

法，但受到不同概念域的隱喻映射作用，到了上古早期就發展出來了。其他介詞也有同樣的情形，當詞彙具有表達處所起點的語意時，空間概念轉移到時間概念，也使得該詞具有引介時間起點的可能性。隱喻結構是局部性的概念轉移 (Lakoff & Johnson 1980)，亦即空間範疇的概念並不會完全映射到時間範疇，因此起始類介詞的空間語意不必然全部轉移到時間語意上，但卻為時間語意的形成提供一個充分的條件。

3.3 同義詞與相因生義

當詞彙在演變過程中發展成多義詞之後，很容易與其他詞彙的某個義項相同而彼此成為同義詞。一旦成為同義詞之後，二個詞彙之間又會互相影響，取得彼此的其他義項，繼而產生更多語意。和王力的看法一致，蔣紹愚 (2001:96) 也認為「同義詞」是在某些意義上同義，而在另一些意義上不相同，他為詞彙發展的途徑提出另一個理論，即詞彙的「相因生義」。當 A 詞原來只和 B 詞的一個義位 B1 相通，透過類推 (analogy) 的作用之後，A 詞又取得了 B 詞的另一個義位 B2，這樣的類推一旦得到社會的認同，新的詞義於是產生。

同樣的機制也可以解釋起始類介詞的發展。馬貝加 (2002:23) 認為：「『從』之成為介詞『自』的同義詞，是從表經由開始的。最早見於春秋末……」。我們認為這個看法值得再商榷。根據 Hsiung (2007) 的分析，「自」在《詩經》僅出現一例表示經由的用法，到了上古晚期才稍微增多，但用例仍不普遍，只出現在《左傳》中，而且很快的這個用法在中古時期以後即消失（參考〈表 2〉）。從使用的時間長度和頻率高低來判斷，表經由並非「自」本身的主要語意，應是受其他詞彙影響才產生的。事實上，「自」和「從」在甲骨文皆可表示處所起點，二個介詞應是在這個時期即已成為同義詞，並且是在成為同義詞之後，「自」才取得「從」表經由的這個義項。

雖然學界對於「從」在甲骨文究竟是表起點、表經由抑或是二者兼具，仍持不同的看法，但我們可以從三方面討論做為判斷的依據：第一、從實詞向虛詞演化的歷程來看，動詞「從」表示「隨行」之意，發展為介詞之後的初始語意極有可能為表經由，而一旦具有這個語意之後，便可以很快或同時的衍生出其他用法。第二、如前文所言，一詞多義是詞彙發展的產物，具有人類使用語言的認知基礎，一個表達在一個基本認知圖式下得以產生不同的語意。因此表經由與表起點的用法皆為「從」的潛在語意，不應排除「從」在甲骨文時期即有引介起點用法之可能性（參考例句 (41) 和 (42)）。第三、「從」在甲骨文時期即可引介時間的

起點，以空間經驗先於時間經驗的認知運作來看，時間起點是從空間起點透過隱喻的機制轉移而來的，所以「從」在這個時期應該已經具有引介處所起點的用法。由此，「從」在甲骨文時期即可同時引介處所起點與經由，應是合理的判斷。也因此「自」和「從」早在上古前期即因皆具有表處所起點的語意而成爲同義詞，並非遲至春秋末期才發生。而當「於」在西周時期也發展出引介起點的用法時，就與「自」及「從」成爲同義詞。「於」和「自」的情形也是一樣的，二個介詞成爲同義詞之後，「自」受到「於」的影響，在上古晚期取得引介處所定點的用法。

4. 結論

從甲骨文時期至今，是一段漢語發展的漫長歷程，起始類介詞在這三千四百年的時間裡歷經許多重要的語意演變。在演變過程中，詞彙因語意的發展形成多義詞，又因多義詞的產生而與其他詞彙成爲同義詞，二個詞彙一旦互爲同義詞之後，又促使語意的進一步發展，產生更多語意。因此，介詞的多義和同義之間關係密切且互相影響，既是造成詞彙演變的因，也是詞彙演變之後的果。

Györi (2002) 指出，語言的歷時變化是共時活動的結果，而且語言的改變與人類認知的普遍原則息息相關，都是爲了達到溝通的目的。人類的認知活動是誘發語意演變的重要機制，在一個語意結構中不同意涵的凸顯，以及二個認知域之間的隱喻轉換都促使詞彙產生多項語意。此外，詞彙的發展不只在單一詞彙的系統內部運作，還涉及到與其他詞彙的互動，同義詞之間因相因生義的類推作用而取得新的義項，也是引起詞彙語意發展的重要原因之一。

詞彙的演變是個複雜的過程，並非單一因素就可以發展至今日的面貌。新詞義的產生是認知機制運作的結果，抑或是詞彙本身相因生義所造成，很難清楚的劃分開來。誠如蔣紹愚 (2001) 所言：「新義的形成必須取得社會的承認」，無論是哪一種因素，都促使語言不斷的發展，也爲語意的變化提供許多可能性。

引用文獻

- Bai, Weiguo (白維國). 1999. Jieci 'da' de lai yuan 介詞“打”的來源 [The origin of the preposition *da*]. *Jindai Hanyu Yanjiu* 近代漢語研究 [Study on Modern Chinese Language], Vol. 2, ed. by Shaoyu Jiang & Lansheng Jiang, 330-339. Beijing: The Commercial Press.
- Brugman, Claudia. 1988. *The Story of Over: Polysemy, Semantics, and the Structure of the Lexicon*. New York: Garland.
- Chen, Changlai (陳昌來). 2002. *Jieci yu Jieyin Gongneng* 介詞與介引功能 [Prepositions and their Introductory Functions]. Hefei: Anhui Education Press.
- Chen, Nianfu (陳年福). 2001. *Jiaguwen Dongci Cihui Yanjiu* 甲骨文動詞詞彙研究 [Studies on Verbs in the Oracle Bone Inscriptions]. Chengdu: Bashu Shushe.
- Djamouri, Reduane. 2001. Shang markers of predication. *Sinitic Grammar: Synchronic and Diachronic Perspectives*, ed. by Hilary Chappell, 143-171. Oxford & New York: Oxford University Press.
- Djamouri, Reduane, and Waltraud Paul. 1997. Les syntagmes prépositionnels en *yu* et *zai* en chinois archaïque. *Cahiers de Linguistique – Asie Orientale* 26.2:221-248.
- Evans, Vyvyan, and Andrea Tyler. 2005. Rethinking English 'prepositions of movement': the case of *to* and *through*. *Adpositions of Movement*, ed. by Hubert Cuyckens, Walter De Mulder & Tanja Mortelmans, 247-270. Amsterdam & Philadelphia: John Benjamins.
- Feng, Chuntian (馮春田). 1991. *Jindai Hanyu Yufa Wenti Yanjiu* 近代漢語語法問題研究 [Studies on the Grammar of Modern Chinese]. Ji'nan: Shandong Education Press.
- Guo, Xiliang (郭錫良). 1997. Jieci 'yu' de qi yuan he fazhan 介詞“于”的起源和發展 [The origin and development of the preposition *yu*]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1997.2:131-138.
- Györi, Gabor. 2002. Semantic change and cognition. *Cognitive Linguistics* 13.2:123-166.
- Haspelmath, Martin. 1997. *From Space to Time: Temporal Adverbials in the World's Languages*. München: Lincom Europa.
- Heine, Bernd, Ulrike Claudi, and Friederike Hünemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Hsiung, Hui-ju (熊慧如). 2006. Duoyi yu tongyi: qishi jieci de lishi yanbian yu renzhi jichu 多義與同義：起始介詞的歷時演變與認知基礎 [Polysemy and synonymy: the diachronic changes of the ablative prepositions and their cognitive foundations]. Paper presented at the 7th Chinese Lexical Semantics Workshop (CLSW-7), May 22-24, 2006. Hsinchu: National Chiao Tung University.

- Hsiung, Hui-ju. 2007. *Evolution des prépositions ablatives en chinois du XIV^e siècle av. J.-C. au XIII^e siècle*. Paris: EHESS dissertation.
- Hsiung, Hui-ju. 2008. The development of the ablative prepositions from Modern to Contemporary Chinese. Paper presented at the XVIIth Biennial Conference of the European Association for Chinese Studies (EACS), August 6-10, 2008. Sweden: Lund University.
- Huang, Weijia (黃偉嘉). 1987. Jia-Jinwen zhong 'zai', 'yu', 'zi', 'cong' si zi jieci yongfa de fazhan bianhua ji qi xianghu guanxi 甲金文中“在、於、自、從”四字介詞用法的發展變化及其相互關係 [The development of the prepositions *zai*, *yu*, *zi* and *cong* in the Scripts on Oracle Bones and Bronze Vessels, and their interrelations]. *Journal of Shanxi Normal University* 1987.1:66-75.
- Jiang, Shaoyu (蔣紹愚). 2001. *Hanyu Cihui Yufashi Lunwenji* 漢語詞匯語法史論文集 [A Collection of Articles on History of Chinese Vocabulary and Grammar]. Beijing: The Commercial Press.
- Jiang, Shaoyu (蔣紹愚). 2005. *Jindai Hanyu Yanjiu Gaiyao* 近代漢語研究概要 [An Outline of the Study of Early Modern Chinese]. Beijing: Peking University Press.
- Lakoff, George. 1987. *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, Vol. 1: *Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2003. Constructional integration, grammaticization, and serial verb constructions. *Language and Linguistics* 4.2:251-278.
- Liu, Danqing (劉丹青). 2003. *Yuxu Leixingxue yu Jieci Lilun* 語序類型學與介詞理論 [Typology of Word Order and Theories of Prepositions]. Beijing: The Commercial Press.
- Lü, Shuxiang (呂叔湘). 1980. *Xiandai Hanyu Babai Ci* 現代漢語八百詞 [Eight Hundred Words in Modern Chinese]. Beijing: The Commercial Press.
- Ma, Beijia (馬貝加). 2002. *Jindai Hanyu Jieci* 近代漢語介詞 [The Prepositions in Modern Chinese]. Beijing: Zhonghua Book Company.
- Mei, Tsu-Lin (梅祖麟). 2004. Jieci 'yu' zai Jiaguwen he Han-Zangyu li de qiyuan 介詞“于”在甲骨文和漢藏語裡的起源 [The origin of the preposition *yu* in the Oracle Bone Inscriptions and in Sino-Tibetan languages]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 2004.4:323-332.
- Pan, Yukun (潘玉坤). 2005. *Xi Zhou Jinwen Yuxu Yanjiu* 西周金文語序研究 [The Study of Word Order in Inscriptions on Bronze Wares of the Western Zhou Dynasty]. Shanghai: East China Normal University Press.

- Peyraube, Alain. 1988. *Syntaxe diachronique du chinois: évolution des constructions datives du XIV^e siècle AV. J.-C. au XVIII^e siècle*. Paris: Collège de France.
- Peyraube, Alain. 1994. On the history of Chinese locative prepositions. *Chinese Languages and Linguistics*, Vol. 2: *Historical Linguistics*, ed. by Paul Jen-kuei Li, Chu-Ren Huang & Chih-Chen Jane Tang, 361-387. Taipei: Institute of History and Philology, Academia Sinica.
- Peyraube, Alain. 1999. Historical change in Chinese grammar. *Cahiers de Linguistique – Asie Orientale* 28.2:177-226.
- Qi, Huyang (齊滄揚). 1998. *Xiandai Hanyu Kongjian Wenti Yanjiu* 現代漢語空間問題研究 [A Study on Space in Contemporary Chinese]. Shanghai: Xuelin Press.
- Qian, Dianxiang (錢奠香). 2000. Tunchang fangyan de jieci 屯昌方言的介詞 [The prepositions in Tunchang dialect]. *Jieci* 介詞 [Prepositions], ed. by Rulong Li & Song Hing Chang, 172-184. Guangzhou: Jinan University Press.
- Sun, Chaofen. 1996. *Word-order Change and Grammaticalization in the History of Chinese*. Stanford: Stanford University Press.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*, Vol. 1 & Vol. 2. Cambridge: MIT Press.
- Tang, Lan (唐蘭). 1995. *Tang Lan Xiansheng Jinwen Lunji* 唐蘭先生金文論集 [Mr. Tang Lan's Articles on Bronze Inscriptions]. Beijing: Forbidden City Publishing.
- Taylor, John R. 1995. *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford & New York: Oxford University Press.
- Traugott, Elizabeth Closs, and Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge & New York: Cambridge University Press.
- Vandeloise, Claude. 1986. *L'Espace en Français: sémantique des prépositions spatiales*. Paris: Seuil.
- Wang, Li (王力). 1991. *Tongyuan Zidian* 同源字典 [A Dictionary of Word Families]. Taipei: Wenshizhe Press.
- Wang, Hongbin (王鴻賓). 2004. Shanggu Hanyu jieci de fazhan yu yanbian 上古漢語介詞的發展與演變 [The development of prepositions in Classical Chinese]. *Journal of Shanghai Normal University (Philosophy and Social Sciences Edition)* 33.5:119-125.
- Wei, Pei-chuan (魏培泉). 1993. Gu Hanyu jieci 'yu' de yanbian lueshi 古漢語介詞「於」的演變略史 [The brief history of the evolution of the preposition *yu* in Ancient Chinese]. *Bulletin of the Institute of History and Philology Academia Sinica* 62.4:717-786.
- Wei, Pei-chuan (魏培泉). 2003. Shanggu Hanyu dao Zhonggu Hanyu yufa de zhongyao fazhan 上古漢語到中古漢語語法的重要發展 [The main evolution of Chinese grammar from Archaic to Medieval times]. *Historical Development of Chinese*

- Language*, ed. by Dah-an Ho, 75-106. Taipei: Institute of Linguistics, Academia Sinica.
- Wu, Fuxiang (吳福祥). 2003. Hanyu bansui jieci yufahua de leixingxue yanjiu 漢語伴隨介詞語法化的類型學研究 [A typological study on the grammaticalization of the Chinese comitative prepositions]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2003.1:43-58.
- Yang, Bojun (楊伯峻), and Leshi He (何樂士). 1992. *Gu Hanyu Yufa ji qi Fazhan* 古漢語語法及其發展 [*Grammar of Ancient Chinese and its Development*]. Beijing: Language and Culture Press.
- Zeng, Yiping (曾毅平). 2000. Shicheng (Longgang) fangyan de jieci 石城（龍崗）方言的介詞 [Prepositions in Shicheng (Longgang) dialect]. *Jieci* 介詞 [*Prepositions*], ed. by Rulong Li & Song Hing Chang, 204-226. Guangzhou: Ji'nan University Press.
- Zhang, Cheng (張楨). 2002. *Hanyu Jieci Cizu Cixu de Lishi Yanbian* 漢語介詞詞組詞序的歷史演變 [*The Historical Evolution of the Word-order of Prepositional Phrases in Chinese*]. Beijing: Peking University Press.
- Zhang, Yujin (張玉金). 1994. *Jiaguwen Xuci Cidian* 甲骨文虛詞詞典 [*Dictionary of the Function Words in the Oracle Bone Inscriptions*]. Beijing: Zhonghua Book Company.
- Zhang, Yujin (張玉金). 2001. *Jiaguwen Yufaxue* 甲骨文語法學 [*Syntax of the Chinese Oracle Bone Inscriptions*]. Shanghai: Xuelin Press.

Polysemy and Semantic Change: A Diachronic and Synchronic Study of Chinese Ablative Prepositions

Hui-ju Hsiung

National Tainan University

The most frequent ablative prepositions in Contemporary Chinese are *yu*, *zi*, *cong*, *you*, *zicong*, *da* and *dacong*. In comparison with *da* and *dacong*, which appeared for the first time in the Late Medieval period, the other five prepositions appeared much earlier and were attested before the Late Archaic period. Most of these prepositions have undergone many significant semantic changes since their first appearance. They not only indicate the spatial and temporal provenance of an action, but also express meanings related to time, space, person and more abstract notions. It is from the mental representations of language users that these polysemous senses are derived. By means of the “Source-Path-Goal” schema, the ablative prepositions can introduce different spatial locations where an action takes place, and through metaphorical projection, the spatial concepts are mapped onto other non-spatial domains of concept. Once a polysemy network is built up, a preposition may potentially acquire senses of another by way of analogy, and the two prepositions consequently become synonyms. The semantic change of the ablative prepositions has a close relation with polysemy and synonymy, which behave as a cause and as an effect of the development.

Key words: ablative preposition, semantic change, polysemy, synonymy, “Source-Path-Goal” schema, metaphor, analogy

試說“連 X+都 VP”構式的語法化*

楊永龍

中國社會科學院

本文探討了作為形式語義匹配的“連 X+都 VP”構式的語法化過程，認為在“連 X+都 VP”構式的語法化過程中，句法形式層面經歷了從 A 式“N1 連 N2+VP”到 B 式“(施)+連 N2+VP”到 C 式“(施)+連 N2+都 VP”的擴展過程；語義層面經歷了特定預設義和特定隱含義的產生過程。語用推理、語境吸收以及語義模式的拷貝在預設義和隱含義的產生中都有重要作用。

關鍵詞：“連 X+都 VP”構式，構式語法化，話題強調構式，預設義，隱含義

1. 引言

“連 X+都 VP”是現代漢語常用構式，許多學者對此有深入研究，如 Tsao (1989)、白梅麗 (Paris 1981)、崔希亮 (1990)、周小兵 (1996)、劉丹青、徐烈炯 (1998)、劉丹青 (2005) 等，在許多方面已經達成共識：“連”是引出話題的介詞；¹ X 是要強調的各種類型的話題。“都”是副詞，但已經沒有實在意義。整

* 謹以本文為貝羅貝先生祝壽。本文寫作中曾與趙長才先生反覆討論，初稿承蒙曹廣順、劉丹青先生提出很好的修改意見，在第五屆漢語語法化問題國際學術研討會（上海師範大學，2009 年 8 月）、山東大學（2009 年 12 月）宣讀時得到與會學者的指教，謹致誠摯謝意。

¹ 一般認為“連”是表示強調的介詞，如呂叔湘 (1999)，但是有時候“連”以外的部分可以獨立成句，而且一定條件下“連”可以省略，由此觀之，“連”已經不同於一般的介詞。所以學界對“連”的定性又有“副詞”說（如黃成一〈談“連”字〉，《中國語文》1956 年第 10 期）、“連詞”說（如李靜遠〈談“連”字〉，《語文知識》1957 年 12 期）、“助詞”說（如張友建〈“連”字是助詞〉，《中國語文》1962 年第 2 期）、“准量詞” (quasi-quantification) 說 (Paris 1981)、 “焦點標記”說（張伯江、方梅 1996）、“話題焦點標記”說（劉丹青、徐烈炯 1998）。從來源看，“連”確實是從介詞發展而來，其主要功能在於引出話題，因為話題可以涉及許多方面，如人、物、時間、處所、動作、過程等，所以“連”引出的不僅可以是體詞性成分，也可以是謂詞性成分甚至小句。漢語詞類系統中沒有“標記”這個類，如果一定要給它歸一個詞類的話，不妨仍然歸為介詞。

個格式的意義包括：(a) 預設義：X 在某一集合內最不可能 VP，(b) 字面義：事實上 X 確實 VP，(c) 隱含義：該集合內的其他成員更可能 VP。例如：

(1) 連小學生都知道這些事。

意思是：小學生最不可能知道這些事，但事實上小學生知道這些事，那麼其他人更可能知道這些事。

漢語史學界也有不少相關研究，如太田辰夫 (1958[2003])、劉堅 (1989)、李思明 (1996)、于江 (1996)、馮春田 (2000)、洪波 (2001)、史金生 (2004) 等。因為虛詞是漢語最重要的語法手段之一，所以在漢語歷史語法研究中，虛詞研究受到了格外關注，也取得了較大成果。以往對該構式的來源的研究，大多也是從虛詞“連”及相關的“和”的語法化入手。人們注意到，“連”原來是動詞，六朝以後有“連帶、包括在內”的意思，如 (2a)，宋代以後語法化為表示強調的“甚至”義介詞，如 (2b)：

- (2) a. 余注此經以來，一千七百餘年，凡傳三人，連子四矣。（葛洪《神仙傳》，引自孫錫信 1992）
b. 今人連寫也自厭煩了，所以讀書苟簡。（《朱子語類》卷 10，引自太田辰夫 1958[2003]）

這類研究解決了“連”的實詞虛化問題，自然是很有意義的。但是如果進一步思考：“連 X + 都 VP”作為一個具有特定形式和意義的完整構式是如何發展來的？其預設義和隱含義是“連”帶來的嗎？如果是，那麼為什麼去掉“連”之後構式義可能沒有改變？如例 (1) 可以說成：“小學生都知道這些事。”此時“小學生”要重讀而“都”不能重讀。如果不是，那麼構式義從何而來？有人把目光再投向構式中的其他虛詞，如考察“都”的虛化，於是發現“都”存在著一個從表總括的範圍副詞到表強調的語氣副詞的語法化過程。但是例 (1) 的“都”和“也”可以互換而構式意義基本不變，於是再考察“也”是如何從表示類同一步步演變為表示強調的。

隨著研究的深入，人們開始注意從作為形式語義匹配 (form-meaning pair) 的整體出發，而不是從其中的虛詞“連”、“都”出發，來探索“連”字句發展演變，如劉丹青 (2005) 從現代漢語共時層面探索了典型“連”字句向非典型“連”字句的進一步發展，此前馮春田 (2000)、洪波 (2001) 也已經開始觸及構式本身。

本文希望在前人研究的基礎上，借鑑構式語法的理論方法，探索作為完整的形式語義匹配的“連 X+都 VP”構式的產生過程，包括“連 X+都 VP”這種構式的句法形式以及與之對應的預設義和隱含義是如何產生的。

2. 現代漢語的“連 X+都 VP”及話題強調構式

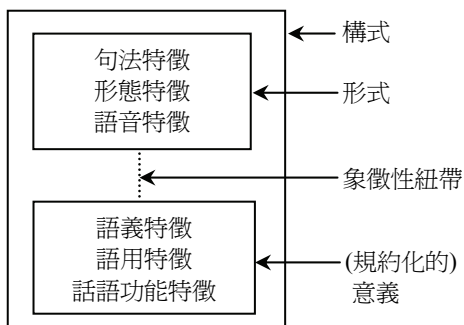
本文把“連 X+都 VP”作為一個代表性構式，為便於說明，稱作甲式。其中 X 是要強調的話題，而話題可能包括不同的句法成分和語義角色，如 (3a)，可以通過“連”把句中的不同成分話題化而加以強調：

- (3) a. 張三考試的時候都不耽誤玩遊戲。
b. 連張三考試的時候都不耽誤玩遊戲。
c. 張三連考試的時候都不耽誤玩遊戲。
d. 張三考試的時候，連玩遊戲都不耽誤。
e. 張三考試的時候，連遊戲都不耽誤玩。

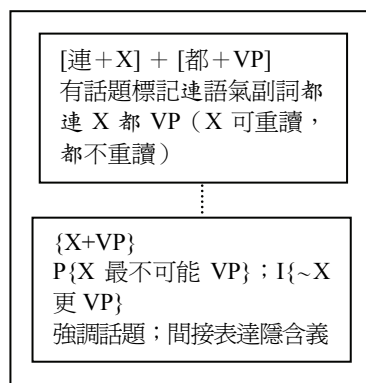
“都”是一個代表字，包括“也”、“還”等，如：

- (4) 連小學生都知道這些事。
連小學生也知道這些事。
連小學生還知道這些事。

構式語法把構式看作由象徵性紐帶聯結在一起的形式和語義的匹配。激進構式語法把語法構式的象徵性結構模式圖示為〈圖 1〉。如果把該結構模式引入具體的漢語語法研究，那麼相關特徵大體上可以明確如下：在形式層面，句法特徵涉及語序和相關的句法關係、論元結構，形態特徵涉及相關虛詞的性質和功能，語音特徵涉及說出來或寫出來的字符串及重音模式。在語義層面，語義特徵涉及構式的字面意思，語用特徵涉及預設義、隱含義等已經規約化的言外之意，話語特徵涉及信息結構，語篇銜接等。據此觀察甲式，其結構模式可圖示為〈圖 2〉。



〈圖 1〉構式的象徵性結構模式
(據 Croft 2001:18)



〈圖 2〉“連 X+都 VP”
構式的結構模式

甲式的隱含義是間接表達的，用於對比的同一集合的成員沒有出現，因此預設義不太明確，隱含義也不太確定，常常會因為語境的不同而不同。如例 (1) 所要表達的意思可能是：{大學生知道、研究生知道，你也知道……}。如果明確顯示對比項，如通過前加“不光/別說……”之類，或後續“更不用說/何況……”及相關疑問句、反問句等，明確說出用於對比的某一集合中的成員，那麼隱含義就會直接顯示出來，變得比較確定。不過這樣一來就變成形式和意義都不相同的另外的構式。如乙式、丙式：

乙式：“前加句，甲式”（如：不光大學生，連小學生都知道這些事。）

丙式：“甲式，後續句”（如：連小學生都知道這些事，何況大學生？）

乙式和丙式形式不同，意義也不同，但處在其中的甲式在預設義和隱含義方面是一致的。如括號中的“連小學生都知道這些事”在兩式中都有 {與大學生比，小學生知道這些事的可能性比較小} 的預設義，以及 {大學生更會知道這些事} 的隱含義。

與甲式相關而略有不同的構式還有：

丁式：“連 X+都 VP”（如：連“的士”都捨不得打。）

丁式與甲式形式相同而意義不同，屬於劉丹青 (2005) 所討論的非典型“連”字

句，該式雖然話題是 X，但對比的不是 X 而是包括 X 在內的整個 {X+VP}。因此，甲式與丁式的預設義是不同的：前者是 {X 最不可能 VP}，突出的是 {X} 的可能性低端，而後者是 {X+VP 最不可能}，突出的是 {X+VP} 的可能性低端。劉丹青 (2005) 指出，非典型“連”字句是由典型的連字句語法化而來，即丁式源於甲式。

丁式可以與甲式一樣作為乙式、丙式的一部分而進入其中，如：

- (5) 不僅不到國外購物了，平時連“的士”都捨不得打。
- (6) 平時連“的士”都捨不得打，哪裡還到國外購物？

為研究便利，我們可以著眼於局部把這類構式放在丁式討論，也可以著眼於整體把 (5) 歸入乙式，把 (6) 歸入丙式。其實，既然構式是形式與意義的匹配，那麼一旦形式或意義發生變化，也就成了不同的構式。因此，嚴格地說這兩例所代表的構式與前述甲、乙、丙、丁分別屬於不同的構式，它們各不相同而又互相關聯，構成一個複雜的構式網絡。該網絡的共同特徵是，都含有特定的字面以外的預設義和隱含義，都有話題強調功能，因此可以統稱為“話題強調構式”。

3. 話題強調構式的前身

前人在討論“連”的虛化時已經指出，表強調的“連”、“和”是從動詞經過 {連帶、包括在內} 義的“連”、“和”虛化而來（劉堅 1989 等）。太田辰夫 (1958[2003]) 用“包括”和“強調”區分兩種不同的用法，後來的學者把相關句式稱之為“包括式”和“強調式”，如李思明 (1996)、馮春田 (2000)。因為考察目標不同，前人對“包括”、“強調”之類沒有嚴格的界定，因此有時對同一例子會有不同的歸類。本文沿用前人既有術語，並在此基礎上略作界定：含有字面以外的特定預設義、隱含義及話題強調功能的構式稱為“話題強調構式”，簡稱“強調式”；作為“強調式”前身的含有這類預設義和隱含義的構式稱為“包括式”。

3.1 唐五代以前的包括式

先看例子：²

² 本文用例有些前人已經舉過，詳見引用文獻。本文對前人舉過的例子有的可能有不同的理解和分析，為避免枝蔓，除影響結論者外一般不再辨析。

- (7) a. 水光連岸動，花風合樹吹。（《先秦漢魏晉南北朝詩·北周詩》）
b. 舊山連藥賣，孤鶴帶雲歸。（李端《聞吉道士還俗因而有贈》，《全唐詩》卷 285）
c. 時挑野菜和根煮，旋斫生柴帶葉燒。（杜荀鶴《山中寡婦》，同上卷 692）
d. 紫芽嫩茗和枝採，朱橘香芭數瓣分。（元稹《貶江陵途中寄樂天》，同上卷 412）

這些例子共同的形式、語義特徵有：

- (a) 字符串是“N1+連+N2+VP”。
- (b) {N1} 與 {N2} 往往連在一起，甚至是同一物體的主體與附件，如“菜”與“菜根”。它們一起充當 VP 的施事（包括與事）或受事，如 (7a) 水光動，岸也動，(7b) 舊山被賣，藥也被賣。
- (c) 但 N1 與 N2 又有主次之別，並非並列關係。其中 N1 是主要陳述的對象，是句子的主語；N2 本來是“連”、“和”、“帶”等動詞的賓語，隨著“連”等虛化為引出連帶對象的介詞，N2 便成了介詞賓語。所以該構式的句法結構是“N1+[[連+N2]+VP]”。
- (d) 整個構式的意思是：{N1 連帶 N2 一起實施 VP 或者受 VP 支配}。如 (7b) 是說 {舊山連帶山上的藥一起被賣掉，孤鶴連帶雲彩一起歸來}。
- (e) 在話語功能和語用方面，該構式主要用於客觀敘事，不含預設義 {X 最不可能 VP} 和隱含義 {~X 更 VP}。

可見，無論從形式層面還是從意義層面看，該式與典型的話題強調構式差距都很大。但是也有值得注意的聯繫：

其一，在形式層面，雖然 N2 在語義上是作為 N1 的連帶物而存在，但在句法層面，[連+N2] 是作為介詞短語來修飾謂語的，因此可以脫離 N1，或者在二者之間加入別的成分。如：

- (8) a. 嘗發所在竹_{N1}篙，有一官長連根_{N2}取之，仍當足。（《世說新語·政事》）
b. 東風吹綻海棠_{N1}開，香榭滿樓臺。香和紅豔一堆堆，又被美人和枝_{N2}折。（張璋等編《全唐五代詞》卷 7，敦煌詞，上海古籍出版社 1986）
c. 須臾得暫時，恰同霜下草_{N1}。橫遭狂風吹，總即連根_{N2}倒。（項楚編《王梵志詩校註》卷 2，上海古籍出版社 1991）

- d. 所食物餘者_{N1}，便和碗_{N2}與犬食。（薛用弱《集異記》）

(8d) 的 N1 與 [和 N2] 之間僅有副詞“便”，其餘三例 N1 則與 N2 間隔較遠，甚至隔有別的句子。如果 N1、N2 是 VP 的受事，那麼 N1 之後、[連 N2] 之前還可能出現施事，如例 (8a) 的“官長”，(8b) 的“美人”。因此，準確地說，該式的句法結構是：(N1)+(施事)+[[連+N2]+VP]。為便於區別，這裡把例 (7) 這類小句中出現 N1 的構式稱作 A 式，把例 (8) 這類小句中不出現 N1 的構式稱作 B 式：

A 式：N1+[[連+N2]+VP]

B 式：(施事)+[[連+N2]+VP]

其二，在意義層面，就 A 式而言，雖然主要用於客觀敘事，而且其中的 N1 是主要陳述對象，但是既然把連帶成分專門說出來，則表明說話人不希望忽略 N2。不希望忽略可以說是一種保守性的強調。在脫離 N1 的 B 式中，這種強調的意思變得更為明顯。

3.2 宋代的包括式

唐五代以前包括式的 VP 前通常沒有表示總括的副詞“都”（或“皆”）和表示類同的副詞“也”（或“亦”），只有個別例子似乎為例外：

- (9) a. 河南北蟲為災，飛則翳日，大如指，食苗草樹葉，連根並盡。
（《朝野僉載》，《太平廣記》卷 474 引）
b. 在生恨你極無量，貪愛之心日夜忙。老去和頭全換卻，少年眼也擬
挽將。（《敦煌變文校注·譬喻經變文》）

可能的解釋是，{N1} 與 {N2} 往往是連在一起的，甚至是同一物體的主體與附件，“並”、“全”傾向於總括同一個體的不同部分，與此相容，所以偶能出現；而“皆”、“都”傾向於總括不同的個體，與此不相容。至於“亦”、“也”，所表示的類同也主要是不同個體之間的類同。

到了宋代，A 式用例漸少，B 式有較大發展：“連”後 N2 擴展到各種不同的個體，不再限於與 N1 有連帶關係的事物，同時，VP 前經常出現“都”（或“皆”）、“也”（或“亦”）等副詞。我們把 VP 前出現副詞的稱作 C 式：

C 式：(施事)+[連+N2] +[都 VP] (“都”是代表字，包括相關副詞)

下面的 (10)、(11)、(12) 分別是宋代的一些 A、B、C 式例子：

- (10) 苦瓠連根苦，甜瓜徹蒂甜。(《五燈會元》卷 9) | 釋迦老子和身放倒，後代兒孫如何接續？(同上，卷 18)
- (11) 尚自待要兩州，我若與你，又是和西京人民存住不得。(《三朝北盟會編》卷 22) | 雲門大師，多以一字禪示人。……雪竇為他一個“關”字，和他三個穿作一串頌出。(《碧岩錄》卷 1) | 才卿問：“‘上老老而民興孝’，恐便是連那‘老眾人之老’說？”曰：“不然。此‘老老’、‘長長’、‘恤孤’，方是就自家身上切近處說。”(《朱子語類》卷 16) | 譬如一株草，剷去而留其根，與連其根剷去，此箇意思如何？(同上，卷 44)
- (12) 撒母與靖相看，曰：“卻是和西京、平、灤都要。靖等來時，只聽得特許燕京六州二十四縣地與南朝，今來卻和西京、平、灤都要，怎生了得？”(《三朝北盟會編》卷 11) | 看來“如好好色，如惡惡臭”一段，便是連那“毋自欺”也說。言人之毋自欺時，便要“如好好色，如惡惡臭”樣方得。(《朱子語類》卷 16) | 若水清，則寶珠在那裡也瑩徹光明；若水濁，則和那寶珠也昏濁了。(同上，卷 59) | 只是前面“體”字說得來較闊，連本末精粗都包在裡面。(同上，卷 36) | 俞亨宗云：“某做知縣，只做得五分。”曰：“何不連那五分都做了？”(同上，卷 112)

可以看出，在構式的形式層面，C 式已經與後世典型的強調式“連 N+都 VP”（甲式）沒有太大區別：(a) 字符串相同，(b) 句法結構相同，(c) 論元關係一致，N2 既可以是受事也可以是施事；區別在於 (d) “連”、“都”的形態特徵，C 式的“連”是典型的介詞，“都”、“也”等是總括副詞或類同副詞。在構式的意義層面，B 式、C 式也向典型強調式更近一層：(a) 除了前面已經說到的對 N2 有保守性強調外，既然將牽涉對象安排到 VP 之前甚至句子的開始，N2 自然在一定程度上具有話題功能。請比較下例畫線部分：

- (13) 兀室雲：“與了地，更要人戶，……大抵地土重於人民，地土已許了，更和人民要，更別無酬答，更無致謝，怎生了得！”(《三朝北盟會編》卷 14)

(b) 典型強調式有隱含的對比項，以及預設義 {X 最不可能 VP}，B 式、C 式有一個隱含的與 N2 同一集合 N1（通常出現在前文中），而且 N2 與 N1 之間有時可能有一定級差，但卻沒有規約化的預設義。(c) 典型強調式有 {~X 更 VP} 隱含義，B 式、C 式沒有。如例 (13) 並不含有 {更要土地} 的意思。根據後面這兩點，以往學者看作強調式的例子，按照本文分析有些仍屬於包括式，如 (11)、(12) 中的一些例子。再如馮春田 (2000) 所舉的較早的強調式例：

- (14) 大杖打又不死，忽若堯王敕知，兼我也遭帶累。（《敦煌變文校注·舜子變》）

如果把“強調”理解為話語功能上的突顯，那麼這裡的“兼”有強調的意思，但整個句子沒有 {我是最不可能遭帶累的人}、{別人更會遭帶累} 之類的預設義和隱含義。

4. 話題強調構式的出現

上述包括式的發展為強調式的產生奠定了基礎，尤其在形式方面，已經替強調式走完了必走之路。當作為包括式的 B 式、C 式具有預設義 {X 最不可能 VP}、隱含義 {~X 更 VP} 時，就成了話題強調構式。它們在宋代也已經出現，為便於說明，可分別記作 Bq 式、Cq 式。

Bq 式與 B 式形式相同，仍然是“(施事)+[[連+N2]+VP]”，但意義不同。如：

- (15) a. 衡陽猶有雁傳書，郴陽和雁無。（秦觀《阮郎歸》四，《全宋詞》463 頁）
b. “他那得似子靜！子靜卻是見得箇道理，卻成一部禪，他和禪識不得。”（《朱子語類》卷 123）

例 (15a) 較淺顯，雁乃信使，沒有信使，自然 {不會有書信}（隱含義）。(15b) 略隱晦，近於劉丹青 (2005) 所說的非典型連字句，是以“識不得禪”與“見得箇道理”相比，不過，因為“見”與“識”同義，因此也可以看作以“禪”和“道理”對比。該例預設義是 {“禪”比“道理”容易理解}，隱含義是 {他更理解不了道理}。

Cq 式與 C 式形式相同，仍是“(施事)+[[連+N2]+都 VP]”，而意義不同。如：

- (16) a. 佗不知“道”，只說“道”時，便不是“道”也。有道者亦自分明，只作尋常本分事說了。《孟子》言“堯、舜性之”，舜“由仁義行”，豈不是尋常說話？至於《易》，只道個“立人之道曰仁與義”，則和“性”字、“由”字也不消道，自己分明。（《河南程氏遺書》卷 1）
- b. 看武侯事迹，儘有駁雜去處；然事雖未純，卻是王者之心。管仲連那心都不好。（《朱子語類》卷 44）
- c. 尹彥明看得好，想見煞著日月看。臨了連格物也看錯了，所以深不信伊川“今日格一件，明日格一件”之說，是看箇甚麼？（《朱子語類》卷 95）

(16a)《孟子》“堯舜性之”意思是堯舜本性好仁，出於自然。其中“性”為動詞，即以之為本性。“由仁義行”意思與此相近，是說舜的仁義生於內心，任由仁義自己出來即可。“由”為動詞，任憑的意思。這裡用於對比的集合是：道，性和由，仁與義。按朱熹的看法，三者一個比一個更為通俗。《易》既然不說“性”和“由”，那麼更用不著說“道”（隱含義）。(16b)前文說到管仲所做的是霸者之事，霸者之事在儒家看來當然不是好事。此處說武侯做的事雖然有不好的，但“卻是王者之心”。比較而言，{心不好比事不好更壞}（預設義），管仲既然心不好，那麼{做的事更不好}（隱含義）。(16c)預設義是，{比較而言“格物”不容易理解錯}，既然把“格物”理解錯了，那麼{別的更有可能理解錯}（隱含義）。

Cq 式其實就是第 2 節所說的甲式：“連 X+都 VP”，其中 X 等於 N2。在 Bq 式和 Cq 式中，其前身中與 N2 關聯的 N1 變成了現在的對比項，該對比項有時在前文中出現，如 (15)、(16a)、(16b)；有時候不出現，如 (16c) 中與“格物”相關 N1。不出現時就成了預設和隱含中的對比項。

宋代甲式的 X 也可以是動詞。如：

- (17) 蓋古人無本，除非首尾熟背得方得。……晁以道嘗欲得《公》、《穀》傳，遍求無之，後得一本，方傳寫得。今人連寫也自厭煩了，所以讀書苟簡。（《朱子語類》卷 10）

該例太田辰夫 ([1958]2003) 以來許多學者舉過，意思是：{相對於熟背來說厭煩抄寫的可能性比較小（預設義），今人厭煩抄寫，自然更厭煩熟背了（隱含義）}。

甲式可以出現前加句或後續句，一起構成前述乙式（“前加句，甲式”）和丙式（“甲式，後續句”）。前者如 (18)，後者如 (19)：

- (18) 但把這底看“巧言令色鮮矣仁”，便見得。且如巧言令色人，盡是私欲，許多有底便都不見了。私慾之害，豈特是仁，和義、禮、智都不見了。（《朱子語類》卷 20）
- (19) 所以告曾子時，無他，只緣他曉得千條萬目。他人連箇千條萬目尚自曉不得，如何識得一貫。（《朱子語類》卷 27）

通過前加句和後接句，既顯示了甲式預設義中 X 的對比項，又直接表達了甲式的隱含義。如 (18) 表明具體對比項是“仁”，預設義是，{“仁”比“義、禮、智”更容易被私欲害沒了}。既然“義、禮、智”都害沒了，可以推知，{“仁”也被害沒了}。{“仁”被害沒了}既是“和義禮智都不見了”的隱含義，也正是前加句的字面意思。同樣，(19) 表明了具體對比項是“一貫”，預設義是：{“千條萬目”比“一貫”容易知道}，語用推理是 {既然不知道“千條萬目”，也就更不知道“一貫”}。{不知道“一貫”}既是甲式的隱含義，也是後續句的字面意思。

值得注意的是，屬於劉丹青 (2005) 所說的非典型連字句（即前述丁式）此時也已經出現。如：

- (20) “不順乎親，不可以為子”是無一事不是處，和親之心也順了。（《朱子語類》卷 56）
- (21) “得乎親”者，不問事之是非，但能曲為承順，則可以得其親之悅。苟父母有做得不是處，我且從之，苟有孝心者皆可然也。“順乎親”，則和那道理也順了，非特得親之悅，又使之不陷於非義，此所以為尤難也。（《朱子語類》卷 56）
- (22) 今做一件好事，便望他功效，則心便兩歧了。非惟是功效不見，連那所做底事都壞了。（《朱子語類》卷 42）

三例形式上是“連 N+都 VP”，但對比的不是 N 而是 {N+VP}。其中 (20)-(21) 是講《孟子》的，大意相同，按朱熹的解釋，“得乎親”指 {讓父母高興}，“順

乎親”指 {讓父母之心合乎正理，不陷於非義}。(20) 沒出現對比項，句意是 {讓父母之心合乎正理是最不容易做好的，既然能讓父母之心合乎正理，那麼別的事更能做好}。(21) 出現了對比項“得親之悅”，隱含義也由後續句具體化了，句意是 {與讓父母高興相比，讓父母之心合乎正理更不容易做好，既然能讓父母之心合乎正理，那麼別的事更能做好，不光能做到讓父母高興}。(22) 預設義是 {與“做好事不見功效”相比，“把做的好事弄壞”不太可能出現}，既然好事都能弄壞，那麼 {做好事不見功效的情況更有可能出現} (隱含義)。

綜上可見，到了宋代，尤其是南宋的《朱子語類》中，無論形式層面還是語義層面，從包括式到強調式的語法化過程都已經完成。在此過程中，有一些值得注意的現象，有的比較容易解釋。如：問題(一)，強調式“連”後的 N 一開始就既有施事也有受事，沒有經歷一個從受到施或相反的擴展過程。為什麼？這是由其前身包括式帶來的，早在“N1 連 N2 + VP”階段 N1 就連同 N2 一起作 VP 的施事或受事了。問題(二)，強調式 VP 前以有副詞“都”等為常，但也可以不出現。怎麼解釋？在其前生包括式的 B 式階段就已如此，加副詞是因為 N1 與 N2 由連帶關係變為分離的個體之後，用“都”等表示總括或類同。變為強調式之後，這種功能雖然還能在 X (N1) 與其對比項 (N2) 之間隱隱約約看到，但隨著對比項潛隱程度的加深，副詞不出現在語義上也沒有影響，所以從強調式產生之日起，直到清代文獻中這種不帶副詞的強調式可以一直延續 (李思明 1996)。當然，“都”等畢竟是祖上 (C 式) 留下來的東西，大量使用使得“連……都”有框式化的傾向，成了該構式形式的一部分，所以文獻中不帶副詞的所占比例較少 (李思明 1996)。問題(三)，“連”後帶動詞的情況出現較早，似乎沒有經歷明顯的從 N 擴展到 V 的過程。如何解釋？“連”字句的語法化過程，其實也是 N2 話題化的過程，該過程在包括式階段就已經開始，至強調式階段完成。這一過程一旦完成，只要是話題就都可以安排在 N2 位置。謂詞性成分所表示的動作、過程等也可以話題化，因此也可以安排在 N2 位置。

也有的需要進一步討論，如：問題(四)，非典型連字句出現得也比較早，似乎也沒有經歷一個明顯的從典型構式到非典型構式的漸變過程。如何解釋？問題(五)，既然包括式和強調式的根本區別在於預設義 {X 最不可能 VP} 和隱含義 {~ X 更 VP} 的有無，那麼強調式的預設義和隱含義是怎麼來的？下面主要討論問題(五)，連帶討論問題(四)及其他問題。

5. 預設義和隱含義的來源

5.1 語義層面的重新分析與語境義的吸收

預設義 {X 最不可能 VP} 的構成是 (a) 要有一個集合，(b) 集合內的成員要有級差，(c) X (N2) 處於該集合中的可能性低端。第一點在包括式時代就有，雖然只有 N1 和 N2 兩項。第二點在包括式時代有一定的基礎，N1 是主體，N2 是附件，主體與附件有一定的級差。但這只是有級差的潛在可能。第三點在包括式時代也有一定基礎，既然說出 {包括 N2}，就是擔心別人認為 {N2 不可能 VP}。這也只是潛在可能，並不意味著說話者認為 {N2 最不可能 VP}。但是，在一定上下文中，潛在可能會實現為實際可能。例如：

- (23) a. 顏子底儘細膩，子路底只是較粗，然都是去得箇私意了，只是有粗細。子路譬如脫得上面兩件麤糟底衣服了，顏子又脫得那近裡面底衣服了，聖人則和那裡面貼肉底汗衫都脫得赤骨立了。（《朱子語類》卷 29）
- b. 而今學者，看來須是先曉得這一層，卻去理會那上面一層方好。而今都是和這下面一層也不曾見得，所以和那上面一層也理會不得。（《朱子語類》卷 62）

先看 (23a)。畫線部分三個小句是遞進關係，所表示的三個事件一個比一個更進一層。強調式預設義的三個要件在這裡都具備了：有對比項、有級差，而且後一項比前一項更不容易做到。但是，這種遞進關係是語境具有的，即使沒有虛詞“又”、“則”、“連”，也不影響這種關係，只是會使這種關係更為隱蔽而已。通過該例可以看到：第一，語境可以表達 {X 最不可能 VP} 義。強調式進入該語境不會影響語境義的表達，有潛在可能性的包括式進入該語境也不影響語境義的表達。正像表層形式不變的句法結構可以重新分析一樣，形式不變的構式在意義層面也可以重新分析。當包括式進入有 {X 最不可能 VP} 意義的語境之後，包括式就可以重新分析為強調式。如“聖人則和那裡面貼肉底汗衫都脫得赤骨立了”，既可以分析為包括式（C 式），意為 {聖人連同汗衫全都脫光}，——此時 {X 最不可能 VP} 意義屬於語境；又可以分析為強調式（Cq 式），義為 {連同汗衫都脫光的情況是不太可能出現的，事實上聖人連同汗衫全都脫光了}，——此時 {X 最不可能 VP} 意義既屬於語境，也屬於該小句。第二，無論把上述第三個小句分析為包括式還是強調式，“連”都具有明顯的引出話題的功能，因為前兩個

小句都是 SVO，而第三個小句變成了“S 連 OV”，通過“連”把 O 提到前面做了話題。第三，三個小句的對比項可以說是相對應的 {N}，也可以說是三個小句所代表的三個事件，即 {N+VP}。此時對比 {N} 與對比包括 {N} 在內的整個 {N+VP} 事件是一致的。其實，當 {VP} 相同相近時，{N+VP} 相對比就等於 {N} 相對比，而當 {VP} 不同時，就是 {N+VP} 的對比。這一點或許能夠解釋上述問題(四)，即為什麼宋代典型連字句剛出現不久就出現了非典型的連字句。

再來看 (23b)，劃線部分是兩個“和 N+VP”小句，構成果關係。雖有對比性集合，但孤立地看，很難看出級差以及 N 是否處於可能性低端。而且從形式上看，“和這下面一層也不曾見得”與“和那上面一層也理會不得”沒有任何區別。不過，前面的句子提供了強調式預設義的所有要件，只要聯繫語境，前一小句就會重新分析為強調式，而後一句仍是包括式。該例同樣表明，包括式一旦進入具有 {X 最不可能 VP} 義的語境，就有可能通過語義上的重新分析而變成強調式。

當包括式經常處於 {X 最不可能 VP} 語境的時候，就會通過語用推理吸收語境義從而使臨時意義成為規約化的預設意義。這是通過估推，從觀察到的現象，援引事理法則而做出的可能的推斷：

現象：包括式經常現身於具有 {X 最不可能 VP} 義的語境

事理：語境的意義是由具體的字符串表達的

推斷：包括式就是表達 {X 最不可能 VP} 義的字符串

5.2 語用推理與隱含義的產生

上面討論預設義的產生時沒有提到隱含義 {~X 更 VP}，其實該隱含義是伴隨著預設義的產生而產生，伴隨著預設義的規約化而規約化的。沒有 {X 最不可能 VP} 這一預設義，就不可能有 {~X 更 VP} 這一隱含義；有了這一預設義，又有與此不一致的字面義 {X 事實上 VP}，就很容易推導出隱含義。如上述 (23a) 有了“連同汗衫都脫光的情況是不太可能出現的”這一預設義後，通過類似於演繹推理的語用推理，很容易推導出“聖人汗衫以外的衣服也都脫光了”的隱含義：

大前提（預設義）：{X 最不可能 VP} 連同汗衫都脫光的情況是最不可能出現的

小前提（字面義）：{X 事實上 VP} 事實上聖人連同汗衫都脫光了

結 論（隱含義）：{~X 更 VP} 聖人更可能把汗衫以外的衣服也都脫光了

就強調式本身而言，隱含義是推導出來的而沒有明確說出的意義，而在實際語言應用中，這個意義可能真的沒有說出，如 (23a) 上下文中就沒有說過“聖人把汗衫以外的衣服也脫光了”之類的話；也有的時候又會在上下文中把這種意義直接或間接、明確或不太明確地表達出來，如 (23b) “和這下面一層也不曾見得”的隱含義是 {上面的一層更理解不了}，接下來的一句說的正是這個意思。假如後面這一句沒有說出來，整個文意也沒有影響。在含有特定隱含義的構式前後再說出一些與隱含義同一個意思的語句，這既是語用推理在具體語言中的實際應用，又能反過來強化該構式的隱含義。

5.3 語義模式的拷貝

在“連 N+都 VP”之前，漢語中早已存在兩類傳統的話題強調構式，既具有 {X 最不可能 VP} 的預設義，也具有 {~X 更 VP} 這一隱含義。一類近似於前述乙式，如例 (24)，一類近似於前述丙式，如例 (25)，兩類都是在含有特定隱含義的構式前後再說出一些與隱含義大體一致的語句：

- (24) a. 非惟百乘之家為然也，雖小國之君亦有之。……非惟小國之君為然也，雖大國之君亦有之。（《孟子·萬章上》）
 b. 天天花里千家住，總為當時隱暴秦；歸去不論無舊識，子孫今亦是他人。（施肩吾《桃源詞》，《全唐詩》卷 494）
- (25) a. 蔓草猶不可除，況君之寵弟乎？（《左傳·隱公元年》）
 b. 玄宗尚且如此，我等寧不傷身？《敦煌變文校注·維摩詰經講經文（三）》

這是自先秦以來十分常見的兩種表示遞進關係的複句，其中的小句之一就是話題強調構式，具有 {X 最不可能 VP} 的所有要件：集合，級差，可能性低端，也具有 {~X 更 VP} 這一隱含義。而且句首的 N 可以是施事也可以是受事，無論施事還是受事都是要強調的話題。小句中 VP 前也有“亦”等副詞，有時仍有表示類同的功能而有時已經相當模糊。而且，所強調的話題不限於名詞或名詞短語，也可以是動詞或動詞短語，甚至介詞結構。如：

- (26) 故王公不致敬盡禮，則不得亟見之。見且由不得亟，而況得而臣之乎？（《孟子·盡心上》） | 臣死且不辭，豈特卮酒乎！（《史記·

樊鄴滕灌列傳》)聞伯夷之風者，貪夫廉而懦夫有立志；聞柳下惠之風者，薄夫敦而鄙夫寬。徒聞風名猶或變節，況親接形面相敦告乎？（《論衡·率性篇》）| 以魯國猶不勝一季氏，況于三季？（《呂氏春秋·察微》）

“臣死且不辭，豈特卮酒乎”是以“死”和“卮酒”對比，而同樣這件事，在《史記》的另一處則是以 {N+VP} 為對比項：

(27) 臣死且不避，卮酒安足辭！（《史記·項羽本紀》）

還有一些例子是前後不出現與隱含義一致的表達。如：

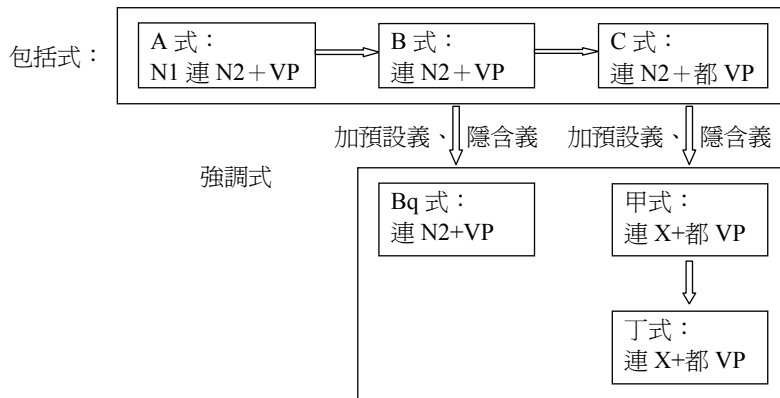
(28) 世之所高，莫若黃帝，黃帝尚不能全德，而戰於涿鹿之野，流血百里。（《莊子·盜跖》）| 吾不忠於君，楚亦何以軫為忠乎？忠且見棄，吾不之楚，何適乎？（《戰國策·秦策》）

這些都是現成的話題強調構式幾乎覆蓋了前述甲、乙、丙、丁各式，既然 B 式、C 式這類包括式可以進入普通的具有 {X 最不可能 VP} 意義的語境，既然包括式與現成的話題強調構式在形式、意義上如此接近，那麼包括式自然也能進入這類遞進複句之中。所以這類遞進複句也是 B 式、C 式重新分析為 Bq 式、Cq 式的最適宜的語境。

不僅如此，這些早已存在的話題強調構式具有現成的語義模式 {X 最不可能 VP，但 X 事實上 VP，所以~X 更 VP}，或者 {X+VP 最不可能，但事實上出現了 X+VP，所以~X+VP 更有可能出現}，具有現成的形式匹配，那麼，“連 N + 都 VP”包括式一旦重新分析為強調式，就可以把現成的語義模式連同形式匹配直接拷貝過來。把語義模式拷貝過來，那麼“連 X + 都 VP”就不必每生成一個句子都需要一個語用推理的過程，非典型“連”字句也不必要每每經歷從典型“連”字句語法化而來的過程；把形式特徵拷貝過來，那麼“連 X + 都 VP”的話題就既可以是施事，又可以是受事；既可是名詞性成分也可以是謂詞性成分，而不必經歷一個漫長的擴展過程。這或許可以進一步補充解釋問題 (一)(三)(四)。

6. 小結

綜上可見，“連 X+都 VP”構式是作為一個完整的形式語義匹配一步步語法化而來，在這個過程中，形式層面經歷了從 A 式“N1 連 N2+VP”到 B 式“(施)+連 N2+VP”到 C 式“(施)+連 N2+都 VP”的擴展過程；語義層面經歷了 N2 的話題化和特定預設義、特定隱含義的產生過程。該過程是通過語義上的重新分析（或稱為重新理解）、語境吸收、語用推理以及語義模式的拷貝實現的。其演化過程可粗略圖示如下：



引用文獻

- Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford & New York: Oxford University Press.
- Cui, Xiliang (崔希亮). 1990. Shilun guanlian xingshi 'lian...ye/dou' de duochong yuyan xinxi 試論關聯形式“連…也/都”的多重語言信息 [The multiple information of conjunctive form *lian...ye/dou*]. *Shijie Hanyu Jiaoxue* 世界漢語教學 [Chinese Teaching in the World] 1990.3:139-144.
- Feng, Chuntian (馮春田). 2000. *Jindai Hanyu Yufa Yanjiu* 近代漢語語法研究 [Research on Early Modern Chinese Grammar]. Ji'nan: Shandong Education Press.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Hong, Bo (洪波). 2001. 'Lian' ziju xudiao “連”字句續貂 [Some issues on the *lian* sentences]. *Yuyan Jiaoxue yu Yanjiu* 語言教學與研究 [Language Teaching and Linguistic Studies] 2001.2:55-63.
- Li, Siming (李思明). 1996. Song Yuan yilai de 'he/lian...' ju 宋元以來的“和/連…”句 [On the *he/lian* construction from Song Dynasty]. *Yuyan Yanjiu* 語言研究 [Studies in Language and Linguistics] 1996.1:88-97.
- Liu, Danqing (劉丹青). 2005. Zuowei dianxing goushi de feidianxing 'lian' ziju 作為典型構式句的非典型“連”字句 [The atypical *lian*-clause as a typical construction]. *Yuyan Jiaoxue yu Yanjiu* 語言教學與研究 [Language Teaching and Linguistic Studies] 2005.4:1-11.
- Liu, Danqing (劉丹青), and Liejiong Xu (徐烈炯). 1998. Jiaodian yu beijing, huati ji Hanyu 'lian' ziju 焦點與背景、話題及漢語“連”字句 [Focus, background, and *lian*-clause in Chinese language]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1998.4:243-252.
- Liu, Jian (劉堅). 1989. Shilun 'he' zi de fazhan, fulun 'gong' zi he 'lian' zi 試論“和”字的發展，附論“共”字和“連”字 [On the development of *he*, *gong* and *lian*]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1989.6: 447-453.
- Lü, Shuxiang (呂叔湘) (ed.) 1999. *Xiandai Hanyu Babai Ci* 現代漢語八百詞 [Eight Hundred Words in Modern Chinese] (revised edition). Beijing: The Commercial Press.
- Ota, Tatsuo (太田辰夫). 1958[2003]. *Zhongguoyu Lishi Wenfa* 中國語歷史文法 [A Historical Grammar of Modern Chinese] (2nd edition), translated by Shaoyu Jiang & Changhua Xu. Beijing: Peking University Press.

- Paris, Marie-Claude (白梅麗). 1981. Hanyu Putonghua zhong de 'lian...ye/dou' 漢語普通話中的“連…也/都” [Some aspects of the syntax and semantics of the *lian...ye/dou* construction in Mandarin], translated by Shenyi Luo. *Guowai Yuyanxue* 國外語言學 [*Contemporary Linguistics*] 1981.3:50-55.
- Qiao, Shihao (喬石豪). 2007. Yuqi fuci 'lian' de lai yuan ji yufahua guocheng 語氣副詞“連”的來源及語法化過程 [The origin and its grammaticalization of adverb *lian*]. *Journal of Anyang Teachers College* 2007.3:97-100.
- Shi, Jinsheng (史金生). 2004. Xiandai Hanyu changyong xuci de yufahua 現代漢語常用虛詞的語法化 [The grammaticalization of common function words in Mandarin Chinese]. Manuscript. Beijing: Chinese Academy of Social Sciences.
- Sun, Xixin (孫錫信). 1992. *Hanyu Lishi Yufa Yaolue* 漢語歷史語法要略 [*Introduction to Historical Grammar of Chinese*]. Shanghai: Fudan University Press.
- Tsao, Feng-fu. 1989. Topic and *lian dou/ye* construction revisited. *Functionalism and Chinese Grammar*, ed. by James H-Y. Tai & Frank F. S. Hsueh, 245-278. South Orange: Chinese Language Teachers Association.
- Yu, Jiang (于江). 1996. Jindai Hanyu 'he' lei xuci de lishi kaocha 近代漢語“和”類虛詞的歷史考察 [A historical research on some function words like *he* in Early Modern Chinese language]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1996.6:457-464.
- Zhang, Bojiang (張伯江), and Mei Fang (方梅). 1996. *Hanyu Gongneng Yufa Yanjiu* 漢語功能語法研究 [*Functional Studies of Chinese Grammar*]. Nanchang: Jiangxi Education Press.
- Zhou, Xiaobing (周小兵). 1996. 'Lian' ziju de shengcheng yu fazhan “連”字句的生成與發展 [The generation and development of *lian* construction]. *Jufa, Yuyi, Pianzhang: Hanyu Yufa Zonghe Yanjiu* 句法・語義・篇章：漢語語法綜合研究 [*A Study of Chinese Grammar: Syntax, Semantics and Discourse*], 170-183. Guangzhou: Guangdong Education Press.

On the Grammaticalization of “*lian* X + *dou* VP” Construction in Mandarin Chinese

Yonglong Yang

Chinese Academy of Social Sciences

This paper is an attempt to research on the grammaticalization of the form-meaning pair of “*lian* X *dou* VP” construction. It argues that there exists, in the process of grammaticalization of this construction, syntactic extensions from “N1 *lian* N2 + VP” to “(agent +) *lian* N2 + VP” and “*lian* N2 + *dou* VP” as well as semantic changes from inclusive construction to emphatic construction through the emergence of specific presuppositions and implications motivated by pragmatic inference, absorption of contextual meaning, and copying of previous semantic patterns.

Key words: “*lian* X + *dou* VP” construction, grammaticalization of construction, emphatic construction, presupposition, implication

Peut-on compter les langues?

Gérard Lenclud

CNRS

Les langues se plient-elles au test du dénombrement? Non, à l'évidence. Le terme de langue n'est pas un authentique nom nombrable. Cela empêche-t-il le linguiste de leur appliquer la grammaire du même et de l'autre? En théorie oui, dans la pratique non. Les langues représenteraient-elles une catégorie d'entités à part? Aucunement: on ne sait pas davantage compter les cultures ou les nations, ni même, comme on va le voir, les arbres, les bateaux ou les personnes. Faudrait-il alors réviser notre ontologie de sens commun et lui substituer une ontologie de parties temporelles selon laquelle êtres et choses seraient tout ce qui leur arrive? Peut-être. C'était, au demeurant, le point de vue de Saussure pour qui une langue est son histoire. L'adoption généralisée d'une telle ontologie entraînerait toutefois des conséquences déconcertantes.

Mots clé: Décompte, identité numérique, identité spécifique, langue, ontologie

J'ai demandé un jour à Alain Peyraube si l'on pouvait compter les langues. Il me répondit sans hésiter que c'était là, à l'évidence, tâche impossible. Pourquoi lui avoir posé la question? Mon intention n'était pas — je le jure — de confronter cet ami linguiste au contenu de maints écrits provenant de sa discipline. Certains de ses collègues donnent le nombre de langues parlées à la surface du globe; selon Grimes & Grimes (1993), il y en avait très exactement 6528 en 1993. D'autres évaluent à quelques unités près celui des langues en danger d'extinction à la façon dont un spécialiste du Muséum peut concevoir de calculer le chiffre total des scarabées pique prune en péril mortel pour cause d'autoroute en construction. Non, si j'ai eu l'idée d'interroger Peyraube sur le sujet, c'est en ayant en tête les problèmes que l'on soulève dans ce domaine de la philosophie auquel on donne parfois le nom de métaphysique descriptive. Ces problèmes sont ceux que suscite notre ontologie de base. Or elle repose sur le schème conceptuel d'identité et ce schème est intimement lié à la division de la référence. «Jusqu'où, explique ainsi Quine, est-ce la *même* pomme que vous avez et à quel point commencez-vous d'en avoir une *autre*?» (Quine 1960). Il faut donc savoir ce qui compte pour une pomme et, partant, pour autant qu'on veut.

Sitôt qu'on s'avise de transporter ces spéculations métaphysiques dans notre univers quotidien, au laboratoire mais aussi bien en croisière, nous voilà face à une énigme. Cette énigme n'est pas celle qu'on croit. Il n'est nul besoin de la métaphysique, après tout, pour constater que nous sommes incapables, dans bien des situations, de dire ce qui compte pour un et, par conséquent, pour deux. Elle réside dans le fait que nous nous accommodons fort bien des difficultés apparemment insurmontables que devrait entraîner cette incapacité principielle. Force sera de nous limiter ici à quelques exemples.

Mais d'abord observons qu'un éminent linguiste affirme qu'il est impossible de compter les langues. Pourtant, il ne saurait s'empêcher, sauf à quitter le métier, d'appliquer aux langues la grammaire du même et de l'autre. Or l'application de cette grammaire exige, nous l'avons vu, de faire la différence entre un et deux, donc a priori de savoir compter les langues. On remarquera, par conséquent, le fait suivant: le concept de langue n'autorise pas à en dénombrer les exemplaires et voilà qui ne constitue aucunement un frein au développement de la linguistique. Certes, ainsi qu'y insistait Roman Jakobson, le commun des mortels vérifie aisément qu'une langue, on la parle ou non; cela fait très empiriquement la différence entre une et deux. Et dire d'un homme qu'il est bilingue, ce n'est pas comme découvrir qu'il est deux en un ou un en deux. Toutefois, s'agissant des langues, un linguiste n'est pas le commun des mortels. C'est pourquoi certains linguistes s'appliquent parfois à raffiner le concept d'identité spécifique de langue. Il s'agit de le doter du pouvoir diviser sa référence; grâce à lui, on pourrait alors émettre des jugements d'identité numérique mieux assurés à propos des langues. En vain, comme Peyraube le confirme; «langue» n'est toujours pas un nom nombrable, au sens de Quine, comme le serait le terme général «pomme».

Serait-ce, par hasard, que les langues constituent un cas à part dans ce que l'on appelle parfois, à grand tort, le mobilier du monde? (A grand tort puisqu'il semble que, justement, l'on soit à même de compter des meubles.) Aucunement. Personne ne sait compter des sociétés, des cultures, des nations, des peuples, des religions, des sciences ou des arts. Cela n'empêche aucunement que toutes ces entités fassent l'objet de travaux savants, y compris comparatifs. Cela n'interdit pas davantage d'émettre des jugements d'identité numérique sur elles, explicites ou déguisés. Un jugement d'identité numérique explicite est celui formulé hasardeusement par Jules Michelet lorsqu'il parle de la France, «une et identique depuis plusieurs siècles». Un jugement d'identité numérique déguisé est celui qui se loge dans toute prédication. Si je dis, par exemple, que la France a pour capitale Paris, j'attribue à la France, et à nul autre pays qu'elle, la propriété d'avoir Paris pour capitale. J'extrait donc la France d'une pluralité d'autres entités, incluant l'univers chez des esprits chauvins, susceptibles d'être les sujets logiques de cette proposition. Ergo je l'individualise. (On prendra garde ici qu'un jugement d'identité spécifique, ouvrant à l'assignation d'une identité numérique, n'a strictement

rien à voir avec l'attribution d'une identité collective. Il faut se souvenir, lorsqu'on s'aventure à évoquer l'identité française, qu'un homme n'appartient pas à sa nation comme il appartient à l'humanité. Une chose est d'être le représentant d'une espèce, dont le nom est exprimée par un terme général; une autre d'être le participant d'une entité, exprimée par un terme collectif. Un être humain se contente d'instancier la première; il est acteur de la seconde, à sa façon qui n'est pas forcément celle de son voisin.) Bref il en est des langues comme de toutes ces entités qu'on ne saurait commuer en individus, au sens logique ou même approché du terme. On ne dispose pas d'un principe de dénombrement; il n'en reste pas moins qu'on leur applique des jugements d'identité numérique.

Il peut sembler, avons-nous dit, que l'on soit à même de compter, du moins en théorie sinon en pratique, d'autres catégories d'entités: celles qui présentent un caractère individuel mieux affirmé que les langues, les cultures ou les nations. Et à ce titre, peut-on penser, elles existeraient plus réellement. Il s'agit, par exemple, des êtres vivants en général (animaux humains, non humains et plantes), des objets matériels ou des personnes. Le propriétaire d'un pavillon de banlieue ne craint pas de dire qu'il y a 3 arbres dans son jardin. L'amiral qui fait manœuvrer son escadre au large des côtes varoises n'ignore aucunement combien de bâtiments il commande. Et lorsqu'il invite à sa table, il n'hésite pas sur le nombre de couverts à dresser: autant que de personnes; il n'y en aura pas entre 19 et 20 mais ou bien 19 ou bien 20.

Hélas, les choses ne sont pas aussi simples. Prenons les arbres pour éviter d'envisager des êtres vivants dont le décompte fait figure de puzzle (un ver de terre coupé en deux est-il un ou deux? Et un bulbe de jonquille se divisant en deux spontanément?). Certes notre banlieusard est en droit d'affirmer qu'il y a 3 arbres dans son jardin. Pourtant, en prélevant des échantillons d'ADN dans différentes branches de mêmes arbres, les botanistes ont souvent constaté des différences génétiques très nettes. Chaque branche peut posséder son propre génome, d'où l'idée que l'arbre n'est sans doute pas un individu mais un «être collectif», une colonie. A en croire la biologie végétale, le jardin de notre banlieusard est bien plus peuplé qu'il ne pense. Chacun de ses arbres est aussi incertain dans son principe (naturel) de composition que l'arbre généalogique au moyen duquel le linguiste représente traditionnellement les langues à classer. Attardons-nous un instant sur l'amiral surveillant des yeux sa flotte. Se pourrait-il vraiment qu'il se trouve, face à ses bateaux, dans la même situation que le linguiste face à ses langues? Assurément un bateau est, selon la formule de Leibniz, un être par agrégation donc pas un *objet* faute d'être *un* objet, mais, après tout, un bateau ne vient au monde que de notre fait (Leibniz 1705). Nous devons donc posséder assez solidement le concept d'espèce «bateau» pour le mettre à contribution et déterminer s'il y en a ici un ou deux. Un amiral n'est pas le plus mal placé à cet égard.

Pour être en mesure de dénombrer des bateaux, il ne faut pas seulement connaître les conditions nécessaires et suffisantes d'appartenance d'un objet à la classe des bateaux. Certes cela implique d'abord, et bien évidemment, de savoir tracer la ligne de démarcation entre ce qui est un bateau (canot ou torpilleur) et ce qui n'en est pas un (tronc flottant ou hydravion). Cela implique également, plus indirectement mais avec autant d'évidence, de savoir délimiter le seuil à partir duquel ce bateau, la *Catherine deux*, cesse d'être celui qu'il est pour en devenir un autre; c'est-à-dire savoir en vertu de quoi un bateau reste identique à lui-même ou encourt un changement d'identité numérique. On ne saurait assigner la possession d'un concept d'espèce à quelqu'un qui ignorerait les conditions mises à l'existence continuée d'un exemplaire de cette espèce. Or tel est bien le cas de chacun d'entre nous, amiral compris.

L'énigme du Bateau de Thésée en fait la démonstration. Offrons-en une version modifiée pour les besoins de la cause. Un bateau de la flotte amirale, l'*Ange-Michel*, quitte le port de Toulon. Il fait escale dans l'île de Monte-Cristo et là, il est entièrement démantelé, pièce par pièce, puis intégralement reconstruit à l'aide de matériaux qualitativement identiques. Appelons ce bateau rénové l'*Ange-Michel Bis*. En même temps, d'autres ouvriers s'occupent à réutiliser les pièces anciennes de l'*Ange-Michel* pour construire un bateau, ce bateau reconstitué qu'on nommera l'*Ange-Michel Ter*. Donc un bateau a quitté Toulon, l'*Ange-Michel*, et deux y ont fait retour, l'*Ange-Michel Bis* et l'*Ange-Michel Ter*. Combien de bateaux cette histoire met-elle en scène? Deux ou trois? Et si c'est deux, l'*Ange-Michel* s'est-il continué dans l'*Ange-Michel Bis* ou *Ter*? Même l'amiral est embarrassé et n'arrive pas à trancher. Le concept d'espèce de bateau ne livre donc pas les conditions de persistance d'un (et le même) bateau. On échoue, par conséquent, à statuer en toute confiance sur l'unité de compte «bateau». Passons vite sur les personnes que nous ne confondons pas, ici du moins, avec les êtres biologiques dont, selon certains, elles ne feraient qu'emprunter le corps. Il est de notoriété publique que nous ignorons quand une personne commence d'être et quand elle finit d'être. Un nourrisson d'un jour, encore privé de tous les attributs dont nous dotons la personne, en est-il une? Un homme en état de coma dépassé, ayant définitivement perdu ces mêmes attributs, en est-il une? (Il se trouve que l'amiral a convié à sa table une personne ayant subi une bissection du cerveau au niveau du corps calleux. Le médecin du bord, qui a lu les travaux à ce sujet, assure que, dans ce cas, chacun des deux hémisphères cérébraux, fonctionnant indépendamment, constitue une sphère séparée de conscience. L'invité serait donc, selon lui, deux personnes appelées à manger dans une seule assiette).

Bref les concepts d'espèce d'arbre, de bateau, de personne, passés au banc d'essai dans des conditions qu'on avouera extrêmes, ne nous livrent pas les conditions nécessaires et suffisantes de préservation de l'identité dans le temps. Ils ne possèdent pas le pouvoir de diviser leur référence. Nous savons qu'êtres et choses changent dans

le cours de leur carrière terrestre; nous sommes incapables, en revanche, de délimiter le seuil de transformation à partir duquel ils cessent d'être ceux qu'ils sont. (Assurément la femme de Loth, changée en statue de sel n'est plus, de l'avis majoritaire, la personne qu'elle était bien qu'elle regrette cruellement dans son for intérieur d'avoir troqué de corps mais il peut sembler que nous hésiterions si elle s'était insensiblement pétrifiée.) En bonne logique, nous devrions, par conséquent, nous abstenir de les compter. Et pourtant, nous leur appliquons la grammaire du même et de l'autre comme Monsieur Jourdain faisait de la prose. En effet, du matin au soir, tout au long de notre vie, nous reconnaissons êtres et choses, affirmant par là que chacun d'eux est resté celui qu'il était, un et le même, en dépit de ses changements manifestes, hêtre drastiquement étété, contre-torpilleur commué en casino flottant, adolescent dissipé métamorphosé en candidat futur à la béatification. Il s'avère donc que le concept d'espèce de langue est en bonne compagnie et les linguistes nullement seuls à identifier leurs objets sans critère rigoureux d'identité. Aucun linguiste ne saurait disposer du certificat de venue au monde d'une langue. «Jamais, disait Saussure, on n'a signalé sur le globe la naissance d'une langue nouvelle.» Aucun d'eux n'est en mesure de déterminer avec quelque exactitude l'instant historique à partir duquel une langue, n'étant plus celle qu'elle était, en est devenu une autre. Cela n'empêche pas les linguistes de faire d'une langue la fille de sa mère ou de dresser le nombre de langues apparentées entre elles.

Plusieurs métaphysiciens ne sont pas satisfaits de notre ontologie de base, notamment pour les raisons que nous venons d'évoquer. Ils proposent de la réformer. C'est ainsi que certains d'entre eux suggèrent de lui en substituer une qui serait mieux en accord avec une version scientifique du monde: une ontologie de «parties temporelles», promouvant les objets dotés d'un corps matériel en êtres quadridimensionnels, des «êtres-processus». Ces objets, un arbre, un bateau, vous ou moi, ramperaient dans le temps, mis à égalité avec l'espace, comme des chenilles sur le sol. Dans le cadre de cette ontologie de «perdurants» et non plus de «continuants», Peyraube, par exemple, serait autant fait de tranches temporelles que de parties spatiales. Il a un tronc, des bras, des jambes; il aurait aussi des parties temporelles, autant qu'il a eu de phases d'existence. Une partie de lui est (et non pas était) dans le Béarn, une autre à Bordeaux, une troisième dans les campagnes chinoises, etc. Peyraube serait non seulement la somme totale de ses parties spatiales mais l'enchaînement complet des événements tissant sa carrière temporelle. Son élection à l'Académie des sciences sociales de Chine lui serait donc ce que lui sont son foie ou ses mollets. Dans le cadre d'une ontologie de parties temporelles, un être n'endure pas des changements; les changements sont incorporés en lui. Un être n'a pas d'histoire; il *est* son histoire. Il *est* tout ce qui lui arrive.

Un linguiste ne saurait se dire étonné par cette ontologie. En effet, le mode d'existence assigné par elle aux êtres et aux choses est très exactement, quoiqu'en

fonction d'attendus différents, celui que Saussure prête aux langues. A le lire, une langue est un être formel et non substantiel, nullement «défini et délimité dans le temps». L'histoire d'une langue est celle de la transformation perpétuelle d'une identité perpétuellement préservée. En vertu de ce principe de «continuité absolue», une langue n'a, à proprement parler, ni passé ni futur; elle se déploie dans un présent constamment renouvelé. Une langue a donc toujours été; elle sera toujours encore. On se souvient que Saussure s'en explique en prenant pour exemple le cas du français et du latin. Le français a succédé au latin, c'est une affaire entendue. Pourtant, dans cette succession, il n'y a pas deux langues: le français *est* le latin transformé. Ce «est» n'est pas de prédication mais d'identité. D'où il résulte que, selon Saussure, il n'y a pas de sens à affirmer que le français descend du latin. Le concept d'espèce de langue, forgé par Saussure, interdit, entre autres, de parler de langues mères ou de langues filles. Partout, à la surface de la terre, il y a un «état de langue» incorporant ses changements au jour le jour. De même que Peyraube, dans une ontologie de parties temporelles, *est* tout ce qui lui arrive, une langue, pour Saussure, *est* son histoire. Si une langue n'a jamais commencé ni fini d'être, celui qui s'essaie à compter les langues risque moins de se tromper que de se rendre coupable d'un contre sens. Comment note-t-on aujourd'hui sa copie?

Il me faudra reprendre ma discussion avec Peyraube. De cette leçon qu'il voudra bien m'accorder, il résultera incidemment que nous pourrons, lui et moi, enrichir d'une unité supplémentaire la somme des parties temporelles qui nous sont communes. Personne ne songe à nier que sa jambe de bois fait partie intégrante du capitaine Achab. Pourquoi en irait-il différemment des relations d'amitié? Elles font autant partie de chacun d'entre nous qu'une prothèse ou même un talon d'Achille.

Références

- Grimes, Joseph E., and Barbara F. Grimes. (compilers) 1993. *Ethnologue Language Family Index*. Dallas: Summer Institute of Linguistics.
- Leibniz, Gottfried. 1705. *Nouveaux Essais sur l'entendement humain* (publié 1765).
- Quine, Willard. 1960[1977]. *Word and Object*. Cambridge: MIT Press. (trad. J. Dopp & P. Gochet, *Le Mot et la Chose*, Paris, Flammarion, coll. «Champs», 1977, rééd. 2010).

Are languages countable?

G rard Lenclud

CNRS

Actually *language* is not a count noun, since no one can identify exactly when exactly Late Latin, for instance, disappeared giving rise to the different Romance languages. And there is notoriously some difficulty in deciding whether the speech of two groups constitute dialects of one language or two distinct languages, since mutual intelligibility is a matter of degree. Therefore, even though linguists do count languages, it seems impossible to say how many different languages are spoken in the same way as one might say how many ships there are anchored in a harbour.

But now is *ship* truly a count noun? The famous paradox of the ship of Theseus, as described by Thomas Hobbes, seems to demonstrate this is not the case. That is why some philosophers have suggested that persisting objects should be considered as “perdurers”, through changes, and not as “endurers”, suffering changes. It implies that such objects as ships or persons do not *have* a history, they *are* their own history. They can be conceptualized as being made up of temporal parts as well as of spatial parts. Maybe Saussure would have agreed to languages being conceived four-dimensionally (not being physical objects). But as far as persons are concerned, it is rather disconcerting to look at them this way. When facing a person, am I facing a person-stage, existing on its own? And would for instance Alain Peyraube’s first years in China be part of him in the same way his right arm or his Achilles tendon actually are ?

Key words: count, digital identity, specific identity, language, ontology

On Principles and Practices of Language Classification

Katia Chirkova (齊卡佳)

CRLAO, CNRS

This article is an overview of issues in language classification, in particular in connection with three subgroups of the Tibeto-Burman language family: Tibetic, Sinitic and Qiangic. First, I discuss the practical application of currently prevalent classifications of Tibetic, Sinitic and Qiangic languages. Then, with reference to insights from classification practices in biology, I review alternatives for a practical classification in linguistics.

Key words: language classification, genetic classification, Sinitic, Tibetic, Qiangic, Tibeto-Burman

1. Introduction

It is a great privilege and pleasure to contribute to the *Festschrift* in honor of Alain Peyraube. I gladly and admiringly join this celebration of his scientific achievements, the diversity of his interests, his renowned erudition, and his foresight on issues of great impact to the field. In this article, I reflect on one of Alain Peyraube's most recent interests: linguistic classification from a general and more specific, East Asian, perspective. His contributions to the issue of classification (Peyraube 2005, 2007) touch upon major trends in the field of East Asian linguistics, namely:

- (1) unabating interest in classification issues, from periodic revisions of the precise composition of traditional groupings of languages in China and East Asia (such as Sino-Tibetan, Austronesian or Austro-Asiatic) to the construction of new macro-families (such as Austric, Austro-Tai, Sino-Tibetan-Austronesian or Proto-East-Asian)
- (2) strong emphasis on interdisciplinary studies, whereby linguistics is coupled with other disciplines (such as genetics, evolutionary biology, archaeology or anthropology) to test relevant phylogenetic hypotheses
- (3) fruitful exchange between biology and linguistics, inspired by a deeply entrenched view of parallelism between the development of living organisms and the development of languages

These major trends guide and shape daily linguistic practice. Since classification is fundamental to all aspects of linguistic research, my main concern in this article is the practical application of existing classifications and the impact of the aforementioned trends on ongoing synchronic investigations. Given my area of expertise, the ensuing discussion will be limited to the field of Tibeto-Burman languages.

2. The theory of evolution and genetic classification in linguistics

In a metaphor that is nearly as old as linguistics itself, language is often viewed as an organism. Ranging from a direct equation¹ to a more conventional symbolic portrayal, the conceptualisation of language as an organism has long accompanied the development of the discipline.² Indeed, parallels between languages and organisms abound (e.g. Janda & Joseph 2003:59-81, Peyraube 2007), allowing for profitable exchange between linguistics and biology.

Fundamentally, both biology and linguistics are grounded in the theory of evolution, which both disciplines hold to be the most adequate and unitary hypothesis to which a great variety of both biological and linguistic phenomena can be related. As a result, the most widespread and influential type of classification in linguistics is genetic, based on the assumption of common ancestry of languages and using basic vocabulary, sound correspondences and, whenever available, grammatical (essentially morphological) evidence as classification criteria. This type of classification is so basic to linguistics that no explicit justification of its validity and legitimacy is normally deemed necessary. Characteristically, the recent edition of the *Encyclopaedia of Languages and Linguistics* (2006) names genetic classification “the most satisfying way to group languages” for “professional linguists and general readers” (Blake 2006:446), but does not explain what is so satisfying about it. One of the few scholars to explicitly address this issue, Michael Noonan (2008:4), points out that genetic classification is not always useful in providing information about the structure of a language, especially on higher taxonomic levels. Instead, he notes, information about where in the world a language is spoken provides more useful information about grammatical structure. On the other hand, Noonan argues, genetic classification has proven a boon to historical linguistics and theories of language change. Ultimately, Noonan concludes, the most satisfying aspect of genetic classification is its sheer familiarity, since “What family does it belong to?” is typically “the first thing a linguist will ask on being told of an unfamiliar language.”

¹ Recently, language as a parasitic species (Kortlandt 1985, 2003, Mufwene 2001) or language as a symbiont (van Driem 2004, 2008).

² See van Driem (2008) for an overview.

In other words, genetic classification is popular due to its familiarity, but its practical value with regard to concrete linguistic data is questionable. Nonetheless, genetic classification is considered superior to other types of linguistic classification, for example, those based on typological features (see §3 for examples). This assumed superiority arguably results from the fundamental assumption that a classification based on evolution, which is in turn responsible for the existence and structure of natural languages, is imbued with explanatory power. Due to the deeply entrenched connection between explanation and theory in science, explanatory (viz. theoretical) analyses and frameworks are routinely valued higher than descriptive ones (cf. Dryer 2006:212-214).³

The dominance of genetic classification is the result of the sweeping success of Indo-European historical linguistics, which, in turn, is due to the clear evidence of a common origin for groups of languages in Europe (e.g. Romance languages) and the extensive amount of written records. The main metaphor that is used by Indo-European linguistics to explain the historical relationships, the genealogical family tree, is based on the assumption that the phylogenetic relationship between groups of languages is analogous to the genealogical relationship between individual languages, i.e. descent along a single genetic line (inheritance or vertical transmission).⁴ The straightforwardness of the model and its appeal in the Indo-European context stimulated its export to other linguistic areas.

Uncritical application of the family tree model to languages outside the Indo-European family has repeatedly engendered criticism, most recently and forcefully in Dixon (1997) or Aikhenvald & Dixon (2001). The conclusion of these two studies is that the family tree model is appropriate and can be proved through reconstruction in many cases (as in the case of Indo-European languages), but is problematic or even unsuitable in those cases where historical evolution is obscured by contact-related phenomena (as in the case of Australian languages). In other words, linguistic convergence and areal diffusion obscure historical development, so that originally genealogically unrelated languages may develop surface resemblances, forming groups which are synchronically similar, but genetically polyphyletic (that is, with multiple ancestors).

Borrowing and contact-induced change (i.e. horizontal transmission) as factors affecting the linear evolution of languages (i.e. vertical transmission) are now commonly taken into account in mainstream historical linguistics. Nonetheless, the impact of contact on linear descent remains subject of heated debate in linguistics, as in biology (cf. Dagan & Martin 2006:118.1). At present, awareness of the importance to enrich the

³ In the words of Karl Popper (1969:103): “The scientist aims at finding a true theory or description of the world [...], which shall also be an explanation of the observable facts.”

⁴ The drawbacks of the group to an individual analogy have been pointed out, among others, by Gilmour (1940:469-470) for systematics and, recently, Mufwene (2001:16-21) for linguistics.

dominant family tree model by taking into account “all the processes that affect language formation and development” (Chappell 2001:354) is on the increase. Much research is being undertaken to find objective criteria for assessing the balance between the effects of vertical and horizontal transmission on language change. Some of the most acute challenges include distinguishing loans from cognates, and common innovations from retentions.⁵ Attempts to “undo” the effects of contact include a methodology for separating innovations resulting from genetic inheritance from those diffused through contact, and a fully articulated theory of the relative diffusibility of features in a contact situation (e.g. Curnow 2001, Dench 2001). To date, all these attempts have been proven futile (Curnow 2001, Dench 2001) and are even argued to be doomed to failure (Thomason 2000).

In those cases where the linear development is known to be obscured by contact-related phenomena, as in the case of many East Asian languages (e.g. Benedict 1972, Matisoff 2001), and where the horizontal dimension of language change is supposedly as important for a coherent assessment of data as the vertical dimension, the emphasis is invariably on the latter (viz. genetic descent). The large number of relevant publications characteristically reflects the dominance of the genetic approach and the associated genetic classification. Also symptomatic of this dominance is the fact that general studies in language classification are limited to genetic classifications (e.g. Ruhlen 1994, Campbell & Poser 2008). Such studies mostly focus on higher taxonomic levels: families and macro-families. While major debates revolve around distant genetic relationships of considerable time depth (such as Nostratic or Amerind), classification issues at more empirically bound lower taxonomic levels have all along attracted less attention. This is remarkable, because it is at these lower taxonomic levels, where data are most voluminous and versatile, that the need for meaningful classifications to organize and store empirical knowledge is at its greatest. I have been confronted with the practical application of these low level classifications of languages (Tibetan and Chinese dialects (or Tibetic and Sinitic languages) and Qiangic languages) and will focus on this issue in the main part of this article.

⁵ A relevant example from the field of Tibeto-Burman linguistics is the Bai language, whose genetic affiliation has long been the subject of scholarly speculation. The ongoing controversy is due to the so far unresolved issue of whether numerous lexical similarities between Chinese and Bai reflect long-term contact between the two languages (loans) or inherited features (cognates).

3. Classifications of Tibetan and Chinese dialects and of Qiangic languages

Classification is a highly practical activity. It provides a practical means by which objects can be identified and compared, and the knowledge about them organized. In addition, a meaningful classification enables the researcher to make inductive generalizations concerning the classified data and to discover new knowledge of the variation and distribution of the examined attributes. In the words of Ernst Mayr (1976:427), a good classification, like a good scientific theory, “has a high predictive power with respect to the assignment of newly discovered species and the pattern of variation of previously unused characters. That classification is the best which is at least affected by such new discoveries”.

How do currently prevalent classifications of Tibetic, Sinitic, and Qiangic languages square with this desideratum for classification? I deal with this question as a practitioner who would like to use existing classifications as a descriptive tool and as a basis for inductive generalizations, while profiting from their predictive power with respect to adding new dialects or languages to the existing classification schemes.

Importantly for the ensuing discussion, the three groups considered (Tibetic, Sinitic, Qiangic) can be divided into two groups, based on whether the genetic relationship among the languages within each group is presupposed or established by comparative linguistics.

The first group, comprising Tibetan and Chinese dialects, consists of languages whose respective genetic relatedness is a matter of oral and written memory, hence the designation “dialects”, even though both Tibetan and Chinese dialects are, respectively, synchronically as diverse as the Romance or Germanic languages within the Indo-European language family. The knowledge of their respective genetic relatedness and even the accepted division into distinct groups by and large derive from non-linguistic sources: self-awareness, culture, history, and geographical distribution of the groups in question. In this context, existing linguistic classifications have mostly sought to translate this established division into linguistic terms. Chinese and Tibetan dialectal studies commonly hold that most substantial differences between dialects lie in phonology and lexicon, whereas syntactic differences are relatively negligible. Hence, the favoured criteria for classification are for the most part phonological. Several classifications have been advanced for both Tibetan and Chinese dialects (the mostly widely accepted classifications are considered in some detail below).

The second group, Qiangic, brings together languages whose genetic relatedness is not known, but hypothesized.

In terms of a genetic argument, all considered cases (Tibetic, Sinitic, Qiangic) are subgroups. That is to say, they are known or hypothesized monophyletic groups. A subgroup admits as evidence only shared innovations as distinguished from retentions or parallel developments (e.g. Harrison 2003:232-239).

The most widespread classification of Tibetan dialects spoken in China divides all dialects into three groups: Dbus-gtsang (or Central), Khams, and Amdo, which correspond to the three namesake historical provinces of Tibet.⁶ In linguistic terms, this classification is based, among others, on the presence of tone, on the simplification of consonant clusters, on the presence of voiced obstruent initials, on the development of Old Tibetan clusters with medials *-w-*, *-y-* and *-r-*, on the number of consonant codas, and on the presence and the number of diphthongs.⁷ (An example of one classification (Hu 1991) is provided in Figure 1. The three groups of dialects spoken in China are indicated in bold.)

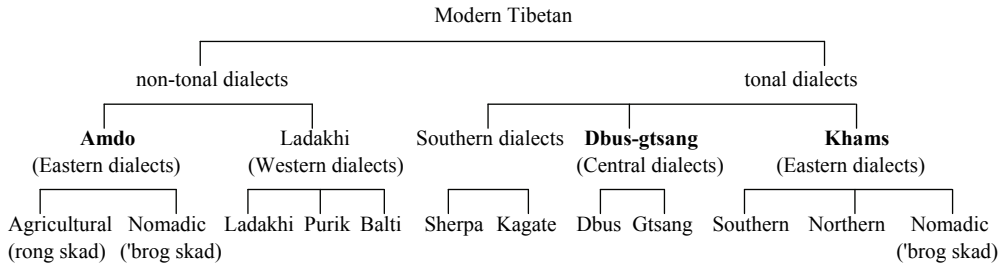


Figure 1: Classification of Tibetan dialects (adapted from Hu 1991:177)

Viewed from the principles of classification outlined above, this classification yields little practical value. Overall, the underlying classification criteria are neither mutually exclusive nor representative for each respective group. As a result, they fail to single out homogeneous groups and to do justice to the actual diversity of Tibetan dialects. Consequently, the resulting classification does not allow for generalizations about the formed groups and it has no predictive power with respect to the assignment of newly

⁶ Cf. Gesang (1964), Qu & Jin (1981), Jin (1983:114-145), Hu (1991:173-177), Zhang (1993, 1996).

⁷ Alternative classifications of Tibetan dialects spoken in China also exist. Targeting the presence of tone as the sole criterion, Hu Tan (1991) arrives at a division of all dialects into just two groups: Khams-gtsang and Amdo (Figure 1). Increasing the number of criteria to five (presence of tone, voiced obstruent initials, consonant clusters, consonant codas and diphthongs), Huang et al. (1994) arrive at four groups: Dbus-gtsang, Khams, Amdo and Yushu 玉樹 (spoken in Zaduo 雜多 *rdza rdo*, Jiegu 結古 *skye rgu*, Nangqian 囊謙 *nang chen* Counties of Yushu Prefecture in Qinghai Province, and Baqing 巴青 *sbra chen*, Dingqing 丁青 *steng chen* and Suo 索 Counties in Tibet).

described varieties. Take, for instance, the Baima language (also known as the Baima Tibetan dialect) of Northern Sichuan and Southern Gansu provinces in the People's Republic of China (PRC). The linguistic affiliation of Baima has long been disputed, due to the popular belief among the Baima people that they are descendents of the Di people, an ancient Tibeto-Burman ethnic group who inhabited roughly the same area until their gradual assimilation into the Han and the Tibetans during the Tang Dynasty (Chirkova 2008c). If judged solely on linguistic grounds, the language of the Baima people is predominantly Tibetan-like in all its linguistic sub-systems, even though it exhibits some non-Tibetan features in its lexicon, morphology, and syntax (Chirkova 2008b). Accordingly, Baima is argued to be a Tibetan dialect (Huang & Zhang 1995:104, Zhang 1997:134-135, 140). The problem, however, is that this new Tibetan dialect cannot be straightforwardly accommodated by the dominant tripartite dialect classification of Tibetan dialects spoken in China.⁸ Based on the exhibited characteristics, Baima is at the same time Khams-like (given the drastic simplification of the syllable structure, the elimination of Old Tibetan syllable codas, and the presence of tones) and Amdo-like (given, for instance, the development of Old Tibetan velar stop clusters with medial *-r-* into alveopalatal affricates and of Old Tibetan dental and labial stop clusters with the same medial into retroflex affricates).⁹ In sum, the dominant classification of Tibetan dialects of China fails to accommodate Baima. In fact, the same holds true for any of the newly discovered dialects in its neighbourhood, such as Zhongu (J. Sun 2003a), Chos-rje (J. Sun 2003b), or the Tibetan dialects of Western Sichuan (Suzuki 2008).

In his discussion of the popular tripartite classification of the Tibetan dialects of China, Jackson T.-S. Sun (2003a:794-797) points out that this classification mixes possibly convergent, but independent developments (such as the presence of tone) with shared retentions (such as the preservation of Old Tibetan voiced obstruent onsets). This, in Sun's analysis, explains its inadequacy. The envisaged solution is to upgrade this routinely accepted (but overall unrevealing) classification into a meaningful (i.e. genetic) classification, which, in turn, should be based on shared innovations and phonological isoglosses to delimit precise dialect boundaries. Sun does not discuss the feasibility of establishing a classification along these lines and, to my knowledge, such a classification has not yet been attempted. However, a close investigation of a possibility of such a *stricto sensu* genetic classification of Chinese dialects by Laurent Sagart (1998) suggests that such a classification is not viable.

The currently widely accepted classification of Chinese dialects is based on one single criterion: the development of the Middle Chinese voiced stops into the modern

⁸ Presumably for this reason, Baima is not included in any of the comprehensive outlines of Tibetan dialects of China (e.g. Zhang 1996).

⁹ Chirkova (2008c, 2008d).

dialects (Li 1937, Yuan et al. 2001).¹⁰ While this criterion works fairly well in the majority of cases, as Norman (1988:181) points out, it is clearly valid only for a relatively recent historical period, given that the voiced initials probably persisted in the North down to at least the tenth or eleventh centuries. This traditional classification of Chinese dialects suffers from the same flaws as its Tibetan counterpart. It fails to do justice to the actual diversity of dialects and to form a reliable basis for generalizations regarding the formed groups. This problem is particularly acute in relation to dialect syntax (e.g. Chirkova 2008a for the Mandarin group).

The insufficiency of one criterion for a meaningful classification of Chinese dialects in general, and Hakka dialects in particular, has been pointed out by Sagart (1998), who further notes that a choice of some other criterion would result in an entirely different classification.

Similar to J. Sun's assessment of the current classifications of Tibetic languages, Sagart views the non-conformity of the traditional classification of Chinese dialects to the standards of genetic classification as one reason for its shortcomings. He notes that the traditional classification of Chinese dialects defines dialect zones whose historical statuses are very different: some are defined by innovation, others by retention, whereas it is generally admitted that only innovations are useful in linguistic classification. However, in stark contrast with J. Sun's conclusion, Sagart finds a *stricto sensu* genetic classification of Chinese dialects impossible and even unnecessary. It is impossible, Sagart argues (1998:299), because innovations may be obliterated or reversed through contact. Hence, dialect groups that are currently recognized in Chinese dialectology are fuzzy entities that owe much of their make-up to contact (i.e. horizontal transmission) as opposed to vertical inheritance.¹¹ This is not to say that a genetic classification of Chinese dialects is altogether unattainable. The linguistic history of the group can be reconstructed to the extent possible from written sources. However, as Sagart argues, a genetic dialect classification is unnecessary, because such a classification has at best a minor role to play in the furthering of knowledge on Chinese dialects. Instead, Sagart points out, geographical approaches to Chinese dialectology hold more promise.

The use of a small number of criteria and the use of common innovations as a valid criterion are disfavoured as basis for a meaningful classification in the case of well-documented dialects or languages *known* to be related (Sinitic). Curiously, these criteria gain in explanatory value in the case of little-known languages that are *not known* to be

¹⁰ Other classifications of Chinese dialects exist, based on different sets of criteria, e.g. Norman (1988) and Wurm et al. (1988). These classifications by and large correspond to the currently widely accepted division of Chinese dialects into groups in the tradition classification by Li (1937).

¹¹ This is a received view in Chinese dialectology, e.g. Pulleyblank (1991, 1998).

related. Take the Qiangic languages of the Chinese South-West as an example.

Qiangic is a putative group of eleven Tibeto-Burman languages of Sichuan and Yunnan provinces in the PRC. After H. Sun (1983, 2001), these eleven languages are often grouped in a northern branch and a southern branch.¹² The Qiangic subgroup in its current understanding was proposed in the 1970s by Chinese linguists.¹³ It brings together geographically adjacent, non-written, and under-researched languages, which are not known to be related. This subgroup relies primarily on shared typology: the widespread use of measure words, directional prefixes and similar case marking, to which some, equally general phonological features such as the presence of uvulars and tones have also been added (H. Sun 1983). Seen from mainstream historical linguistics, these are type-identifying features (Nichols 1996:48), in other words, features that are found in enough unrelated language families and that are consequently low in identificational value as to the genetic relationship. Overall, the percentage of common vocabulary among the languages of the Qiangic subgroup (the feature which, in contrast to the typological characteristics above, could provide more reliable support for the hypothesis of a common origin of the languages) is relatively small. Altogether it makes up less than 20% between any two Qiangic languages (Huang 1991:355). Notably, the type-identifying features in Qiangic are not found in their totality in languages from the neighboring subgroups. At the same time, languages of these neighboring (and phylogenetically better-understood) subgroups spoken in the Qiangic area, appear to share with Qiangic languages many features that are held specific of the Qiangic subgroup. An example of such a language is Kami, a dialect of Tibetan spoken in the southern end of the Qiangic area. Kami has typically Qiangic and at the same time strikingly non-Tibetan features, such as directional prefixes, measure words, or uvular phonemes (Chirkova forthcoming). The presence of these ‘Qiangic’ features in a Tibetic language suggests that Qiangic is rather a *Sprachbund*: a geographically delimited area that includes languages that are genetically unrelated, but share a fair number of highly distinctive traits.

Even though the grounds for its classification are demonstrably weak, the Qiangic subgroup has found its way in all present classifications of the Tibeto-Burman language family. This is perhaps because even despite their diversity, Qiangic languages are possibly more closely related to each other than to anything else in Tibeto-Burman (Bradley 1997:36).

Formulated on such loose grounds, the Qiangic subgroup fails the test of a meaningful classification: it cannot serve as basis for generalizations or single out

¹² Northern Qiangic languages include Qiang, rGyalrong (Jiarong), Ergong, Prinmi (Pumi) and Minyak (Muya). Southern Qiangic languages comprise Zhaba, Choyo (Queyu), Ersu, Namuyi, Xumi (a.k.a. Shixing) and Guiqiong.

¹³ For the history of the Qiangic subgroup, see J. Sun (1992).

features specific for the group. In fact, the precise composition of this subgroup has been debated since its very proposal. In the 1980s, the aforementioned Baima language was argued to be Qiangic, because it exhibits some features that are held to be typical of this subgroup, such as the presence of directional prefixes (H. Sun 1980). Conversely, in the 1990s, one of the Qiangic languages proper, Namuyi, was claimed to be genetically related to the Ngwi languages instead (Lama 1994, Huang 1997:13-15), based on a large amount of basic vocabulary shared between Ngwi and Namuyi.

Recent years have witnessed an upsurge of interest in Qiangic languages and linguistics. Many Qiangic languages are currently being investigated. Attempts are also undertaken to place this methodologically problematic grouping on a scientific footing: by focusing on common innovations, lexical comparisons and, eventually, phonological reconstructions.¹⁴ Viewed in the context of the above discussion, such a classification is bound to be based on few criteria, which are furthermore spurious, as no objective criteria to distinguish common features (in lexicon or elsewhere) from accidental similarity in form, borrowing, and/or genetic drift have yet been found (cf. Thomason & Kaufman 1988:6).¹⁵

In fact, there is not necessarily a correlation between the number of underlying criteria and the quality of the resulting classification. Even one character may be sufficient, if it is known to be correlated with many others of varied nature. If the correlation has been established, the use of the one feature will amount to the use of all correlated features together. Unfortunately, to my knowledge, lexical sharings are not correlated with any other linguistic feature. The use of this one single criterion is hence bound to result in a biased classification, whose validity is furthermore contingent on the acceptance of lexical sharings as probative of genetic relatedness.

What would be the gains of the genetic approach in the case of Qiangic languages, if pursued rigorously? A genetic argument in the subgroup context can ultimately lead to either of these two possibilities:

First, the comparative method could show that the Qiangic subgroup is polyphyletic. The fact that Qiangic languages have grown to share a number of salient features is then to be explained through effects of convergence and hybridisation. Notably,

¹⁴ The establishment of a genetic relationship in the Tibeto-Burman context must rely—*faute de mieux*—on shared derivational morphology and lexical comparisons, in spite of the fact that both can be borrowed, given a sufficient degree of contact. Because few inflectional paradigms can be identified across the family, shared and especially irregular inflectional paradigms cannot be relied on to prove the existence of the family (Handel 2008:435).

¹⁵ Subgrouping is known in mainstream comparative linguistics to be methodologically and factually problematic. However, its limitations are argued away on the basis of the practical necessity of subgrouping assumptions for the application of the comparative method (Harrison 2003:239).

the same conclusion can be directly arrived at by a closer investigation of empirical data, without resorting to the formal apparatus of lexical comparisons, loanword stratification, search for cognate etyma, establishment of sound correspondences, and phonological reconstructions.

Second, the comparative method could show that the Qiangic group is monophyletic. In that case, the extreme synchronic diversity of the languages forming the group, and the relatively meagre amount of shared cognate vocabulary, would have to be explained. The comparative method cannot provide this explanation and does not even posit such among its explicit goals. Consequently, a closer investigation of empirical data will again be required if any sense is to be made of the linguistic situation and the historical development of the relevant languages.

The following three conclusions can be made on the basis of the three cases reviewed in this section.

First, the dominant status of genetic classification and the expectations associated therewith are at odds with the actual practical merits of genetic classification. After all, genetic classification reflects known or hypothesized lines of descent, and in so doing, it restricts the available data in a way that excludes much information.

Second, in those cases where the line of descent is unknown (e.g. Qiangic), to proceed with an explanatory approach (genetic classification) before assembling a representative set of empirical data, runs the risk of applying potentially subjective criteria and of involving circular reasoning.

Third, as suggested by studies of well-documented languages that are known to be related (Sinitic, Tibetic), it is impossible to combine information on genetic descent with maximum information on synchronic similarities between the classified varieties into one classification scheme. Therefore, at least two approaches (and classifications) are warranted to organize these two types of information, for example, one genetic and one geographical (cf. Sagart for Sinitic).

Overall, given the little practical value of existing classifications of Tibeto-Burman languages, genetic or otherwise, and given the overall pressure to reflect on phylogenies, a linguist working with Tibeto-Burman languages cannot but get involved in classification work. Indeed, the lion's share of production in Tibeto-Burman linguistics deals with genetic classification issues: from individual languages to subgroups, subgroups to families, and families to macro-families (cf. Handel 2008). While existing linguistic classifications are far from satisfactory and no immediate linguistic solutions are in sight, it may be worthwhile, given the success of the "language as an organism" metaphor, to turn to biology for inspiration and possible solutions.

4. From classifications in biology to classifications in linguistics

In its simplest form, classification in biology is the ordering of entities into groups on the basis of their relationships, that is, of their associations by contiguity, similarity, or both (cf. Simpson 1961:9, quoted from Sokal & Sneath 1963:3).¹⁶ Notably, the term “relationship” may imply relationship by ancestry or, alternatively, it may indicate the overall similarity (or affinity) as judged by the characters of the organisms without any implication as to their relationship by ancestry. It has been argued that no classificatory scheme is able simultaneously to yield information on the degree of synchronic similarity and genetic descent (Michener & Sokal 1957, Sokal & Sneath 1963:55-57). Consequently, two major types of classification exist: empirical (synchronic) and genetic (diachronic). The empirical classification is often referred to as “natural”, i.e. denoting a system that occurs naturally in the empirical world. The two types of classification, natural and genetic, differ in scope, in purpose, and in the number of underlying criteria.

A natural classification is one that endeavors to utilize many attributes of the classified entities and to accurately reflect most of the various natures of the objects. While it is not explanatory as to the biological phenomena that it describes, a natural classification forms solid basis for generalizations concerning the described data and it is useful for a wide range of purposes (e.g. Gilmour 1940:472). An example of an empirical classification in biology, that groups animals and kinds according to their shared physical characteristics, is the familiar Linnaean taxonomy widely used in the biological sciences.¹⁷

A natural system of classification in biology is contrasted with more or less ‘artificial’ systems, which have no natural occurrence prior to their construction by the researcher. Such an ‘artificial’ classification is based on a limited range of attributes, selected with an eye to a specific classification purpose. Consequently, it is useful for limited purposes (Blackwelder 1967:186).

Given that evolution is responsible for the existence and structure of the natural system, classification in biology (as in linguistics) is possible because evolution has produced diverse kinds (or languages). From this it is possible to generalize that classification is based on evolution, but it must be recognized that this means that it is based

¹⁶ The compilation in this section is based on Huxley (1940), Sokal & Sneath (1963), Blackwelder (1967), Jardine & Sibson (1971), Salthe (1972), Mayr (1976, 1988), Bailey (1994), and Pigliucci & Kaplan (2006).

¹⁷ A phonetic view of classification in biology, according to which the sole aim of classifications is to reflect as accurately as possible the relative similarities or dissimilarities of populations, is frequently said to be pioneered by Michel Adanson in his *Familles des Plantes* (1763) (cf. Jardine & Sibson 1971:136-137).

on the results of evolution: empirical facts, synchronically observable concrete living organisms. Evolution can naturally be seen as the central cause in biology. Nonetheless, it cannot be used as an explanatory concept or an independent basis for classification. The main reason for that is that common ancestry is never known as a fact, but merely hypothesized. Furthermore, the phylogeny is based, on the one hand, on previous classifications and, on the other hand, on known empirical facts, so that there are no examples of classifications that are actually based on what were taken to be phylogenetic facts as distinct from the features of living organisms (Sokal & Sneath 1963:56-57, Blackwelder 1967:191-193). The genetic type of classification, which focuses on providing a summary of existing knowledge of phylogeny, is hence a subsidiary classification (e.g. Gilmour 1940:473).

To return to the field of linguistics, we can observe that linguistics does not have in its arsenal any analogy to natural classification in biology. Instead, genetic classification has come to be used as an explanatory concept and a general linguistic classification. It is, however, a secondary classification, because it selects few criteria (essentially basic vocabulary in the domain of Tibeto-Burman languages) and it pursues one specific domain of inquiry (the common ancestry of languages). As such, it falls short of accounting for the versatility of empirical data, and it is unable to accommodate new data, as discussed above on the basis of Tibetic, Sinitic, and Qiangic languages. While a significant and intriguing facet of linguistic inquiry, indispensable for the purposes of historical linguistics, it is unsuitable as general linguistic classification.

Successful in those cases where written records (linguistic “fossil record”) are available and linguistic similarities self-evident (through history or mutual intelligibility), genetic classification reaches its limits in those cases where relationships between languages are less evident and written records are scarce or absent. Since no reliable method to ascertain characters probative of historical relatedness has been developed, much caution is required when using genetic arguments for classification purposes. Of course, there is no harm in speculating about possible relationships between languages in the latter type of situation, but in linguistics, as in biology (cf. Sokal & Sneath 1963:8), it appears more informative to concentrate on collection and analysis of the maximum of empirical facts, because facts inform and advance progress in linguistic theory and historical linguistics.

5. Discussion

To return to my departing point as a practitioner in need of a practical linguistic classification, I cannot but conclude that a genetic classification might not be “the most satisfying way to group languages” for “professional linguists and general readers”, as it

is praised to be.

There is need for another general linguistic classification, which will be satisfying in the sense that it can give justice to the versatility of the subject matter of linguistics, while providing a practical means to organize our knowledge of various facets of the continuously growing body of data. To serve this purpose, this classification is bound to be empirical, i.e. based on a variety and quantity of features.

On a more abstract level, in classification as elsewhere in linguistics, there is no fundamental contrast between description and explanation, as there is naturally need for both descriptive and explanatory approaches (Dryer 2006). While there is no intrinsic incompatibility in describing a language and then explaining the things described, it appears desirable that both types of approaches enjoy equal footing, so that explanatory approaches do not channel descriptive ones. The primary goal in linguistics, as in any area of scientific inquiry, remains first and foremost to describe a set of facts without any particular theoretical implication. This set of facts is further open to synthesis, hypothesis generation and interpretation, including, for instance, grading of observed similarities in terms of their respective historical status (retentions, innovations, parallel developments), and assessing the balance between horizontal and vertical lines of transmission in language change. All in all, my claim is that we need both empirical and genetic classifications, as well as many other types of classification (for instance, typological) for various specific purposes.

In terms of close parallels between linguistics and biology, the two disciplines continue to go hand in hand. Both owe their success to the theory of evolution with the tree model as its central concept. This theory and the tree concept have all along exerted strong influence on the development of the two disciplines. For over a century, linguistics and biology were mainly concerned with fleshing out a universal tree in their respective disciplines. The ongoing revision of this concept in biology, where evidence begins to mount up that a substantial component of the evolutionary process is non-tree-like (Doolittle 1999, Dagan & Martin 2006, Lawton 2009), is bound to have a major impact on linguistics. In fact, linguistics (where the need to incorporate horizontal transmission into ideas about linguistic evolution has long been evident) is already undergoing a similar paradigm shift.

In sum, further progress is possible only if a broad range of varied data is taken into consideration, and this broad range of data requires well-grounded theories and adequate classifications to be well-understood.

References

- Aikhenvald, Alexandra Y., and R. M. W. Dixon. (eds.) 2001. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford & New York: Oxford University Press.
- Adanson, Michel. 1763. *Familles des Plantes*. Paris: Vincent.
- Bailey, Kenneth D. 1994. *Typologies and Taxonomies: An Introduction to Classification Techniques*. Thousand Oaks, London & New Delhi: SAGE.
- Benedict, Paul K. 1972. *Sino-Tibetan: A Conspectus*. Cambridge: Cambridge University Press.
- Blackwelder, Richard E. 1967. *Taxonomy: A Text and Reference Book*. New York & London: John Wiley & Sons.
- Blake, Barry J. 2006. Classification of languages. *Encyclopedia of Languages and Linguistics*, ed. by Keith Brown et al., 446-457. Amsterdam: Elsevier.
- Bradley, David. 1997. Tibeto-Burman languages and their classification. *Tibeto-Burman Languages of the Himalayas*, ed. by David Bradley, 1-72. Canberra: The Australian National University.
- Campbell, Lyle, and William J. Poser. 2008. *Language Classification: History and Method*. Cambridge & New York: Cambridge University Press.
- Chappell, Hilary. 2001. Contact and diffusion in Sinitic. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 328-357. Oxford & New York: Oxford University Press.
- Chirkova, Katia. 2008a. *Gěi* 'give' in Beijing Mandarin. *Cahiers de Linguistique – Asie Orientale* 37.1:3-42.
- Chirkova, Katia. 2008b. On the position of Baima within Tibetan: a look from basic vocabulary. *Evidence and Counter-evidence: Essays in Honour of Frederik Kortlandt*, Vol. 2, ed. by Alexander Lubotsky, Jos Schaeken & Jeroen Wiedenhof, 69-91. Amsterdam & New York: Rodopi.
- Chirkova, Katia. 2008c. Baima Zangzu wei Dizu shuo zhiyi [The Baima Tibetans and the Di people of Chinese historical records: challenging the link]. *Bulletin of Chinese Linguistics* 3.1:167-180.
- Chirkova, Katia. 2008d. Baimayu shizheng fanchou ji qi yu Zangyu fangyan de bijiao [Evidentials in Baima and Tibetan dialects compared]. *Minzu Yuwen [Minority Languages of China]* 2008.3:36-43.
- Chirkova, Katia. (forthcoming). Phonological profile of Kami, the Tibetan dialect of Muli. Special issue of *Language and Linguistics on lesser known dialects of Tibetan*, ed. by Jackson T.-S. Sun.

- Curnow, Timothy Jowan. 2001. What languages features can be ‘borrowed’? *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 412-436. Oxford & New York: Oxford University Press.
- Dagan, Tal, and William Martin. 2006. The tree of one percent. *Genome Biology* 7.10:118.1-118.7.
- Dench, Alan. 2001. Descent and diffusion: the complexity of the Pilbara situation. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 105-133. Oxford & New York: Oxford University Press.
- Dixon, R. M. W. 1997. *The Rise and Fall of Languages*. Cambridge & New York: Cambridge University Press.
- Doolittle, W. Ford. 1999. Phylogenetic classification and the universal tree. *Science* 284: 2124-2128.
- Driem, George van. 2004. Language as organism: a brief introduction to the Leiden theory of language evolution. *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on his Seventieth Birthday*, ed. by Ying-chin Lin, Fang-min Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang & Dah-an Ho, 1-9. Taipei: Institute of Linguistics, Academia Sinica.
- Driem, George van. 2008. The language organism: parasite or mutualist? *Evidence and Counter-evidence: Essays in Honour of Frederik Kortlandt*, Vol. 2, ed. by Alexander Lubostky, Jos Schaeken & Jeroen Wiendenhof, 101-112. Amsterdam & New York: Rodopi.
- Dryer, Matthew. 2006. Descriptive theories, explanatory theories, and basic linguistic theory. *Catching Language: Issues in Grammar Writing*, ed. by Felix Ameka, Alan Dench & Nicholas Evans, 207-234. Berlin & New York: Mouton de Gruyter.
- Gesang, Jumian (skal bzang 'gyur med). 1964. *Zangyu Fangyan Gaiyao / Bod kyi yul skad gces bsds* [Essentials of Tibetan Dialects]. Beijing: Minzu University of China.
- Gilmour, J. S. L. 1940. Taxonomy and philosophy. *The New Systematics*, ed. by Julian Huxley, 461-474. Oxford: Oxford University Press.
- Handel, Zev. 2008. What is Sino-Tibetan? Snapshot of a field and a language family in flux. *Language and Linguistics Compass* 2.3:422-441.
- Harrison, Shelly P. 2003. On the limits of the comparative method. *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph & Richard D. Janda, 213-243. Malden: Blackwell.
- Hu, Tan. 1991. Zangyu [The Tibetan language]. *Han-Zangyu Gailun [An Introduction to Sino-Tibetan Languages]*, Vol. 2, ed. by Xueliang Ma, 109-207. Beijing: Peking University Press.

- Huang, Bufan. 1991. Qiangyuzhi [Qiangic languages]. *Han-Zangyu Gailun [An Introduction to Sino-Tibetan Languages]*, Vol. 2, ed. by Xueliang Ma, 208-369. Beijing: Peking University Press.
- Huang, Bufan. 1997. Tongyuanci bijiao cibiao de xuanci fanwei he biao zhun—yi Zang-Mianyu tongyuanci bijiao cibiao de zhiding wei li [On the scope and criteria of word selection for comparative cognacy lists, exemplified by a draft list of cognate words in Tibeto-Burman languages]. *Minzu Yuwen [Minority Languages of China]* 1997.4:10-16.
- Huang, Bufan, Suonan Jiangcai, and Minghui Zhang. 1994. Yushu Zangyu de yuyin tedian he lishi yanbian guilü [Characteristics in Yushu Tibetan phonology and its rules of historical change]. *Tibetology in China* 1994.2:111-134.
- Huang, Bufan, and Minghui Zhang. 1995. Baimahua zhishu wenti yanjiu [A study of the genetic affiliation of Baima]. *Tibetology in China* 1995.2:79-118.
- Huxley, Julian. (ed.) 1940. *The New Systematics*. Oxford: Oxford University Press.
- Jardine, Nicholas, and Robin Sibson. 1971. *Mathematical Taxonomy*. London: John Wiley & Sons.
- Jin, Peng. 1983. *Zangyu Jianzhi [Outline of the Tibetan Language]*. Beijing: Nationalities Press.
- Janda, Richard D., and Brian D. Joseph. 2003. On language, change, and language change — or, of history, linguistics, and historical linguistics. *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph & Richard D. Janda, 3-180. Malden: Blackwell.
- Kortlandt, Frederik H. H. 1985. A parasitological view of non-constructive sets. *Studia linguistica diachronica et synchronica: Werner Winter, sexagenario anno MCMLXXXIII*, ed. by Ursula Pieper & Gerhard Stickel, 477-483. Berlin: Mouton de Gruyter.
- Kortlandt, Frederik H. H. 2003. The origin and nature of the linguistic parasite. *Language in Time and Space: A Festschrift for Werner Winter on the Occasion of his 80th Birthday*, ed. by Brigitte L. M. Bauer & Georges-Jean Pinault, 241-244. Berlin & New York: Mouton de Gruyter.
- Lama, Ziwo. 1994. Namuyiyu zhishu yanjiu [On the genetic affiliation of the Namuyi language]. *Minzu Yuwen [Minority Languages of China]* 1994.1:50-60.
- Lawton, Graham. 2009. Why Darwin was wrong about the tree of life. *New Scientist* 2692:34-39.
- Li, Fang-Kuei. 1937. Chinese languages and dialects. *The Chinese Yearbook 1937*, 59-65. Shanghai: The Commercial Press. Reprinted in *Journal of Chinese Linguistics* 1.1:1-13, 1973.

- Matisoff, James A. 2001. Genetic versus contact relationship: prosodic diffusibility in South-East Asian languages. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 291-327. Oxford & New York: Oxford University Press.
- Mayr, Ernst. 1976. *Evolution and the Diversity of Life: Selected Essays*. Cambridge: Belknap Press of Harvard University Press.
- Mayr, Ernst. 1988. *Toward a New Philosophy of Biology: Observations of an Evolutionist*. Cambridge: Belknap Press of Harvard University Press.
- Michener, Charles D., and Robert R. Sokal. 1957. A quantitative approach to a problem in classification. *Evolution* 11:130-162.
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge & New York: Cambridge University Press.
- Nichols, Johanna. 1996. The comparative method as heuristic. *The Comparative Method Revisited: Regularity and Irregularity in Language Change*, ed. by Mark Durie & Malcolm Ross, 39-71. Oxford & New York: Oxford University Press.
- Noonan, Michael. 2008. Genetic classification and language contact. Heidelberg: Bibliothek des Südasien-Instituts der Universität Heidelberg. Online-Ressource. Electronic publications by Michael Noonan; 23: <http://archiv.ub.uni-heidelberg.de/savifadok/volltexte/2008/215>.
- Norman, Jerry. 1988. *Chinese*. Cambridge & New York: Cambridge University Press.
- Peyraube, Alain. 2005. L'homme aux six mille langues. *Aux origines des langues et du langage*, ed. by Jean-Marie Hombert, 450-461. Paris: Editions Fayard.
- Peyraube, Alain. 2007. Languages and genes in China and in East Asia. *Bulletin of Chinese Linguistics* 2.1:1-15.
- Pigliucci, Massimo, and Jonathan Kaplan. 2006. *Making Sense of Evolution: The Conceptual Foundations of Evolutionary Biology*. Chicago: University of Chicago Press.
- Popper, Karl R. 1969. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge & Kegan Paul. First published in 1963.
- Pulleyblank, Edwin G. 1991. Chinese dialect studies. *Languages and Dialects of China*, ed. by William S-Y. Wang, 431-456. Berkeley: Project on Linguistic Analysis.
- Pulleyblank, Edwin G. 1998. *Qieyun* and *Yunjing*: the essential foundation for Chinese historical linguistics. *The Journal of the American Oriental Society* 118.2:200-216.
- Qu, Aitang, and Xiaojing Jin. 1981. Zangyu fangyan de yanjiu fangfa [Methods in the study of the Tibetan dialects]. *Journal of Southwest University for Nationalities* 1981.3:76-83.
- Ruhlen, Merritt. 1994. *On the Origins of Languages: Studies in Linguistic Taxonomy*. Stanford: Stanford University Press.

- Sagart, Laurent. 1998. On distinguishing Hakka and non-Hakka dialects. *Journal of Chinese Linguistics* 26.2:281-302.
- Salthe, Stanley N. 1972. *Evolutionary Biology*. New York: Holt, Rinehart & Winston.
- Simpson, George Gaylord. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.
- Sokal, Robert R., and Peter H. A. Sneath. 1963. *Principles of Numerical Taxonomy*. San Francisco: W. H. Freeman & Company.
- Sun, Hongkai. 1980. Baimaren de yuyan [The language of the Baima]. *Baima Zangren Zushu Wenti Taolunji* [A Collection of Discussions on the Question of the Ethnic Classification of the Baima Tibetans], ed. by Sichuan Nationalities Research Institute, 15-26. Chengdu: Sichuan Nationalities Research Institute.
- Sun, Hongkai. 1983. Liujiang liuyu de minzu yuyan ji qi xishu fenlei [Minority languages of the six river valley and their genetic classification]. *Minzu Xuebao* [Scholarly Journal of Nationalities] 3:99-273.
- Sun, Hongkai. 2001. Lun Zang-Mian yuzu zhong de Qiangyuzhi yuyan [On the Qiangic branch of the Tibet-Burman language family]. *Language and Linguistics* 2.1:157-181.
- Sun, Jackson T.-S. 1992. Review of *Zang-Mianyu Yuyin he Cihui* [Tibeto-Burman Phonology and Lexicon]. *Linguistics of the Tibeto-Burman Area* 15.2:73-113.
- Sun, Jackson T.-S. 2003a. Phonological profile of Zhongu: a new Tibetan dialect of northern Sichuan. *Language and Linguistics* 4.4:769-836.
- Sun, Jackson T.-S. 2003b. Qiuji Zangyu de yuyin tezheng [Phonological profile of Chos-rje Tibetan]. *Minzu Yuwen* [Minority Languages of China] 2003.6:1-6.
- Suzuki, Hiroyuki. 2008. Nouveau regard sur les dialectes tibétains à l'est d'Aba: Phonétique et classification du dialecte de Shangkhog [Songpan-Jiuzhaigou]. *Linguistics of the Tibeto-Burman Area* 31.1:85-108.
- Thomason, Sarah Grey. 2000. On the unpredictability of contact effects. *Estudios de Sociolinguística* 1.1:173-182.
- Thomason, Sarah Grey, and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Wurm, S. A. et al. 1988. *Language Atlas of China*. Hong Kong: Longman.
- Yuan, Jiahua et al. 2001. *Hanyu Fangyan Gaiyao* [Outline of Chinese Dialects]. Beijing: Language and Culture Press.
- Zhang, Jichuan. 1993. Zangyu fangyan fenlei guanjian [Opinions on the classification of Tibetan dialects]. *Minzu Yuwen Lunwenji* [Papers on Minority Languages], ed. by Qingxia Dai, Meizhen Luo & Yingxin Yang, 297-309. Beijing: Minzu University of China Press.

- Zhang, Jichuan. 1996. A sketch of Tibetan dialectology in China: classifications of Tibetan dialects. *Cahiers de Linguistique – Asie Orientale* 25.2:113-144.
- Zhang, Jichuan. 1997. Particularités phonétiques du Baima. *Cahiers de Linguistique – Asie Orientale* 26.1:131-153.

The Classification of Sinitic Languages: What Is “Chinese”?*

Victor H. Mair (梅維恒)
University of Pennsylvania

Chinese as the name of a language is a misnomer. It has been applied to numerous dialects, styles, and languages from the middle of the 2nd millennium BC. Sinitic is a more satisfactory designation for covering all these entities.... The present-day spoken languages are not mutually intelligible (some are further apart than Portuguese and Italian), and neither are the major subdivisions within each group. Søren Egerod (1988)

I think we should refer to the complex of linguistic forms now called Chinese “dialects” as the Sinitic language family. Jerry Norman (2009a)

If one’s terms are imprecise, one’s language will be illogical. 名不正, 言不順. Analects, 13.3

Terminological imprecision, particularly with regard to the Chinese word *fangyan* and its translation by the English word “dialect,” has resulted in a situation whereby Sinitic language taxonomy may variously be described as chaotic, impenetrable, or functionally absent. For such a large, diverse agglomeration of languages as Sinitic, this is an unacceptable state of affairs. Through rigorous definition and careful analysis, it is possible to arrive at a clearer understanding of the nature of the relationships among the constituent languages of the Sinitic Language Group/Family (SLG/F).

Key words: Sinitic languages, languages versus dialects, topolects, *Putonghua*

* For assistance in researching various aspects of this study, I am deeply grateful to Jerry Norman, W. South Coblin, Tsu-Lin Mei, Wolfgang Behr, William H. Baxter, Guillaume Jacques, Bai Weiguo, Huang Heqing, Xu Wenkan, Yao Dehuai, Sanping Chen, Jidong Yang, David Prager Branner, Thomas S. Mullaney, Mark Elliott, Robert Lindsay, Linda Chance, Alban Kojima, S. Robert Ramsey, William C. Hannas, Neil Kuster, Minkyung Ji, and Daniel Sou. The author alone, however, is accountable for all facts and interpretations presented herein.

1. Preliminaries

The present study is an extension of the author's "What is a Chinese 'dialect/topolect'? Reflections on some key Sino-English linguistic terms" (Mair 1991). This is an updating of that paper and should be read in conjunction with it. We may summarize the major findings of that article as follows:

i. Loose usage of terms such as *fangyan* 方言 (universally, but mistakenly, rendered as 'dialect' in English), Chinese, and *Hanyu* 漢語 'Han language' has drastic implications for scientific, linguistic classification.

Corollary: Only with precise terminology can we have accurate, exacting classification.

ii. A more precise translation of *fangyan*, one that effectively neutralizes its deleterious consequences for the scientific classification of the languages of East Asia, is 'topolect.' This is a word which means exactly the same thing as *fangyan*, no more and no less, and which has been accepted by the *American Heritage Dictionary of the English Language* and other authorities.

iii. The number of *fangyan* 'topolects' and their taxonomic status is uncertain.

iv. The number of so-called *da fangyan* 大方言 ("major topolects") and their taxonomic status is uncertain.

v. The demographically huge aggregation of speech forms loosely and collectively known as *Hanyu*¹ 'Han [People's] Languages[s]' is fully deserving of rigorous classification comparable to that which has been applied to Indo-European, Tibeto-Burman, Afro-Asiatic (Semitic-Hamitic), Amerindian, Austronesian, Austro-Asiatic, and virtually all the other major language stocks on earth. It is inconceivable that *Hanyu* has been and is taxonomically undifferentiable throughout time and space. Yet the most authoritative reference works in China, such as the *Yuyan wenzi* 語言文字 [Language and Script] volume of the *Zhongguo da baike quanshu* 中國大百科全書 [Great Encyclopedia of China], fail to divide *Hanyu* into constituent subdivisions, treating it as a single entity. This is in stark contrast to their treatment of much smaller groups such as Tibeto-Burman, Miao-Yao, and Zhuang-Dong, which they divide into numerous branches, sub-branches, and languages.

¹ To avoid an excess of what might be interpreted as scare quotes, all occurrences of '*Hanyu*' in this paper should be understood as prefaced by 'so-called.' The reason for this is that, in its current undifferentiated, unanalyzed state, '*Hanyu*' is nugatory in discussions of the scientific classification of Sinitic languages. Due to similar developmental and historical reasons, the same holds for '*Han*,' '*Hanzu*,' and '*Han minzu*.' See §3.

vi. Ideally, linguists should strive for a logical division of *Hanyu* into branches, sub-branches, languages, dialects, and sub-dialects, just as they do for all other major language groupings.

vii. Traditional understandings of *fangyan*, which are still operative in contemporary discussions of the nature of *Hanyu*, are incompatible with modern, linguistic classification.

viii. ‘Mandarin’ is an accurate and appropriate English translation of *guanhua* 官話, the forerunner of *guoyu* 國語 ‘National Language’ as used in Taiwan,² *huayu* ‘Hua [People’s] Language’ as used in Singapore, and *Putonghua*, ‘Common Speech’ as used in People’s Republic of China.

ix. ‘Mandarin’ is not equivalent to ‘Chinese,’ whatever the latter may signify.

x. The fact that there are approximately two dozen different ways to refer to Mandarin in Chinese, including several that end in *wen* 文 ‘writing’³ is a clear indication of the nomenclatural imbroglio surrounding this language (or congeries of languages and dialects).

xi. The various speech forms of *Hanyu* significantly differ among themselves in respect to phonology, morphology, lexicon, idiomatic usage, grammar, syntax, and all other aspects of linguistic composition and construction.

xii. *Hanyu* simultaneously means two very different things: Modern Standard Mandarin (MSM) and the totality of all languages spoken by the so-called Han peoples. Because this multivalence of *Hanyu* has not hitherto been sufficiently emphasized, much less recognized, it leads to subconscious confusion concerning the interrelationships among the components of the Sinitic Language Group/Family (SLG/F).⁴ Cf. §3.1.

xiii. Cantonese and Mandarin are separate languages. Cantonese is not a ‘dialect’ of Mandarin or of *Hanyu*, and it is grossly erroneous to refer to it as such. Since Cantonese and Mandarin are separate languages (or, perhaps more accurately, separate branches), it is wrong to refer to them as ‘dialects.’ The same holds for Hokkien, Shanghainese, and so forth.⁵

² For a discussion of the history of the term *guoyu*, see Mair (1994).

³ For a brilliant, but little known, disquisition on the history of the relationship between writing and language in the EAH, see Lu Xun (1934, reprint 2005).

⁴ If efforts to link Sinitic with other major language groups continue to be as unconvincing as they have been to date, it may well be that Sinitic will end up being classified as a family unto itself. Because it remains to be determined whether Sinitic is a group or a family, I provisionally style it the Sinitic Language Group/Family (SLG/F).

⁵ N.B.: Mair (1991) and other works by the author cited in the Bibliography serve to document diverse unreferenced statements herein. See also Norman (1988), Ramsey (1987), DeFrancis (1984), and Zhou (2003).

2. Presuppositions

i. Language groups are classifiable, both internally and externally (Lass 1997, Campbell 1997).

ii. A lect⁶ is a discrete, distinct form of speech that is identified with a particular population, whether such a population be determined by place, ethnicity, occupation, or other criterion.

iii. There are countless lects in China, many of which are completely or partially unintelligible to each other.

iv. Mutual (un)intelligibility⁷ is determinable and quantifiable (Casad 1974, Cheng 1982, 1991, 1998, and Tang & van Heuven 2009).

v. Mutual (un)intelligibility is significant and should be taken into consideration in the classification of the SLG/F.

vi. Phonological, morphological, lexical, and even syntactical criteria may also be used for genetic classification, with phonological criteria being particularly relevant in the case of Sinitic (Coblin 2009).

3. Problems

i. There are more than two dozen synonyms for the national language of China (Mair 1991:11). The two most frequently encountered nowadays are *Hanyu* and *Putonghua*. What is the relationship between *Hanyu* and *Putonghua*? Unmistakably, *Putonghua* constitutes a subset of *Hanyu*, since the former term was adopted by the government of the People's Republic of China government after 1950 as a designation for the contemporary lingua franca of the nation. That is to say, *Putonghua* is a particular type of *Hanyu*; it is the 'common speech' of the nation in contrast to all of the other forms of *Hanyu* which are considered to interfere with or inhibit communication among the Han

⁶ Although 'lect' usually refers to social or regional varieties within a speech community, I accept this convenient usage of SIL because it is the most neutral designation I can think of for a separately identifiable pattern or form of speech, even more so than topolect. SIL (formerly known as the Summer Institute of Languages), which runs the influential, authoritative online list of languages entitled *Ethnologue*, has been charged with the task of assigning ISO numbers to lects. The assignment of an ISO number to a lect means that it has been officially recognized by the world linguistic community as a separate language.

⁷ The reason for the formulation 'mutual (un)intelligibility' is that 'intelligibility' and 'unintelligibility' correspond to each other in an obverse fashion. For example, if lect A is considered to be 40% intelligible to speakers of lect B, it may also said to be 60% unintelligible to speakers of lect B.

peoples. In contrast to *Putonghua*, *Hanyu* is recognized diachronically not only as having a much lengthier reach backward in time (supposedly to the very beginning of the SLG/F), but also synchronically beyond the limits of the lingua franca to embrace all the local and regional varieties of related speech forms that are more or less mutually unintelligible. Indeed, the very *raison d'être* for *Putonghua* is to serve as a means of communication for speakers of various lects of *Hanyu* who would otherwise not be able to talk to each other unconstrainedly (or at all). But *Hanyu*, in the sense of a comprehensive entity that includes all forms of language presumably spoken by individuals of Han ethnicity, throughout the time of existence of the Han people, is a modern construct.

ii. *Hanyu*, as employed in current cultural and political discourse, is a thoroughly modern, nationalist concept. Fundamentally, *Hanyu* signifies the language(s) spoken by the *Hanzu* 漢族 ‘people[s] of Han ethnicity’, while *Hanzu* is short for *Han minzu* 漢民族 ‘the Han nation/people/ethnos’. Both *Hanzu* and *minzu*—with their current signification of nationhood—are expressions that cannot be found before around the turn of the 20th century and are the result of nationalistic discourse that came to East Asia from the West. More specifically, they are likely to have been calqued in Japan or under Japanese influence to cope with new concepts concerning nationhood and ethnicity (Wilkinson 2000:708, Fogel 2005, Doak 2007, Huang 2009a, Xu 2009, and Yao 2009).

The hallowed appellation ‘Han’ was originally the self-designation for two celebrated dynasties: the Former/Western Han (202 BC-AD 23) and the Later/Eastern Han (25-220), separated by Wang Mang’s Xin (‘New’) dynasty (9-23). Naturally, the subjects of the two Han dynasties were contemporaneously referred to as *hanren* ‘Han person[s]’, but, with the collapse of the Eastern Han, such terms fell out of use. In the whole of the massive *Zhu Zi yulei* 朱子語類 [Classified Conversations of Master Zhu], the leading Neo-Confucian scholar of his age, Zhu Xi (朱熹; 1130-1200) employs the term ‘Han person(s)’ twenty-seven times, but in each case he means ‘people of the Han Dynasty’, not an ethnically distinct group (Zhao 2009). This is particularly revealing in light of the fact that *Zhu Zi Yulei* has a considerable proportion of vernacular passages. Although the name Han was sporadically resurrected by several ephemeral political entities in the centuries that followed, the last dynasty to call itself Han fell in 1364, after less than four years of existence (Beck 1986:371-373).

As a matter of fact, in the centuries from the demise of the Eastern Han up to modern times, terms such as *han* 漢, *hanzi* 漢字, *hanren* 漢人, and *han'er* 漢兒 were used (often derogatorily⁸) by the northern peoples (e.g., Särbi, Khitans, Jurchens,

⁸ The derogatory connotations of ‘Han’ have continued into modern times. Amazingly, when not used in its neologistical ethnic sense, *han* in contemporary usage frequently conveys a pejorative sense, as in the following common expressions: *chunhan* 蠢漢 ‘dull person’, *benhan* 笨漢 ‘stupid man’, *shahan* 傻漢 ‘foolish man’, *lanhan* 懶漢 ‘lazy man’, *wulaihan* 無賴漢 ‘rascal’,

Mongols) who established control over all or large parts of what is now called ‘China’ to refer to their indigenous subjects. Hence, in *Bei shi* 北史 [History of the Northern Dynasties], we find such expressions as *gouhan* 狗漢 ‘dog Han’ (92.3053; also in *Bei Qi shu* 北齊書 [History of the Northern Qi] 50.692-693) and *kongtouhan* 空頭漢 ‘empty-headed Han’ (54.1967). An especially lackluster whitewashing of such a deprecatory term by a later editor occurs in the basic annals of emperor Wenxuan 文宣 (r. 550-559) in *History of the Northern Dynasties* and in *History of the Northern Qi* where the emperor refers to an official that he despises as a *chihan* 痴漢 ‘silly Han’. Sima Guang 司馬光 (1019-1086), in *Zizhi tongjian* 資治通鑑 (Comprehensive Mirror for Aid in Government), 166.5149-5150 changes this to *chiren* 痴人 ‘silly person’. The renowned synthesizer’s systematic, stubborn bowdlerization of any references to *han* in the official histories is almost laughably documented at 190.5959 of the *Comprehensive Mirror*, which has the Tang emperor Gao Zu declare, “This son has long been occupied with military affairs outside and was taught by a bookish student. He’s no longer the lad of old.” 此兒久典兵在外, 爲書生所教, 非復昔日子也. In *Jiu Tang shu* 舊唐書 [Old Tang History] 64.2415-2416, this was: “This son has already been exclusively occupied with military affairs outside for a long time and was taught by a book-reading Han. He’s no longer my lad of old.” 此兒典兵既久, 在外專制, 爲讀書漢所教, 非復我昔日子也. In *Xin Tang shu* 新唐書 [New Tang History] 79.3540, the equivalent passage reads thus: “[This] son has long been occupied with military affairs and was corrupted by a Confucian student. He’s no longer my son of old.” 兒久典兵, 爲儒生所誤, 非復我昔日子. The lengths to which the Confucian historian went to avoid mention of any sort of Han becomes truly comical at 194.6096 which relates an expostulation of the Tang emperor Tai Zong: “When the emperor had departed from the court, he angrily said, ‘It may be necessary to kill this old country bumpkin.’” 上嘗罷朝, 怒曰: 會須殺此田舍翁. Compare this to an event as recorded in *Da Tang xinyu* 大唐新語 [New Accounts of the Great Tang] 1.13: “When the emperor had departed from the court, he said to himself, ‘I’ll have to kill off this yokel-Han.’” 太宗嘗罷朝, 自言: 殺卻此田舍漢. The parallel passage in *Sui Tang jiahua* 隋唐佳話 [Fine Tales from the Sui and Tang] A7 has:

manghan 莽漢 ‘boorish fellow’, *menwaihan* 門外漢 ‘outsider’, *liulanghan* 流浪漢 ‘wastrel’, and *zuihan* 醉漢 ‘drunkard’. Positive expressions incorporating *han*, such as *haohan* 好漢 ‘brave man’, *tiehan* 鐵漢 ‘strong man’, *yinghan* 硬漢 ‘man of fortitude’, and *nanzihan* 男子漢 ‘real man’, are encountered much less frequently than those with negative connotations, have historically often been ironically employed by non-Sinitic overlords with reference to their autochthonous subjects, and usually are related to macho or outlaw characteristics. In fact, so accustomed are Chinese to referring to disreputable persons as ‘such-and-such a *han*’ that *-han* has become a standard suffix in the media for a male person with negative associations, e.g., *bulianghan* 不良漢 ‘good for nothing guy’, *jiejihan* 劫機漢 ‘airplane hijacker’, etc. (Chen 2008a).

“Having departed from the court, Taizong angrily said, ‘It may be necessary to kill this yokel-Han.’”⁹ 太宗曾罷朝，怒曰：會須殺此田舍漢。Sima Guang’s allergy to any expression involving *han* in the mouth of someone with northern associations even caused him to nullify a famous, colorful statement of Empress Wu Zetian 武則天 (625-705). Admittedly with a degree of arch ambiguity, the empress had said, “I need a good Han. Are there any?” 朕要一好漢，有乎？(*New Accounts of the Great Tang* 6.94) Sima Guang’s wordy rendering of the Empress’s inquiry is innocuous and awkward: “I would like to obtain an excellent scholar-official that I can use. Who can do it?” 朕欲得一佳士用之，誰可者？(*Comprehensive Mirror* 207.6551, first year of the Jiushi [Lasting Vision] reign period 700). Also in *Fine Tales from the Sui and Tang* B33, we find the Tang emperor Gao Zong (r. 650-683) referring to a harsh official who has caned a favored but miscreant eunuch as being “this savage Han” 此漢獐。In Feng Yan’s 封演 (*jinshi* 756) *Feng shi wenjian ji* 封氏聞見記 [Things Seen and Heard by Mr. Feng], 10.88, 10.93, a certain Wang Yanguang 王嚴光 declares, “Take this [ignorant] Han to use as turtle bait!” 取此等[蒙]漢以充鰲餌。(Wang Yanguang is also identified as a turtle catcher in *Lei shuo* 類說 [Classified Conversations], 1.6.26a.) And in Zhuang Chuo’s 莊綽 12th century *Ji lei bian* 雞肋編 [Chicken Rib Chapters], B70, a Song emperor derides a fish merchant with the words, “This Han is malicious [literally, ‘toxic’]!” 這漢毒也。In all of these instances, ‘Han’ may be interpreted as ‘fellow’, or worse. At best, toward the end of the Six Dynasties and continuing through the early Tang (while northern influences at court were still very strong), *Han* implied an educated, effete person, including Sinitically acculturated northerners (Chen 2008b). In Zen/Chan texts, *Han* occurs regularly as a colloquial term meaning ‘fellow’ (Broughton 1999:146).

In the “Benji” 本紀 [Basic Annals] of Shi Zu 世祖 of *Yuan shi* 原史 [History of the Yuan], 13.268, Sinitic-speaking Khitan and Jurchen individuals living in certain areas were legally classified as *hanren*. This shows clearly that *han* in the Mongol period was not thought of as an ethnic category in the same way it is today. The ethnolinguistic indeterminacy of *han* in pre-modern times is also borne out by a passage in the *Taiping huanyu ji* 太平寰宇記 [Universal Geography of the Taiping Era [976-983]], 35.11a (Wenhai chubanshe reprint (this passage is inexplicably missing from some editions, even in the *Siku quanshu* 四庫全書 [Complete Library of the Four Treasuries])), citing a Sui period (581-618) text, that refers to the Buluoji 步落稽 (associated by some with the Bulgars) as *hutou hanshe* 胡頭漢舌 ‘barbarian-headed Han-tongued’.

⁹ It is remarkable how many different vernacular suffixes were available in medieval times for addition to the old classical expression *tianshe* 田舍 (literally, “fields and cottages”), ranging all the way from *gong* 公 and *lang* 郎 to *ren* 人, *zi* 子, *er* 兒, *weng* 翁, *niang* 娘 (N.B.), *nu* 奴, and *han* 漢, with Han evidently being toward the derisive end of the scale.

Before its modern reincarnation as the language of the ethnic Han nation, *Hanyu* had an earlier life as a designation for the language of the people of the Han Dynasty as shown by a search of large data bases such as CHANT (Chinese Ancient Texts) and CBETA (Chinese Buddhist Electronic Text Association). A CBETA search reveals that the term *Hanyu* did occur during the Six Dynasties, but primarily in contrast to *fanyan/you/wen* 梵言/語/文 (Sanskrit), *huyan/you* 胡言/語 ‘West Central Asian language’, etc. to describe the translation of Buddhist scriptures from Indic languages or the familiarization of foreign monks and translators with Sinitic languages. Moreover, terms such as *jinyu* 晉語 ‘language of the Jin people’ occur in similar contexts to distinguish Sinitic translation from Indic original. In other words, *Hanyu* in this early period meant ‘language of the Han Dynasty’ just as *Jinyu* meant ‘language of the Jin Dynasty’. In some instances, *Hanyu* meant ‘tales/stories/accounts of the Han period’, similar to the usage of in the title of the Warring States period work, *Guoyu* 國語 [Discourses of the States {of the Spring and Autumn Period}] or in the Liu Song period work entitled *Shishuo xinyu* 世說新語 (New Account of Tales of An Era) by Liu Yiqing 劉義慶 (403-444) concerning personalities who lived in China between about AD 150 and 420.

The term *Hanzu* did occur rarely before the end of the 19th century, but not with its modern meaning of ‘Han nation/people/ethnos’. The earliest occurrence of the term *Hanzu* is in an ironic context, namely, the memorial epitaph of the Son (Zen/Chan) Buddhist master Jin’gam (744-850), written by the famous Silla poet, Choe Chiwon 崔致遠 (857-928) upon the request of the Silla king Heongang: “Yu-Dang Silla’guk Go Jirisan Ssang’gyesa Gyosi Jin’gam Seonsa Bi’myeong” 有唐新羅國故知異山雙谿寺教諡眞鑑禪師碑銘 (Stele Inscription for the Late Meditation [Son] Master Whose Posthumous Name Is Jin’gam, of the Ssang’gye Temple on Jirisan in the Tang Tributary State of Silla).¹⁰ The epitaph is preserved at the Ssang’gye temple as Korean National Treasure No. 47. The sentence in question occurs near the beginning of the inscription and states that Jingam’s ‘ancestors were *hanzu*.’ 其先漢族. Choe Chiwon (MSM Cui Zhiyuan) was a Korean, he arrived in China at the age of 12 and passed the *jinshi* (advanced scholar) examination in 874. Several things about this early, isolated occurrence of the term *hanzu* are worth noting: a. it was used by a Korean with reference to another Korean, b. it was not taken up by other writers in the following centuries as a fixed term, and c. it was not intended as a blanket designation for Sinitic ethnicity. After that single medieval occurrence in a most intriguing context, *hanzu* is not in evidence again until near the end of the Qing dynasty when it comes to be employed by ethnically conscious Chinese, first in contrast to the Manchu overlords,

¹⁰ Ssanggyesa, located on the southern slopes of Jirisan in Hadong County of Gyeongsangnam-do, in the southeastern part of South Korea, is the head temple of the Jogye Order of Korean Buddhism.

and then by nationalists like Sun Yat-sen in contrast to the other peoples of East Asia and Southeast Asia (Mengguzu 蒙古族 [Mongols], Malaizu 馬來族 [Malays], Ribenzu 日本族 [Japanese], and Manzu 滿族 [Manchus]).

The rarity of ‘Han’ as an umbrella designation for an East Asian state, people, or language after the third century is borne out by the fact that—in pre-modern times—foreign peoples referred to what we now know as “China” by such terms as Cina, Kitai, Seres, Tamghaj, Thin, and so forth, not anything resembling ‘Han’. There is no premodern citation for Hanzu in the *Hanyu da cidian* (HDC).¹¹ Considering the supreme importance and conspicuous frequency of the term in modern political rhetoric from the PRC, this could hardly be an accident.

iii. The relationships among *Hanyu*, *Hanzu*, and *Putonghua* are neatly—but somewhat perplexingly—summarized in the following formulation by Wang Li 王力 (1900-1986), an ardent proponent of *Putonghua*: ‘*Putonghua* is the standard language of Modern *Hanyu*; it is the common language of the *Han minzu*.’¹² 普通話是現代漢語的標準語，是漢民族共同語 (HDC 5.777a). From around this time, *Hanyu* and *Putonghua* came to replace *Guoyu* ‘National Language’ on the mainland, but were not exactly distinguished, one from the other, nor were they precisely differentiated from the innumerable more or less mutually unintelligible Sinitic lects spoken throughout China.

iv. As with *Hanyu* and *Hanzu*, similar observations may be made concerning the relatively recent appearance of *hanzi* in the sense of ‘sinogram’ or ‘Chinese character’. It is noteworthy that there are no premodern citations for *hanzi* in the HDC. The earliest known occurrences (1610, 1726) of *hanzi* are in the works of Matteo Ricci and his Jesuit successors in reference to Europeans familiarizing themselves with Sinitic languages and the Chinese script (Huang forthcoming). It is both curious and telling that these latter day occurrences of *Hanyu* and the apparently new coinage of *hanzi* are to be found in the context of foreigners becoming acquainted with Sinitic languages and the Chinese script. This parallels exactly what had occurred many centuries earlier when Indian and

¹¹ The abbreviations used in this paper are:

EAH: East Asian Heartland

EEAH: Extended East Asian Heartland

HDC: *Hanyu Da Cidian* 漢語大詞典 [Unabridged Word Dictionary of Sinitic]

MSM: Modern Standard Mandarin

¹² This statement is from Wang Li’s ‘Tuiguang Putonghua de san ge wenti’ 推廣普通話的三個問題 [Three Questions Concerning the Promotion of Putonghua]. Lu Xun (1881-1936), the greatest writer of the 20th century, had earlier used the expression *Putonghua*, but not as a proper noun, merely in the sense of “common language” (HDC 5.777a). It was only after the middle of the century that Putonghua came to be the official designation for the national language of the People’s Republic of China.

Central Asian monks and translators arrived to live and work in East Asia.

v. In premodern times, Sinitic speakers normally referred to their language simply by such terms as *yan* 言, *yu* 語, *hua* 話, and *wen* 文.¹³ There rarely was a need to specify it as their own language in contrast to the language of some alien people who spoke a completely different language. It was only in modern times, when the tempo and quantity of linguistic contacts with peoples from abroad (especially those coming from a great distance) increased dramatically, that Sinitic speakers were obliged to clearly identify their own languages as distinct from the languages of the foreigners.

vi. Until the rise of modern nationalistic consciousness, it was usually foreign speakers who felt the need to cap *yan*, *yu*, *hua*, and *wen* with a modifier to distinguish them as different from their own languages. Such, apparently, was the case with the earliest occurrences of *hanhua* 漢話, *zhongyu* 中語, *zhongguoyu* 中國語, *zhongguohua* 中國話, *tangyu* 唐語, *tangrenhua* 唐人話, *qingyu* 清語, *zhinayu* 支那語, *huayu* 華語, and so forth (Huang 2009b).

vii. All of the problems discussed in this section will become much clearer when the papers of the pathbreaking “Critical Han Studies Conference and Workshop” held at Stanford University from April 25-27, 2008 are published.¹⁴

viii. It has been proposed that *Hanyu* be straightforwardly rendered into English as Hanic or Hannic (Aymat 2009). For all of the difficulties alluded to in this section, however, it would seem prudent to avoid this very recent appellation for the entire language group/family in favor of Sinitic, which has the advantage of being a neutral designation that is easily combined with other group names if necessary (Sino-).

¹³ In periods of intense interaction between Sinitic and non-Sinitic speakers, already in antiquity distinguishing modifiers might be added to self-referential attributes. For example, when the Xiongnu (Huns) impinged heavily upon the EAH, terms such as *Qinren* 秦人 ‘Qin person/people’ and *qinyu* 秦語 ‘Qin language’ arose. Such terms, however, typically did not last much beyond the time of most intense interaction. It was only in modern times, when massive, sustained, cultural encroachment was operative, that the need was felt for a distinguishing set of autonomous terms.

¹⁴ Among the most important papers delivered at the conference and discussed at the workshop is Mark Elliott’s revelatory “*Hushuo* 胡說: The Northern Other and Han Ethnogenesis,” which independently corroborates many of the points made in this paper. The title of Elliott’s paper speaks for itself.

4. Protocols

i. With regard to linguistic classification, “Chinese” contradictorily means at least two quite different things:

- a. the national language of the People’s Republic of China (i.e., *Putonghua*)
- b. the entirety of all non-Tibeto-Burman, non-Austronesian, non-Austro-Asiatic, non-Hmong-Mien (Miao-Yao), non-Tai-Kadai (Kra-Dai), non-Indo-European, non-Turkic, non-Mongolic, non-Tungusic lects in China

ii. To avoid confusion, in this study “Chinese” (when used at all) is reserved for the modern national language of the People’s Republic of China (i.e. MSM), whereas “Sinitic” shall be used to refer to the mass of hitherto inadequately differentiated (yet presumably genetically related) lects specified in 4.i.b.

iii. Sinitic may be divided temporally (all ranges are rough approximations) into Modern Sinitic (MS, 13th c.-present), Middle Sinitic (MS, 3rd c.-12th c.), Old Sinitic (OS, 10th c. BC-2nd c. AD), and Archaic Sinitic (AS, 11th c. BC and before), with the following subdivisions: Early Modern Sinitic (EMoS, 13th c.-19th c.), Late Modern Sinitic (LMoS, 20th c.-present); Early Middle Sinitic (EMiS, 3rd c.-6th c.), Late Middle Sinitic (LMiS, 7th c.-12th c. AD); Early Old Sinitic (EOS, 10th c.-3rd c. BC), Late Old Sinitic (LOS, 2nd c. BC-2nd c. AD); Early Archaic Sinitic (EAS, 14th c. BC and before), Late Archaic Sinitic (LAS, 13th c.-11th c. BC).

iv. Our comprehension of the historical development of the regional subgroupings of Sinitic is still very limited, but eventually we should aim to divide them into Old, Middle, and Modern phases, as well as to determine the relationship of these stages to the overall development of the SLG/F as a whole.

v. In this study, “dialect” is divorced from *fangyan*; the English word and the Chinese word are not considered as coterminous.

5. Propositions

i. The SLG/F is not an undifferentiable monolith.

I adopt SLG/F as a neutral designation for the mass of genetically related lects specified in §4.i.b and remain agnostic about whether the SLG/F is actually a family unto itself or whether it is more or less closely linked to some other group(s)—such as Tibeto-Burman or Austronesian—in a family (see 5.ii). As a matter of fact, tremendous progress has been made in the analysis of the branches of the SLG/F by Jerry Norman, W. South Coblin, Anne Yue-Hashimoto, Richard Van Ness Simmons, David Prager Branner, Laurent Sagart, William H. Baxter, and others. Their work, unfortunately, is partially vitiated by inadequate terminology, since they refer to all cladistic levels of the SLG/F

as “dialects”, making it difficult to distinguish clearly the hierarchical relationships among the constituent branches, languages, and genuine dialects and sub-dialects of the group.

Corollary: The SLG/F is internally classifiable.

Caveat: There is no consensus on the internal classification of the SLG/F.

ii. The SLG/F is not isolated from all other language groups.

Proposals for linking the SLG to Tibetan, Tibeto-Burman, Tai, Austronesian, Austro-Asiatic, Caucasian, Na-Dene, and other families and groups have been made (Wang 1995, Handel 2008, Behr 2005, van Driem 2007, Sagart 1994, Sagart, Blench & Sanchez-Mazas 2005, Schuessler 2007, Thurgood 2008, and Thurgood & LaPolla 2003), but none of these proposals has been argued with sufficient rigor and supported by adequate data to gain broad acceptance. I maintain that, until the internal classification of the SLG/F is worked out, it is both premature and impractical to speculate on the external classification of the SLG/F. The fact that the proposed external genetic relationships for Sinitic vary so wildly is a good indication of how far we are from achieving a convincing solution to the problem of the affinities of the SLG/F. Whether the Sinitic languages turn out to constitute a family of their own or a group belonging to a larger family, they surely have some sort of relationships—be they genetic or contact—with languages in other families or groups, but it is futile to attempt to determine the external affinities of Sinitic until the tasks outlined in §4.iv are resolved.

Corollary: The SLG/F is externally classifiable.

Caveat: There is no consensus on the external classification of the SLG.

6. Proposals

i. The significance of mutual (un)intelligibility should not be ignored in discussions of the classification of the SLG/F.

Corollary: If two lects are completely or largely mutually unintelligible, it makes little sense to refer to them as dialects of each other or of a higher-level entity.

ii. The scientific classification of languages should not be held hostage to extra-linguistic political and cultural prejudices.

Corollary: Dogmatic assertions of identity or minor, insignificant variations are unacceptable in the face of a high degree of lexical, phonological, grammatical, and other types of dissimilarity.

iii. In the scientific classification of modern Sinitic languages, as much as possible, data should be drawn from the strikingly different spoken varieties, not from standard written forms. Writing is a second-order linguistic phenomenon. Since most speakers of

Sinitic throughout prehistory and history have been illiterate, the nonessentiality of writing for the existence of the SLG/F is self-evident.

iv. A better Sinitic translation of ‘dialect’ than *fangyan* 方言 would be *xiangyan* 相言 ‘mutual lect’ or *tongyan* 通言 ‘intelligible lect’, since the basic meaning of the English word derives from Greek *dialegesthai* ‘to converse’ < *dia-* ‘one with another’ + *legesthai*, the middle voice of *legein* ‘to tell’.

v. In traditional usage, perhaps *xiangtan* 鄉談 ‘village talk’¹⁵ may be considered as a rough, functional equivalent to ‘dialect’. Other somewhat comparable terms are *suyu* 俗語 ‘common/popular saying’, *liyu* 里語 ‘neighborhood speech’, *liyu* 俚語 ‘rustic/vulgar speech’, *tuhua* 土話 or *tuyu* 土語 ‘patois’, and so forth. However, like *fangyan*, none of these terms are suitable for use in modern, scientific, linguistic classification.

vi. The neologisms *yuxi* 語系 ‘family’, *yuzu* 語族 ‘group’ and *yuzhi* 語支 ‘branch’ are perfectly serviceable for modern, scientific analysis.

vii. It is becoming increasingly common for Western linguists to make the following sort of disclaimer in their publications: “Although ‘Chinese’ is commonly thought of as a single language made up of countless ‘dialects,’ linguistically it is more accurate to describe ‘Chinese’ as an extremely complex and diverse group of related *languages* that is comparable to Romance or Germanic in time, depth and scope. These languages that make up the ‘Chinese’ group may be classified into various subgroups; in the modified traditional scheme to which I adhere, these subgroups are called ‘major dialects.’ Languages belonging to the different ‘major dialects’ are mutually unintelligible, and even within the ‘major dialects’ there may be a high degree of internal diversity. A few scholars refer to ‘Chinese’ as Sinitic because they rightly recognize it as a group (or family) of languages rather than a single language. However, in this book/article/paper, I shall continue to follow common convention—based on Chinese cultural and political beliefs—that the various languages of the group are ‘dialects.’” While this represents a definite improvement over the time when scholars were so cowed by tradition that they would not dare to question the dogma that, for China alone, language and dialect are taxonomically undifferentiable, it is still tantamount to saying that, although one realizes languages are not conterminous with dialects, one will doggedly persist in the false assumption that they are identical. One of the chief purposes of the present study is to help my colleagues confront linguistic reality more directly and unabashedly.

viii. If Sardinian and Italian—which are written with the same script, found in the same country, and belong to the same group (Romance)—are considered to be two different languages, all the more so Mandarin and Cantonese—one of which is written

¹⁵ This expression is found already in Zhu Xi’s *Zhu Zi yulei* [Classified Conversations of Master Zhu], 140 and in the great Ming vernacular novel, *Shuihu zhuan* 水滸傳 [Water Margins], 61. See *HDC*, 10.670b.

with Chinese characters and the other is normally not written (but may be with a combination of the regular character set, over a thousand special characters, and conspicuous use of letters from the Roman alphabet), are found in the same country, and belong to the same group (Sinitic)—should be considered as two different languages. The differences between spoken Cantonese and spoken Mandarin are greater than those between spoken Sardinian and spoken Italian. Similar criteria obtain for Catalan and Spanish, and for Scots and English.

7. Clarifications

i. A pervasive fallacy asserts that “all the [Sinitic] ‘dialects’ are the same when written down.” Nothing could be further from the truth. As a matter of fact, the vast majority of the countless varieties of Sinitic have never received a written form, and, indeed, are difficult—if not impossible—to write with the standard set of characters (even in its largest incarnation). When, as a sort of tour de force, they exceptionally do get written down in their full, unadulterated form, as has happened with Cantonese and Taiwanese from time to time, the results are intelligible only to individuals who are literate according to the special conventions of the relevant language.

ii. The tremendous differences between the various Sinitic lects and the standard written language is borne out by the absence of suitable characters for writing numerous basic morphemes, even among the so-called Mandarin dialects. The traditional concept of *benzi* 本字 (“original characters”) for missing morphemes constitutes a flawed policy that fails to take into account borrowings from non-Sinitic origins, substrate languages, and other unknown sources.

iii. Throughout the history of the SLG/F, there have basically been only two types of written language:

- a. Classical/Literary (divinatory, inscriptional, documentary, etc.—starting from the 13th c. BC and continuing up to modern times)
- b. Vernacular (beginning with snatches of the *koine* in medieval times and culminating with full acceptance in the early 20th c.)

Neither of these types of written Sinitic directly and fully reflects any particular speech form tied to a specific time and place, although Vernacular Sinitic (VS) is closer to certain forms of spoken language than is Classical Sinitic (CS) and Literary Sinitic (LS), which tend to be highly elliptical and aggressively terse.

iv. The oft-quoted quip that the difference between a language and a dialect is that the former has an army and a navy while the latter does not is simply silly and should not be adduced in any serious discussion of linguistic classification. There are countless counterexamples that may be brought forward to negate it: French in Quebec; French,

German, Italian, and Rumantsch in Switzerland; French and Flemish in Belgium; Amerindian languages in the United States—all are recognized as independent languages but all lack armies and navies of their own; English and French as the languages of numerous sovereign states; Portuguese in Portugal and Brazil; Spanish in Spain and in many Central American and South American nations—all are dialects of major languages, yet all belong to countries that possess armies and navies of their own.

v. “Missionaries in the 19th century produced Bible translations in Amoy, Foochow, Kienow, Kienyang, and Shaowu (perhaps others as well), showing that one could not use a single ‘Min’ translation that would work for all these different speech areas. Actually, there are possibly twenty or thirty non-mutually intelligible forms of Min alone” (Norman 2009a). North of Fuzhou in Fuan 福安 county, the Catholic missionaries had a romanized form of the local language, which they called in Spanish *Foganes* (Norman 2009b). Church Romanization has been widely and successfully utilized for Hokkien (Amoy and Taiwanese) (A. Lin 1999, and C. Lin 1999). If Sinitic is accepted as a family, then Southern Min, Northern Min, and Eastern Min would count as groups within it, and they would be divisible into branches; if Sinitic is accepted as a group, then Southern Min, Northern Min, and Eastern Min would count as branches within it, and they would be divisible into languages. Similar observations would hold for Wu and other major clusters of Sinitic languages.

vi. The contentious, non-scientific nature of the debate over the SLG/F is manifest in the circumlocutions used to designate its constituent members: “speech forms,” “varieties,” “styles,” “regionalects,” “dialects” (no matter how far up or down the taxonomic scale one may go), and so forth. At the same time, scholars openly admit that the main reasons why they do not use normal linguistic terminology (family, group, branch, language, dialect) in dealing with the SLG/F are due to sociopolitical and cultural factors. The fallacy of such a bizarre approach is evident when one considers that all nations have special sociopolitical and cultural circumstances, yet an impartial analytical outlook does not allow such circumstances to interfere with pure linguistic research.¹⁶

¹⁶ Of course, it will be objected that “there is no such thing as ‘pure linguistic research’” or an impartial, objective outlook. Granted, but I maintain that we should strive to our utmost to come as close to these desiderata as possible, not to throw all standards to the winds because the phenomena we study are complicated and challenging. Think of how complex the human genome is (not to mention the genomes of many other organisms). Nonetheless, dedicated scientists have mapped the entire genomes of human beings and other organisms. We should be no less dedicated in our own research on Sinitic languages, their history, their affinities, and their nature. The microscopic existence of continua does not negate the macroscopically classifiable units (comparable to phyla, genera, and species).

8. Conclusions

i. The classification of the SLG/F, both internally and externally, is still in the beginning stages; much difficult work remains to be done.

ii. ‘Chinese’, ‘dialect’, and other terms in broad popular usage should be employed with extreme caution in technical discussions of the countless varieties of speech forms that currently exist and that have existed at various periods and places during the past in the East Asian Heartland (EAH) and Extended East Asian Heartland (EEAH) (Mair 2005a).

iii. The approximately 1.4 billion people of China do not all share the same language. Rather, they speak a complex mixture of many different tongues, not all of which belong to the SLG/F.

iv. The billion or so members of the *Hanzu* do not all share the same language. Rather, they speak a combination of numerous different languages and dialects that are more or less mutually (un)intelligible.

v. The 867,200,000 individuals who are claimed to be speakers of Mandarin (when examined on the ground in specific locales—not in the abstract as imagined speakers of MSM) actually speak a Babel of tongues that are more or less mutually (un)intelligible. No attempt has been made to determine the number of speakers of MSM, but it is unlikely to constitute as much as a third or half of the total number claimed for Mandarin.

vi. The linguistic landscape of China is every bit as diverse and complex as that of India. Both are modern nation states with national languages (MSM and Hindi) and both have rich cultural histories, including extensive development of their classical languages (Sanskrit and LS). Politically, although one is the world’s largest democracy and the other is the world’s largest country ruled by a single party, they are both sovereign states. The whole world acknowledges that there are many Indic languages in India (Hindi, Bengali, Marathi, Gujarati, Oriya, Assamese, Dogri, Konkani, Maithili, Nepali, Punjabi, Sindhi, and Urdu), not to mention numerous non-Indic languages (Tamil, Telegu, Kannada, Malayalam, Bodo, Manipuri, Santali, and Kashmiri), just as there are numerous non-Sinitic languages in China. The differences among the separate languages of the SLG/F (Cantonese, Taiwanese, Shanghainese, Sichuanese, Hakka, etc.) are every bit as great as those among the Indic group of languages, if not more so, yet scholars and laypersons alike persist in calling the various Sinitic languages ‘dialects’, while readily admitting that Hindi, Bengali, Marathi, etc. are separate languages. There is no linguistic justification for this singularly peculiar usage.

vii. It is unthinkable that an allegedly genetically related linguistic assemblage of such great geographical scope, chronological depth, and burgeoning number of speakers as Sinitic would have only a single form throughout time and space, with but minor

variations ('dialects').

viii. There is an urgent need for the classification of the SLG/F, but this cannot be accomplished satisfactorily without precise, linguistically justifiable terminology.

ix. The people of China have a right to conceive and speak of the languages of their country however they wish; linguists of the world have a duty to study the languages of China according to universal principles. If linguists abandon their scientific duty, the current chaos and lack of consensus concerning the nature of Sinitic will continue, much to the detriment of our understanding not only of the languages of China, but to linguistics as a whole.

x. It is time for linguists to openly declare that there exists a multiplicity of Sinitic languages (not just dialects) and to act upon this reality in their analysis and classification, both internally and externally to the language group/family.

References

- Aymat, Ivan. 2009. Personal communication. March 9.
- Beck, B. J. Mansvelt. 1986. The fall of Han. *The Cambridge History of China*, Vol. 1: *The Ch'in and Han Empires, 221 B.C. - A.D. 220*, ed. by Denis Twitchett & Michael Loewe, 317-376. Cambridge: Cambridge University Press.
- Behr, Wolfgang. 2005. Genealogie, Phonologie und Morphologie des Antikchinesischen. *Antikchinesisch: Ein Lehrbuch in drei Teilen*, by Robert H. Gassmann & Wolfgang Behr, 408-445. Bern & New York: Peter Lang.
- Broughton, Jeffrey L. 1999. *The Bodhidharma Anthology: The Earliest Records of Zen*. Berkeley & Los Angeles: University of California Press.
- Campbell, Lyle. 1997. *American Indian Languages: The Historical Linguistics of Native America*. Oxford & New York: Oxford University Press.
- Casad, Eugene H. 1974. *Dialect Intelligibility Testing*. Norman, Oklahoma: Summer Institute of Linguistics at the University of Oklahoma.
- Chao, Yuen Ren. 1983. *Tongzi Fang'an [A Project for General Chinese]*. Beijing: The Commercial Press.
- Chappell, Hilary. (ed.) 2001. *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford & New York: Oxford University Press. Republished in 2004 by Oxford University Press as *Chinese Grammar: Synchronic and Diachronic Perspectives*.
- Chen, Sanping. 2008a. Personal communication. April 14.
- Chen, Sanping. 2008b. Personal communication. April 10.
- Cheng, Chin-Chuan. 1982. A quantification of Chinese dialect affinity. *Studies in the Linguistic Sciences* 12.1:29-47.

- Cheng, Chin-Chuan. 1991. Quantifying affinity among Chinese dialects. *Languages and Dialects of China*, ed. by William S-Y. Wang, 78-112. Berkeley: Journal of Chinese Linguistics.
- Cheng, Chin-Chuan. 1998. Extra-linguistic data for understanding dialect mutual intelligibility. Paper presented at the 1998 Annual Conference of the Pacific Neighborhood Consortium (PNC), May 15-18, 1998. Taipei: Academia Sinica.
- Coblin, W. South. 2009. Personal communication. January 17.
- DeFrancis, John. 1984. *The Chinese Language: Fact and Fantasy*. Honolulu: University of Hawai'i Press.
- Doak, Kevin M. 2007. *A History of Nationalism in Modern Japan: Placing the People*. Leiden & Boston: Brill.
- Driem, George van. 2007. The diversity of the Tibeto-Burman language family and the linguistic ancestry of Chinese. *Bulletin of Chinese Linguistics* 1.2:211-270.
- Egerod, Søren. 1988. Sino-Tibetan Languages. *The New Encyclopaedia Britannica*, Vol. 22, 721-731. Chicago: Encyclopaedia Britannica.
- Fogel, Joshua A. (ed.) 2005. *The Teleology of the Modern Nation-State: Japan and China*. Philadelphia: University of Pennsylvania Press.
- Gassmann, Robert H., and Wolfgang Behr. 2005. *Antikchinesisch: Ein Lehrbuch in drei Teilen*. Bern & New York: Peter Lang.
- Handel, Zev. 2008. What is Sino-Tibetan? Snapshot of a field and a language family in flux. *Language and Linguistics Compass* 2.3:422-441.
- Huang, Heqing. 2009a. Personal communication. February 4.
- Huang, Heqing. 2009b. Personal communication. February 10.
- Huang, Heqing. (forthcoming). 'Hanren,' 'Hanzu,' 'Hanyu,' 'Hanzi' ciyuan luekao [A brief study on the etymology of 'Han Person,' 'Han Nationality,' 'Han Language,' and 'Hanogram'].
- Lass, Roger. 1997. *Historical Linguistics and Language Change*. Cambridge & New York: Cambridge University Press. See especially Chapter 3: "Relatedness, Ancestry, and Comparison."
- Lin, Alvin. 1999. Writing Taiwanese: The Development of Modern Written Taiwanese. *Sino-Platonic Papers* 89.
- Lin, Christine Louise. 1999. The Presbyterian Church in Taiwan and the Advocacy of Local Autonomy. *Sino-Platonic Papers* 92.
- Lu, Xun. 1934[2005]. An outsider's chats about written language, trans. by Victor H. Mair. *Hawai'i Reader in Traditional Chinese Culture*, ed. by Victor H. Mair, Nancy S. Steinhardt & Paul R. Goldin, 617-641. Honolulu: University of Hawai'i Press.
- Mair, Victor H. 1991. What is a Chinese 'dialect/topolect'? Reflections on some key Sino-English linguistic terms. *Sino-Platonic Papers* 29:1-31.

- Mair, Victor H. 1994. Buddhism and the rise of the written vernacular in East Asia: the making of National Languages. *Journal of Asian Studies* 53.3:707-751.
- Mair, Victor H. 2001. Introduction: the origins and impact of literati culture. *The Columbia History of Chinese Literature*, ed. by Victor H. Mair, 1-15. New York: Columbia University Press.
- Mair, Victor H. 2005a. The north(west)ern peoples and the recurrent origins of the 'Chinese' state. *The Teleology of the Modern Nation-State: Japan and China*, ed. by Joshua A. Fogel, 46-84. Philadelphia: University of Pennsylvania Press.
- Mair, Victor H. 2005b. Introduction: a constantly shifting mosaic of peoples and cultures. *Hawai'i Reader in Traditional Chinese Culture*, ed. by Victor H. Mair, Nancy S. Steinhardt & Paul R. Goldin, 1-7. Honolulu: University of Hawai'i Press.
- Norman, Jerry. 1988. *Chinese*. Cambridge & New York: Cambridge University Press.
- Norman, Jerry. 2009a. Personal communication. January 29.
- Norman, Jerry. 2009b. Personal communication. March 3.
- Ramsey, S. Robert. 1987. *The Languages of China*. Princeton: Princeton University Press.
- Sagart, Laurent. 1994. Proto-Austronesian and Old Chinese evidence for Sino-Austronesian. *Oceanic Linguistics* 33.2:271-308.
- Sagart, Laurent, Roger Blench, and Alicia Sanchez-Mazas. (eds.) 2005. *The Peopling of East Asia: Putting together Archaeology, Linguistics and Genetics*. London & New York: RoutledgeCurzon.
- Sanders, Robert M. 1987. The four languages of 'Mandarin'. *Sino-Platonic Papers* 4:1-14.
- Schuessler, Axel. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- Tang, Chaoju, and Vincent J. van Heuven. 2009. Mutual intelligibility of Chinese dialects experimentally tested. *Lingua* 119.5:709-732.
- Thurgood, Graham. 2008. Han-Zang yuxi: puxi chuancheng, wenhua jiechu, yiji lishi bianhua [The Sino-Tibetan languages: genetic inheritance, external contact, and change]. *Minzu Yuwen [Minority Languages of China]* 2008.2:3-15.
- Thurgood, Graham, and Randy J. LaPolla. (eds.) 2003. *The Sino-Tibetan Languages*. London: Routledge.
- Wang, William S-Y. (ed.) 1991. *Languages and Dialects of China*. Berkeley: Journal of Chinese Linguistics.
- Wang, William S-Y. (ed.) 1995. *The Ancestry of the Chinese Language*. Berkeley: Journal of Chinese Linguistics.
- Wilkinson, Endymion P. 2000. *Chinese History: A Manual*. Revised and Enlarged. Cambridge: Harvard University Asia Center.
- Xu, Wenkan. 2009. Personal communication. February 11.

- Yao, Dehuai. 2009. Personal communication. February 11.
- Zhao, Gang. 2009. *A Marginal Identity and its Modern Transformation: Inquiring into Han Identity*. Manuscript.
- Zhou, Youguang. 2003. *The Historical Evolution of Chinese Languages and Scripts*. Bilingual edition of *Zhongguo Yuwen de Shidai Yanjin* (trans. by Zhang Liqing). *Pathways to Advanced Skills*, Vol. VIII. Columbus: National East Asian Languages Resource Center, The Ohio State University.

On Isomorphism and Formulas of Equivalence in Language Contact^{*}

Bernd Heine
Universität zu Köln

A survey of a number of documented cases of intense language contact suggests that looking for ways of establishing formulas of equivalence between the structures of languages in contact provides one of the motivations of people aiming at communicating successfully in bilingual situations. Equivalence manifests itself in corresponding structures of different languages (or dialects) that are conceived of and/or treated as being the same. The paper is concerned with problems that people experience when trying to adjust the structure of one language to that of another language.

Key words: contact-induced change, formula of equivalence, isomorphism, language contact, replica language, translational equivalence

Alain Peyraube's academic work is centrally concerned with the dynamics of language use and language change, and the present paper is devoted to one aspect of those dynamics, namely contact-induced language change. A survey of a number of documented cases of language contact suggests that looking for ways of establishing formulas of equivalence between the structures of the languages in contact provides one of the motivations of people aiming at communicating successfully in bilingual situations.

1. On equivalence

In an article on language contact involving Irish and English, Bliss observes:

“It is a striking fact that there is an almost complete correspondence between the uses of the dependent ending -(e)ann in early Modern Irish and the uses of the auxiliary do in English: every use of the auxiliary do in English requires the use of the dependent form in early Modern Irish [...]”. (Bliss 1972:78-79)

^{*} I wish to express my gratitude to a number of colleagues for their cooperation when I was working on this paper, in particular to Walter Breu, Claudine Chamoreau, Hilary Chappell, Tania Kuteva, and Regina Martinez Casas.

The question that one may wish to ask is what induced early Modern Irish speakers to establish a connection between an Irish suffix and a verbal auxiliary in English? The literature on language contact abounds with similar examples, where speakers relate structurally contrasting elements of two languages in contact to one another. Flores Farfán (2004:91-92) found in his analysis of language contact between Nahuatl and Spanish in Mexico that under heavy influence of Spanish, speakers of Nahuatl extended the use of their future marker *-s* to function as an equivalent of the Spanish infinitive, and other studies on languages in contact show that speakers tend to treat infinitive markers in one language as being equivalent to markers of nominalization in another language (see Heine & Kuteva 2005: Ch. 6).

Equivalence is a central notion of both contact linguistics and translation theory (see below). It manifests itself in corresponding structures of different languages (or dialects) that are conceived of and/or described as being the same; for example, when speakers regularly identify nouns in language R (the replica language) with nouns in language M (the model language)¹ then they establish what—following Keesing (1991)—we will call a formula of equivalence between two languages in contact. Establishing equivalence is a process that can be of two kinds: it may simply mean that speakers select a given entity Rx of the replica language to correspond to an entity Mx of the model language, but it may also concern a more complex process whereby speakers modify, that is, change, existing material of the replica language to achieve equivalence with the model language.

What is called here ‘equivalence’ has been described in a number of different ways in the literature and referred to variously with terms such as “connection”, “correspondence”, “isogrammatism”, “mutual isomorphism”, or “similarity”. As these terms suggest, the term equivalence, or formula of equivalence, may refer to a range of different things. Heine & Kuteva (2005: §6.1) propose to distinguish two main notions of equivalence. One notion, called structural isomorphism² (henceforth referred to as simply *isomorphism*), rests on the linguist’s theoretical constructs of categories. ‘Equivalent categories’ then means that there is a category Mx in language M and a category Rx in language R which are in some sense taken to be structurally the same. What ‘structurally the same’ stands for exactly is contingent upon the analyst’s descriptive framework—that is, on

¹ Concerning the terms “replica language” and “model language”, see Heine & Kuteva (2005, 2006).

² Note that the term ‘isomorphism’ has received a range of applications in linguistics; it is used especially in language-internal analysis for one-to-one mappings between linguistic form and meaning (see, e.g., Croft 2003). The way the term is used in contact linguistics in general and in the present paper in particular differs from the former applications in that it is strictly comparative in nature, relating to comparisons between languages (or dialects) in contact.

how categories are defined in that framework.

The second notion is *translational equivalence*. It concerns the bilingual behavior of speakers (or writers) and hearers (or readers) in the translation of concepts or contents from one language into another. Equivalent categories are those that are regularly used in translation work as corresponding to one another between the two languages concerned. Accordingly, if we find that speakers regularly translate category Mx of language M by using category Rx in language R, then we will say that this is an instance of translational equivalence between Mx and Rx—irrespective of the grammatical structure of the categories concerned.

1.1 Equivalence in translation work

Obviously, the notion “translation(al) equivalence” is relevant in particular to translation theory, and in fact quite some research has been done on it (see Catford 1965, Uwajeh 2007, and Singh 2008), even if in some of the works the term is discussed controversially (see Leonardi 2000). Important for our purposes is that the information sent from the source language to the target language be “equal in value”, and that a number of levels of equivalence are distinguished, where each level targets a different “degree of equivalence” (Uwajeh 2007):

- (a) Conceptual equivalence (sameness of conceptual units),
- (b) Propositional equivalence (sameness of thought patterns),
- (c) Thematic equivalence (sameness of subject matter), and
- (d) Contextual equivalence (sameness of context variables).

There is reason to assume that the conceptual and communicative behavior underlying translation work, i.e. that of a translator, is similar to that characterizing the behavior of speakers in language contact who aim at establishing equivalence between the model and the replica language. Let us refer to these speakers as the contact speakers. Both concern language use and, hence, are accessible to analysis via established linguistic methodology, and both relate to the interaction of different languages (more precisely, of people speaking different languages). Accordingly, the analysis of translations offers a rich field to the contact linguist—a field that has so far not yet really been exploited.

But there are also differences. First, the setting of participants involved in the two kinds of situations is different:³ in translation work, the translator acts as a kind of “filter” in bridging a communication gap between sender and receiver; in contact

³ Our concern here is with typical situations; it goes without saying that there is a range of variations where our generalizations do not necessarily apply.

situations, by contrast, there are only two participants, namely sender and receiver. Second, translators are bound to produce a translation of some specific information from source language to target language while contact speakers are not: the latter may, but need not, establish equivalence, while they are not bound to produce some specific translation. And third, whereas translation work is planned activity, the activity of contact speakers may, but need not, be planned.

What this suggests with reference to the subject matter discussed in this paper is that research on contact-induced replication can benefit greatly from findings made in “translatology” (i.e. translation studies) but that the two should be kept apart since they are concerned with clearly contrasting processes.

1.2 Searching for equivalence

We may illustrate the behavior of people aiming at achieving equivalence between two Indo-European languages in contact with the following example taken from Verschik (2008:72). The example relates to the situation of language contact in Estonia, where the minority language, Russian, represents the weak code or, as we will say here, the replica language and the national language, Estonian, the strong code⁴ or model language. Speakers of the Slavic language Russian in Estonia have created a number of replica constructions on the model of the Baltic language Estonian, and the following is one instance of such constructions (Verschik 2008:73). In Estonian, nominal modifiers precede their nominal heads, that is, there is modifier-head (possessor-possessee) order, whereas Russian has the opposite order. There appears to be a fairly stable pattern that Estonian Russians were found to use in responding to the contrasting order in their version of Russian: since adjectives precede their head nouns in Russian, Estonian Russians tend to draw on their [adjective - noun] construction to present the head noun as a relative adjective modifier, thereby replicating the modifier-head order of nominal possession in Estonian. Thus, where a monolingual Russian (R) might say something like (1a), Estonian Russians (ER) would use (1b) on the model of Estonian (E) (1c). While being compatible with Russian syntax, (1b) would not be used by monolingual Russians.

⁴ The terms “weak code” and “strong code” are employed in the insightful discussion in Verschik (2008:83ff.).

- (1) a. R⁵ pereryv dlja kofe
 break for coffee
 ‘coffee break’ (Verschik 2008:73)
- b. ER kofej-n- aja pauza
 coffee-ADJ- F.NOM.SG break
- c. E kohv- i- paus
 coffee- GEN- break

Accordingly, there appears to be some kind of formula of equivalence where (2a) of Estonian Russian corresponds to (2b) based on word order correspondence.

- (2) a. ER modifier noun-ADJ -head noun
 b. E modifier-GEN -head noun

But there also is another kind of equivalence where word order is ignored in favor of morphological isomorphism: In the following example, the nominal modifier is constructed in the genitive in both languages:

- (3) a. E saate- kava ‘broadcasting program’
 broadcasting.GEN- program
- b. ER programma peredač- ø
 program broadcasting- GEN.PL (Verschik 2008:72)

That these Russians do not draw on the structure of (2a) by using a relative adjective in (3b), is due to specific norms and habits, as Verschik (2008:72) argues: the choice of translational equivalents is often dependent on conventions, and a relative adjective **peredáčnyj* ‘of broadcasting, having to do with broadcasting’ (cf. *kofejnyj* ‘of coffee’) from the noun *peredáč* ‘broadcasting’ is not used in monolingual Russian.

1.3 Contact-induced change

The example just presented may have shown on the one hand that establishing isomorphic formulas of equivalence in language contact is a flexible activity that is

⁵ Abbreviations: ACC = accusative; ADJ = adjective; DIM = diminutive; E = Estonian; ER = Estonian Russian; F = feminine; GEN = genitive; IMP = imperative; M = model language; NOM = nominative; PL = plural; R = replica language; *R = (reconstructed) pre-contact replica language; SG = singular; R = Russian.

influenced—at least to some extent—by the structures that are accessible to speakers. In the present example, it is syntactic and morphological considerations that appear to have influenced speakers' behavior, but it may as well involve a range of alternative considerations in addition. On the other hand, the example also suggests that aiming at equivalence in language contact may lead to linguistic change: the kind of Russian used by the Estonian Russians referred to above is no longer what it used to be, namely the one characterizing monolingual Russians—in other words, the language has changed. Conversely, this observation can also be used to reconstruct motivations underlying language change; in fact, it constitutes a cornerstone in the methodology used by scholars of contact linguistics, and it is also the methodology used here. By analyzing instances of concluded contact-induced change, it is possible to understand both the motivations responsible for and the mechanism leading to change.

2. A typology

In the present paper we wish to determine how people in language contact achieve equivalence, more precisely (structural) isomorphism. To this end we will look at documented cases of contact-induced linguistic change with a view to finding answers to the following questions:

- (i) What do people do to establish formulas of equivalence between languages in contact?
- (ii) What is the outcome of the process?

In order to answer these questions we will draw on a number of different language contact situations. The bulk of the data presented in this paper, however, is taken from Molise Croatian or Molisean Slavic, in short Molisean, the language of a community of Croatian speakers from the Herzegovinian Neretva Valley who emigrated around AD 1500 because of the Turkish invasion of the Balkans, settling in areas of southeastern Italy that were sparsely inhabited due to earthquakes and epidemics. Today, Molisean is spoken only in two villages, Acquaviva and Montemitro, located in the Molise Region in the Campobasso Province of Italy. After contact both with the local varieties and with Standard Italian over a period of half a millennium, their language has been massively influenced by this Romance language (for a survey and more details, see Breu 1998, see also Breu 1999, 2003a, 2003b, 2003c, 2004).⁶

⁶ We are grateful to Walter Breu for having provided us with the data on which the present analysis rests.

In order to secure comparability across a larger corpus of data, the approach used is multiply reductionist. First, our concern will be exclusively with lexical and grammatical replication, that is, with the contact-induced transfer of structures and meanings but *not* of forms (see Heine & Kuteva 2005: Ch. 1); in other words, we will have to ignore the transfer of forms (= phonological borrowing) or form-meaning units (= borrowing of lexical or grammatical material). And second, we will have to reduce some, at times, complex contact phenomena to specific linguistic reflexes, looking at these reflexes from one particular perspective while ignoring alternative perspectives. We will propose a catalogue of formulas of equivalence, where we use the following symbols in formulaic representations:

M	= model language
R	= replica language
*R	= (reconstructed) pre-contact replica language
[a], [b]	= property “a”, “b” of a category in the model or replica language
[(a)]	= property “a” whose use is optional
M[a]	= a category of the model language having property “a”
∅	= zero form or meaning

Most of the data that exist on grammatical replication have been collected by comparing language R with another language M, and grammatical change is described with reference to some earlier system of R. Hence, description involves reconstruction work, tracing some state in the modern replica language back to some pre-contact state of the same language; we use an asterisk (*R) to signal that the latter is a reconstructed entity—even if in some cases it is still accessible to present-day speakers and hearers.

We will describe grammatical replication and its products in terms of a few basic cognitive operations that speakers and hearers are hypothesized to perform in order to handle the linguistic codes that they are confronted with. A cognitive operation will be said to be maximally simple if, with reference to a given formula of equivalence, no noteworthy cognitive effort is required to move from one code to another, since both the categories and their properties are, at least on the surface, identical in M and R. Conversely, a maximally complex operation is required whenever there is a formula of equivalence where both the categories and their properties are different in every detail of the codes concerned.

These contrasting cases, but also the problems inherent in the approach proposed here, can be illustrated with the following example from the language contact situation between the Slavic language Molisean and the Romance language Italian. The formula in (4) is suggestive of a maximally simple operation since no cognitive effort is required

to move from one code to another: There is total isomorphism in that all the three codes distinguished above have identical structures. For example, the meaning ‘earth’ is expressed in all three codes by a lexical item that can be said to be equivalent and isomorphic, namely *terra* in the Italian model language (M) and *zemlj^a* in both pre-contact Molisean (*R) and modern Molisean (R).

(4) Model language	Pre-contact		Replica language
			replica language
M[a] -	*R[a]	>	R[a]

But a complex operation is required once ‘earth’ is not seen in isolation but rather as part of a semantic network that also includes a second lexeme, which in regional Italian is *fango* ‘wet dirt’ and in pre-contact Molisean *kaš^a* ‘mush’. The process triggered by language contact is described by Breu (2003b:357-358) thus:

Since the hypernym *terra* ‘earth’ could be used regionally for *fango*, *kaša* was also treated as an equivalent of *terra*. For reasons of semantic adaptation, this pertained not only to its special meaning ‘wet dirt’, but also to the whole extent of the meaning of *terra*, therefore to ‘ground’ and to ‘earth’ or ‘world’ as well. On the other hand, Slavic *zemlja* had existed previously as the general word for ‘earth’, and was now used as a synonym for *kaša* in the meanings ‘ground’ and ‘earth’. It is however still distinct from *kaša* in not meaning ‘wet dirt’. Surprisingly, speakers show no tendency to resolve this complicated situation. (Breu 2003b:358)

The semantic differentiation in the three codes is summarized in (5); in a formulaic representation, the bilingual speaker of Molisean is faced with the situation depicted in (6). The complexity of this situation can be summarized thus: (a) There is no equivalence between any of the codes (we are ignoring pre-contact Molisean since it is no longer accessible to speakers); (b) The speaker is confronted with the task of having to establish equivalence relations between two sets of meaning clusters in the two codes; and (c) These clusters are all structurally different from one another: whereas *fango* is monosemous, its ‘equivalent’ *kaš^a* is triply polysemous, and whereas *terra* is associated with three meanings, its ‘equivalent’ *zemlj^a* has only two.

(5)

Model language: Italian	*Pre-contact replica language: Pre-contact Molisean	Replica language: Modern Molisean
<i>fango</i> 'wet dirt'	<i>kaš^a</i> 'mush'	<i>kaš^a</i> 'wet dirt' 'ground', 'earth'
<i>terra</i> 'wet dirt' 'earth', 'ground'	<i>zemljⁱ</i> 'earth'	<i>zemljⁱ</i> 'earth' 'ground'

(6)

Model language	Pre-contact	Replica language replica language
$M_1[a]:$	$*R_1[a]$	$> R_1[a + b + c]$
$M_2[a + b + c]:$	$*R_2[b]$	$> R_2[b + c]$

How bilingual speakers carry out such complex operations is an issue that is beyond the scope of this paper. We will not deal here with such complex equivalence patterns but rather will interpret them as consisting of combinations of elementary patterns, and—due to size limitations—we will also not be able to analyze what is described by Heine & Kuteva (2005) as evolutionary equivalence, which requires a separate treatment. Rather, we will be confined to a basic typology of equivalence patterns as they surface in research on grammatical replication with a view to trying to understand the nature of the cognitive burden that is on the bilingual when regularly confronted with handling two or even more linguistic codes simultaneously.

In the following paragraphs, a catalogue of cases is discussed. These cases are meant to illustrate the most common types of grammatical and lexical replication that we have come across in documented situations of contact-induced change, in an attempt to understand the behavior of bilingual speakers searching for formulas of equivalence in moving from one language to another. The basis of the classification presented will rest on the end product of the processes to be described rather than the processes themselves.

2.1 Total isomorphism

To start with, our interest is with processes leading to isomorphic formulas of equivalence where the model and the replica languages exhibit the same kind of structure.

2.1.1 Equivalence without structural change

The first case concerns what we called above a simple cognitive operation in that the same kind of category is found in all three codes, and speakers and hearers can simply move from one code to another without any special effort simply by equating an item in one code with an item belonging to the same kind of paradigm in another code, as graphically represented in (4), re-printed here for convenience as (7); an example was provided above, involving Italian *terra* (M) and *zemlj^a* in both pre-contact Molisean (*R) and modern Molisean (R).

$$(7) \quad M[a] \qquad *R[a] \qquad > \qquad R[a]$$

While the case just discussed does not involve linguistic change, all other instances leading to total equivalence entail some kind of structural change.

2.1.2 Unmarking

One kind of such change is pragmatic in nature. What surfaces from much of the work that has been carried out on language contact is that it is discourse pragmatics that is likely to be affected first. The result is that speakers tend to use certain structures of R more frequently and/or to extend them to new contexts. A paradigm case of this process is discussed by Heine & Kuteva (2005, 2006) with reference to pattern extension. For example, there may be a major use pattern or construction in the model language (M[a]) while the pre-contact replica language has only minor use pattern or some otherwise marginal equivalent (M[(a)]) of the model situation. Frequently, though not necessarily, this marginal structure is not recognized by speakers or grammarians as a distinct linguistic structure while the major pattern of the model language usually is. In language contact, the minor use pattern or marginal construction of the replica language may then develop into a major use pattern or construction immediately matching the situation of the model language, as sketched in the formula of (8).

$$(8) \quad M[a] \qquad *R[(a)] \qquad > \qquad R[a]$$

One effect that a change from minor to major use pattern may have is that it leads from what tends to be described as a marked structure, confined to specific contexts, to an unmarked structure no longer showing contextual restrictions. We may illustrate this case with the change in West Rumelian Turkish from verb-final (SOV) to verb-medial word order. Turkish is commonly portrayed as a verb-final language, but there are also pragmatically marked structures where the verb precedes its complement (= SVO). For

example, Friedman (2003:66) notes that a sentence such as (9a), where the verb occurs in non-final position, would have a pragmatically marked meaning like ‘It is Erol who is the good student’ in Standard Turkish. Turkish varieties spoken on the Balkans have been strongly influenced by SVO Balkan languages, and in West Rumelian Turkish dialects spoken in Macedonia, the above sentence would be an unmarked sentence equivalent of English ‘Erol is a good student’—corresponding to the unmarked Macedonian sentence (9c). This suggests that under the influence of Macedonian, and perhaps of other Balkanic verb-medial (SVO) languages, speakers of West Rumelian Turkish dialects developed one of their pragmatically marked minor use patterns into an unmarked major pattern—thereby establishing syntactic equivalence with the language or languages of their Balkanic neighbors, as depicted in (10). Consequently, in these Turkish dialects of Macedonia the verb occurs far more frequently in a non-final position than it does in Standard Turkish.⁷

- (9) Turkish and Macedonian (Friedman 2003:66; no glosses provided)
- a. Erol’ dur iyi öğrenci.
‘It is Erol who is the good student.’ Standard Turkish
 - b. Erol’ dur iyi öğrenci.
‘Erol is a good student.’ West Rumelian Turkish
 - c. Erol e dobar učenik.
‘Erol is a good student.’ Macedonian

(10)

<i>Model language:</i> Macedonian	<i>*Pre-contact replica language:</i> Standard Turkish	<i>Replica language:</i> West Rumelian Turkish
SVO	SOV	
	(SVO)	SVO

2.1.3 Addition

Another common process leading to total equivalence in the form of isomorphism can be seen in cases where a category in R lacks a feature that exists in M and speakers

⁷ Whereas Macedonian is a Southern Slavic language, Turkish belongs to the Turkish branch of the Altaic family. In more general terms, Johanson (1992:255) observes: “In türkischen Sprechsprachen des balkanischen Areals hat das ständige Kopieren die Wortstellung so umgestaltet, dass z.B. Komplemente und freie Angaben gewöhnlich dem Prädikatskern folgen. Besonders das Gagausische ahmt die Rechtsverzweigkeit der umgebenden slawischen Sprachen und des Rumänischen nach [...]”

add that feature to establish isomorphism with the category in M, as in the structural change of (11). An example is provided by the regional Italian verb *portare* in (12). This verb is polysemous, meaning both ‘carry’ and ‘drive a car’ (M[a]), while pre-contact Molisean *nosit* only meant ‘carry’ (*R[ø]). Modern Molisean has replicated the Italian polysemy by adding a second meaning to *nosit*, which now also means both ‘carry’ and ‘drive a car’ (R[a]).

$$(11) \quad M[a] \quad *R[\emptyset] \quad > \quad R[a]$$

(12)

Model language: Italian	*Pre-contact replica language: Pre-contact Molisean	Replica language: Modern Molisean
<i>portare</i> (a) ‘carry’ (b) ‘drive a car’	<i>nosit</i> ‘carry’	<i>nosit</i> (a) ‘carry’ (b) ‘drive a car’

A similar example of lexical replication is provided by language contact on the Channel island Guernsey between the French/Norman dialect Guernésiais and English, the official language of Guernsey. Replication in this contact situation went in both directions; in the present case, Guernésiais speakers replicated the polysemy of the English verb *run* by extending the meaning of their verb *courir* ‘to run (= progressing quickly on foot)’ to also express ‘to manage, to conduct operations’ on the model of English ‘run’:

- (13) Guernésiais (Jones 2002:149)
- | | | | | | |
|------------|-----|--------|---------|-----|-----------|
| sL’ éghise | est | couraë | par | la | paraesse. |
| the church | is | run | through | the | parish |
- ‘The church is run by the parish.’

This is a ubiquitous process in grammatical replication, being a manifestation of context extension: One of the most common casualties in language contact is that an existing form or construction (*R[ø]) is applied to new contexts (> R[a]) on the model of another language (M[a]).

That the replica language creates a new category (R[a]) for which there was no exact equivalent in the pre-contact language (*R[ø]) by grammaticalizing a category corresponding to a model category (M[a]) can be demonstrated with a wider range of examples; suffice it to mention just a couple of cases. The Quichua language of Ecuador has grammaticalized its verb *carana* ‘give’ to a modal auxiliary with the main verb encoded in the gerund form (= M[a]), cf. (14a). The function of the modal is to soften requests (translatable in English by *do me the favor of*). This grammaticalization has

been replicated both semantically and structurally in Spanish as spoken in the Ecuadorian Sierra (= R[a]), as in (14b). As the description by Hurley (1995:48-49) suggests, there was no equivalent for this category in this variety of Spanish (= *R[ø]).

- (14) a. Ecuadorian Quichua (Quechuan; Hurley 1995:48)
 Papa- cu- ta randi- shpa cara- hua- y.
 potato- DIM- ACC buy- ing give- me- IMP
 ‘Do me the favor of buying me some potatoes.’
- b. Spanish of the Ecuadorian Sierra (Hurley 1995:48)
 Déme haciendo un sanduche.
 give.me making a sandwich
 ‘Do me the favor of making me a sandwich.’

In the following example, it is a grammatical form that is affected, where the form of a functional category acquires some extra substance on the model of another language. This applies e.g. to situations where a language using a single negation marker replicates the structure of double negation to be found in the other language. Yiddish (M) uses a double negative (M[a]), as in (15a), while English (*R) has a single negation marker (*not*) (*R[ø]). But in their English variety (R), Yiddish speakers living at The Beach in Venice along the U.S. coast next to Los Angeles also tend to use a double negative (R[a]) as in (15b). While contact in such cases adds a (redundant) grammatical form, the category itself is not really affected by this change.

- (15) Yiddish speakers in Venice, Los Angeles (Rayfield 1970:70)
- a. Yiddish: *ikh hob nit keyn gelt.*
 I have not no money
- b. Yiddish English: *I don’t have no money.*
- c. Standard English: I don’t have any money.

The following example of addition concerns the use of definite articles. Even in a pronounced article language such as English there are some restrictions on the use of articles, and in contact with the French/Norman dialect Guernésiais, the English definite article spread to contexts where it would not be used in England but is used in this Norman dialect, in particular before names of languages (16a), adverbials of direction and position (e.g., street-names) (16b), adverbials of time expressing a regular repetition (16c), plural nouns with generic reference (16d), or nouns such as *school* and *bus* in generic uses (16e).

- (16) Replication in Guernsey English on the model of Guernésiais (Ramisch 1989:113-116, Jones 2002:146)
- a. They never did the Guernsey French at school.
 - b. He's got a chain of h'm shops in the, in the Fountain Street.
 - c. And we go the Saturday evening like—old time dancing.
 - d. As a whole I believe the Guernsey people—are h'm friendly and they work together.
 - e. It was always by the bus we went.

Extension of the English definite article as a result of contact with Celtic languages has also been reported for Irish English (e.g., *I had a few jars over the Christmas*.) and English spoken in the Gaelic-speaking area of Scotland (e.g. *have porridge for the dinner*; Ramisch 1989:117). Furthermore, the extension of definite articles to new contexts is an ordinary grammaticalization process that could equally well be triggered language-internally. For these two reasons, the present case need not be due to language contact. That, nevertheless, in the case of Guernsey English, contact was a contributing factor is suggested by the fact that there are corresponding article uses in Guernésiais, that is, English speakers use the article exactly in those contexts where it would be used in Guernésiais, as can be seen in the following examples of (17) corresponding to the English examples in (16).

- (17) Guernésiais (Ramisch 1989:113ff.)
- a. mō per save lə bwō frāse e lōgje e lə patwa.
'My father knew (the) good French and English and (the) *patois*.'
 - b. nu va a la vil pur ʃəpai.
'We go to (the) town for shopping.'
 - c. nuzi va lə samdi o ser.
'We go (there) the Saturday (at.the) evening.'
 - d. le zεfā aprō vit ẽ lōgaʒ.
'The children learn a language quickly.'
 - e. nuze tuʒur alai dō la bōs.
'We always went by (the) bus.'

2.1.4 Loss

Rather than adding a structural feature in language contact, speakers may eliminate a feature in order to establish isomorphism, that is, to conform to the structure of the model language, in accordance with the formula in (18):

$$(18) \quad M[\emptyset] \quad * (R[a]) \quad > \quad R[\emptyset]$$

One way of establishing equivalence consists in fact of simply omitting a form for which there is no equivalent in the model language. The Uto-Aztecan language Nahuatl of Mexico has an imperative prefix *š-* or *x-* ($* (R[a])$), for which there is no equivalent in the contact language Spanish ($M[\emptyset]$). Accordingly, in Chontal Nahuatl the imperative prefix tends to be deleted ($> R[\emptyset]$), where (19b) is used instead of the more standard form (19a):⁸

(19) Chontal Nahuatl (Aztecán; Flores Farfán 2004:93)

- a. *x-* *tleko!*
 IMP- ascend
 ‘Go up!’
- b. *tleko!*
 ‘Go up!’

The process thus leads to the loss of a form in the replica language where there is no equivalent for it in the model language. It might suggest that there are what one may wish to call vulnerable categories which can disappear in situations of language contact in cases where there is no corresponding model category.

The English present perfect appears to be such a vulnerable category: when being used regularly by speaker communities whose model language (in most cases their L2, but sometimes also their L1) does not have an appropriate corresponding aspect category, the present perfect tends to be eliminated in the English variety concerned. For example, in Guernsey English (R), the present tense has been extended to contexts where it also covers what in English is expressed by the present perfect, namely for actions that began in the past and continue up to the present. The authors (Ramisch 1989:150, and Jones 2002:147) attribute this to Guernésiais influence (M). Thus, a structure illustrated by (20a) appears to have been responsible for the Guernsey English structure in (20b) without there being any present perfect ($*R[a]$).

- (20) a. Guernésiais (Jones 2002:148)
 I y a v’ chin quâsi mille aens qué nou-s-est Britanniques.
 [Lit. ‘It is nearly thousand years that we are British’]
 ‘We have been British for nearly a thousand years.’

⁸ When discussing this example with Nahuatl experts in Mexico City however, their reaction was that this example must be taken with care.

b. Guernsey English (Ramisch 1989:150)

There's nearly a thousand years we are British.

Essentially the same contact-induced process has been observed in English varieties of Yiddish speakers in the USA: In the absence of a present perfect in Yiddish, the speakers have extended the English present tense to also cover the present perfect ("uncompleted action continuing into the present") on the model of Yiddish, e.g., Yiddish English *I don't teach since 1937* (Rayfield 1970:67).

Another example of loss is provided by the English local genitive, e.g., *at Bill's, at the chemist's*. There is no equivalent of this form in Guernésiais, and older speakers of Guernsey English tend to drop the genitive clitic, saying *He bought it to Creasey* instead of Standard English *at Creasey's* on the model of Guernésiais *si krisi* (= Standard French *chez Creasey*) (Ramisch 1989:152-153).

Molisean provides yet another example, as shown in (21). A history of roughly 500 years of contact with Italian has led to the restructuring of its gender system: Slavic languages distinguish a neuter category (*R[a]) in addition to masculine and feminine, whereas the model language Italian lacks a neuter (M[ø]), having only masculine and feminine. As Breu (2003b:364) is able to demonstrate convincingly, one of the effects of contact was that Molisean speakers gave up their neuter gender as a nominal category,⁹ redistributing neuter nouns over the masculine and feminine categories, thereby establishing an equivalence relation between the replica language (R[ø]) and the model language.

(21)

<i>Model language:</i>	<i>*Pre-contact replica language:</i>	<i>Replica language:</i>
Italian	Pre-contact Molisean	Modern Molisean
masculine	masculine	masculine
feminine	feminine	feminine
—	neuter	—

2.1.5 Replacement

Another way of establishing total equivalence in the form of isomorphism is by replacing an item *R[b] of the pre-contact replica language with an item R[a] corresponding to an item of the model language (M[a]), that is, via the process depicted in (22):

⁹ Note however that the neuter did not disappear entirely; on the contrary, it was restructured in the areas of pronouns and nominalization, giving rise to new kinds of structures, once again influenced by the model language, Italian (Breu 2003c).

$$(22) \quad M[a] \quad *R[b] \quad > \quad R[a]$$

An example from the contact situation between English and Guernésiais is the following. In this French/Norman dialect (M), like in Standard French, the preposition *a* denotes not only destination ('to') but frequently also static location ('at'), e.g., [ʒə vai a sai pjɛr pɔʁ] 'I go to Saint Peter's Port' vs. [ʒə dmær a sai pjɛr pɔʁ] 'I live in Saint Peter's Port' (Ramisch 1989:137) (M[a]). Guernsey English speakers (R) appear to have replicated this case polysemy by using their destination preposition *to* in such contexts instead of their prepositions of static location *in* or *at* (*R[b] > R[a]):

- (23) Replication of Guernésiais in Guernsey English (Jones 2002:147)
- a. There's some of them candles to the Forest Museum.
 - b. He is to town. 'He is in town.'

2.1.6 Change in paradigm

Another fairly common process leading to total isomorphic equivalence is one where R has a lexical distinction (*R₁[a], *R₂[b]) corresponding to a distinction made within one single lexical item in M: (M[a + b]). In the process leading to isomorphism, shown in (24), the meaning of one of the forms in R is generalized to also cover the other meaning (*R₁[a > b]), and the second form may disappear from the language (*R₂ > ∅).

$$(24) \quad M[a + b] \quad \begin{matrix} *R_1[a] \\ *R_2[b] \end{matrix} \quad > \quad R[a + b]$$

Typical cases of this process concern situations where there is a polysemous category in the model language while the pre-contact replica language has two distinct forms for each of the meanings of the model category. Replication has the effect that in the replica language one of the forms is grammaticalized to acquire a second meaning, thereby creating the same polysemy as exists in the model language. Examples can be found in western and central European languages, where Romance and Germanic languages have a case polysemy, using the same form for comitative ('together with') and instrumental participants ('by means of') (M[a + b]), cf. English *with*.

This process has been illustrated with a number of examples in Heine & Kuteva (2006), suffice it to give another example not contained therein. Guernésiais distinguishes traditionally between the two, using *dauve* 'with' as a comitative and *atou* as an instrumental preposition. Contact with English was presumably a contributing factor for a process towards polysemy, whereby the use of *dauve* was generalized and *atou* tends to be given up in favor of *dauve*; Jones (2002:157) found *atou* to be used only in 3% of

the contexts in which her informants had to respond using instrumental ‘with’.

An example of a similar kind relates to clause subordinating conjunctions as used by Yiddish-English bilinguals in Venice along the U.S. coast next to Los Angeles (Rayfield 1970:61). In Yiddish, which is the first language as well as the model language (M) of these speakers, there does not appear to have been a distinction between a temporal (‘when’) and a conditional conjunction (‘if’); rather, *ven* ‘when, if’ (M[a + b]) was used for both. English, the replica language (*R), distinguishes between the two, using *when* (*R₁[a]) for temporal and *if* (*R₂[b]) for conditional sentences. In their English variety (R), these speakers appear to have replicated the use of Yiddish *ven* by grammaticalizing English *when* to a marker of conditional protasis,¹⁰ thereby matching the situation in their model language (> R[a + b]), where the two are not morphologically distinguished, e.g., Yiddish English *When a lot of people will come, you’ll need a bigger place*.

2.2 Partial isomorphism: optional structures

A fairly widespread development towards isomorphism can be observed in a process captured in (25), where the replica language lacks a category (= *R[∅]) that is found in the model language (M[a]), and R speakers grammaticalize a corresponding category, which however does not attain the fully grammaticalized status of the model, remaining a weakly grammaticalized replica (R[(a)]). ‘Weakly grammaticalized’ means that, compared to the M category, the R category is used less frequently and/or in fewer contexts, or its use is optional whereas that of the M category is obligatory.

$$(25) \quad M[a] \quad *R[\emptyset] \quad > \quad R[(a)]$$

A number of examples of this process are provided by the rise of possessive perfects in Europe (Heine & Kuteva 2006: Ch. 4). Romance and Germanic languages generally have full-fledged possessive perfects using the verb for ‘have’ as an auxiliary and the past perfect participle to encode the main verb (M[a]). Such a perfect category does not exist traditionally in Celtic or Slavic languages or in Basque (*R[∅]), but speakers of languages such as Basque, Irish, Bulgarian and a number of other Slavic languages have replicated the possessive perfect of Romance or Germanic languages; however, they did not proceed beyond stage 2, that is, the replicated perfect can occur with transitive verbs and human subjects but essentially not with intransitive verbs nor with inanimate subjects (see Heine & Kuteva 2006:144-145).

¹⁰ The grammaticalization from temporal to conditional conjunction is a crosslinguistically common process; see Heine & Kuteva (2002).

(27)

Model language: Italian	*Pre-contact replica language: Earlier Molisean	Replica language: Modern Molisean
<i>tavola</i> ‘board, table’	<i>dask^a</i> ‘board’	<i>dask^a</i> ‘board’
	<i>storc^a</i> ‘table’	<i>storc^a</i> ‘table’

A contrast in paradigms of the kind captured in (26b) that is retained in language contact is illustrated with example (28), where it was the model language rather than the replica language that had a lexical differentiation: Italian distinguishes lexically between *mano* ‘hand’ and *braccio* ‘arm’, while pre-contact Molisean had one form, *ruka*, for both ‘hand’ and ‘arm’, and Modern Molisean simply retained this distinction. Once again, no change took place and no isomorphism was achieved, in that both M and R retain their contrasting structures (Breu 2003b:359).

(28)

Model language: Italian	*Pre-contact replica language: Earlier Molisean	Replica language: Modern Molisean
<i>mano</i> ‘hand’	<i>ruka</i> ‘hand, arm’	<i>ruka</i> ‘hand, arm’
<i>braccio</i> ‘arm’		

2.3.2 Coexistence

With the term ‘coexistence’, Heine & Kuteva (2005: Ch. 4) describe a process where M (= M[a]) and R (= *R[b]) have an equivalent grammatical category but R replicates the category of M, with the effect that R now has two categories—the old and the new one—and the two coexist side-by-side, as depicted in (29). The result of this process is a paradigm case of non-isomorphism, in that the situation in R is neither that of the pre-contact replica language (*R) nor of the model language (M).

(29) Type 4: M[a] *R[b] > R₁[b]
 R₂[a]

The example in (31) illustrates the process represented schematically in (29). Molisean has inherited a Southern Slavic ‘want’-future tense, based on the grammaticalization of a volition verbal (‘want’ + main verb) (= *R[b]). In the Italian dialects surrounding the Molisean area there is a future tense (M[a]) based on the

grammaticalization of a possession verb ('have' + main verb), and Molisean speakers used replica grammaticalization for creating a second future tense by means of their possession verb *imat* 'have, must' ($R_2[a]$), thereby acquiring a 'have'-future like the Italian model. The two futures are however not identical in meaning, differing in the modality expressed: whereas the 'want'-future ($R_1[b]$) expresses probability, cf. (30a), the 'have'-future ($R_2[a]$) is a necessity future, as in (30b). To conclude, the resulting situation in the replica language is one which differs not only in its semantics but also in its morphosyntax both from the model language and from the situation found prior to language contact in the replica language (Breu 2003b:369-370).

- (30) Modern Molisean (Slavic; Breu 2003b:370)
- | | | | |
|----|---------------|-----------------------------------|--------------------|
| a. | ču po rabit. | 'I will (probably) go to work.' | Probability future |
| b. | mam po rabit. | 'I will (definitely) go to work.' | Necessity future |

(31)

Model language: Italian of Molise	*Pre-contact replica language: Earlier Molisean	Replica language: Modern Molisean
	'want'-future	'want'-future: probability
'have'-future		'have'-future: necessity

2.4 Complex non-isomorphism

But non-isomorphism can go further in that it may involve more than one distinguishing feature, and with the term 'complex non-isomorphism' we refer to a product of contact-induced change where M and R differ from one another in two or more features.

One kind of pragmatic manipulation commonly to be observed in situations of language contact concerns the rise of new but optional structures, in most cases serving to achieve disambiguation. A typical case, captured in the formula of (32), appears to involve two stages, where the first one is in accordance with a change in paradigm (see above): speakers establish non-isomorphic equivalence between a lexical contrast in the model language ($M_1[a]$, $M_2[b]$) and a polysemy contrast in the replica language ($*R[a + b]$). This may then have the effect that speakers add an optional disambiguating phrase (x) to one of the meanings in R to distinguish between the two polysemous readings.

- (32) $M_1[a]$ $*R[a + b]$ > $R[a + b(x)]$
 $M_2[b]$

The formula in (32) can be illustrated with the example in (33). The Molisean noun *neb^a* is polysemous, meaning both ‘sky’ and ‘palate’, while Italian has different nouns for these meanings, namely *cielo* ‘sky’ and *palato* ‘palate’. Now, modern Molisean speakers retain their polysemy but, being aware that in their model language Italian the two are lexically distinguished, they also introduce a distinction by optionally adding a genitival phrase (*neb^a do usti* ‘sky of the mouth’) to differentiate the meanings ‘sky’ and ‘palate’ (Breu 2003b:358-359).

(33)

Model language: Italian	*Pre-contact replica language: Pre-contact Molisean	Replica language: Modern Molisean
<i>cielo</i> ‘sky’	<i>neb^a</i> ‘sky, palate’	<i>neb^a</i> ‘sky’
<i>palato</i> ‘palate’		<i>neb^a (do usti)</i> ‘palate’

The process underlying (32) can be interpreted as being in the direction towards isomorphism. The resulting structure, however, differs from that of both the model and the replica languages, and the distinction vis-à-vis M is twofold: on the one hand there is a lexical contrast in M but not in R; on the other hand there is an optional constituent in R which has no equivalent in M. Hence one may argue that rather than simplification, the process was one leading towards structural complexification.

The following example represents a variation on the previous one, where R on the one hand retains a formal distinction (*R₁[a], *R₂[b]) corresponding to a polysemy set in M (M[a + b]), but on the other hand replicates the polysemy pattern of M (*R₁[a > b], in accordance with (34).

(34)

$$\begin{array}{ccc}
 M[a + b] & \begin{array}{l} *R_1[a] \\ *R_2[b] \end{array} & > \begin{array}{l} R_1[a + b] \\ R_2[b] \end{array}
 \end{array}$$

An example is found in (35): Italian *prima* is polysemous, having ‘first’ and ‘earlier’ among its meanings, for which pre-contact Molisean had a separate word for each: *prvo* ‘first’ and *prije* ‘earlier’. Replication meant that Molisean *prvo* extended its meaning to ‘first, earlier’—thus being isomorphic with Italian *prima*, in accordance with (11); but Molisean *prije* has been retained in its earlier meaning ‘earlier’ (Breu 2003b:355). Thus, there was change leading on the one hand to full equivalence in one lexical form; on the other hand, Molisean retained a lexical distinction for which there is no equivalent in the model language. For the replica language there is a new complexification in that one meaning, ‘earlier’, is now expressed by two different lexemes.

(35)

<i>Model language:</i> Italian	<i>*Pre-contact replica language:</i> Pre-contact Molisean	<i>Replica language:</i> Modern Molisean
<i>prima</i> ‘first, earlier’	<i>prvo</i> ‘first’	<i>prvo</i> ‘first, earlier’
	<i>prije</i> ‘earlier’	<i>Prije</i> ‘earlier’

For another, even more complex case of complex non-isomorphism, see (5) above.

2.5 Isomorphism versus translational equivalence

This paper is concerned exclusively with isomorphism, which is immediately accessible via standard methods of linguistic analysis. Accordingly, we had nothing to say about translational equivalence, whose analysis would require a pragmatically-based approach. That the two are based on perspectives that may lead to contrasting conclusions can be illustrated with the following example, once more relating to the contact situation in the Molise Region of southeastern Italy. The model language Italian has full-fledged definite ($M_1[a]$) and indefinite ($M_2[b]$) articles, while pre-contact Molisean presumably had neither (= $*R[\emptyset]$). Molisean acquired an indefinite article ($R_1[a]$) via grammaticalization of its numeral ‘one’, like Italian, but it did not grammaticalize a definite article. The resulting situation, summarized in (36), is one where the indefinite article exhibits isomorphism between the two languages while definite-article marking shows non-isomorphism since there is zero marking in the replica language corresponding to an overt article in the model language, as depicted in (37a). Thus, overall there is no isomorphism in the article system of Italian and Molisean.

The situation is different however when it comes to translational equivalence. As Breu (2003a, 2003b:368, 2005) found out, modern Molisean speakers have established a systematic correspondence whereby the Italian distinction of indefinite vs. definite article is equated with the Molisean distinction indefinite article vs. zero article. Accordingly, in the opposition between the indefinite and the definite article, Molisean zero has the significance of a definite article ($R_2[b]$) and there appears to be a translation pattern of total equivalence, as in (37b). If this interpretation is correct, this would suggest that there is every reason to keep the two notions of equivalence apart.

(36)

<i>Model language:</i> Italian	<i>*Pre-contact replica language:</i> Earlier Molisean	<i>Replica language:</i> Modern Molisean
Indefinite article	No articles	Indefinite article
Definite article		Zero definite article

- (37) a. $M_1[a] \quad *R[\emptyset] \quad > \quad R_1[a]$
 $M_2[b] \quad \quad \quad R_2[\emptyset]$
 b. $M_1[a] \quad *R[\emptyset] \quad > \quad R_1[a]$
 $M_2[b] \quad \quad \quad R_2[b]$

3. Conclusion

Issues discussed in this paper have to some extent also been addressed in studies on second language acquisition, and it would be possible to apply some of the techniques of analysis that have been developed there also to the contact phenomena examined here. The acquisition of functional categories as a process involving the re-configuration or re-assembling of formal and semantic feature bundles in the grammar of second languages (Lardiere 2007) might especially provide further clues on how speakers in language contact approach the problem of equivalence. Nevertheless, one has to be aware that second language acquisition and contact-induced replication (or borrowing) are different processes, or at least different components of the same process, in that the latter entails linguistic change, which is not, or not necessarily the case in the former process.

In accordance with the subject matter of the present paper we had to reduce a complex issue to one specific manifestation of it. One kind of reduction meant that we had to focus on inter-lingual interaction and therefore were unable to do justice to how language contact affects the internal structure of a given language, or a given construction of that language; speakers tend to replicate from another language components of a category but not necessarily the internal structure of that category, or the category as a whole.

It is also hard to tell what exactly the observations on isomorphism made in this paper mean with regard to the cognitive tasks that bilinguals are facing in situations of language contact such as the ones described here. This is an empirical question that we could not tackle here but that can be answered by means of an appropriate research program. For the time being we hypothesize that an ideal situation would be one where there is complete isomorphism, that is, where the languages in contact exhibit total equivalence in all their structures, and that any kind of non-isomorphism is likely to require some extra effort for the bilingual speaker to move from one language to another.

The typology proposed in §2 is far from exhaustive. We were confined to a few salient processes that can be observed in situations of language contact, and we were also confined to just one kind of interpretation of the examples presented; for quite a number of these examples, there are additional alternative interpretations.

Given the fact that the data basis that exists for the phenomena looked at in this paper is fairly small, it would be premature to attempt any quantitative generalizations. Nevertheless, it would seem that at least one generalization is possible: a survey of documented cases of contact-induced language change suggests, for example, that speakers are more likely to add new structural material in one language in order to establish formulas of equivalence with another language than to eliminate existing material. This suggests overall that in situations of language contact of the kind analyzed in this paper, languages are less likely to lose than to gain in substance.

The observations made in the paper suggest that searching for ways of establishing formulas of equivalence characterized by isomorphism appears to be at least one of the motivations to be observed in the behavior of speakers in language contact. As we saw in §2.2 through §2.4, however, speakers were not always successful in this endeavor. That aiming at establishing patterns of equivalence may actually have the opposite effect, leading to non-equivalence, can be illustrated with the following case reported by Haase (1992) on Basque. At an earlier stage in its development, Basque was using *hi* as a second person singular pronoun and *zu* for second person plural. Since Roman times, the Romance languages have a second person distinction between an informal/familiar form (etymologically derived from Latin *tu*) and a formal/polite form (*vos*), which is the result of the extension of the second person plural pronoun (*vos*) to also serve as a honorific/formal/polite second person singular pronoun. Early Basque speakers appear to have replicated this distinction by extending the meaning of their plural pronoun *zu* to uses as a polite second person singular pronoun, and a new plural pronoun *zu-ek* was formed by adding the plural marker *-ek*. But the meaning of the Basque pronouns changed: *hi* became restricted to highly familiar second person reference while *zu* turned into the general marker of second person address—with the effect that there no longer is any isomorphism between Basque and the Romance languages, as can be seen in (38): *hi* clearly does not correspond to the modern reflexes of Romance *tu*, nor does *zu* correspond to either *tu* or *vos*; rather, *zu* can be used for both in many contexts (Haase 1992:134).

- (38) Basque second person pronouns and grammatical replication (based on Haase 1992)

	<i>Model language:</i> Romance languages	<i>*Pre-contact replica language:</i> Pre-contact Basque	<i>Replica language:</i> Modern Basque
2.SG familiar	<i>*tu</i>	<i>hi</i>	<i>hi</i>
2.SG general			<i>zu</i>
2.SG polite	<i>*vos</i>		
2.PL			<i>zu-ek</i>

This Basque example also sheds some light on an issue just touched upon, namely on whether language contact contributes to simplification, complexification, or any other situation in the languages concerned. It has frequently been argued that one major effect of contact-induced language change is that it leads to a simplification in the typological outfit of the languages concerned. This is also the conclusion reached e.g. by Gumperz and Wilson in their study on language contact in the Indian village Kupwar. These authors argue that almost all changes in Kupwar “can be interpreted as reductions or generalizations that simplify surface structure in relation to underlying categories and relationships” (Gumperz & Wilson 1971:164). In contrast, Heath (1978:125) concludes with reference to the language contact situation of Arnhem Land, northern Australia that he studied in some detail that “no significant simplification has resulted from indirect diffusion”. Our findings on this issue are in line with those made by Heath, in that none of the various definitions that have been proposed for “simplification” of language structure appears to be relevant to understanding the nature of processes such as the ones we were concerned with in this paper. That simplification is not an appropriate notion to understand contact-induced change is suggested in particular by an observation made above, namely that there appear to be distinctly more changes leading to the addition than to loss of grammatical categories or features in the replica language. The result is that the grammars of the languages concerned are enriched but certainly not, or not necessarily, simplified (cf. Johanson 1992:201).

References

- Bliss, Alan Joseph. 1972. Languages in contact: some problems of Hiberno-English. *Proceedings of the Royal Irish Academy. Section C: Archaeology, Celtic Studies, History, Linguistics, Literature*, 63-82. Dublin: Royal Irish Academy.
- Breu, Walter. 1990. Sprache und Sprachverhalten in den slavischen Dörfern des Molise (Süditalien). *Slavistische Linguistik 1989*, ed. by Walter Breu, 35-65. München: Otto Sagner.
- Breu, Walter. 1992. Das italo-kroatische Verbsystem zwischen slavischem Erbe und kontaktbedingter Entwicklung. *Slavistische Linguistik 1991*, ed. by Timann Reuther, 93-122. München: Otto Sagner.
- Breu, Walter. 1994. Der Faktor Sprachkontakt in einer dynamischen Typologie des Slavischen. *Slavistische Linguistik 1993*, ed. by Hans Robert Mehlig, 41-64. München: Otto Sagner.
- Breu, Walter. 1996. Überlegungen zu einer Klassifizierung des grammatischen Wandels im Sprachkontakt (am Beispiel slavischer Kontaktfälle). *Sprachtypologie und Universalienforschung* 49.1:21-38.
- Breu, Walter. 1998. Romanisches Adstrat im Moliseslavischen. *Die Welt der Slaven* 43:339-354.
- Breu, Walter. 1999. Die Komparation im Moliseslavischen. *Des racines et des ailes: théories, modèles, expériences en linguistique et en didactique*, ed. by René Métrich, Albert Hudlett & Heinz-Helmut Lüger, 37-63. Nancy: A.N.C.A.
- Breu, Walter. 2003a. Der indefinite Artikel in slavischen Mikrosprachen: Grammatikalisierung im totalen Sprachkontakt. *Slavistische Linguistik 2001*, ed. by Holger Kuße, 27-68. München: Otto Sagner.
- Breu, Walter. 2003b. Bilingualism and linguistic interference in the Slavic-Romance contact area of Molise (Southern Italy). *Words in Time: Diachronic Semantics from Different Points in View*, ed. by Regine Eckardt, Klaus von Heusinger & Christoph Schwarze, 351-373. Berlin: Mouton de Gruyter.
- Breu, Walter. 2003c. Impersonales Neutrum im Moliseslavischen. *Rusistika, Slavistika, Lingvistika: Festschrift für Werner Lehfeldt*, ed. by Sebastian Kempgen, Ulrich Schweier & Tilman Berger, 57-71. München: Otto Sagner.
- Breu, Walter. 2004. Der definite Artikel in der obersorbischen Umgangssprache. *Slavistische Linguistik 2002*, ed. by Marion Krause & Christian Sappok, 9-57. München: Otto Sagner.
- Campbell, Lyle. 1987. Syntactic change in Pipil. *International Journal of American Linguistics* 53.3:253-280.
- Catford, John Cunnison. 1965. *A Linguistic Theory of Translation: An Essay in Applied Linguistics*. Oxford: Oxford University Press.

- Croft, William. 2003. *Typology and Universals* (2nd edition). Cambridge & New York: Cambridge University Press.
- Csató, Éva Ágnes, Bo Isaksson, and Carina Jahani. (eds.) 2005. *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*. London & New York: RoutledgeCurzon.
- Filppula, Marku. 2003. The quest for the most ‘parsimonious’ explanations: endogeneity vs. contact revisited. *Motives for Language Change*, ed. by Raymond Hickey, 161-173. Cambridge & New York: Cambridge University Press.
- Flores Farfán, José Antonio. 2004. Notes on Nahuatl typological change. *Sprachtypologie und Universalienforschung* 57:85-97.
- Friedman, Victor A. 2003. *Turkish in Macedonia and Beyond: Studies in Contact, Typology and Other Phenomena in the Balkans and the Caucasus*. Wiesbaden: Harrassowitz.
- Gumperz, John J., and Robert Wilson. 1971. Convergence and creolization: a case from the Indo-Aryan/Dravidian border in India. *Pidginization and Creolization of Languages: Proceedings of a Conference Held at the University of the West Indies, Mona, Jamaica, April 1968*, ed. by Dell Hymes, 151-167. Cambridge: Cambridge University Press.
- Haase, Martin. 1992. *Sprachkontakt und Sprachwandel im Baskenland: die Einflüsse des Gaskognischen und Französischen auf das Baskische*. Hamburg: Buske.
- Haase, Martin. 1997. Gascon et basque: bilinguisme et substrat. *Sprachtypologie und Universalienforschung* 50.3:189-228.
- Haspelmath, Martin, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible. (eds.) 2001. *Language Typology and Language Universals: An International Handbook*, Vol. 2. Berlin & New York: Walter de Gruyter.
- Heath, Jeffrey. 1978. *Linguistic Diffusion in Arnhem Land*. Canberra: Australian Institute of Aboriginal Studies.
- Heine, Bernd. 2012. On polysemy copying and grammaticalization in language contact. *Dynamics of Contact-Induced Language Change*, ed. by Claudine Chamoreau & Isabelle Légise, 125-166. Berlin & New York: Mouton de Gruyter.
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge & New York: Cambridge University Press.
- Heine, Bernd, and Tania Kuteva. 2003. Contact-induced grammaticalization. *Studies in Language* 27.3:529-572.
- Heine, Bernd, and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge & New York: Cambridge University Press.
- Heine, Bernd, and Tania Kuteva. 2006. *The Changing Languages of Europe*. Oxford & New York: Oxford University Press.

- Hurley, Joni Kay. 1995. The impact of Quichua on verb forms used in Spanish requests in Otavalo, Ecuador. *Spanish in Four Continents: Studies in Language Contact and Bilingualism*, ed. by Carmen Silva-Corvalán, 39-51. Washington, D.C.: Georgetown University Press.
- Jenkins, Rebecca Sue. 2002. *Language Contact and Composite Structures in New Ireland, Papua New Guinea*. Ann Arbor: UMI Dissertation Services.
- Johanson, Lars. 1992. Strukturelle Faktoren in türkischen Sprachkontakten. *Sitzungsberichte der Wissenschaftlichen Gesellschaft an der Johann Wolfgang Goethe-Universität Frankfurt am Main* 29.5:169-299.
- Johanson, Lars. 2002. *Structural Factors in Turkic Language Contacts*. Richmond: Curzon.
- Jones, Mari C. 2002. Mette a haout dauve la grippe des Angllais: convergence on the Island of Guernsey. *Language Change: The Interplay of Internal, External and Extra-linguistic Factors*, ed. by Mari C. Jones & Edith Esch, 143-168. Berlin & New York: Mouton de Gruyter.
- Keesing, Roger M. 1991. Substrates, calquing and grammaticalization in Melanesian Pidgin. *Approaches to Grammaticalization*, Vol. 1, ed. by Elizabeth Closs Traugott & Bernd Heine, 315-342. Amsterdam & Philadelphia: John Benjamins.
- Kuteva, Tania. 1998. Large linguistic areas in grammaticalization: auxiliation in Europe. *Language Sciences* 20.3:289-311.
- Lardiere, Donna. 2007. Acquiring (or assembling) functional categories in second language acquisition. *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, ed. by Alyona Belikova, Luisa Meroni & Mari Umeda, 233-244. Somerville: Cascadilla Proceedings Project.
- Leonardi, Vanessa. 2000. Equivalence in translation: between myth and reality. *Translation Journal* 4.4, <http://accurapid.com/journal/14equiv.htm>.
- Menz, Astrid. 1999. *Gagausische Syntax: eine Studie zum kontaktinduzierten Sprachwandel*. Wiesbaden: Harrassowitz.
- Milroy, James. 2003. On the role of the speaker in language change. *Motives for Language Change*, ed. by Raymond Hickey, 143-157. Cambridge & New York: Cambridge University Press.
- Ramisch, Heinrich. 1989. *The Variation of English in Guernsey/Channel Islands*. Frankfurt: Peter Lang.
- Rayfield, Joan Rachel. 1970. *The Language of a Bilingual Community*. The Hague: Mouton.
- Reichl, Karl. 1983. Syntactic interference in Afghan Uzbek. *Anthropos* 78:481-500.
- Schmid, Monika S., Jennifer R. Austin, and Dieter Stein. (eds.) 1998. *Historical Linguistics 1997: Selected Papers from the 13th International Conference on*

- Historical Linguistics, Düsseldorf, 10-17 August 1997*. Amsterdam & Philadelphia: John Benjamins.
- Singh, Anil Kumar. 2008. Review of Uwajeh (2007). *Linguist List* Vol-19-1935, 19 June, 2008.
- Stolz, Thomas. 1991. *Sprachbund im Baltikum? Estnisch und Lettisch im Zentrum einer sprachlichen Konvergenzlandschaft*. Bochum: N. Brockmeyer.
- Thomason, Sarah Grey. 2001. *Language Contact: An Introduction*. Washington, D.C.: Georgetown University Press.
- Uwajeh, M. K. C. 2007. *Translation Equivalence: An Essay in Theoretical Linguistics*. München: Lincom Europa.
- Verschik, Anna. 2008. *Emerging Bilingual Speech*. London & New York: Continuum International.
- Weinreich, Uriel. 1953[1964]. *Languages in Contact*. London, The Hague & Paris: Mouton.
- Zavala, Roberto. 2002. Calcos sintácticos en algunos complejos verbales Mayas y Mixe-Zoques. *Pueblos y Fronteras* 4:169-187.
- Zimmer, Stefan, Rolf Ködderitzsch, and Arndt Wigger. (eds.) 1999. *Akten des Zweiten Deutschen Keltologen-Symposiums*. Tübingen: Niemeyer.

Pan-Sinitic Object Marking: Morphology and Syntax^{*}

Hilary Chappell (曹茜蕾)

EHESS

In Chinese languages, when a direct object occurs in a non-canonical position preceding the main verb, this SOV structure can be morphologically marked by a preposition whose source comes largely from verbs or deverbal prepositions. For example, markers such as *kā* 共 in Southern Min are ultimately derived from the verb ‘to accompany’, *pau*¹¹ 幫 in many Huizhou and Wu dialects is derived from the verb ‘to help’ and *bǎ* 把 from the verb ‘to hold’ in standard Mandarin and the Jin dialects. In general, these markers are used to highlight an explicit change of state affecting a referential object, located in this preverbal position. This analysis sets out to address the issue of diversity in such object-marking constructions in order to examine the question of whether areal patterns exist within Sinitic languages on the basis of the main lexical fields of the object markers, if not the construction types. The possibility of establishing four major linguistic zones in China is thus explored with respect to grammaticalization pathways.

Key words: typology, grammaticalization, object marking, disposal constructions, linguistic zones

1. Background to the issue

In the case of transitive verbs, it is uncontroversial to state that a common word order in Sinitic languages is for direct objects to follow the main verb without any overt morphological marking:

^{*} This is a “cross-straits” paper as earlier versions were presented in turn at both the Institute of Linguistics, Academia Sinica, during the joint 14th Annual Conference of the International Association of Chinese Linguistics and 10th International Symposium on Chinese Languages and Linguistics, held in Taipei in May 25-29, 2006 and also at an invited seminar at the Institute of Linguistics, Chinese Academy of Social Sciences in Beijing on 23rd October 2006. I would like to express my thanks to both audiences for their comments and suggestions, and in particular, Cao Guangshun, Hung-nin Samuel Cheung, Dong Kun, Lai Huei-ling, Li Lan, Li Ming, Chinfu Lien, Meng Pengsheng, Patricia Müller-Liu, Alain Peyraube, Qin Xiongyuan, Yang Yonglong and Zhao Changcai, as well as the two anonymous reviewers for this volume.

Subject – Verb_{transitive} – Object

Thus, when a direct object occurs in a non-canonical position preceding the main verb, this can be signalled by a special marker generally preposed to it, for example, *bǎ* 把 in standard Mandarin.¹ An example from a spoken narrative follows.

Subject – [*Marker* – Object] – Verb_{transitive}

(1) Standard Mandarin object marking construction with *bǎ* 把²

(NP_{CAUSE/SUBJECT}) – [OBJECT MARKER_[bǎ 把] + NP_{DIRECT OBJECT}] – VP

天 黑 了， 我 就 能 把 帽子 摘 了，
tiān hēi le wǒ jiù néng bǎ màozi zhāi le
 sky dark INC I SG then able OM cap doff INC

把 辮子 放在 兜裡...
bǎ biànzi fàng-zai dōu-li
 OM plait place-at pocket-in

‘When night falls, I can take off my cap and put my plait in my pocket.’

(Zhang & Sang 1987:488 *Shijie zhi zui* 世界之最)

I would also like to acknowledge that research leading to these results has received funding from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013): ERC Advanced Grant agreement No. 230388: ‘The hybrid syntactic typology of Sinitic languages’ (2009-2012).

¹ Note, however, that preposed direct object arguments are not required by the grammar to take such marking—morphologically unmarked OV constructions are common in Chinese languages, where the direct object noun has given information content (but is not necessarily an affected patient).

² Abbreviations used in the glossing of examples are as follows: ACC = accusative marker preceding affected object noun, ACH = achievement aspect marker, ADV = marker of adverb formation, AGT = agentive marker, ASST = assertive modality particle, CLF = classifier, COMP = completive aspect marker, COMPR = comparative marker, CONT = continuative aspect marker, COP = copula, CRS = sentence-final marker of a currently relevant state of affairs, DEM = demonstrative, DIMN = diminutive suffix, DIR = directional aspect marker encliticized to verbs, EXT = extent, marker of a postverbal complement indicating the extent of an action or its result state: ‘so X that’, GEN = genitive marker, INC = inchoative aspect marker, LIG = marker of ligature and dependency for attributive phrases, also for relative clauses, LOC = locative, NAME = proper name, NEG = negative adverb, NEG: IMP = negative imperative modal verb, NOM = nominalizer, PFV = perfective aspect marker, PL = plural, PRT = modal or discourse particle, Q = quantifier, SG singular.

Due to its highly transitive nature, this type of syntactic construction is known as the ‘disposal construction’ or *chǔzhìshì* 處置式 in Chinese linguistics. It has been well-studied for standard Mandarin, in particular with regard to its semantic and syntactic constraints which can be summarized as follows:

- (i) the requirement of a referential direct object
- (ii) the constraint that the direct object code a semantic patient that undergoes an explicit change of state (also described as a causativity feature; cf. Chappell 1992a, Bisang 1992)

The direct object often codes referents that represent given or old information (cf. Chafe 1987) in the sense that they have already been mentioned earlier in the conversation or text. In fact, this constraint is often described less precisely in terms of ‘definiteness’ of the direct object. In example (1), the narrator’s very long, luxuriant hair is in fact the topic of the story—and how it has made her famous—while another fact mentioned earlier in the narrative, that she hides her plait under her cap during the day while at work, means that both ‘hair’ or related terms such as ‘plait’ and ‘cap’ (coreferential for their possessor) can all be potential candidates for morphological marking.

The second causativity constraint is realized in the complexity of the predicate: completive types of aspect marking, resultative verbs and postposed locative prepositional phrases all represent ways of expressing this change of state. The variety of syntactic means available is described in Liu (1997). Again, in example (1), the predicate *fàng-zai dōu-li* ‘to put in my pocket’ codes explicit displacement for the object ‘plait’ as too does *zhāi* ‘take off’. Note that such constraints may not necessarily be in operation for all Sinitic languages, as much variation has already been observed in specific dialectal studies. This is an important task for future research in this area, but remains outside the scope of the present discussion.³

The present study represents an enlargement of an earlier survey of this construction type which used a single representative variety for seven Sinitic languages (see Chappell 2006) to ten Chinese dialect groups and an expanded database of secondary reference materials numbering over 200 dialects. It is possible to identify the basic syntactic configuration for the *bǎ* 把 construction in standard Mandarin as the most common kind of disposal construction cross-linguistically in Sinitic, being attested in all ten main dialect groups, (though note that it may not be the native strategy in all of these). For example, the same syntactic configuration, but often with quite different sources for the object marker, is found in Xiang, Gan, Hui, Wu, Min, Yue, Jin, Pinghua and Hakka

³ See Teng (1982) for a detailed discussion of these constraints in Southern Min, or Lai (2003a, 2003b) for a variety of Hakka spoken in Taiwan.

dialects, not to mention in the unclassified Waxiang language (Wu & Chappell 2008). Here is an example from the Xiang dialect of Chenxi, whose object marker is also used synchronically as the lexical verb ‘to help’:

- (2) Chenxi Xiang basic disposal construction with *pau*⁴⁴ 幫 ‘help’:

(NP_{SUBJECT}) – [MARKER_{OM} + NP_{DIRECT OBJECT}] – VERB PHRASE

我 幫 月毛毛 放 嚟 床 上 好 嗎?
 ηo^{33} *pau*⁴⁴ nye^{213} mau^{213} mau^{213} fau^{214} ta^{31} $dzau^{213}$ sa^{31} xau^{31} ma^{44}
 1SG OM baby put at bed on OK Q
 ‘May I put the baby to bed?’ (Wu 2005:204)

In terms of diachronic syntax, this construction type evolves from a serial verb construction found in vernacular texts of the Late Medieval Chinese period (7th-13th centuries) which have the following form:

(NP_{SUBJECT}) – VERB₁ – NP_{DIRECT OBJECT (i)} – VERB₂ (– Pro_{DIRECT OBJECT (i)}) >

(NP_{SUBJECT}) – [PREPOSITION – NP_{DIRECT OBJECT (i)}] – VERB₂ (– Pro_{DIRECT OBJECT (i)})

In this construction type, the first verb is reanalyzed as a preposition signalling that the following noun or pronoun is the direct object (Peyraube 1989, 1996). Depending on the Chinese dialect, the redundant coreferential pronoun in clause-final position may eventually be omitted (see §4 on Cantonese and Hakka for retention of this pronoun).

In contrast to the historical evolution of the object-marking construction, at least five types of “disposal” construction are identified in this present synchronic study combined with three main semantic domains as the sources for the object markers. In terms of structural types, several southern Sinitic languages use more than one strategy for coding affected objects, including configurations with clause-initial objects followed by their object marker, as in certain Min and Wu dialects, not to mention hybridized forms, with double marking, as in Southern Min. It is shown that verbs of taking such as *BĀ* 把 and *NÁ* 拿, well-described for Mandarin, are far from being the sole source of object markers in Chinese languages.⁴ In addition, two new lexical sources are uncovered which involve respectively (i) comitatives and (ii) verbs of giving and helping. This is the

⁴ Small capitals, for example, *NÁ* 拿, with the standard Mandarin *pīnyīn* romanization are used for quoting an ‘archimorpheme’ or ‘allofam’ (to use the term coined by James Matisoff) adopted here as a way of citing a morpheme in terms of its general usage, in the case where I am referring to a large number of cognate morphemes found widely distributed across Sinitic languages with the same function.

first treatment of this particular cluster of construction types from within a pan-Sinitic typological perspective. The analysis explores the possibility of establishing four major linguistic zones in China on the basis of preferred grammaticalization pathways.

In the following section, I examine linguistic diversity from the angle of the principal lexical fields for the sources of these object markers in Sinitic languages.

2. Sources for object markers

There are three main sources for object markers in Sinitic languages, broadly defined as (i) verbs of taking and holding, (ii) verbs of giving and helping and (iii) comitatives (Chappell 2006). Importantly, the classification of object markers in this survey was carried out according to lexical field, uncovering quite a large number of different verb forms in these three semantic domains (see Table 1 and Appendix 1).

Take verbs superficially appear to be the most common source in Sinitic languages and also crosslinguistically, as is the case in many West African Benue-Kwa languages of the Niger-Congo family (see Lord 1993, Heine & Kuteva 2002) and also in many creoles (Muysken & Veenstra 1995:298).⁵ A sample of principal exponents for this category in Sinitic follows:

2.1 Verbs of taking and holding > Object markers, e.g.:

- cognates and synonyms of *bǎ* 把 ‘to take’ as in Standard Mandarin, Zhongyuan Mandarin; the Jin dialects (N.B. the verbal use has been lost in Mandarin)
- *jiāng* 將 ‘to take, lead’ as evidenced in more formal registers of the Southern Sinitic languages, Hakka, Min and Cantonese

⁵ I use the *pīnyīn* romanization system for the Mandarin examples, a system adopted in 1958 for transcription and language pedagogy purposes by the Chinese government. Chinese characters in traditional *fántīzì* form are provided, and where feasible, also for the other Sinitic languages treated in this study. Where a demotic character is not known or attested, the use of an empty box, thus □, is conventional practice in Chinese linguistics.

Tone marks are indicated on all syllables, apart from unstressed ones, and without indicating tone sandhi. For the Southern Min examples, I use the Church Romanization system, as exemplified in the Carstairs Douglas dictionary (1990). Meixian Hakka examples follow the romanization devised by MacIver (1991), unless examples from Rey (1988) are being quoted which use a francophone system. The Cantonese examples from Cheung (1992) use the Yale system while my own data employs the Sidney Lau system. The Shanghainese, Xiang, Huizhou, and Gan examples are rendered in the International Phonetic Alphabet.

- *ná* 拿 ‘to take, hold’ in the central Wu, Xiang and Gan dialects;
e.g. *no*⁵³ in Shanghainese
- *laq*⁷ 掬 ‘to take’ in Gan dialects; also *pa*³ 把 as a borrowed form
from Mandarin
- *tæ*⁴⁴ 擔 ‘to carry, take’ in Xiang dialects (e.g. Dongkou)

A second frequent source is represented by verbs of giving, which appears to be extremely unusual in the languages of the world as a source for object markers. For example, this possibility is neither identified nor discussed in the major crosslinguistic study of *give* verbs presented in Newman (1996), nor similarly in Heine & Kuteva’s lexicon of grammaticalization pathways (2002).

2.2 Verbs of giving and helping > Object markers, e.g.:

- cognates and synonyms of *gěi* 給 ‘to give’ as in Beijing Mandarin, Southwestern,
Jianghuai and Zhongyuan Mandarin
- *bǎ* 把 ‘to give’ in many Hunan Xiang dialects, E’dong
Jianghuai Mandarin, Nanchang Gan
- *te*⁵ 得 ‘to give’ in Xiang dialects,
tet 得 in Dabu Hakka
- *bāng* 幫 ‘to help’ in Wu, Huizhou and Xiang dialects;
e.g. Jinhua (Wu)
- *dei*¹¹ 代 ‘to help’ in Wenzhou (Wu)

The same observation applies to comitatives for which grammaticalization into object markers rarely receives a description in the relevant crosslinguistic literature.

Crosslinguistic documentation shows that comitatives typically develop into instrumentals, allatives, manner and even ergative markers, but never into accusative or object markers (*pace* Stolz 2001 on the instrumental and ergative association; see also Heine & Kuteva 2002 for a similar listing of comitative targets to Stolz). Sinitic languages thus show a special pathway in this respect which has heretofore not been attested in other language families (see Chappell 2000 for this source in Southern Min). Listed below are the dialect groups and individual dialects, uncovered to date, which show this pathway for Sinitic.

2.3 Comitatives > Object markers, e.g.:

- cognates and synonyms of *kā* 共 in Min dialects;
- *t’ung*¹¹ 同 and *lau*¹¹ 佬 in Hakka dialects;

- *tseŋ⁴⁵* 則 (? < 著) in Shaoxing Wu;
- *kən⁴²* 跟 in Jiang-Huai Mandarin dialects of Jiangsu province (Shuyang, Huaiyin);
Southwestern Mandarin in Hubei province (Suixian);
- *ke⁵⁵* 跟 in Waxiang (Hunan).

These markers all have the source meanings in the semantic domain of ‘to accompany’, ‘to mix’ or ‘to follow’ (see Wu Fuxiang 2003).

3. Pathways of grammaticalization for object markers in Sinitic

A brief explanation is in order regarding how these different verbal and comitative sources develop into object markers:

3.1 Verbs of taking and holding

TAKE/HOLD > instrumental > direct object

Verbs of manipulation including ‘take’ and ‘hold’ semantically bleach into object markers, often, but not always, via an instrumental stage (cf. *yǐ* and *jiāng* in Medieval Chinese). This is also the source of object markers in languages of Southeast Asia, including Hmong, Vietnamese, Thai and Khmer (Bisang 1992). Further afield, this pathway has been well-described for Benue-Kwa languages (Lord 1993) and for creoles, as mentioned earlier. Hence, I will not dwell on the grammaticalization process here.

3.2 Verbs of giving and helping

GIVE/HELP > beneficiary > direct object

I have decided to merge the two semantic domains of verbs of giving and helping for the reason that their grammaticalization pathways are most likely parallel: It is well-known that verbs of giving develop into markers of the dative or the beneficiary role (Newman 1996). More specifically, in the first stage of grammaticalization, this semantic class of verbs grammaticalizes into prepositions marking the beneficiary ‘for’, ‘on behalf of’ in V_1 position of a serial verb construction in the case of Sinitic languages. From this stage, they develop further, in the given Sinitic languages, into object markers.

Although this is not common, it can be explained in terms of the closely related semantic change from DATIVE > ACCUSATIVE in Indo-European languages, described in

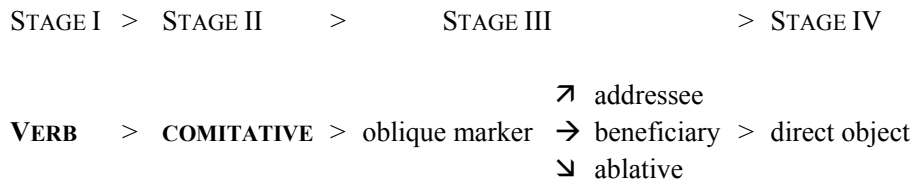
Heine & Kuteva (2002:103, 37), for example, the use of dative *a* as an accusative preposition with Spanish animate nouns; also for the diachronic change between the Old English and Modern English periods for the pronouns *hire*_{DAT} > *her*_{DAT>ACC}, *him*_{DAT} > *him*_{DAT>ACC} which involves reanalysis of a former dative as an accusative/dative; and finally for Persian *râ* (Heine, Claudi & Hünemeyer 1991:165-168).

3.3 Comitative source

Object markers generated by the morpheme *KĀ* 共 in most Min dialects evolved from an original verb meaning ‘to gather, share’, while *t’ung*¹¹ 同 in many Hakka dialects can be traced back to the basic lexical meanings of ‘to be the same as’, ‘to accompany’ in Archaic Chinese. In certain Hakka dialects, *lau*¹¹ 搵 is still used as a verb ‘to mix’ alongside its comitative and other grammatical functions. These verbs grammaticalize first into adverbials with an adjunct function, next into comitative markers ‘with’, and later into a coordinative conjunction ‘and’ in certain Sinitic languages (Liu & Peyraube 1994 on the history of *gòng* 共 from Archaic Chinese to Early Modern Mandarin and Wu Fuxiang 2003 on the diachronic and synchronic dialectal situation for the coordinative conjunction pathway). In other Sinitic languages, such as the Min group, a somewhat different pathway is followed whereby *KĀ*, having reached its comitative marker stage ‘with’, further grammaticalizes into an oblique marker of addressee, benefactive and ablative case roles, then specializes into an accusative or object marker via the beneficiary meaning (Chappell 2000).

The morpheme 跟 GEN used as an object marker in certain Mandarin dialects and in Wuxiang is associated with the verb ‘to follow’ while *tseŋ*⁴⁵ 則 in Shaoxing Wu is likely to have 著 *zhuó* ‘to stick together’ as its source and not 則 (see Xu & Tao 1999:139). Hence, the lexical verbs which occur as sources for the comitative all share the notion of being together, if not being in the same place.

This pathway shares the last stages of its grammaticalization with *give* and *help* verbs, that is, from beneficiary to object marker.



Two Meixian Hakka examples follow which show firstly the comitative and verbal use; and secondly the object marking function of *lau*¹¹ 搵 (see also Lai 2003a, 2003b):

- (3) Meixian Hakka – Comitative and verbal use of *lau*¹¹ 搵

NP – [*lau*¹¹ + comitative NP] – Verb₁ – (Verb₂)

糯米酒 搵 葡萄酒 搵 唔 合
nó mǐ tsiòu laŋ p'ou t'aô tsiòu laŋ mǝ kâp
 rice:wine COM grape:wine mix NEG together

‘le vin de riz mêlé au vin de raisin n’est pas bon’

[‘Rice wine and grape wine don’t mix well together.’] (Rey 1926:479)

- (4) Meixian Hakka – Object use of *lau*¹¹ 搵

(NP_{AGENT}) – [*lau*¹¹_{OM} + NP_{DIRECT OBJECT}] – Verb₁ – (Verb₂)

我 搵 屋 買 到 (了)
ngai lau vuk mai tô le
 1SG OM house buy COMP CRS

‘I (successfully) bought the house.’ (Lin 1990:79)

Possible exceptions:

Before concluding, it behoves us to point out that there are necessarily exceptions to this proposed typology: In my survey of current literature on the topic of disposal constructions, several object markers turned up that do not appear to belong to any of the three proposed lexical fields; their sources are either obscure or not stated. These are:

- (i) 招 [tʂao⁵³] in Xinjiang (新絳), Shanxi province (山西), thus possibly a Jin dialect (affiliation not given). The lexical source is not indicated apart from what might be inferred from the character chosen, ‘to beckon’ in Mandarin. Again, this may more likely be related to 著 *ZHUÓ, ZHÁO* ‘to adhere to’ a verb source which possesses multiple grammaticalization pathways, such as those of complementizers, aspect, causative and passive markers in certain Sinitic languages. Without further proof, this hypothesis can only remain speculative at this point of time.
- (ii) 對 [tui²¹⁴] in the Chaozhou dialect of Southern Min (潮州話), according to Zhan Bohui et al. (1991). The gloss is not provided at all so it can only be remarked that in Chaozhou and other Min dialects, this lexeme also has the meanings of ‘to face’ and ‘from’. The construction appears, however, to be used alongside those with other markers such as *KĀ* 共 < COM discussed above.
- (iii) 到 [tʂe³²⁴] in the Yixian dialect (黟縣話) of Huizhou (徽州) according to Hirata (1998:280). Note that this marker is used alongside [pei³¹] 界 ‘to give’. Again, the lexical source is not indicated, apart from what might be inferred

from the character chosen which means ‘to arrive’ or ‘to’ in Mandarin.⁶

- (iv) Qian (2001) observes that in the Ji-Lu dialects of Shandong province, in addition to the principal use of *bǎ* 把, *lái* 來 ‘to come’ and *lián* 連 ‘coordinative conjunction’ are also used. This is certainly worthy of further research.
- (v) 叫 *JIAO* < causative verb ‘to make’ in certain Shandong dialects (Qian 2001) and also in Central Mandarin, for example, in Henan province (Cao 2008, Map 93). This is related to the intriguing issue of passives and disposal constructions using identical markers (see Wu 1999, 2013).

These remain unsolved for the present, awaiting further research. However, given the relatively small number and their low frequency, I believe it is plausible to maintain the overall trends and patterns for the three principal lexical sources of the object markers. These three main sources have been more recently independently confirmed in a larger study using data from over 650 dialects⁷ and also coincide to some extent with Endo (2004),⁸ and the data found in the dialect map for disposal markers, Map 93, in Cao Zhiyun (2008).

4. Construction types

At least five types of disposal construction can be identified in Sinitic languages, defined in terms of their syntactic configuration. Even though the word order varies considerably, all construction types share the feature of explicit morphological marking of the direct object. This supplies the justification for the combined treatment as

⁶ Similar examples can be found in Cao Zhiyun (2008) [Volume on Grammar], Map 93, and in Endo (2004) that also appear to have a low frequency, including conjunctions such as [LIAN 連]; also [ZHAO 找], [ZHUO 著] and [ZHAO 招] (no source meaning explicitly given) and two other markers which may also be used as passive markers: [JIAO 叫] and [TING 聽]. The latter three are the subject of work in progress on identical agent and patient markers in Sinitic languages.

⁷ This project is entitled ‘The disposal construction in Sinitic languages and related issues’ which I am currently undertaking with Li Lan who has established a database of materials for over 650 dialects.

⁸ I thank Christine Lamarre for bringing this important analysis to my attention. Endo (2004) uses a database of over 450 dialects and independently classifies disposal markers (that is, object markers) into two main types: TAKE/HOLD and GIVE. He does not use a lexical field classification, however, as in the present analysis, but rather classifies each form in terms of descending frequency in his database. Comitative forms are classified in with the beneficiary use of GIVE verbs (see his Table 2) rather than being given separate treatment.

object-marking or disposal constructions.

Each structure has been given a descriptive label, to be regarded purely as a mnemonic device to aid in distinguishing one from another:

I. Common disposal construction

$(NP_{\text{SUBJECT}}) - [MARKER_{\text{OM}} + NP_{\text{DIRECT OBJECT}}] - \text{VERB PHRASE}$

II. “Medieval” disposal construction with a resumptive pronoun following verb

$(NP_{\text{SUBJECT}}) - [MARKER_{\text{OM}} + NP_{\text{DIRECT OBJECT}(i)}] \text{ VERB}_1 - (\text{VERB}_2) - \text{PRONOUN}_{(i)}$

III. Disposal construction with clause-initial object and its resumptive pronoun introduced by the object marker

$NP_{\text{DIRECT OBJECT}(i)} - [MARKER_{\text{OM}} + \text{PRONOUN}_{(i)}] - \text{VERB PHRASE}$

IV. “Archaic Chinese” disposal construction with clause-initial object followed by the object marker and a zero anaphor

$NP_{\text{DIRECT OBJECT}} - [MARKER_{\text{OM}} + ___] \text{ VERB PHRASE}$

V. Hybrid disposal construction with two object markers

$(NP_{\text{SUBJECT}}) - [MARKER_{\text{OM}(i)} - NP_{\text{DIRECT OBJECT}(i)}] - MARKER_{\text{OM}(ii)} - \text{PRONOUN}_{(i)} - \text{VERB PHRASE}$

4.1 Common disposal construction

$(NP_{\text{SUBJECT}}) - [MARKER_{\text{OM}} + NP_{\text{DIRECT OBJECT}}] - \text{VERB PHRASE}$

All the Sinitic languages examined in this brief survey possess at least one type of disposal construction, which I have labelled ‘the common disposal construction’ for convenience.⁹ In other words, it is found in Mandarin, Jin, Xiang, Gan, Wu, Huizhou,

⁹ Endo (2004) also surveys construction types for the disposal construction in Chinese dialects. He recognizes two main types, the narrow and the broad disposal constructions (狹義處置式和廣義處置式). The ‘narrow’ type corresponds more or less to what I am calling the ‘common disposal construction’, that is, the basic type found in all Sinitic languages surveyed. His ‘broad’ type is then further subdivided into two constructions which have either a postverbal pronominal anaphor coding the direct object (our construction II, the ‘Medieval’ type, as found in Yue dialects) or a preverbal one (our construction types III and V, both common in Min with III also occurring in Wu; see Endo’s Map 5 with similar trends). For Endo, the determining

Min, Yue, Pinghua and Hakka. As this has already been exemplified above for Mandarin and Xiang, just one further example is provided below from Taiwanese Southern Min.¹⁰

- (5) Taiwanese Southern Min disposal construction with *kā*:

(NP_{SUBJECT}) – [*KA*_{OM} + NP_{DO}] – VERB PHRASE

啊 汝 共 汝 的 氣力 攞 用去 啊
a lì kā lì ê khuì-lat lóng iòng-khì a
 PRT 2SG OM 2SG GEN strength all use-DIR PRT
 ‘You used up all your strength.’ (Jesse’s Story:823)¹¹

4.2 “Medieval” disposal construction with a resumptive pronoun following verb

The second type of disposal construction appears to mainly form positive imperatives. I have located it in Meixian Hakka, Cantonese Yue, Shanghainese Wu, Jianghuai Mandarin (Yingshan and Chaoxian) and Southwestern Mandarin (Gong’an). It is possibly a reflex of an isomorphic structure found in Medieval Chinese which is discussed in depth in Peyraube (1996), the structure with a postverbal resumptive pronoun being common in Tang dynasty vernacular texts (see also Chappell 2006):

(NP_{SUBJECT}) – [MARKER_{OM} + NP_{DIRECT OBJECT(i)}] VERB₁ – (VERB₂) – PRONOUN_(i)

- (6) Northeastern Hakka: Jiayingzhou or Meixian Hakka

(NP_{SUBJECT}) – [*TSIONG*_{OM} + NP_{DIRECT OBJECT(i)}] VERB₁ – (VERB₂) – PRONOUN_(i)

將 裡 隻 雞 拿來 食帛 佢
tsiōng lì tchâc kē nā-loi chīt-p’êt kî
 OM this-CLF chicken bring eat-COMP 3SG

factor is thus the position of the resumptive pronoun, while in my analysis it is the overall syntactic configuration which is at issue.

Map 92 in Cao Zhiyun (2008) looks at variation in disposal constructions in terms of the areas in which (i) the common disposal construction is used versus (ii) those which add the verbal prefix *gěi* to the predicate: 給 + VP versus (iii) areas or dialects where the disposal construction is not used at all. As such, it is not strictly comparable with our study, apart from confirming the use of the common disposal construction all over China.

¹⁰ Chappell (2006) discusses the construction types in more depth, according to each particular Sinitic language in the sample. Hence, this will not be repeated here.

¹¹ This narrative was recorded in Taipei in the 1990s. Jesse recounts various adventures as a young boy growing up on the outskirts of Taipei in the postwar period.

De cette poule, n'en laissez rien.

['Eat up all this chicken.'] (example from Rey 1926:III)

- (7) Hong Kong Cantonese disposal construction with *jēung* and a postverbal pronoun:

(NP_{SUBJECT}) – [JEUNG_{OM} + NP_{DIRECT OBJECT(i)}] VERB₁ – (VERB₂) – PRONOUN_(i)

千祈 唔好 將 D 頭髮 染黑 佢
chinkèih m̃h.hóu jēung dī tàuhfaat yihm-hāk kéuih
 be:sure NEG:IMP OM CLF_{PL} hair dye-black 3SG
 'Be sure not to dye your hair black.' (example from Cheung 1992:286)

4.3 Disposal construction with clause-initial object and its resumptive pronoun introduced by the object marker

In a third construction type, the direct object is placed in clause-initial position and an anaphoric or resumptive pronoun follows the object marker. It appears to be restricted geographically to the central and southeast of China, specifically to certain Min and Wu dialects: Taiwanese, Xiamen, Shantou, Chaozhou and Hainan (Southern Min), Fuzhou (Mindong), Min dialects of southern Zhejiang, Wenzhou (Wu) and other Zhejiang Wu dialects including Shanghainese, Shaoxing, Zeguo (Yuan 1960:286, Xu & Tao 1999, Huang 1996).

NP_{DIRECT OBJECT(i)} – (NP_{SUBJECT}) – [MARKER_{OM} + PRONOUN_(i)] – VERB PHRASE

- (8) Taiwanese Southern Min accusative *kā* construction with a clause-initial object:

NP_{OBJECT(i)} – [KA_{OM} + PRONOUN_(i)] – VP

門 共 伊 關 起來
m̃ŋg kā yī kuiⁿ khì-lâi
 door ACC 3SG close INC

'Close the door.'

[more literally: the door, take it and close] (example from Tsao 1994:383).

- (9) Wenzhou dialect accusative *dei¹¹* construction with a clause-initial object:

NP_{OBJECT(i)} – [DEI¹¹ + PRONOUN_(i)] – VP

蘋果 代 渠 吃 交
beŋ³¹ku³⁵ dei¹¹ gei³¹ ts^hɿ³¹³ huɔ^o
 apple OM 3SG eat PRT

'Eat up the apple!' [more literally: apple, take it and eat]

(example from Xu & Tao 1999:143)

In both construction types II and III, the pronoun is necessarily coreferential with the direct object preposed to it, not to mention, typically in third person singular form. This contrasts strikingly with the case for standard Mandarin which does not possess either structural subtype for its *bǎ* construction.

4.4 “Archaic Chinese” disposal construction with clause-initial object followed by the object marker and a zero anaphor

NP_{DIRECT OBJECT(i)} – [MARKER_{OM} + ____] VERB PHRASE

In a similar fashion, a fourth structural type places the direct object into clause-initial position preceding the object marker. However, there is no ‘double marking’ of this object by means of a resumptive pronoun: in fact, the object marker directly precedes the predicate. In my survey of the literature on this topic, it was identified just for the small number of following dialects: Dongkou (Xiang); Huaiyin (Jianghuai Mandarin); and Xiuzhuan Hakka.

(10) Dongkou Xiang

NP_{DIRECT OBJECT(i)} – [MARKER_{OM} + ____] VERB PHRASE

衣衫 担 脱 咖
i⁴⁴scē²¹ tã⁴⁴ tʰɔ⁴⁴ kua²¹
 clothes OM take.off-ASP

‘Take off your clothes.’

(example from Wu 2005:205; data recorded by Tang Zuofan in Wu 1996: 472-473)

This does not appear to be a case of the development of a passive prefix as in Taiwanese Southern Min (Cheng & Tsao 1995) since the object marker may also precede intransitive verbs such as ‘to go’; see example (11) below. In other words, it can be expanded into a type of serial verb construction where the verbs share the same object.

(11) 書 担 去 賣 咖
ɕy⁴⁴ tã⁴⁴ tɕʰɿ²⁴ mai⁴⁴ kua²¹
 book OM go sell ASP

‘(Go and) sell these books.’ (example from Wu 2005:205)

This construction type is attested as early as texts from the Archaic period whence its label (see Peyraube 1988:77 for examples from the *Shi Jing* 詩經 or Book of Odes).

4.5 Hybrid disposal construction with two object markers

(NP_{SUBJECT}) – [MARKER_{OM(i)} – NP_{DIRECT OBJECT(i)}] – MARKER_{OM(ii)} – PRONOUN_(i) – VP

A fifth structure found so far only in Southern Min involves hybridization, whereby both the vernacular and literary markers of the disposal construction co-occur, as well as the use of a resumptive pronoun. The double marking strictly follows the order of the literary marker first, that is, *chiong*, followed by the vernacular marker *kā*. I have only found it described for the two Southern Min dialects of Taiwanese and Chaozhou.

- (12) Taiwanese Southern Min hybrid form with two object markers:

(NP_{SUBJECT}) – [*CHIONG*_{OM} + NP_{DO(i)}] – [*KĀ*_{OM} + PRONOUN_{DO(i)}] – VP

將	門	共	伊	關	起來
<i>chiong</i>	<i>mîg</i>	<i>kā</i>	<i>yī</i>	<i>kuiⁿ</i>	<i>khi-lâi</i>
OM ₁	door	OM ₂	3SG	close	INC

‘Close the door.’ (example from Tsao 1994:383)

[more literally: take the door, take it and close:up]

This ends the overview of lexical sources and construction types for object-marking or disposal constructions in Sinitic languages. The fifth and final part of this description concentrates on the typological analysis of Sinitic languages, according to the source of the object marker in the basic disposal construction for which the most data are available.

5. Linguistic zones in China

According to a preliminary survey of approximately 200 Sinitic languages and dialects from secondary references (in addition to the main sources quoted above), four linguistic areas can be discerned, and tentatively established, according to the main source of the object marker:

- 1) **NORTHERN ZONE:**
Mandarin and Jin, Northern Wu.
- 2) **CENTRAL CHINA TRANSITIONAL ZONE:**
Xiang, Gan, Hui and Southern Wu, many central and southern Mandarin dialects including Zhongyuan, Jianghuai and Southwestern Guanhua areas.
- 3) **SOUTHEASTERN ZONE:**
Min and certain Hakka and Wu dialects, and a non-contiguous area comprised of several Jianghuai and Southwestern Mandarin dialects, also Waxiang (Hunan).

- 4) **SOUTHERN ZONE:**
Yue and Hakka dialects.

5.1 Discussion

For the Jin, Northern and Northwestern Mandarin dialect groups, *take* verbs are in the majority, as is the case for the Northern Wu dialect region (the Taihu cluster). In the centre of China, however, we find a swathe of *give* and *help* verbs used as object markers as the primary lexical source. This is particularly apparent in the transitional zones of the Xiang, Gan, Hui and Southern Wu dialects for areas south of the Qiantang River (see Xu & Tao 1999 on Wu dialects). It also holds for the non-northern Mandarin dialect groups (see further remarks below).

In the central south provinces of Hunan and Jiangxi, Xiang and Gan languages show similarities in having *take* verbs as a second, less frequent, source. Often it is difficult to identify the basic lexical meaning of the verb associated with the object marker from the reference materials, since in these central transitional zones, depending on the dialect, ‘give’ can mean ‘take’ or ‘take’ can mean ‘give’ (temporarily using the Mandarin meaning as the departure point for this observation). For example, the meaning of *na*² 拿 and *laq*⁷ 搵 in Gan dialects of Nanchang alternates between ‘give’ in some or ‘take’ in others, whereas the first verb can only mean ‘hold’, ‘take’ in standard Mandarin (the second verb is not used in Mandarin). Similarly, TE 得 in certain Guangdong Hakka and Southwestern Mandarin dialects of Hunan means ‘to give’ and not ‘to obtain’ as it does in standard Mandarin.

It is interesting to note that many Mandarin dialects of the Zhongyuan (Central Plains), Jianghuai and Xinan (Southwestern) groups also use object markers based on ‘give’, and not ‘take’, unlike the northern group of Mandarin dialects which includes the standard language, *pǔtōnghuà* 普通話. An exception to this generalization would however be the vernacular form of the Beijing dialect (Northern Mandarin) in which *gěi* 給 ‘give’ is the object marker, not *bǎ* 把 < ‘hold’ (see Wang 2004, Chirkova 2008; the example from her transcribed conversations which follows interestingly also makes use of a resumptive pronoun after its object marker).

- (13) Colloquial Beijing dialect

一個 小 廟 給 它 換成 黃 琉璃瓦 嗎？

yíge xiǎo miào gěi tā huànchéng huáng liúliwǎ ma

one:CLF little temple OM 3SG change:into yellow tile Q

‘Otherwise, how would a tiny little temple get covered with yellow glazed tiles?’

Help verbs are the predominant source for the Huizhou dialects while this source is

common alongside *give* in Southern Wu.

Pinghua and Yue dialects, including Cantonese, but also many Hakka dialects form a second area of *take* verbs in the south but, in the case of Hakka with widely varying sources compared with the Northern area: The Yue dialects show a predilection for the marker *jiāng* 將 and its cognates whereas a range of verbs is attested for Hakka dialects. However, in the literature consulted, many scholars note that morphologically unmarked OV or SVO constructions are preferred to the disposal construction in Yue and Hakka. Furthermore, they observe that the disposal construction with *jiāng* 將 belongs to a more literary or formal register (書面語 *shūmiànyǔ*). For example, see Cheung (1992). Hence, it does not appear to originally have been a ‘native’ strategy for coding objects but has possibly been introduced by the court language in the Medieval period. In Chappell (2006), I discuss some other examples of *take* verbs found in colloquial Hong Kong Cantonese narratives that are used to introduce objects in serial verb constructions.

Out on a limb, we find the Min dialects which make use of a comitative source, namely *KĀ* 共 and its cognates, as mentioned in §2 and §3. This is the same semantic field as found for the source of object markers in certain Hakka, Wu and even Mandarin dialects. Hakka dialects, including Meixian 梅縣 and some of those located in Taiwan, make use of *t’ung*¹¹ 同 < ‘to accompany’ and *lau*¹¹ 搵 < ‘to mix’, as explained in §2 above. A small segment of Wu dialects, including Shaoxing 紹興 and Ningbo 寧波, have what was originally a comitative marker as their source for object markers (搭), as do a smattering of Jianghuai and Southwestern Mandarin dialects, not to mention the Waxiang language of Hunan, as yet unclassified. The latter groups all use GEN 跟 < ‘to follow’ and its cognates (see Chappell, Peyraube & Wu 2011 for a detailed discussion).¹²

6. Conclusion

Firstly, I have shown that, regardless of the large number of languages and dialects which belong to the Sinitic taxon, there are just three main sources for object markers:

- verbs of taking and holding;
- verbs of giving and helping and
- comitatives

On this basis, and this basis only—the source of object markers—four linguistic zones

¹² A summary form of the data for the four linguistic zones is presented in Table 1 below for ease of reference. A fuller listing of object markers and the source of information is provided in Appendix 1 while a dialect map of China representing the four linguistic zones is given in Appendix 2.

may be tentatively set up: Northern, Central, Southern and Southeastern (see Table 1).

Secondly, at least five different kinds of object marking constructions can be identified in which overt morphological marking is used to code the direct object. These can be distinguished on the basis of their word order, constituency and the position of the direct object with respect to its marker. This was treated in §4.

Table 1: Principal sources of object markers in the four linguistic zones

LINGUISTIC ZONE	EXAMPLES
<p>(1) NORTHERN ZONE</p> <p>Mandarin, Jin and Northern Wu</p> <p><i>Take</i> verbs (Zone I)</p>	<ul style="list-style-type: none"> - Jin dialects 晉: BA 把 - Northern Mandarin dialects 北方官話: BA 把, NA 拿 - Northwestern Mandarin dialects 西北官話: Lanzhou 蘭州, Qinghai 青海 (西北部): BA 把 - Northern Wu: Suhu cluster 蘇戶小片, Shaoxi cluster 蓆溪小片: NA 拿
<p>(2) CENTRAL CHINA TRANSITIONAL ZONE</p> <p>1. <i>Give/help</i> verbs – primary source</p> <p>2. <i>Take</i> verbs – secondary source</p>	<ul style="list-style-type: none"> - Central, Jianghuai and SW Mandarin dialects, 江淮, 中原和西南官話: GEI 給, BA 把 = ‘give’ - Xiang 湘: BA 把 = ‘give’ - Gan 贛: <i>na</i>² 拿, <i>laq</i>⁷ 搵 = ‘give’ or ‘take’ - Huizhou 徽州: BANG 幫 ‘help’ - Southern Wu 吳 (南部): DAI 代 ‘help’, YUE 約 ‘give’ - Hakka dialects 客家話: <ul style="list-style-type: none"> 1. <i>give</i> verbs NA 拿 DE 得 2. a range of <i>take</i> verbs: TI 提, ZHUO 捉, DOULAI 兜來, BAI 擺, YAN 揸, DA 搭
<p>(3) SOUTHEASTERN ZONE</p> <p>Min and certain Hakka and Wu dialects; a sprinkling of Jianghuai and SW Mandarin; Waxiang</p> <p><i>Comitatives</i></p>	<ul style="list-style-type: none"> - Min dialects: KĀ 共 e.g.: Fuzhou 福州, Taiwanese 台灣, Xiamen 廈門, Chaozhou 潮州, Dongshan 東山, Quanzhou 泉州, Hainan 海南 N.B. use of JIANG 將 in formal registers - Hakka dialects: Meixian 梅縣, Taiwan 台灣: LAU 佬, TUNG 同 - Wu dialects 吳 (Taihu cluster): Shaoxing 紹興: ZE 則 (=著); Ningbo 寧波: TA 搭 - Jianghuai and SW Mandarin dialects - Waxiang (unclassified): GEN 跟
<p>(4) SOUTHERN ZONE</p> <p>Yue; Hakka dialects spoken in Guangdong</p> <p><i>Take</i> verbs (Zone II)</p>	<ul style="list-style-type: none"> - Yue 粵 and Hakka 客家 dialects: JIANG 將 - Yue 粵: Serial verb constructions with partly grammaticalized <i>take</i> verbs N.B. use of JIANG 將 and BA 把 corresponds to more formal registers

Apart from the “common” disposal construction, found in all ten Sinitic languages, a determining feature of some of these constructions was the linking of the direct object with a following resumptive pronoun (types II, III, V). A particularly clear case was the “Medieval-style” construction (II) with a resumptive pronoun following the main verb. Clause-initial direct objects (III) or those constructions with double-marking, construction V (with the co-occurrence of two different object markers as in certain Southern Min dialects) also reveal the use of such a resumptive pronoun but one that is introduced by the object marker in this case. In contrast to this, the “Archaic Chinese” clause-initial object construction was followed by its overt marker and a zero anaphor, that is, it did not make use of any resumptive pronoun (IV).

The lack of data on most of these construction types, except for the common one, does not allow for any reliable classification into zones or determining of regional patterning at this point of time, as opposed to the type and semantic source of object markers which allows a broad division into four zones (as discussed in §5 above). All that can be said is that the degree of variation is particularly prominent in the Central and the majority of Southern Sinitic languages, that is, everywhere but in the north which uses the one main common disposal construction type. Postverbal resumptive pronouns common in imperative form are found in Wu, Hakka and Yue dialect groups but not in Min where the preference is to prepose the full lexical form of the object into clause-initial position with the pronominal reprise directly following the object marker (including the hybridized form). This type is found in Wu dialects as well.

The extensive range of structural possibilities and semantic constraints which shape these different disposal constructions highlights the fact that it is an erroneous exercise to assume a broad syntactic isomorphism between Mandarin and other Sinitic languages, as has often been the case in many earlier grammatical studies. From this it can be concluded that researching Sinitic languages *as a whole* (and not just standard Mandarin) can make an important contribution to typology and grammaticalization theory, since, for example, comitatives and verbs of giving/helping are not widely known or attested as sources for object markers in the relevant literature.

A comprehensive survey of this construction type would also aid in determining the relevant intra-Sinitic typological parameters that pertain to syntax and morphology and thereby refine the work of pioneers in typology and dialectology such as Li Fang-Kuei, Yuan Jiahua and his team of linguists, Mantaro Hashimoto, Anne Yue(-Hashimoto) and Jerry Norman *inter alia*. There is probably not just a north-south divide, as proposed by Hashimoto, nor even a north-central-south trichotomy, as proposed by Norman, but at least four major zones, identified on the basis of shared morphosyntactic properties. The reason for enlarging the number of linguistic areas is the fact that the Min dialect group typically shows different syntactic strategies to the Hakka and Yue groups of Southern Sinitic.

Further support for the diversity which abounds in the object-marking strategies of Sinitic languages can be found in typological studies of other constructions or morpho-syntactic categories: for example, Lamarre (2001) on verb complementation; Li Lan (2003) on comparative constructions; Liu Danqing (2003) on basic word order differences; and my own work on aspect and evidential markers (Chappell 1992b, 2001c) which similarly show that the Min dialect group patterns differently from the rest of Sinitic—it uses different lexical sources for its verb complementizers, comparative and evidential markers, and undoubtedly for many other phenomena. A recently completed study on complementizers grammaticalized from quotative verbs (Chappell 2008) similarly reveals that these verbs in Yue and Hakka are at a much younger stage of development than in Southern Min, while the central zone, represented by certain Wu, Xiang and Gan dialects, does not turn up this development at all.

Finally, these linguistic zones or grammaticalization areas crosscut the traditional dialect group boundaries, such as those set up by Li Fang-Kuei (1973) and Yuan Jiahua (1960) and which were largely based on phonological and morphological criteria. Further in-depth research is needed however on a much larger set of grammatical constructions before we can claim these zones exist in any confident way. This small study represents an initial step in this direction, hinting at the possibility of using a typology based on grammaticalization pathways, as one crucial diagnostic.

Appendix 1

SURVEY OF OBJECT MARKERS IN SINITIC LANGUAGES BY DIALECT GROUP (from South to North)

1. Yue dialect group 粵方言:

TAKE verbs

JIANG 將 is the most common marker—the case for 41/45 dialects surveyed in Guangdong province with the caveat that many scholars remark on the preference for OV or for SVO structures where Mandarin uses BA 把.

Other markers listed in surveys: BA 把, □ [ŋja⁵⁵]

Sources: Zhan & Cheung (1988) on Pearl River Delta dialects

Zhan & Cheung (1994) on Yue dialects in Northern Guangdong

Chang (2000); Huang (1996:238-239)

2. Hakka dialect group 客家方言:

TAKE verbs

Hakka dialects, more than any other group, show a tendency to pattern according to their geographical area and neighbours. Thus, even though most of the dialects represented in our survey use a variety of TAKE verbs as their object markers, others in Jiangxi, for example, use a GIVE verb, while yet others use a comitative, as in Taiwan.

(i) Hakka in Guangdong province: 5/5 use JIANG 將

Similar to Yue dialects in this area, they may not use a disposal construction at all. If they do, it is likely to be JIANG 將. Two studies turn up an approximately equal usage of either SOV or OV structures as opposed to S *Marker* OV (Li & Chang 1992, Liu 1999:743).

Source: Zhan & Cheung (1988)

(ii) Hakka in Jiangxi, Hunan, Fujian and Sichuan provinces: 12/17 can use BA 把 ‘take’.

Source: Li & Chang (1992)

(iii) Hakka in Jiangxi province: 5/11 use NA 拿 ‘give’.

Source: Liu (1999:633, 743)

- (iv) Hunan Hakka: e.g. Rucheng uses NA 拿 ‘give’

Source: Huang (1996:662)

- (v) *Other markers listed in surveys:* other attested TAKE verbs in Hakka dialects:

TI 提, ZHUO 捉, DOULAI 兜來, BAI 擺, YAN 攪, DA 搭.

Source: Li & Chang (1992)

- (vi) Comitative:

Note that such surveys using elicitation techniques on the basis of standard Mandarin do not turn up object markers with a comitative source at all, namely, *t'ung*¹¹ 同 and *lau*¹¹ 撈 (Chappell 2006). For further detailed descriptions of *lau*¹¹ 撈 as an object marker, see Lin (1990) and Lai (2003a, 2003b). JIANG 將 is also listed as the main disposal marker in many grammars of Hakka but does not appear to be the native strategy.

3. Gan dialect group 贛方言:

GIVE verbs and TAKE verbs

In general, the Gan dialects show a preference for verbs of giving as the source for their object markers, and secondly, verbs of taking, reflecting their status as central transitional zone dialects.

- (i) Gan dialects in Jiangxi, Hunan, Fujian, Hubei and Anhui provinces:

14/17 use a GIVE verb, particularly BA 把 with the meaning of ‘give’ (9/17). Two other principal verbs of giving in this survey are *ma*³ 𠵿 (listed with the meaning of ‘give’ on p.386) (3/17) and *pei*³ 𠵿 (3/17). The two verbs *na*² 拿 and *laq*⁷ 撈 ‘to hold’ are also sources of object markers; these can mean either ‘take’ or ‘give’, depending on the dialect.

Source: Li & Chang (1992)

- (ii) In Jiangxi province taken alone, verbs of taking appear to prevail but a caveat applies:

8/21 Gan dialects use 把 = ‘take’ while another 4 /21 had NA 拿 and 1/21, TI 提 (total 13/21 dialects surveyed). Note that a further 3/21 of Gan dialects with BA 把 as the OM use this as their *give* verb, while 5/21 Gan dialects had BAI 擺 as their OM similarly meaning ‘give’).

Source: Liu (1999:743). See also Liu (1999:633) for a list of GIVE verbs in Gan dialects.

4. Xiang dialect group 湘方言:

GIVE verbs and TAKE verbs

Similarly, the Xiang dialects show a preference for verbs of giving as the source for object markers, and secondly, verbs of taking. A further source, third in frequency is found in verbs of helping.

Overall, GIVE is the main source of disposal markers in a survey carried out by Wu (2005: Ch. 6) of 98 Hunan dialects including Xiang, Gan/Hakka and Mandarin: 74/98. In this function, it is the verb BA 把 ‘give’ which prevails: 62/74. Restricting the discussion to just the 33 Xiang dialects in her survey, we obtain the following figures:

- (i) GIVE: 把, 給
27/33 localities 26/27 BA 把 and 1/27 GEI 給.
- (ii) TAKE: 5/33 NA 拿, DAN 擔
In this category, 3/5 localities use NA 拿 and 2/5 use DAN.
- (iii) HELP: 1/33 BANG 幫
Source: Wu (2005: Ch. 6)

5. Wu dialect group 吳方言:

GIVE, HELP and TAKE verbs, and the COMITATIVE

According to Xu & Tao (1999), there are three main sources for object markers which correspond to clearly delineated geographical areas in Wu:

- (i) TAKE verbs:
TAKE verbs as OM are spread over the region north of the Qiantang river (錢塘江) including Shanghai, Suzhou (Suhu cluster) and Shaoxi clusters.
Examples: NA 拿, ZHUO 捉, BA 把
- (ii) GIVE and HELP verbs:
This is the most widespread source for Wu dialects, and is found across the entire region except for the Taihu subgroup: WUZHOU 婺州, CHUQU 處衢片, OUJIANG 甌江
GIVE: BO 撥, YUE 約
HELP: BANG 幫, DAI 代

(iii) COMITATIVE

This source is located within the Taihu subgroup, namely, Shaoxing and Ningbo in central eastern Zhejiang.

Shaoxing 紹興: ZE 則

Ningbo 寧波: DA 搭

Source: Xu & Tao (1999)

Qian (1992)

6. Huizhou dialect group 徽州方言:

HELP / GIVE verbs

According to Hirata Shoji's description of 8 Huizhou dialects (1998), the HELP verb BANG 幫 is found in 4/8, while GIVE verbs and TAKE verbs have two each (2/8). This means that the combined category of HELP/GIVE is possibly the preponderant one.

Source: Hirata (1998)

7. Min dialect group 閩方言:

COMITATIVES par excellence

- (i) In most dialects surveyed, the marker used is the cognate of GONG 共, which also serves as the comitative (Chappell 2000). This includes the Mindong dialect of Fuzhou as well as Southern Min where its use is widespread.

GONG 共:

Fuzhou *ky*¹ (Chen 1997)

Xiamen *ka*²~*kā*^η (Zhou 1991)

Chaozhou *ka*²³²

Quanzhou *ka*^η (Lin 1993)

Dongshan *kā*³ (Nakajima 1977)

Hainan *ka*² (Nakajima 1979)

- (ii) JIANG 將 is also used in more formal registers, but is not considered as the native strategy (see Tsao 1994, Cheng & Tsao 1995).

Fuzhou *tsy*^o^η⁵⁵ ~ *tsia*^η¹

Quanzhou *tsi*^o^η¹

Chaozhou *tsia*^η¹

Hainan *tsia*^η¹ (*cf.* references as above)

Putian *tsy*^æ^η⁵⁵

Xiamen *tsi*^o^η¹

Chaoyang *tsa*^η³

8. Mandarin dialect group 北方話:

TAKE verbs and GIVE verbs

(i) TAKE verbs: BA 把

Jianghuai group: BA 把 (Chaoxian, Anhui; Yingshan, Hubei), NA 拿 (Taixing, Jiangsu)

Lanyin group: BA 把 (Lanzhou, Gansu)

Zhongyuan group: BA 把 (Qinghai; Weinan, Shaanxi)

(ii) GIVE verbs: GEI 給

Beijing Mandarin (see Chirkova 2008)

Zhongyuan group (Luoyang, Henan; Suqian, Jiangsu)

Jianghuai group (E'dong, Hubei, Yangzhou, Jiangsu BA 把 = 'give')

Southwestern Mandarin group (Wuhan, Hubei BA 把 = 'give')

(iii) HELP verbs

Southwestern Mandarin: BANG 幫 (Heqing, Yunnan)

Source: Huang (1996)**9. Jin dialect group** 晉方言:

TAKE verbs BA 把

According to Hou & Wen (1993), all 6 Jin dialects in their survey of Shanxi province make use of BA 把 in the disposal construction; as does Huhehot in Inner Mongolia; Jiaocheng, Shanxi (Huang 1996)

Sources: Hou & Wen (1993)**10. Guangxi Pinghua** 廣西桂北平話:

TAKE verbs: Mainly JIANG 將 in more formal registers.

GIVE verbs

[dɔ²¹] Guibei Pinghua Quanzhou Wenqiao patois (Zheng & Lin 2005)

Nanning Pinghua 許 [həi] (Qin 2000 and p.c.)

11. Hunan Waxiang 湖南瓦鄉:

COMITATIVE

Hunan Yuanling Waxiang: 跟 [kɛ⁵⁵] (Yang 1999)Hunan Guzhang Waxiang: 跟 [kai⁵⁵] (Wu & Chappell 2008)

Appendix 2



References

- Bisang, Walter. 1992. *Das Verb im Chinesischen, Hmong, Vietnamesischen, Thai und Khmer*. Tübingen: Gunter Narr Verlag.
- Cao, Zhiyun. 2008. *Hanyu Fangyan Dituji* [*Linguistic Atlas of Chinese Dialects*], Yufa Juan [Volume on Grammar]. Beijing: The Commercial Press.
- Chafe, Wallace. 1987. Cognitive constraints on information flow. *Coherence and Grounding in Discourse*, ed. by Russell Tomlin, 21-51. Amsterdam & Philadelphia: John Benjamins.
- Chang, Song-Hing. 2000. Xianggang Yueyu de jieci [Prepositions in Hong Kong Cantonese]. *Jieci* [*Prepositions*], ed. by Rulong Li & Song-Hing Chang, 236-244. Guangzhou: Ji'nan University Press.
- Chappell, Hilary. 1992a. Causativity and the *bǎ* construction in Chinese. *Partizipation: Das sprachliche Erfassen von Sachverhalten*, ed. by Hansjakob Seiler & Walfried Premper, 509-530. Tübingen: Gunter Narr Verlag.
- Chappell, Hilary. 1992b. Towards a typology of aspect in Sinitic languages. *Chinese Languages and Linguistics*, Vol. 1: *Chinese Dialects*, 67-106. Taipei: Institute of History and Philology, Academia Sinica.
- Chappell, Hilary. 2000. Dialect grammar in two early modern Southern Min texts: a comparative study of dative *kit*, comitative *cang* and diminutive *-guia*. *Journal of Chinese Linguistics* 28.2:247-302.
- Chappell, Hilary. (ed.) 2001a. *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford & New York: Oxford University Press.
- Chappell, Hilary. 2001b. Language contact and areal diffusion in Sinitic languages. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 328-357. Oxford & New York: Oxford University Press.
- Chappell, Hilary. 2001c. A typology of evidential markers in Sinitic languages. *Sinitic Grammar: Synchronic and Diachronic Perspectives*, ed. by Hilary Chappell, 56-84. Oxford & New York: Oxford University Press.
- Chappell, Hilary. 2006. From Eurocentrism to Sinocentrism: the case of disposal constructions in Sinitic languages. *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix Ameka, Alan Dench & Nicholas Evans, 441-486. Berlin & New York: Mouton de Gruyter.
- Chappell, Hilary. 2008. Variation in the grammaticalization of complementizers from verba dicendi in Sinitic languages. *Linguistic Typology* 12.1:45-98.
- Chappell, Hilary, Alain Peyraube, and Yunji Wu. 2011. A comitative source for object markers in Sinitic languages: *kai*⁵⁵ in Waxiang and *kang*⁷ in Southern Min. *Journal of East Asian Linguistics* 20.4:291-338.

- Chen, Zeping. 1997. *Fuzhou Fangyan Yanjiu* [Research on the Fuzhou Dialect]. Fuzhou: Fujian People's Publishing House.
- Cheng, Ying, and Feng-fu Tsao. 1995. Minnanyu 'ka' yongfa zhijian de guanxi [The relationship between uses of *ka* in Southern Min]. *Taiwan Minnanyu Lunwenji* [Papers from the 1994 Conference on Language Teaching and Linguistics in Taiwan, Vol. 1: South Min], ed. by Feng-fu Tsao & Meihui Tsai, 23-45. Taipei: Crane.
- Cheung, Hung-nin Samuel. 1992. The pretransitive in Cantonese. *Chinese Languages and Linguistics*, Vol. 1: *Chinese Dialects*, 241-303. Taipei: Institute of History and Philology, Academia Sinica.
- Chirkova, Katia. 2008. *Gei* 'give' in Beijing and beyond. *Cahiers de Linguistique – Asie Orientale* 37.1:3-42.
- Douglas, Carstairs. 1990. *Chinese-English Dictionary of the Vernacular or Spoken Language of Amoy, with the Principal Variations of the Chang-Chew and Chin-Chew Dialects*. Supplement by Thomas Barclay. Taipei: SMC Publishing Inc. Original edition, London: Trübner & Co., 1873.
- Endo, Masahiro. 2004. Hanyu fangyan chuzhishi biaoqi de dili fenbu yu gezhong chuzhiju [The geographical distribution of disposal markers in Chinese dialects and several types of disposal structures]. *Chuugoku Gogaku* 251:1-15.
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge & New York: Cambridge University Press.
- Heine, Bernd, Ulrike Claudi, and Friederike Hünemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Hirata, Shoji. 1998. *Huizhou Fangyan Yanjiu* [Research on the Huizhou Dialect]. Tokyo: Kobun.
- Hou, Jingyi, and Duanzheng Wen. 1993. *Shanxi Fangyan Diaochaoyanjiu Baogao* [A Report on the Survey and Research into Shanxi Dialects]. Taiyuan: Shanxi United Universities Press.
- Huang, Borong. (ed.) 1996. *Hanyu Fangyan Yufa Leibian* [Concordance of Chinese Dialect Grammar]. Qingdao: Qingdao Press.
- Lai, Huei-ling. 2003a. Hakka *lau* constructions: a constructional approach. *Language and Linguistics* 4.2:353-378.
- Lai, Huei-ling. 2003b. The semantic extension of Hakka *lau*. *Language and Linguistics* 4.3:533-561.
- Lamarre, Christine. 2001. Verb complement constructions in Chinese dialects: types and markers. *Sinitic Grammar: Synchronic and Diachronic Perspectives*, ed. by Hilary Chappell, 85-120. Oxford & New York: Oxford University Press.
- Li, Fang-Kuei. 1973. Languages and dialects of China. *Journal of Chinese Linguistics* 1.1:1-13.

- Li, Lan. 2003. Xiandai Hanyu fangyan chabiju de yuxu leixing [Word order typology of comparative constructions in modern Chinese dialects]. *Fangyan [Dialect]* 2003.3: 1-21.
- Li, Rulong, and Song-Hing Chang. (eds.) 1992. *Ke-Gan Fangyan Diaocha Baogao [A Report on a Survey of Kejia and Gan Dialects]*. Xiamen: Xiamen University Press.
- Li, Rulong, and Song-Hing Chang. (eds.) 2000. *Jieci [Prepositions]*. Guangzhou: Ji'nan University Press.
- Lin, Liantong. 1993. *Quanzhou Fangyanzhi [Quanzhou Dialect Gazetteer]*. Beijing: China Social Sciences Press.
- Lin, Yingjin. 1990. Lun Keyu fangyan zhi 'pun' yu 'lau' [On *pun* and *lau* in Hakka]. *Cahiers de Linguistique – Asie Orientale* 19.1:61-89.
- Liu, Danqing. 2003. *Yuxu Leixingxue yu Jieci Lilun [Word Order Typology and the Theory of Prepositions]*. Beijing: The Commercial Press.
- Liu, Feng-Hsi. 1997. An aspectual analysis of *ba*. *Journal of East Asian Linguistics* 6.1:51-99.
- Liu, Jian, and Alain Peyraube. 1994. History of some coordinative constructions in Chinese. *Journal of Chinese Linguistics* 22.2:179-201.
- Liu, Lunxin. 1999. *Ke-Gan Fangyan Bijiao Yanjiu [Comparative Research on Kejia and Gan Dialects]*. Beijing: China Social Sciences Press.
- Lord, Carol. 1993. *Historical Change in Serial Verb Constructions*. Amsterdam & Philadelphia: John Benjamins.
- MacIver, Donald. 1991. *A Chinese-English Dictionary Hakka-Dialect as Spoken in Kwang-Tung Province*. Revised by M. C. MacKenzie. 2nd edition. Taipei: Southern Materials Center. Original edition, Shanghai: Presbyterian Mission Press, 1926.
- Muysken, Pieter, and Tonjes Veenstra. 1995. Serial verbs. *Pidgins and Creoles: An Introduction*, ed. by Jacques Arends, Pieter Muysken & Norval Smith, 289-301. Amsterdam & Philadelphia: John Benjamins.
- Nakajima, Motoki. 1977. *Bingo Tōzantō Hōgen Kiso Goishū [A Study of the Basic Vocabulary of the Min Dialect in the Tung Shan Island]*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Nakajima, Motoki. 1979. *Fukken Kango Hōgen Kiso Goishū [A Comparative Lexicon of Fukien Dialects]*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Newman, John. 1996. *Give: A Cognitive Linguistic Study*. Berlin & New York: Mouton de Gruyter.
- Peyraube, Alain. 1985. Les formes en *ba* en chinois vernaculaire médiéval et moderne. *Cahiers de Linguistique – Asie Orientale* 14.2:193-213.

- Peyraube, Alain. 1988. *Syntaxe diachronique du chinois: évolution des constructions datives du 14^e siècle av. J.-C. au 18^e siècle*. Paris: Collège de France, Institut des Hautes Etudes Chinoises.
- Peyraube, Alain. 1989. Zaoqi ‘ba’ ziju de jige wenti [Some problems concerning the early *ba* forms]. *Yuwen Yanjiu* [Linguistic Researches] 1989.1:1-9.
- Peyraube, Alain. 1991. Syntactic change in Chinese: on grammaticalization. *Bulletin of the Institute of History and Philology Academia Sinica* 59.3:617-652.
- Peyraube, Alain. 1996. Recent issues in Chinese historical syntax. *New Horizons in Chinese Linguistics*, ed. by C.-T James Huang & Y.-H. Audrey Li, 161-214. Dordrecht: Kluwer.
- Qian, Nairong. 1992. *Dangdai Wuyu Yanjiu* [Research on Contemporary Wu Dialects]. Shanghai: Shanghai Education Press.
- Qian, Zengyi. 2001. *Shandong Fangyan Yanjiu* [Research on Shandong Dialects]. Ji’nan: Qilu Shushe.
- Qin, Yuanxiong. 2000. Nanning Pinghua de jieci [Prepositions in Nanning Pinghua]. *Jieci* [Prepositions], ed. by Rulong Li & Song-Hing Chang, 227-235. Guangzhou: Ji’nan University Press.
- Rey, Charles. 1988. *Dictionnaire Chinois-Français: Dialecte Hakka*. Taipei: Southern Materials Center. Original edition, Hong Kong: Imprimerie de la Société des Missions-Etrangères, 1926.
- Sagart, Laurent. 1999. Notes on the Nanchang dialect (Sketch grammar and transcriptions). Manuscript. Melbourne: La Trobe University.
- Stolz, Thomas. 1999. Comitatives vs. instrumentals vs. agents. *Aspects of Typology and Universals*, ed. by Walter Bisang, 153-174. Berlin: Akademie Verlag.
- Stolz, Thomas. 2001. On Circum-Baltic instrumentals and comitatives: to and fro coherence. *Circum-Baltic Languages*, Vol. 2: *Grammar and Typology*, ed. by Maria Koptjevskaja-Tamm & Östen Dahl, 591-612. Oxford & New York: Oxford University Press.
- Teng, Shou-hsin. 1982. Disposal structures in Amoy. *Bulletin of the Institute of History and Philology Academia Sinica* 53.2:331-352.
- Tsao, Feng-fu. 1994. On the mechanisms and constraints in syntactic change: evidence from Chinese dialects. *Chinese Languages and Linguistics*, Vol. 2: *Historical Linguistics*, ed. by Paul Jen-kuei Li, Chu-Ren Huang & Chih-Chen Jane Tang, 389-419. Taipei: Institute of History and Philology, Academia Sinica.
- Wang, Jian. 2004. ‘Gei’ zi biaoshi chuzhi de lai yuan [The origin of ‘gei’ sentences expressing disposal]. *Yuwen Yanjiu* [Linguistic Researches] 2004.4:9-13.
- Wu, Fuxiang. 2003. Hanyu bansui jieci yufahua leixingxue yanjiu [A typological study of the grammaticalization of comitative prepositions in Chinese]. *Zhongguo Yuwen* [Studies of the Chinese Language] 2003.1:43-58.

- Wu, Yunji. (ed.) 1996. *Hunan Fangyan de Dongtai Zhuci* [*The Aspectual Markers in the Hunan Dialects*]. Changsha: Hunan Normal University Press.
- Wu, Yunji. 1999. An etymological study of disposal and passive markers in Hunan dialects. *Journal of Chinese Linguistics* 27.2:90-123.
- Wu, Yunji. 2005. *A Synchronic and Diachronic Study of the Grammar of the Chinese Xiang Dialects*. Berlin & New York: Mouton de Gruyter.
- Wu, Yunji. 2013. Hunan nanbu chuzhi he beidong tong biaoji xianxiang de zai sikao [Some further thoughts on the identical disposal and passive markers in Southern Hunan]. *Breaking Down the Barriers: Interdisciplinary Studies in Chinese Linguistics and Beyond*, ed. by Guangshun Cao, Hilary Chappell, Redouane Djamouri & Thekla Wiebusch, 995-1010. Taipei: Institute of Linguistics, Academia Sinica.
- Wu, Yunji, and Hilary Chappell. 2008. Xiangxi Waxianghua he Xinan Guanhua biao chuzhishi de 'gen' suo yinfa de yiwen [The unusual case of [gēn] used as a disposal marker in Waxiang and the Southwestern Mandarin of Western Hunan, China]. *Studies in Chinese Linguistics* 26:1-13.
- Xu, Baohua, and Huan Tao. 1999. Wuyu de chuzhiju [Disposal constructions in the Wu dialects]. *Hanyu Fangyan Gongshi yu Lishi Yufa Yantao Lunwenji* [*Synchronic and Diachronic Perspectives on the Grammar of Sinitic Languages*], ed. by Yunji Wu, 135-167. Guangzhou: Ji'nan University Press.
- Yang, Wei. 1999. *Yuanling Xianghua Yanjiu* [*Research on the Yuanling Waxiang Dialect*]. Changsha: Hunan Education Press.
- Yuan, Jiahua. (comp.) 1960[1989]. *Hanyu Fangyan Gaiyao* [*An Outline of Chinese Dialects*]. Beijing: Language Reform Press.
- Zhan Bohui, and Yat-Shing Cheung. 1988. *A Survey of Dialects in the Pearl River Delta*, Vol. 2: *Comparative Lexicon*. Guangzhou: Guangdong People's Publishing House.
- Zhan, Bohui, and Yat-Shing Cheung. 1994. *Yuebei Shi Xian Shi Yue Fangyan Diaocha Baogao* [*A Survey of Yue Dialects in North Guangdong*]. Xiamen: Xiamen University Press.
- Zhan, Bohui et al. 1991. *Hanyu Fangyan ji Fangyan Diaocha* [*Chinese Dialects and Dialect Investigation*]. Wuhan: Hubei Education Press.
- Zhang, Xinxin, and Ye Sang. (eds.) 1987. *Beijing Ren: Yibai ge Putongren de Zishu* [*Beijing People: Narratives of 100 Ordinary People*]. Shanghai: Shanghai Literature and Art Publishing Group.
- Zhou, Changji. 1991. *Minnanhua yu Putonghua* [*The Southern Min Dialect and Standard Chinese*]. Beijing: Language and Culture Press.
- Zheng, Zuoguang, and Yi Lin. 2005. *Guibei Pinghua yu Tuiguang Putonghua Yanjiu* [*Guibei Pinghua and Research on Promoting Putonghua*]. Nanning: Guangxi Nationalities Press.

Hilary Chappell

Zhu, Guanming. 2005. Hubei Gong'an fangyan de jige yufa xianxiang [Several grammatical phenomena in the Gong'an dialect of Hubei]. *Fangyan* [*Dialect*] 2005.3:253-257.

The Development of VOC to VCO in Mai-Cantonese of Sanya in Hainan^{*}

Andy C. Chin (錢志安) and Benjamin K. Tsou (鄒嘉彥)

The Hong Kong Institute of Education

This paper examines an on-going change of the resultative verb construction (結果補語) in the Mai (邁話) dialect spoken in Sanya (三亞), Hainan Province. Our fieldwork data show that two patterns co-exist for the resultative verb construction which differ in terms of the relative word order of the complement (C) and the object (O): VOC vs. VCO (such as *da laoshu si* 打老鼠死 vs. *dasi laoshu* 打死老鼠 ‘beat the mouse to death’). After considering a number of linguistic and extra-linguistic factors as well as the results of our fieldwork conducted among native Mai speakers of different ages, we argue that VOC is an older pattern while the VCO pattern is used more often by younger speakers, which can be primarily attributed to the influence of *Standard Chinese* as a result of language contact.

Key words: resultative verb construction, languages in Sanya of Hainan Province, language contact, diachronic change

1. Introduction

The Mai dialect investigated in this study is now mainly spoken in Sanya, at the southernmost tip of Hainan Province. In the past, Hainanese (海南話) and the Li (黎話) and Huihui (回輝) languages have been recognized as the three major linguistic varieties spoken in Hainan Province (see, for example, *The Language Atlas of China*, maps A4 and B8). In fact, according to the recent sociolinguistic surveys conducted by the authors’ team (Tsou et al. 2007b), Sanya has at least eight distinct linguistic varieties spoken as a first language by the inhabitants. Some of them are Sinitic languages such as Standard Chinese, Hainanese (海南話, a variety of the Min dialects), Mai (邁話, related to the Yue dialects according to Ouyang et al. 2005), Jun (軍話, literally ‘army speech’, a northern dialect brought by the army during the Ming dynasty, see Qiu 2005),

^{*} The work reported in this paper is supported by a Competitive Earmarked Research Grant (CERG) from the Research Grant Council, Hong Kong [Project number: 9041068 — *From the Language Atlas of China to Language and Ethnicity: Research on Linguistic Demography*]. An earlier version of this paper was presented at the 14th Annual Conference of the International Association of Chinese Linguistics (IACL) in 2006.

Danzhou (儋州話, see Ting 1986) and Danjia (蜆家, a variety of the Cantonese dialects spoken by boat people). Other non-Sinitic languages include Li (黎話, of the Tai-Kadai family, see Ouyang & Zheng 1980, 1983), Huihui (回輝話, an Austroasiatic language related to the Cham language of Vietnam and Cambodia, see Zheng 1997) and Miao (苗話, a member of the Hmong-Mien family).

Approximately 15,000 people speak Mai (Ouyang et al. 2005) with about one-third living in the village of Yanglan (羊欄村), in Phoenix Township (鳳凰鎮), near the Sanya international airport. There is also a large population of Mai speakers in the village of Shuinan (水南村), Yacheng (崖城), roughly 50 kilometers west of Sanya.

Little research has been conducted on the Mai dialect, especially in terms of syntax. One of the reasons for this is that the linguistic affiliation of the Mai dialect is still unclear. Liang (1984) considers Mai to be a variety of Cantonese, while Huang & Li (1987), by comparing the Mai dialect with the dialects of Danzhou (儋州), Wenchang (文昌), Taishan (台山) and Linchuan (臨川), claim that Mai contains features of Cantonese, Min and Gan. According to their extensive fieldwork, Jiang, Ouyang & Tsou (2007) observe that Mai has two vowels which basically correspond to the long and short vowels [a] and [ə] in Standard Cantonese. Furthermore, Mai has three entering tones (i.e. 上陰入, 下陰入 and 陽入). On the other hand, Mai also possesses some features such as the implosive initials [ʔb] and [ʔd] that are typical of the Min dialect of Hainan. The reflexes of Middle Chinese voiced stops and affricate initials become voiceless aspirated stops in all tones in Mai. This feature is typical across most Gan and Hakka dialects. These phonological traits strongly suggest that Mai has assimilated phonological features from different dialects in the course of development.

In terms of syntax, Mai is a typical Southern dialect. For example, the comparative construction has the structure [NP₁ V 過 NP₂]. The sentence ‘he is taller than me’ is rendered as [ki⁵⁵ kou¹³ kua³³ ŋɔ³³] 佢高過我 ‘s/he-high-surpass-I’. The double-object construction has the direct object preceding the indirect object, such as [ŋɔ³³ vɔn¹³ aʔ⁵ bun³³ θi¹³ vɔn¹³ ki⁵⁵] 我分一本書分佢 ‘I-give-a-CL-book-give-s/he, I gave a book to him/her’.¹ The gender marker comes after the head noun it modifies. Examples include [ŋau⁵⁵ kuŋ¹³] 牛公 ‘bull’ and [kɔi¹³ na³³] 雞乸 ‘hen’.

Another syntactic structure deserving our attention is the resultative verb construction (結果補語).² According to our fieldwork on Mai in 2005, we found that the resultative verb construction has the word order of verb-object-complement, or VOC for short. Consider (1):

¹ The use of 分 as the double-object verb may be an influence of the Hakka dialect.

² Besides the resultative-verb construction, there are other types of verb-complement structures such as extent complement, phase complement and potential complement. In this paper, focus is placed only on the resultative complement.

- (1) ki⁵⁵ tshe³³ ŋɔ³³ ʔdi³³ θan¹³ ʔbɔŋ¹³
 渠 扯 我 的 衫 崩
 s/he tear I DE shirt broken
 ‘S/he tore apart my shirt.’

In (1), [tshe³³] 扯 ‘to tear’ is the main verb (V), [ʔbɔŋ¹³] 崩 ‘broken’ is the resultative complement (C) and [ŋɔ³³ ʔdi³³ θan¹³] 我的衫 ‘I-DE-shirt, my shirt’ is the object (O). The sentence has the word order of **VOC**.

At the same time, it is noted that sentence (1) can also be expressed with the word order of **VCO**. Consider (2):

- (2) ki⁵⁵ tshe³³ ʔbɔŋ¹³ ŋɔ³³ ʔdi³³ θan¹³
 渠 扯 崩 我 的 衫
 s/he tear broken I DE shirt
 ‘S/he tore apart my shirt.’

From the point of view of communication and cognitive processing, it can be argued that it is more effective and economical for a language to have one surface form for each syntactic construction. The co-existence of two word order patterns for the resultative verb construction in Mai thus raises some interesting questions. For example, which pattern (VOC or VCO) is native in Mai? Why and how did the alternate pattern come into existence?

2. VOC and VCO in the Mai dialect

In the following sections, we will trace the relative chronology of the two patterns (VOC and VCO) in the Mai dialect. We claim that VOC is an older pattern while VCO is a new pattern borrowed from Standard Chinese as a result of language contact. Our argument is based on linguistic and extra-linguistic factors. These include (a) iconicity; (b) the historical development of the resultative verb construction in the Chinese language; (c) the migratory history of the Mai people from the mainland to Hainan; and (d) language contact between Mai and non-Sinitic languages such as Li.

2.1 VOC and VCO in terms of iconicity

When we compare VOC and VCO, we note that the former reflects the temporal sequence of the action and the consequence observed in the real world: an action (V) is instigated by an agent on the object (O) which then reaches a resultant state (C). The

VCO pattern, on the other hand, does not fully conform to the temporal sequence of the action because the resultative complement (C) precedes the object (O).

It has been claimed that human language is not totally arbitrary in the sense that an iconic relationship between language and human experience can exist. In other words, our grammar can be shaped by iconicity. This suggests that the surface form of a syntactic construction usually begins in an order that parallels our experience in the real world (see Greenberg 1966, and Haiman 1980, 1985). In this regard, we have support for the postulation that VOC predates VCO in Mai.

In addition to the resultative verb construction, there are other syntactic constructions in Chinese that are iconically motivated according to the temporal sequence of the action (see, for example, Tai 1993). Consider the following example:

- (3) 張三 從 香港 經過 珠海 到 澳門
 Zhang San cong Xianggang jingguo Zhuhai dao Aomen
 Zhang-San from Hong Kong pass Zhuhai arrive Macau
 ‘Zhang San went to Macau from Hong Kong through Zhuhai.’

It is clear that in (3), the sequence of the three locations follow the actual route that Zhang San took, i.e. starting from Hong Kong, then to Zhuhai and finally to Macau. It is interesting to note that such a parallel to the temporal sequence is not found in English.

If iconicity plays a role in shaping Chinese grammar, we would then expect that the same phenomenon (i.e. VOC predating VCO) could also be found in other dialects of the Chinese language. In the next section, we examine the diachronic development of the resultative verb construction in the Chinese language.

2.2 The diachronic development of the resultative verb construction in Chinese

The resultative verb construction in most modern Chinese dialects has the VCO order. Consider the following examples:

- (4) 他 [打]_v [破]_c 了 [杯子]_o (Standard Chinese)
 ta da po le beizi
 s/he hit break ASP cup
 ‘S/he broke the cup.’

- (5) 佢 [打]_v [爛]_c 咗 [隻 杯]_o (Hong Kong Cantonese)
 k'ey¹³ ta³⁵ lan²² tso³⁵ tsek³ pui⁵⁵
 s/he hit break ASP CL cup
 'S/he broke the cup.' (他打破了杯子)
- (6) 快滴 [洗]_v [淨]_c [啊手]_o (Meixian 梅縣 Hakka)
 quick ADV wash clean CL hand (Li & Chang 1992:455)
 'Wash your hands clean quickly.' (快點把手洗乾淨)
- (7) 他 眼睛 一鼓 噠, [嚇]_v [死]_c [人]_o 的 (Changsha 長沙 Xiang)
 he eye once open-wide scare die person SFP (Bao et al. 1999:311)
 'Once he opens his eyes wide, it can scare people to death.' (他的眼睛瞪得老大, 嚇死人)
- (8) 衣裳 [脫]_v [落]_c [它]_o (Wuxi 無錫 Wu)
 clothes strip off it (Qian 1992:1010)
 'strip off the clothes' (脫掉衣服)

However, it has been reported that the VCO pattern only came into existence around the Southern and Northern dynasties (南北朝) (see, for example, Jiang 2000, Yue 2001 and Shi & Li 2001).³ Around the same time, VOC was another pattern expressing the resultative verb meaning. For example,

- (9) 今 當 [打]_v [汝 前 兩 齒]_o [折]_c (*Xian yu jing* j. 11 賢愚經, 卷 11)
 jin dang da ru qian liang chi zhe
 now should beat you front two teeth broken
 'Now someone should break your two front teeth.'
- (10) 即 便 以 嘴 [啄]_v [雌鴿]_o [殺]_c (*Bai yu jing* 百喻經, *Erge yu* 二鴿喻)
 ji bian yi zui zhao cige sha
 immediately then use beak peck female pigeon die
 'Then peck the female pigeon with its beak which causes it to die'
- (11) 周 仲智 [飲]_v [酒]_o [醉]_c (*Shishuo xinyu*, *Yaliang* 世說新語, 雅量)
 Zhou Zhongzhi yin jiu zui
 Zhou Zhongzhi drink liquor drunk
 'Zhou Zhongzhi drank and got drunk.'

³ Notice that the complement (C) is also a verb (stative or intransitive). When tracing the historical development of VCO, scholars distinguished VCO from the V₁V₂O structure which has two concatenating verbs followed by an object.

Starting in the Song dynasty, the use of the VOC pattern began to decline and the VCO pattern become dominant (see Shi & Li 2001 and Jiang & Cao 2005). In other words, VOC was a major pattern for the resultative verb construction until the Song dynasty, after which point VCO replaced VOC. The historical development of the resultative verb construction in the Chinese language supports our claim that VOC predates VCO.

2.3 The development of the Mai dialect

Besides language-internal factors, there are extra-linguistic factors which can account for the dominance of VOC in Mai.

The migratory history of Mai speakers can also give us some clues on the relative chronology of VOC and VCO. According to the genealogical record of a Mai speaking informant's family, their ancestors migrated from the mainland to Wenchang (文昌), Hainan, around the late 10th century. They subsequently moved further southward along the east coast and then settled in the Sanya area. When combining the demographic history of the Mai people and the diachronic development of the resultative verb construction in the Chinese language, as outlined in §2.2, we can postulate that the language of the early Mai people brought to Hainan still used the VOC order.⁴ The language then became isolated from other Sinitic languages and did not follow mainstream developments. Therefore, VOC continued to be used as the dominant pattern and did not undergo any change in word order to VCO, as in Standard Chinese.

2.4 Language contact

As discussed in §1, there are at least eight linguistic varieties spoken in Sanya. In other words, most people are multilingual and mutual influence among these languages is expected (see, for example, the various phonological features of Mai described in §1). Therefore, there has been on-going language contact among these linguistic varieties. In particular, the Li language (黎話) of the Tai-Kadai family deserves our attention. The Li people are understood to be the earliest settlers on Hainan.⁵ It is interesting to note that Li also uses a VOC resultative verb construction. Consider (12) and (13):

⁴ It is irrelevant to our argument whether or not a distinct Mai dialect was in existence when the Mai ancestors moved to Hainan.

⁵ The first settlers of other languages or dialects such as Jun, Huihui and Min went to Hainan much later.

- (12) 吃 草 飽 那 一 隻 水牛 (Ouyang & Zheng 1980:100)
tei⁵ kə:n⁴ khəm¹ kɛ⁴ ku² puən⁴ tshei¹
eat grass full that one CL buffalo
'the buffalo which has been well fed with grass'
- (13) 我 做 工 完 也 休息 (Ouyang & Zheng 1983:568)
hou¹ vu:k⁷ koŋ¹ ba:i³ ŋan¹ tshu:k⁷
I do work finish also rest
'I will rest after finishing the work.'

In other words, contact between the Li language and Mai has facilitated the use of VOC.

2.5 Summary

We have considered different factors to work out the relative chronology of the development of VOC and VCO patterns in Mai. All of these factors have combined to suggest that VOC predates VCO. The development of the resultative verb construction in the Mai dialect can be outlined as follows:

When the Mai people migrated to Hainan in the late 10th century, the dominant pattern of the resultative verb construction in the language they brought with them was VOC. Their language became isolated and did not follow the language in the mainland to develop into VCO. Furthermore, the contact with the Li language, which also uses VOC, has contributed to maintaining VOC as the dominant pattern.

3. The increasing use of VCO in Mai

While VOC is claimed to be the native pattern in the Mai dialect, it is found that VCO is also used, especially among young and educated speakers. It is worth exploring the reason for the increasing use of VCO, and an obvious factor deserving consideration could be the influence of Standard Chinese, or Modern Standard Chinese, in which VCO is used.⁶

⁶ A similar situation of language contact is also found in the Cantonese dialects of Nanning (南寧) and Baise (百色) as well as in Pinghua (平話) of the Central Guangxi region where both VOC and VCO patterns also co-exist. According to Kwok (2010), the VOC pattern was introduced from the Zhuang language. For details on the use of VOC and VCO in the Nanning Cantonese dialect, see Lin & Qin (2008).

The fast-changing sociolinguistic situation of Sanya in the last two decades has given rise to a situation in which Standard Chinese has become a relatively important linguistic variety in the community because it is the official language used by the government and in schools. One significant consequence is that some native speakers of Mai (such as school children) have shifted their home language to Standard Chinese so that a growing number of young Mai speakers are using Standard Chinese more frequently than their mother tongue, especially with their siblings and parents (Tsou et al. 2007a, 2007b, 2010, and Chin, Tsou & Mok 2008).

In addition, the syntax of Mai has also undergone significant change due to the impact of Standard Chinese. Aside from the resultative verb construction, we also note that some Mai speakers have begun to use the double-object construction with the indirect object preceding the direct object such as [ŋɔ³³ vɔn¹³ ki⁵⁵ aɿ⁵ bun³³ θi¹³] ‘我分佢一本書’ for ‘I gave him/her a book’.⁷

4. A survey on the use of VOC and VCO in Mai

To gain a better understanding on the use of VOC and VCO in Mai, we did a preliminary study with three native Mai speakers in 2005. The background information of the three speakers is summarized below.

	Speaker 1	Speaker 2	Speaker 3
Age	70	55	25
Education level	Junior secondary	Junior secondary	Junior secondary
Occupation	Retired (former accounting clerk)	Village secretary	Shop assistant
Mother tongue	Mai	Mai	Mai
Knowledge of other languages or dialects	Standard Chinese, Hainanese	Standard Chinese, Hainanese	Standard Chinese, Hainanese
Language used at home	Mai	Mai	Mai
Languages used outside home/village	Standard Chinese, Hainanese	Standard Chinese, Hainanese	Standard Chinese, Hainanese

⁷ The Li language is also found to be influenced by the Standard Chinese resultative verb construction. For example, the resultative verb construction can also be expressed with the VCO pattern (see Example (12) above): 我做完工也休息 (Ouyang & Zheng 1983:568).

While Mai is the mother tongue of these three speakers, they all also use Hainanese with other locals as well as Standard Chinese when they are outside of their local communities, in areas such as downtown and when using public transportation. All three also regularly watch TV and most programs are broadcast in Standard Chinese.

The survey was based on a list of about 30 sentences containing the resultative verb construction. The speakers were asked to produce the sentences in the Mai dialect first. If the VOC or the VCO pattern was used, the speaker would then be checked with the usage of the alternate pattern. Some examples are given below:

- (14) 火燒死三個人 hua³³ θiu¹³ θi³³ θan¹³ kɔ²¹ nɔn⁵⁵ [VCO]
 火燒三個人死 hua³³ θiu¹³ θan¹³ kɔ²¹ nɔn⁵⁵ θi³³ [VOC]
 ‘The fire burnt three people to death.’
- (15) 剖開西瓜 pɔ⁵⁵ hai¹³ θɔi¹³ kua¹³ [VCO]
 剖西瓜開 pɔ⁵⁵ θɔi¹³ kua¹³ hai¹³ [VOC]
 ‘Cut open the water melon.’
- (16) 渠扭斷狗的項骨 ki⁵⁵ nau³³ thuan²¹ kau³³ ʔdi³³ huan²¹ kut³ [VCO]
 渠扭狗的項骨斷 ki⁵⁵ nau³³ kau³³ ʔdi³³ huan²¹ kut³ thuan²¹ [VOC]
 ‘S/he twisted and broke the dog’s neck.’
- (17) 渠砍斷木幹 ki⁵⁵ khɛ^{ʔ5} thuan²¹ muk³ kuan²¹ [VCO]
 渠砍木幹斷 ki⁵⁵ khɛ^{ʔ5} muk³ kuan²¹ thuan²¹ [VOC]
 ‘S/he chopped off the piece of wood.’
- (18) 渠的車碰死人 ki⁵⁵ ʔdi³³ tsha¹³ phuŋ³³ θi³³ nɔn⁵⁵ [VCO]
 渠的車碰人死 ki⁵⁵ ʔdi³³ tsha¹³ phuŋ³³ nɔn⁵⁵ θi³³ [VOC]
 ‘His/her car ran over and killed a man.’
- (19) 渠打傷先生 (i.e. 老師) ki⁵⁵ ʔda³³ θiŋ¹³ θen¹³ θeŋ¹³ [VCO]
 渠打先生傷 ki⁵⁵ ʔda³³ θen¹³ θeŋ¹³ θiŋ¹³ [VOC]
 ‘S/he beat the teacher.’

The distribution of VOC and VCO used by the three speakers is summarized below.⁸

⁸ The total number of sentences obtained from each speaker differs because some sentences were not necessarily rendered with a resultative-verb construction. For example, the sentence 他打開了門 was expressed as 他開了門.

	Speaker 1	Speaker 2	Speaker 3
VOC only	12	13	4
VCO only	0	1	10
VOC or VCO	14	8	4

Some key observations are obtained on the use of the two patterns. First, the youngest speaker (i.e. speaker 3) used the VOC the least, at a rate equivalent to only one-third that of the other two older speakers. It is interesting and important to note that for some sentences, speakers 1 and 2 allow the use of either VOC or VCO. This represents a transitional stage in which the shift of VOC to VCO is ongoing. However, their acceptance of VCO as the *exclusive* pattern is still low when compared with speaker 3.

5. Concluding remarks

In this paper, we examine the possible origin and the development of the resultative verb construction in the Mai dialect. We argue that the VOC and VCO patterns represent two syntactic strata in the dialect.

The development of VOC to VCO involves the displacement of the resultative complement. Displacement of constituents is not only found in the resultative verb construction. Other examples include the locative phrase and the dative construction. Peyraube (1996:181) has argued that the word order change of the dative construction from V1 + V2 + IO + DO to V1 + DO + V2 + IO was probably analogical to the serial verb construction with the pattern V1 + O1 + V2 + O2, which could also be the origin of the resultative verb complement construction.⁹

Although Peyraube (1996) mainly explained the historical development of Chinese syntax by means of language-internal factors, such as analogy, reanalysis and grammaticalization, he also emphasized that external borrowing should be taken into consideration as well. In this paper, we take an integral approach to show how both language-internal and language-external factors can enhance our understanding of the development of the resultative verb construction in the Mai dialect.

⁹ For more details on the relationship between the serial verb construction and the resultative-verb construction, see Li (1987).

References

- Bao, Houxing, Zhenhua Cui, Ruoyun Shen, and Yunji Wu. 1999. *Changsha Fangyan Yanjiu* [A Study of the Changsha Dialect]. Changsha: Hunan Education Press.
- Chin, Andy Chi-on, Benjamin K. Tsou, and Kenny Mok. 2008. Guangxi Nanningshi he Hainansheng Sanyashi Yue zuqun yuyan zhuanyi xianxiang bijiao [A comparison of language shift between the Cantonese speakers in Nanning of Guangxi and Sanya of Hainan]. Paper presented at the 13th International Conference on the Cantonese and Yue Dialects, December 18-20, 2008. Hong Kong: City University of Hong Kong.
- Greenberg, Joseph H. 1966. Some universals of grammar, with particular reference to the order of meaningful elements. *Universals of Language*, ed. by Joseph Greenberg, 73-113. Cambridge: MIT Press.
- Haiman, John. 1980. The iconicity of grammar. *Language* 53.2:312-328.
- Haiman, John. 1985. *Iconicity in Syntax*. Amsterdam & Philadelphia: John Benjamins.
- Huang, Gudan, and Rulong Li. 1987. Hainandao de Maihua: yi zhong hunhexing de fangyan [The Mai dialect in Hainan: a hybrid type dialect]. *Zhongguo Yuwen* [Studies of the Chinese Language] 1987.4:268-275.
- Jiang, Di, Jueya Ouyang, and Benjamin K. Tsou. 2007. Hainansheng Sanyashi Maihua yinxi [The phonological system of the Mai dialect of Sanya in Hainan Province]. *Fangyan* [Dialect] 2007.1:23-34.
- Jiang, Shaoyu. 2000. *Hanyu Cihui Yufashi Lunwenji* [A Collection of Essays on the Historical Development of Chinese Lexicon and Syntax]. Beijing: The Commercial Press.
- Jiang, Shaoyu, and Guangshun Cao. 2005. *Jindai Hanyu Yufashi Yanjiu Zongshu* [A Summary of the Research on the Historical Development of Syntax in Pre-modern Chinese]. Beijing: The Commercial Press.
- Kwok, Bit-chee. 2010. Grammatical change in language contact: on the origin of the “verb + object + complement” in Nanning Yue. *Diachronic Change and Language Contact: Dialects in South East China*, ed. by Hung-nin Samuel Cheung & Song-Hing Chang, 201-216. Berkeley: Project on Linguistic Analysis.
- Li, Ping. 1987. Shishuo Xinyu he Baiyujing zhong de dongbu jieyou [The resultative verb construction in *Shishuo Xinyu* and *Baiyujing*]. *Yuyanxue Luncong* [Essays on Linguistics] 14:129-157. Beijing: The Commercial Press.
- Li, Rulong, and Song-Hing Chang. 1992. *Ke Gan Fangyan Diaocha Baogao* [A Report on the Hakka and Gan Dialects]. Xiamen: Xiamen University Press.
- Liang, Yougang. 1984. Guangdongsheng Hainandao Hanyu fangyan de fenlei [The classification of the Chinese dialects spoken in the Hainan Island of Guangdong Province]. *Fangyan* [Dialect] 1984.4:264-267.

- Lin, Yi, and Fengyu Qin. 2008. *Guangxi Nanning Baihua Yanjiu* [A Study of the Cantonese Dialect Spoken in Nanning of Guangxi]. Guilin: Guangxi Normal University Press.
- Ouyang, Jueya, and Yiqing Zheng. 1980. *Liyu Jianzhi* [A Brief Description of the Li Language]. Beijing: Nationalities Press.
- Ouyang, Jueya, and Yiqing Zheng. 1983. *Liyu Diaocha Yanjiu* [A Study of the Li Language]. Beijing: China Social Sciences Press.
- Ouyang, Jueya, Benjamin K. Tsou, and Di Jiang. 2005. Sanya Maihua de guishu wenti [The affiliation of the Mai dialect of Sanya]. Paper presented at the 10th International Conference on the Yue Dialects, December 12-14, 2005. Hong Kong: The Chinese University of Hong Kong.
- Peyraube, Alain. 1996. Recent issues in Chinese historical syntax. *New Horizons in Chinese Linguistics*, ed. by C.-T. James Huang & Y.-H. Audrey Li, 161-214. Dordrecht: Kluwer.
- Qian, Nairong. 1992. *Dangdai Wuyu Yanjiu* [A Study of the Modern Wu Dialects]. Shanghai: Shanghai Education Press.
- Qiu, Xueqiang. 2005. *Junhua Yanjiu* [A Study of the Jun Dialect]. Beijing: China Social Sciences Press.
- Shi, Yuzhi, and Charles N. Li. 2001. *Hanyu Yufahua de Licheng: Xingtai Jufa Fazhan de Dongyin he Jizhi* [A History of Grammaticalization in Chinese: Motivations and Mechanisms of Evolution of Chinese Morpho-syntax]. Beijing: Peking University Press.
- Tai, James H.-Y. 1993. Iconicity: motivations in Chinese grammar. *Principles and Prediction: The Analysis of Natural Language*, ed. by Mushira Eid & Gregory K. Iverson, 153-174. Amsterdam & Philadelphia: John Benjamins.
- Ting, Pang-Hsin. 1986. *Danzhou Cunhua* [The Dialect of Danzhou]. Taipei: Institute of History and Philology, Academia Sinica.
- Tsou, Benjamin K., Andy Chi-on Chin, Jueya Ouyang, Carol To, Kenny Mok, and Weihui Yang. 2007a. Critical multilingual shift in Sanya, China: accelerated urbanization and possible sociolinguistic repair. *Proceedings of the Eleventh Conference of the Foundation for Endangered Languages*, ed. by Maya David, Nicholas Ostler & Caesar Dealwis, 67-74. Kuala Lumpur: University of Malaya.
- Tsou, Benjamin K., Jueya Ouyang, Di Jiang, and Andy Chi-on Chin. 2007b. Sanya Maihua zuqun de yuyan shiyong zhuangkuang [Language use by the Mai speakers of Sanya]. *Di Shiyi Jie Guoji Yue Fangyan Yantaohui Lunwenji* [Proceeding of the 11th International Conference on the Yue Dialects], ed. by Yi Lin & Jin Yu, 356-364. Nanning: Guangxi People's Publishing House.

- Tsou, Benjamin K., Andy Chi-on Chin, Kenny Mok. 2010. Accelerated urbanization, triglossia and language shift: a case study of Sanya of China's Hainan Province. *Industrialization and the Restructuring of Speech Communities in China and Europe*, ed. by M. van den Berg & Daming Xu, 269-282. Cambridge: Cambridge Scholars Publishing Company.
- Wurm, Stephen A., Benjamin K. Tsou, and David Bradley. (eds.) 1987. *Language Atlas of China*. Hong Kong: Longman.
- Yue, Anne O. 2001. The verb-complement construction in historical perspective with special reference to Cantonese. *Synchronic and Diachronic Perspectives on the Grammar of Sinitic Languages*, ed. by Hilary Chappell, 232-265. Oxford & New York: Oxford University Press.
- Zheng, Yiqing. 1997. *Huihuihua Yanjiu [A Study of the Huihui Language]*. Shanghai: Shanghai Far East Publisher.

南方民族語言比較句語序的演變和變異

吳福祥

中國社會科學院

中國南方的民族語言（侗台、苗瑤、南亞及南島）裡，差比句基準與形容詞的語序有“Adj-M-St”、“M-St-Adj”和“Adj-M-St/M-St-Adj”三種類型。本文從接觸語言學和語言類型學角度論證南方民族語言比較句的固有語序是 Adj-M-St，部分語言出現的 M-St-Adj 模式則是接觸引發的“Adj-M-St > M-St-Adj”演變和變異的產物。觸發這種演變和變異的“模式語”是漢語，而演變的主要機制是“語序重組”。

關鍵詞：差比式，語序演變，語法複製，語序重組

1. 引言

本文基於接觸語言學架構和類型學視野討論中國南方民族語言的比較句（差比句）語序受漢語影響而發生的演變或變異。¹ 著重探討的問題是：(i) 哪些語言的比較句語序發生了演變或變異，哪些語言還保持固有語序？(ii) 這種語序演變或變異的機制是什麼？(iii) 如何判定這種語序演變或變異是與漢語接觸而導致的？²

本文以侗台、苗瑤、南亞及南島等語系或語族的四十餘種語言（見〈表1〉）為考察對象，材料取自相關的調查報告、語言志、參考語法以及研究文獻。這些語言在類型上跟漢語大同小異：(a) 小句語序為 SVO；(b) 使用前置詞而

¹ 本文的“（中國）南方民族語言”範圍限於侗台、苗瑤、南亞及南島等語系或語族的語言，不包括藏緬語。另一方面，本文的“比較句”指的是差比句。

² 比較句作為一種跨語言普遍可見的結構式，由比較主體（記作 S）、基準（記作 St）、形容詞（記作“Adj”）和比較標記（記作“M”）四個成分構成。類型學家對比較句語序的跨語言比較和研究主要基於兩個句法參項，即基準、形容詞“成分對”（pair of elements）的語序和基準、比較標記成分對的語序。本文對南方民族語言比較句語序的考察限於基準、形容詞成分對，換言之，本文所謂南方民族語言裡比較句語序的演變和變異，實際上指的是比較句基準和形容詞語序的演變和變異。這是因為，本文所考察的語言均為 SVO 語言，而在這樣的語言裡基準、比較標記成分對的語序均為“比較標記-基準”（M-St）。

非後置詞；(c) 具有運動式結構；(d) 形態上屬於分析性語言，缺少屈折變化；(e) 有體範疇而無時範疇；(f) 有量詞範疇而無（嚴格意義上的）數範疇；(g) 大多具有聲調 (contour tones) 範疇。

本文的研究方法是：給定一個“成分對”(pair of elements) A 和 B，如果它在同一個語系、語族或語支的不同語言裡具有“A-B”、“A-B/B-A”和“B-A”三種語序類型，而後者(B-A)正好跟漢語的模式一致，那我們先假設，這些語言固有的語序模式是“A-B”，“B-A”是受漢語影響而發生語序演變的產物，而“A-B/B-A”則顯示處在“A-B > B-A”演變之中的變異階段；換言之，我們把上述三種語序類型分別視為“未變(A-B)”、“在變(A-B/B-A)”和“已變(B-A)”三個演變階段(第 2 節)。然後，從類型學的語序關聯和發生學關係分布模式等角度來對上述假設進行證明(第 3 節)。最後，我們基於接觸語言學的新近成果討論上述語序演變或變異的機制(第 4 節)。

2. 南方民族語言比較句語序的演變和變異

世界語言中，比較句基準與形容詞的語序有 Adj-St [形容詞-基準] 和 St-Adj [基準-形容詞] 兩種類型。比如英語、法語的比較句採用 Adj-M-St 模式，日語、漢語則分別使用 St-M-Adj、M-St-Adj 模式（日語和漢語比較標記的位置雖有不同，但基準均在形容詞之前）：

- (1) 英語：you are more pretty than your sister.
你 是 更 漂亮 比較標記 你的 姐妹
Adj M St
你比你姐姐漂亮。
- (2) 法語：tu es plus jolie que ta sœur.
你 是 更 漂亮 比較標記 你的 姐妹
Adj M St
你比你姐姐漂亮。
- (3) 日語：Tokyo-wa Sapporo-yori atatakai.
東京-話題標記 札幌-比較標記 暖和
東京比札幌暖和。
- (4) 漢語：你 比 你姐姐 漂亮。
M St Adj

在我們考察的南方民族語言裡，比較句基準與形容詞的語序有“Adj-M-St”、“M-St-Adj”和“Adj-M-St/M-St-Adj”三種類型。³我們假定，這些民族語言比較句固有的語序模式同於英語、法語（即 Adj-M-St）而異於漢語，而“M-St-Adj”和“Adj-M-St/M-St-Adj”則是受漢語影響而發生語序演變或變異的結果。⁴

2.1 侗台語

侗台語中，只有壯傣語支的傣語和仡央語支的拉基語比較句的語序未發生演變和變異，換言之，這兩個語言比較句的基準和形容詞的語序保持固有的 Adj-M-St 模式。例如：

傣語（王均等 1984）

Adj-M-St:

- (5) a. van²ni⁸ hən⁴ lə¹ van²suən². (西雙版納傣語)
 今天 熱 比 前天
 今天比前天熱。
 b. xau¹tsau³ jai⁵ lə¹ to¹xa³ sən¹ pi¹. (西雙版納傣語)
 他 大 比 我 兩 歲
 他比我大兩歲。

拉基語（李雲兵 2000）

Adj-M-St:

- (6) a. pi³⁵ŋe⁴⁴ ze³⁵ a⁵⁵ vua³⁵ na⁴⁴nu³⁵.
 今年 稻 好 過 去年
 今年的稻子比去年的好。
 b. vu⁴⁴ ki⁵⁵ to³¹ sei³⁵ laŋ⁵⁵ la⁴⁴ m⁵⁵.
 碗 我 大 過 個 的 你
 我的碗比你的大。

³ 本文對這些語言裡比較句語序類型的判定（Adj-M-St、M-St-Adj 還是 Adj-M-St/M-St-Adj），主要依據相關語言描寫者的判斷和舉例。

⁴ 如附注 2 所述，在本文所涉及的語言裡比較標記和基準的語序均為“M-St”。因此，為討論的方便，下文在描述南方民族語言比較句基準和形容詞的語序模式時，我們加上基準這一成分。

另一方面，侗水語支的標話、拉珈語以及仡央語支的普標語、仡佬語（平壩）、木佬語受漢語影響，比較句的基準相對於形容詞的位置已由 Adj-M-St 變為 M-St-Adj，也就是說，這些語言比較句的基準和形容詞的語序通常已不再採用固有的 Adj-M-St 模式。例如：

標話（梁敏、張均如 2002）

M-St-Adj:

- (7) a. to³kək⁹ kɛ⁶ jo⁶ pi¹ to³ tsak⁷ kɛ⁶ jo⁶ pai².
山腳 的 田 比 山 側 的 田 肥
山腳的田比山坡的田肥。
b. tsia¹ kɛ⁶ pi¹ man² kɛ⁶ lɛ¹ li³ fɔ³ ni³!
我 的 比 他 的 好 得 多 呢
我的比他的好得多呢！

普標語（梁敏、張均如、李雲兵 2007）

M-St-Adj:

- (8) a. kau⁵³ pi⁴⁵ mi³³ qha:ŋ³³, ku⁵⁵ pi⁴⁵ kau⁵³ ken²¹³ qha:ŋ³³.
我 比 你 高 他 比 我 更 高
我比你高，他比我更高。
b. pa²¹³ mu⁵³ ku⁵⁴ nai⁴⁵ pi⁴⁵ tau⁵³ pa²¹³ ŋa⁴⁵ nin³³.
隻 豬 他 這 比 三 隻 那 肥
他這隻豬比那三隻肥。

木佬語（薄文澤 2003）

M-St-Adj:

- (9) a. tɕhi²⁴tshe³³ pi³³ tho³³la³³tei³³ tsau⁵³.
汽車 比 拖拉機 快
汽車比拖拉機快。
b. lai⁵³ŋo⁵³ pi³¹ lai⁵³my²⁴ luŋ²⁴.
馬 比 豬 大
馬比豬大。

最後，有些語言比較句的基準和形容詞的語序有 Adj-M-St 和 M-St-Adj 兩種模式，也就是說，在這些語言裡比較句的基準相對於形容詞的位置正處在“Adj-

M-St > M-St-Adj”演變之中。具體說來，這種語序類型約有以下三種情形。⁵

第一，有些語言以固有的 Adj-M-St 為主要或優勢模式，M-St-Adj 為少見或次要模式；即呈現的是“Adj-M-St/m-st-adj”類型。⁶ 黎語支的黎語即屬於這種情形：

黎語（王均等 1984，李旭練 1999）

Adj-M-St:

- (10) a. meʊ¹ phe:k⁷ dua³ hou¹.
 你 高 過 我
 你比我高。
 b. pai³gu:ŋ¹ leŋ¹mu:n¹ dua³ lu:k⁷khau³.
 妹妹 漂亮 過 姐姐
 妹妹比姐姐漂亮。（李旭練 1999）

M-St-Adj:

- (11) pai³gu:ŋ¹ bi¹ lu:k⁷khau³ leŋ¹mu:n¹.
 妹妹 比 姐姐 漂亮
 妹妹比姐姐漂亮。（李旭練 1999）

第二，有些語言裡 M-St-Adj 和 Adj-M-St 均為常見模式，不易確定或未有資料顯示哪種模式為優勢或主要模式，即呈現“M-St-Adj/Adj-M-St”類型。屬於這種情形的語言有：壯傣語支的壯語、布依語、臨高語，侗水語支的侗語、仫佬語、水語、毛難語、茶洞語、佯僜語、莫語，以及仡央語支的布央語。例如：

壯語（王均等 1984）

Adj-M-St:

- (12) a. sa:ŋ¹ kva⁵ muŋ².
 高 過 你
 比你高。

⁵ 跟 Adj-M-St、M-St-Adj 和 Adj-M-St/M-St-Adj 三種類型的判定一樣，我們對下面三種變異情形的判斷也主要基於相關語言描寫者的斷言和舉例。由於資料的局限，本文對個別語言比較句語序變異類型的判斷可能未盡準確，但這並不影響本文的基本結論。重要的是，所有相關的語言，其比較句的語序均已發生變異（處於演變過程中），差別只是變異程度的不同。

⁶ 如果某個語言中“成分對”A 和 B 的語序具有交替形式 (AB/BA) 且其間存在優勢和非優勢之別，類型學家通常將標記優勢模式的字母大寫，標記非優勢模式的字母小寫。

吳福祥

M-St-Adj:

- b. pei³ muŋ² sa:ŋ¹.
比 你 高
比你高。

臨高語（梁敏、張均如 1997）

Adj-M-St:

- (13) a. en³ kua³ lian¹.
好吃 過 蜜糖
比蜜甜。

M-St-Adj:

- b. hau² bi⁴ mə² haŋ¹.
我 比 你 高
我比你高。

侗語（王均等 1984）

Adj-M-St:

- (14) a. oŋ¹ məi⁴ nai⁶ pha:ŋ¹ ta⁶ oŋ¹ ta⁵.
棵 樹 這 高 過 棵 那
這棵樹比那棵樹高。

M-St-Adj:

- b. ma:u⁶ pi³ ŋa² jaŋ^{3'}.
他 比 你 強壯
他比你強壯。

莫語（楊通銀 2000）

Adj-M-St:

- (15) a. pəu⁴ man¹ te⁵ ta⁶ nəi⁴ man¹.
父 他 大 過 母 他
他父親比他母親年紀大。

M-St-Adj:

- b. pəu⁴ man¹ pi³ nəi⁴ man¹ te⁵.
父 他 比 母 他 大
他父親比他母親年紀大。

布央語（李錦芳 1999）

Adj-M-St:

- (16) a. kɛ⁵⁴ va:ŋ³¹² ma²⁴ ku⁵⁴.
 他 高 比 我
 他比我高。

M-St-Adj:

- b. a⁰ta:p¹¹ pa:i⁵⁴ ni¹¹ pi³³ a⁰ta:p¹¹ ŋa¹¹ qa⁰naŋ¹¹.
 擔 穀 這 比 擔 那 重
 這擔穀子比那擔重。

第三，有的語言以 M-St-Adj 為主要或優勢模式，Adj-M-St 則為少見或次要模式。換言之，這類語言固有的 Adj-M-St 模式正處在被外來模式 M-St-Adj 淘汰之中。黎語支的村語大約屬於這種情形：

村語（歐陽覺亞 1998）

Adj-M-St:

- (17) a. kə⁵ mɛi¹ kuə³ mɔ⁵.
 我 高 過 你
 我比你高。

M-St-Adj:

- b. kə⁵ bɛi³ mɔ⁵ mɛi¹.
 我 比 你 高
 我比你高。
 c. mɔ⁵ bɛi³ na⁵ khon¹ fu³ ken⁵.
 你 比 他 重 三 斤
 你比他重三斤。

2.2 苗瑤語

跟上舉侗台語不同，苗瑤語中所有語言的比較句語序均在不同程度上發生了演變或變異。其中巴哼語、炯奈語、畚語、勉語、優諾語的比較句均只採用 M-St-Adj 模式，換言之，這些語言比較句業已完成“M-St-Adj > Adj-M-St”的語序演變。例如：

巴哼語（毛宗武、李雲兵 1997）

Adj-M-St:

- (18) a. vɸɿ³¹ pi³¹ mɸɿ³³ ljo³⁵.
我 比 你 大
我比你大。
b. pɸɿ³³ tɸɿ⁴⁴ pi³¹ pɸɿ³³ kwɔ³⁵ jɿ⁵⁵ jɸɿ³³.
花 紅 比 花 白 好 看
紅花比白花好看。

畬語（毛宗武、蒙朝吉 1986）

Adj-M-St:

- (19) a. nuŋ⁴ pji⁵ vaŋ⁴ hin¹ pa¹ tshuŋ³.
他 比 我 高 三 寸
他比我高三寸。
b. vaŋ⁴ ŋjuŋ⁴ pji⁵ nuŋ⁴ ŋjuŋ⁴ ŋɔŋ⁵ tu⁵ u⁵.
我 的 比 他 的 好 得 多
我的比他的好得多。

優諾語（毛宗武、李雲兵 2007）

Adj-M-St:

- (20) a. nu³³ koŋ¹³ no²² pi²² nu³³ koŋ¹³ ni³¹ sən³³.
個 山 這 比 個 山 那 高
這座山比那座山高。
b. vɔ²² pi³³ naŋ²² liu³³, naŋ²² pi³³ ŋ¹³ le³⁵.
我 比 他 大 他 比 你 小
我比他大，他比你小。

勉語（毛宗武、蒙朝吉、鄭宗澤 1982）

Adj-M-St:

- (21) a. mwei² pei³ je¹ laŋ¹.
你 比 我 高
你比我高。
b. kla⁴² kla³¹ pe³⁵ ka⁴² laŋ³⁵ kan³³.
桃子 比 李子 甜
桃子比李子甜。（李雲兵 2008:165）

另一方面，其他苗瑤語的比較句語序則有 Adj-M-St 和 M-St-Adj 兩種模式，也就是說，在這些語言裡比較句基準和形容詞的位置正處在 “M-St-Adj > Adj-M-St” 演變過程中。析而言之，這種語序類型約有下面兩種情形。

第一，有些語言 Adj-M-St 和 M-St-Adj 均為常見模式，不易確定或未有資料顯示哪種模式為優勢或主要模式。巴那語即屬於這種情形：

巴那語（李雲兵 2008:166）

Adj-M-St:

- (22) a. le²⁴ te³¹ ljou²⁴ kua³⁵ le²⁴ tei⁴⁴ u²⁴ teo⁵⁵.
 個 兄 大 過 個 弟 二 年
 哥哥比弟弟大兩歲。

M-St-Adj:

- b. va²² za⁴⁴ pi⁴⁴ ni²² za⁴⁴ ŋũ³⁵.
 我 的 比 他 的 好
 我的比他的好。

第二，有的語言以 M-St-Adj 為主要或優勢模式，Adj-M-St 為少見或次要模式，即呈現 “M-St-Adj/Adj-M-St” 類型。換言之，這類語言固有的 Adj-M-St 模式正處在被外來模式 M-St-Adj 淘汰之中，苗語和布努語大約屬於這種情形。例如：

苗語（黔東）（王輔世 1985）

Adj-M-St:

- (23) nen⁵⁵ xhi³³ ɕhaŋ⁴⁴ vi¹¹.
 他 高 過 我
 他比我高。

M-St-Adj:

- (24) a. vi¹¹ paŋ³¹ phaŋ³³ u³⁵ noŋ³⁵ pi⁵⁵ moŋ⁵⁵ paŋ³¹ ta³⁵.
 我 的 件 衣 這 比 你 的 長
 我的這件衣服比你的長。
 b. teo⁵⁵ ə³³ noŋ³⁵ pi⁵⁵ teo⁵⁵ ə³³ ɛ³³ ljhə³³ ne⁴⁴.
 條 河 這 比 條 河 那 大 多
 這條河比那條河大得多。

2.3 南亞語

大多數南亞語比較句的語序未發生演變或變異，換言之，這些語言比較句的語序保持固有的 Adj-M-St 模式。例如：

德昂語（陳相木、王敬榴、賴永良 1986）

Adj-M-St:

- (25) a. ʔo dza:r doʔ mɔi.
我 高 比 你
我比你高。
b. bah ŋɔih doʔ bɔt.
酒 貴 比 煙
酒比煙貴。

克蔑語（陳國慶 2005）

Adj-M-St:

- (26) a. kho⁵⁵ khon³¹ mo^{ʔ53} pɤm³⁵ lɤ⁵⁵ kho⁵⁵ se^{ʔ53}.
橋 石頭 結實 比 橋 木材
石橋比木橋結實。
b. ɣn³⁵ tɛn³¹ lɤ⁵⁵ ɔ^{ʔ53}.
他 矮 比 我
他比我矮。

京語（歐陽覺亞等 1984）

Adj-M-St:

- (27) a. ta:u¹ ka:u¹ hə:n¹ mai², nɔ⁵ lai⁶ ka:u¹ hə:n¹ ta:u¹ nuə³.
我 高 過 你 他 又 高 過 我 再
我比你高，他比我更高。
b. kai⁵ nai² naŋ⁶ hə:n¹ kai⁵ kiə¹ hai¹ kən¹.
個 這 重 過 個 那 兩 斤
這個比那個重兩斤。

布芒語（刀潔 2007）

Adj-M-St:

- (28) a. da⁵⁵ jau⁵¹ le²⁴ ku³³.
 我 高 比 她
 我比她高。
 b. phən⁵⁵ i⁵¹ ǎm⁵⁵ le²⁴ phən⁵⁵ ni⁵⁵.
 件 這 好 比 件 那
 這件比那件好。

另一方面，克木語受漢語影響，比較句語序已由 Adj-M-St 演變為 M-St-Adj，也就是說，這個語言的比較句語序通常已不再採用固有的 Adj-M-St 模式：

克木語（陳國慶 2002）

M-St-Adj:

- (29) a. ʔah bɛʔ bi ʔah swaŋ lɿʔ mah.
 肉 羊 比 肉 豬 好 吃
 羊肉比豬肉好吃。
 b. ʔoʔ bi gu dziau ŋɛʔ.
 我 比 他 瘦 一點兒
 我比他瘦一些。

最後，其他南亞語的比較句語序則有 Adj-M-St 和 M-St-Adj 兩種模式，也就是說，在這些語言裡比較句基準和形容詞的位置正處在 “M-St-Adj > Adj-M-St” 演變過程中。概而言之，這種語序類型約有下面兩種情形。

第一，有的語言裡 Adj-M-St 和 M-St-Adj 均為常見模式，不易確定或未有資料顯示哪種模式為優勢或主要模式。俵語即屬於這種情形：

俵語（李旭練 1999）

Adj-M-St:

- (30) a. vi⁴ thau¹ ŋam² ndo⁵ mbi³ nam¹.
 哥 大 過 弟 兩 歲
 哥哥比弟弟大兩歲。
 b. tɛə² ciu⁵ ʔo⁵tsɔ³ kəŋ⁵ ŋam² tɛə² zu¹ tai⁵.
 時候 收 糧食 冷 過 時候 種 玉米
 秋天比春天冷。

M-St-Adj:

- (31) a. vi⁴ pi³ ndo⁵ thau¹ mbi³ nam¹.
哥 比 弟 大 兩 歲
哥哥比弟弟大兩歲。
b. tɕə² ɕiu⁵ ʔɔ⁵tsɔ³ pi³ tɕə² zu¹ tai⁵ kəŋ⁵.
時候 收 糧食 比 時候 種 玉米 冷
秋天比春天冷。

第二，有的語言以 M-St-Adj 為主要或優勢模式，Adj-M-St 為少見或次要模式，即呈現“M-St-Adj/Adj-M-St”類型。換言之，這類語言固有的 Adj-M-St 模式正處在被外來模式 M-St-Adj 淘汰之中。莽語和布庚語大約屬於這種情形。例如：

莽語（高永奇 2003）

Adj-M-St:

- (32) ʔu⁵¹ gɣaŋ⁵¹ dɔŋ⁵¹ mi³¹.
我 高 多 你
我比你高。

M-St-Adj:

- (33) a. ʔa³¹ʔu⁵¹ pi⁵⁵ mi³¹ to³¹.
我 比 你 大
我比你大。
b. mə³¹θa⁵¹ θua³⁵ ʔə³¹min⁵¹ pi⁵⁵ mə³¹θa⁵¹ θua³⁵ ʔə³¹ʔy⁵¹ me⁵⁵.
件 衣服 這 比 件 衣服 那 新
這件衣服比那件新。

布庚語（李雲兵 2005）

Adj-M-St:

- (34) mu⁴⁴ nen⁴⁴ ndza³¹ tsei³¹ mu⁴⁴ nen⁴⁴.
一 年 好 過 一 年
一年好於一年。

M-St-Adj:

- (35) a. i³¹ pi³¹ ɔ³¹ tho³¹ mi⁴⁴ nen⁴⁴.
他 比 我 大 五 歲
他比我大五歲。

- b. mu³¹ pi³¹ o³¹ ɣo⁴⁴, i³¹ pi³¹ o³¹ keŋ²⁴ ɣo⁴⁴.
 你 比 我 高 他 比 我 更 高
 你比我高，他比我更高。

2.4 南島語

海南三亞回族人使用的回輝話屬於南島語系馬來-波利尼西亞語族 (Malayo-Polynesian) 占語支 (Chamic)，與越南境內的拉德語 (Rade)、羅格萊語 (Roglai) 以及藩朗占語 (Phan Rang Cham) 等占語支語言具有極近的發生學關係。回輝話裡 M-St-Adj 和 Adj-M-St 兩種語序模式均為常見（即呈現 “Adj-M-St/M-St-Adj” 類型），也就是說，這個語言比較句語序業已發生變異。例如：

回輝話（鄭貽青 1997）

Adj-M-St:

- (36) a. lu⁴³ kau³³ pion³² la:u³² lu⁴³ ha³³.
 碗 我 大 過 碗 你
 我的碗比你的大。
 b. thun³³ ni³³ sa³³ tha:i²¹ na:i³² la:u³³ thun³³ hi¹¹.
 今年 的 稻 好 過 去年
 今年的稻子比去年好。
 c. ʔa¹¹ ko³³ phai⁴³ kham²¹ na:i³² la:u²² ʔa¹¹ thai¹¹.
 哥 讀 書 好 過 弟
 哥哥學習比弟弟好。

M-St-Adj:

- (37) a. kau³³ pi¹¹ ha³³ tsat²⁴ tso³³ kiə³³ sun³³.
 我 比 你 矮 三 寸
 我比你矮三寸。
 b. nau³³ pi¹¹ kau³³ khion¹¹ piai³³ ta:ŋ³².
 他 比 我 高 許多
 他比我高得多。
 c. thun³³ ni³³ pi¹¹ thun³³ hi¹¹ tsan³³ kia³³ thua¹¹ pha:n³².
 今年 比 去年 增加 兩 倍
 今年比去年增加兩倍。

以上的描寫和分析可概括成〈表 1〉。

〈表 1〉南方部分民族語言中比較句語序的演變和變異

語言		演變和變異階段					資料來源
名稱	系屬	未變 A-M-St	演變中 (A-M-St/M-St-A)			已變 M-St-A	
			A-M-St/ m-st-a	A-M-St/ M-St-A	a-m-st/ M-St-A		
傣語	壯傣 (侗台)	+					王均等 1984
拉基	仡央 (侗台)	+					李雲兵 2000
黎語	黎語支 (侗台)		+				王均等 1984
壯語	壯傣 (侗台)			+			王均等 1984
布依	壯傣 (侗台)			+			王均等 1984
侗語	侗水 (侗台)			+			王均等 1984
仡佬	侗水 (侗台)			+			王均等 1984
水語	侗水 (侗台)			+			王均等 1984
毛難	侗水 (侗台)			+			王均等 1984
茶洞	侗水 (侗台)			+			李錦芳 2007a
佯僂	侗水 (侗台)			+			薄文澤 1997
臨高	壯傣 (侗台)			+			梁、張 1997
莫語	侗水 (侗台)			+			楊通銀 2000
布央	仡央 (侗台)			+			李錦芳 1999
村語	黎語支 (侗台)				+		歐陽覺亞 1998
拉珈	侗水 (侗台)					+	毛、蒙、鄭 1982
標話	侗水 (侗台)					+	梁、張 2002
蔡家	侗水 (侗台)					+	薄文澤 2007
普標	仡央 (侗台)					+	梁、張、李 2007
仡佬	仡央 (侗台)					+	張濟民 1993
木佬	仡央 (侗台)					+	薄文澤 2003
巴那	苗語支 (苗瑤)			+			陳其光 2007
苗語	苗語支 (苗瑤)				+		王輔世 1985
布努	苗語支 (苗瑤)				+		毛、蒙、鄭 1982
巴哼	苗語支 (苗瑤)					+	毛、李 1997
炯奈	苗語支 (苗瑤)					+	毛、李 2002
畚語	苗語支 (苗瑤)					+	毛、蒙 1986
勉語	瑤語支 (苗瑤)					+	毛、蒙、鄭 1982
優諾	苗語支 (苗瑤)					+	毛、李 2007
佉語	孟高棉 (南亞)	+					周、顏 1984

德昂	孟高棉 (南亞)	+					陳、王、賴 1986
布朗	孟高棉 (南亞)	+					李、聶、邱 1986
克蔑	孟高棉 (南亞)	+					陳國慶 2005
京語	孟高棉 (南亞)	+					歐陽等 1984
布興	孟高棉 (南亞)	+					高永奇 2004
布芒	孟高棉 (南亞)	+					刀潔 2007
俅語	孟高棉 (南亞)			+			李旭練 1999
莽語	孟高棉 (南亞)				+		高永奇 2003
布庚	孟高棉 (南亞)				+		李雲兵 2005
克木	孟高棉 (南亞)					+	陳國慶 2002
回輝	占語支 (南島)			+			鄭貽青 1997

3. 判定接觸性語法演變的參數

在上面的討論中，我們假設中國南方民族語言比較句的基準和形容詞的固有語序是 Adj-M-St，而部分語言使用的 M-St-Adj 模式以及大多數語言具有的 Adj-M-St/M-St-Adj 交替模式，則是與漢語接觸而發生的語序演變和變異的結果。下面我們從語言類型學、接觸語言學等角度來對這個假設進行證明。我們著重要回答的問題是：

- (i) 既然有些南方民族語言只使用 Adj-M-St，而另外一些語言只見 M-St-Adj，為什麼不可以假設後者是南方民族語言固有的語序，而前者是後者發生演變的結果？
- (ii) 即便有證據顯示南方部分民族語言確已發生或正在發生“Adj-M-St > M-St-Adj”演變，為什麼不可以假設這是部分民族語言獨立發生的平行性或普遍性演變而非接觸引發的語法演變？
- (iii) 即使可以證明南方民族語言“Adj-M-St > M-St-Adj”確是接觸引發的語法演變或變異，那麼觸發這種演變或變異的“模式語”(model language)為什麼不可以是某個（或某些）具有 M-St-Adj 模式的南方民族語言而非漢語？

3.1 語言類型學的語序關聯

世界語言的小句語序主要有 SVO、SOV 和 VSO 三種模式，此外極少數語言

是 OVS、OSV、VOS。語言學家把這六種語序分別概括為 **VO** (SVO, VSO, VOS) 和 **OV** (SOV, OVS, OSV) 兩種類型。考察表明，世界語言裡動詞和賓語的這兩類語序關聯於某些另外的語序模式（參看 Greenberg 1963, Lehmann 1973, 1978, Vennemann 1974, 1976, Hawkins 1983, Dryer 1991, 1992, 2007, 2009）。

所謂“語序關聯”，簡單地說，是指不同語序之間存在蘊涵 (implicational) 關係。Dryer 的定義是：如果一個“成分對”X 和 Y 呈現這樣的情形：X 位於 Y 前這種形式在 VO 型語言裡遠比在 OV 型語言裡常見，那麼 <X, Y> 則是一個“關聯對” (correlation pair)，其中 X 是動詞模件 (verb patterner)，Y 是賓語模件 (object patterner) (Dryer 1992:87, 2009)。下面的〈表 2〉是 Dryer (1992, 2007, 2009) 發現的無爭議的語序關聯對：

〈表 2〉語序關聯對 (Correlation Pairs) (Dryer 1992:108, 2009)

動詞模件 (verb patterner)	賓語模件 (object patterner)	示例 (example)
動詞	賓語	<i>ate + the sandwich</i>
附置詞 (adposition)	NP	<i>on + the table</i>
名詞	關係小句	<i>movies + that we saw</i>
冠詞	N'	<i>the + tall man</i>
係詞	表語 (verb predicate)	<i>is + a teacher</i>
“想”義動詞 (‘want’)	VP	<i>wants + to see Mary</i>
時-體助動詞	VP	<i>has + eaten dinner</i>
表否定的助動詞	VP	
標補詞 (complementizer)	小句 (S)	<i>that + John is sick</i>
疑問性小詞 (question particle)	小句 (S)	
狀語從句	小句 (S)	<i>because + Bob has left</i>
表複數的小詞 (plural word)	N'	
名詞	屬格定語 (genitive)	<i>father + of John</i>
形容詞	比較基準	<i>taller + than Bob</i>
動詞	附置詞短語 (PP)	<i>slept + on the floor</i>
動詞	方式副詞	<i>ran + slowly</i>

據 Dryer (1992, 2007, 2009)，〈表 2〉所列“成分對”跟動詞和賓語的語序有單向關聯和雙向關聯之分。單向關聯是指動詞、賓語成分對和相關“成分對”之間的單向蘊涵關係，比如動詞和賓語的語序跟名詞和關係小句語序之間就是一種單向關聯：

$$VO \supset NRel; \quad RelN \supset OV$$

這兩個蘊涵式讀為：若一個語言是 VO，那麼它的關係小句和名詞的語序是 NRel，反之不然；若一個語言的關係小句和名詞的語序是 RelN，那麼該語言是 OV，反之不然。

單向關聯的前提是：在兩個成分對所具有的四種邏輯上可能的模式中，三種常見，一種罕見。比如動詞、賓語成分對和關係小句、名詞成分對的匹配模式中，OV&RelN、OV&NRel 和 VO&NRel 常見，而 VO&RelN 極其罕見。據此，我們可以得出蘊涵性概括：“若是 VO，則應 NRel”，或者“若是 RelN，則應 OV”。但反之則不然，比如我們不能作出“若是 NRel，則應 VO”這樣的斷言，因為具有 NRel 類型的語言有些是 VO，有些則是 OV。同樣，我們也不能作出“若是 OV，則應 RelN”這樣的預測，因為對於 OV 語言來說，RelN 和 NRel 同樣都是可能的。

另一方面，雙向關聯是指動詞、賓語成分對和相關成分對之間的雙向蘊涵關係，比如動詞和賓語的語序跟名詞和附置詞的語序之間就是一種雙向關聯：

$$OV \Leftrightarrow Po; \quad VO \Leftrightarrow Pr$$

這兩個蘊涵式讀為：若一個語言是 OV，那麼它通常使用後置詞 (Po)，反之亦然；若一個語言是 VO，那麼它通常使用前置詞 (Pr)，反之亦然。

跟單向關聯不同，雙向關聯的條件是，在兩個成分對所具有的四種邏輯上可能的模式中，只有兩種常見，另兩種罕見。比如動詞、賓語成分對和附置詞（前置詞/後置詞）、名詞成分對的匹配模式中，OV&Po 和 VO&Pr 常見，OV&Pr 和 VO&Po 罕見。在這種情形下，我們可以對上述兩種成分對的匹配模式作出雙向性的斷言和預測：“若是 OV，則應 Po”以及“若是 VO，則應 Pr”；反之亦然：“若是 Po，則應 OV”以及“若是 Pr，則應 VO”。所以說，附置詞的位置（前置詞抑或後置詞）與動詞和賓語的語序之間是一種雙向關聯。

需要說明的是，Dryer (1992, 2007, 2009) 上述語序關聯的概括並非演繹而來，而是基於跨地區和跨語系 625 種語言的定量分析作出的。比如他的“比較句基準和形容詞的語序與動詞和賓語語序雙向關聯”這一結論就是基於〈表 3〉中動詞、賓語成分對與比較句基準和形容詞成分對的匹配資料而得出的：

〈表 3〉比較句中形容詞和基準的語序 (Dryer 1992:92)

	非洲	歐亞	東南亞及 大洋洲	澳洲及 新幾內亞	北美州	南美洲	總計
OV&StAdj	3	15	3	0	5	3	29
OV&AdjSt	2	1	0	0	1	3	7
VO&StAdj	0	0	1	0	0	0	1
VO&AdjSt	5	7	8	0	7	4	31

〈表 3〉中阿拉伯數字表示的是“語組”(genus)的數量而不是具體語言的數量。語組是基於發生學關係對毗鄰地區不同語言所做的歸類，每個語組包含數量不等的語言。由〈表 3〉可以看出，在世界語言中 OV&StAdj 和 VO&AdjSt 兩種匹配模式普遍可見，OV&AdjSt 相對罕見；而 VO&StAdj 則極其罕見，在 Dryer (1992, 2007) 的語種庫裡僅見於漢語語組。因為在上述四種匹配模式中 OV&AdjSt 和 VO&StAdj 罕見，所以 VO 和 AdjSt 之間以及 OV 和 StAdj 之間是雙向關聯。

就本文的論題而言，Dryer (1992, 2007, 2009) 語序關聯的發現給我們的啓示是：

(a) 既然 VO 型語言通常只見 AdjSt 模式，我們有理由相信，本文所考察的南方民族語言（VO 型語言），其固有的語序模式應是 AdjSt 而非 StAdj。前述的傣語、德昂等語言正是在這個語法參項上完全保持固有語序 (AdjSt)，而標話、巴哼等語言則可視為在相關語序參項上完成了由 AdjSt 到 StAdj 的語序演變。

(b) 既然類似漢語 VO&StAdj 這種匹配模式在世界語言中極為罕見，那麼部分民族語言業已完成或正在發生的“AdjSt > StAdj”演變，就不大可能是這些語言獨立發生的平行性或普遍性演變。因為，如果某種語法結構是跨區域、跨語言罕見的模式，那麼這種語法結構在同一區域若干語言中的存在，極有可能是語言接觸導致的區域擴散的結果，而不大可能是語言話語或歷史演變的普遍原則或平行演變作用的產物（吳福祥 2008）。

(c) 既然 VO&StAdj 這類跨語言罕見的匹配模式只見於漢語，那麼觸發中國南方部分民族語言發生“AdjSt > StAdj”演變或變異的“模式語”就只能是漢語。

3.2 發生學關係的分布模式

識別和判定接觸性演變的一個重要方法是觀察和比較某一特徵在相關語言的親屬語或姐妹語中的分布模式。基本的假設是：給定某個語言特徵 P 為語言 M 和語言 R 所共有，而這兩個語言地理上緊鄰且處於長期密切的接觸之中；如果特徵 P 也見於與語言 M 有發生學關係的語言但不見於與語言 R 有發生學關係的語言，

那麼可以推斷這是接觸引發的語言特徵遷移的實例，更明確地說，P 是由語言 M 移入語言 R 的（參看 Heine & Kuteva 2005, 2007）。

根據上述假設，如果我們將上述南方民族語言的 M-St-Adj 模式歸因於漢語的影響，那麼最好的證據是這些語言的境外親屬語或姐妹語，其比較句中基準和形容詞的語序採用的是 Adj-M-St 而非 M-St-Adj。而這樣的證據並不難得到：在我們所能觀察到的與國內侗台語及孟高棉語（南亞語）發生學關係極近的境外語言中，比較句基準和形容詞的語序均為 Adj-M-St。請看〈表 4〉：

〈表 4〉境外若干侗台語和孟高棉語比較句的語序

語言			比較句語序		資料來源
名稱	系屬	使用國家	Adj-M-St	M-St-Adj	
標準泰語	侗台	泰國	+		Iwasaki & Ingkaphirom 2005
老撾語	侗台	老撾	+		Enfield 2007
儂語 (Nung)	侗台	越南	+		Saul & Wilson 1980
土語 (Tai-Tho)	侗台	越南	+		Morev 1988
白泰語 (White Tai)	侗台	越南	+		Morev 1988
拉哈語 (Laha)	侗台	越南	+		Morev 1988
Tai-Muey 語	侗台	老撾	+		Morev 1988
普泰語 (Phutai)	侗台	泰國	+		Morev 1988
泰陽語 (Tai-Yang)	侗台	泰國	+		Morev 1988
撣語 (Sha/Shan)	侗台	緬甸	+		Morev 1988
南部撣語 (S. Shan)	侗台	緬甸	+		Morev 1988
Phake 語	侗台	印度	+		Morev 1988
坎梯語 (Khamti)	侗台	緬甸，印度	+	+	Morev 1988
阿含語 (Ahom)	侗台	印度	+		Morev 1988
越南語	孟高棉	越南	+		Stassen 1985, Ansaldo 2004
柬埔寨語	孟高棉	柬埔寨	+		Jacob 1968
Brôu 語	孟高棉	越南	+		Miller 1964

下面是相關語言的例子：

標準泰語 (Iwasaki & Ingkaphirom 2005)

(38) a. dǝk-bua sǝy kwàa dǝk-kùlǝp. (94)

蓮花 美麗 比 玫瑰

蓮花比玫瑰好看。

- b. théksîi wîŋ rew kwàa rót-mee. (95)
的士 跑 快 比 巴士
計程車比公共汽車跑得快。

老撾語 (Enfield 2007)

- (39) a. khòj⁵ suung³ kuaø caw⁴. (256)
我 高 比 你
我比你高。
b. caan³ khòj⁵ tèk⁵ kuaø caan³ caw⁴. (261)
盤子 我 破 比 盤子 你
我的盤子比你的盤子破。

儂語 (Nung) (Saul & Wilson 1980)

- (40) a. cáu lái heng hò'n mu'hng. (75)
我 多 強壯 比 你
我比你強壯。
b. vahng pì bəo khong hò'n vahng nǝng ɔ. (103)
男 哥哥 窮 比 男 弟弟
哥哥比弟弟窮。

土語 (Tai-Tho) (Morev 1988)

- (41) a. slɔŋ¹ nɔi⁵ kwaa² saam¹. (95)
二 少 比 三
“二”比“三”少。
b. Dia⁵ luong¹ qua² moi⁴ can⁵ Dia. (99)
人 名 大 比 每 人
比每個人都高。

白泰語 (White Tai) (Morev 1988:98)

- (42) mak²-təng¹ waan¹ se¹ mak²-keng¹. (98)
甜瓜 甜 比 鳳梨
甜瓜比鳳梨甜。

拉哈語 (Laha) (Morev 1988)

- (43) a. lən² zen⁶ kwaang² sya lən² e⁶. (97)

房子 我 寬敞 比 房子 他

我的房子比他的房子寬敞。

- b. kən¹ kwaang² sya¹ lon² e⁶. (98)

多 寬敞 比 房子 那

比那個房子寬敞。

Tai-Muey (Morev 1988)

- (44) nam² nii⁵ vai³ lyyn⁵ nam³-khong¹. (97)

河 這 快 比 湄公河

這條河的水流比湄公河快。

普泰語 (Phutai) (Morev 1988)

- (45) a. həə³ lam³ nii³ cop⁵ kwaa³ həə³ lam³ tan⁵. (95)

船 條 這 好 比 船 條 那

這條船比那條船好。

- b. cop⁵ kwaa³ mət⁵ muu³. (99)

好 比 所有 組

比各組都好。

泰陽語 (Tai-Yang) (Morev 1988)

- (46) a. myy⁶-nni⁶ naaw¹ ləə¹ myy⁶-waa⁴. (97)

今天 冷 比 昨天

今天比昨天冷。

- b. phæng⁴ kaai¹ phəə. (99)

貴 比 什麼/誰

比什麼都貴。

撣語 (Sha/Shan) (Morev 1988)

- (47) tə-bee⁴ sai¹ njiu² tə-məw¹. (97)

羊 乾淨 比 豬

羊比豬乾淨。

南部譚語 (S. Shan) (Morev 1988:98)

- (48) man⁴ li⁴-ngaam⁴ se¹ saaw¹ tai⁴.
她 美麗 比 姑娘 泰國
她比泰國姑娘漂亮。

Phake (Morev 1988:94)

- (49) chaw³ khən phi² haŋ⁴ khom³san⁴.
人名 多 胖 比 人名
Chaw 比 Khomsan 胖。

坎梯語 (Khamti) (Morev 1988:98)

- (50) khən³ sung⁴ lym-shi a-nan².
多 高 比 那
比那個高。

Ahom (阿含語) (Morev 1988:94)

- (51) man nəŋ-man khən suŋ luk man nəy-yən.
他 哥 多 高 比 他 姐
他哥哥比他姐姐高。

越南語 (Stassen 1985, Ansaldo 2004)

- (52) a. vang qui hon bac. (Stassen 1985:43)
金子 值錢 過 銀子
金子比銀子值錢。
b. anh ây tốt hơn tôi. (Ansaldo 2004:5)
他 好 過 我
他比我好。

柬埔寨語 (Jacob 1968:140)

- (53) a. ni:əŋ nih lɔ: ci:əŋ ni:əŋ nùh.
姑娘 這 漂亮 比 姑娘 那
這姑娘比那姑娘漂亮。
b. kət ba:n səse:(r) siənpəhəu craən ci:əŋ lò:k.
他 得 寫 書 多 比 您
他寫的書比您多。

Brôu 語 (Miller 1964)

(54) alic o clou'i nheq... (72)

豬 好 比 所有的

比其他的豬都好的這頭豬...

一個可能例外是坎梯語，這個語言的比較句也可以採用“M-St-Adj”語序：

坎梯語 (Morev 1988)

(55) kon³ a-nai² mai kon³ a-nan² khən³ suŋ⁴. (94)

男人 這 比 男人 那 多 高

這個男人比那個男人高。

但 Morev (1988:94) 認為，坎梯語這種“M-St-Adj”模式是毗鄰的東印地語 (East Indian languages) 句法擴散的產物。在當地的東印地語裡，比較句的標記採用離格及其他旁格形式來表達，而坎梯語的 *mai* 正是旁格介詞。實際上，在坎梯語比較句裡程度詞 *khən³* 的使用是強制性的。因此 Morev (1988:94) 認為，坎梯語固有的比較句形式應是“S（主語）+ M（模態詞，即程度詞 *khən³*）+ F（特徵詞，即謂語形容詞）”。

3.3 語法描寫者的觀察和判斷

在很多南方民族語言的調查報告和研究文獻裡，語法描寫者在不同程度上對相關語言中 M-St-Adj 模式的出現或來源從漢語影響這個角度進行了分析和解釋，這類觀察和判斷是本文立論的一個重要證據。特別是有些語法描寫者同時是對象語言的母語者，其觀察和判斷尤其值得關注和信從。下面是一些語法描寫者針對特定語言中比較句受漢語影響發生語序演變或變異所作的觀察和描述：

茶洞語：

“由介詞 *ta⁶* ‘過’ 引介比較對象，放在形容詞後，與在形容詞前的被比較者進行對比。例如 *ni⁴fai⁵*（太陽）*lək⁸*（大）*ta⁶*（過）*njeŋ²kaŋ¹*（月亮）‘太陽比月亮大’。受漢語影響，這一句也可念成 *ni⁴fai⁵*（太陽）*pi³*（比）*njeŋ²kaŋ¹*（月亮）*lək⁸*（大）。”（李錦芳 2007a:1336）

莫語：

“比較句中，形容詞謂語所帶的數量補語和被比較物都應在謂語之後，但借入漢語的 pi³ ‘比’ 字結構以後，被比較物放到了形容詞謂語之前”。(154)

“以形容詞作謂語的比較句，目前流行和漢語一樣的句式 ‘A 比 B…’ ……看來形容詞謂語比較句的句式正在漢語影響下產生變化。”（楊通銀 2000: 105-106）

俵語：

“俵語固有的比較句語序是：甲＋形容詞＋比較詞＋乙。在漢語的影響下，產生了‘甲＋比較詞＋形容詞＋乙’的新語序。兩種語序並存並用”。（李旭練 1999:174）

“俵語和壯侗語一般都有兩種比較句式，一種是較古老的句式，結構為‘形容詞＋介賓短語’，一種是在漢語的影響下產生的新句式，結構為‘介賓短語“比……”＋形容詞’”。（李旭練 1999:192）

回輝話：

“表示比較關係時，一般在形容詞與後面的名詞或代詞之間加介詞 lau³² ‘過’ 來表示。……由於受漢語的影響，現在也常常按漢語的次序，用 pi¹¹ ‘比’ 來表示……。”（鄭貽青 1997:75）

3.4 語言接觸的歷史和現狀

接觸引發的語法演變，顧名思義這種語法演變過程是由語言之間的接觸誘發和導致的。因此，判定接觸性語法演變的先決條件是須有證據顯示，相關的語言之間存在或曾經存在接觸關係。比如我們要證明中國南方民族語言 M-St-Adj 模式的出現是漢語句法擴散的結果，那麼先決條件是漢語跟這些民族語言具有接觸關係，而這方面的證據可謂俯拾即是：大量的人類學、考古學、語言史和社會語言學的資料和成果顯示，漢語和中國南方民族語言具有數千年的接觸歷史（參看 LaPolla 2001, Matthews 2007, 吳安其 2008），特別是近數十年以來，這種接觸尤其廣泛而深刻。

3.5 平行的語法複製過程或其他特徵遷移的實例

在語言接觸及由此導致的語言演變中，一個特別常見的情形是，語言 A 某一特徵 F 的存在並不是一種孤立的接觸引發的語言演變現象，相反，在這個語言裡

通常還有其他“語言 B > 語言 A”特徵遷移的實例。特別是，如果這些假定的特徵遷移的實例之間是互相獨立並且涉及不同的遷移種類（比如語法複製、語法借用以及詞彙、語音成分的借用等），而相反的特徵遷移實例（即“語言 A > 語言 B”）未曾發現或相對罕見；那麼，語言 A 裡 F 的存在應該來自語言 B 對應範疇的遷移。在本文考察的南方民族語言裡，有大量的語言特徵來自漢語。首先，所有南方民族語言都存在數量不等的漢語借詞（實義詞）；其次，幾乎所有南方民族語言都具有漢語的語法借用成分（如連詞、副詞、助動詞、介詞、語氣詞、結構助詞、體標記等）；再次，很多語言從漢語複製了相關的結構式，最典型的是漢語正反問句“V 不 V”（吳福祥 2008）；最後，很多南方民族語言複製了漢語“‘獲得’義動詞 > 補語標記”、“‘拿’義動詞 > 工具介詞 > 處置式標記”等語法化過程（吳福祥 2009）。但是，相反的語法特徵遷移的實例則殊為罕見。

最值得注意的是，在很多使用 M-St-Adj 語序模式的語言裡，其比較標記正是借自漢語標準語或方言的介詞“比”，如標語 *pi*¹、普標語 *pi*⁴⁵、木佬語 *pi*³³、黎語 *bi*¹、侗語 *pji*³、仡佬語 *pi*³、水語 *pi*³、布依語 *pi*⁶、毛難語 *pi*³、茶洞語 *pi*³、佯僂語 *bi*²、臨高語 *bi*⁴、布央語 *pi*²⁴、布努語 *pi*⁴³、巴哼語 *pi*³¹、炯奈語 *pi*⁵³、畚語 *pji*⁵、莫語 *pi*³、優諾語 *pi*³³、巴那語 *pi*⁴⁴、苗語 *pi*⁵⁵、克木語 *bi*、俚語 *pi*³、莽語 *pi*⁵⁵、布庚語 *pi*³¹、回輝話 *pi*¹¹ 以及拉珈語 *pei*³、仡佬語（平壩）*phei*⁵⁵、壯語 *pei*³、村語 *bei*³、勉語 *pei*³ 等。既然這些語言的比較標記是漢語的借用成分，那麼其相應的語序模式 M-St-Adj 的產生自然是漢語句法擴散的產物。⁷

4. 演變的機制

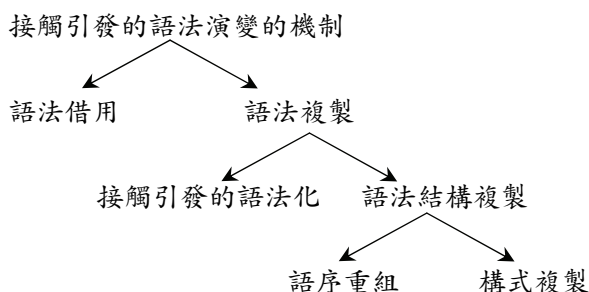
拙作（吳福祥 2008）參照 Heine (Heine 2007, Heine & Kuteva 2003, 2005, 2006, 2007, 2008) 的相關論述，把接觸引發的語法演變的機制分為“語法借用” (grammatical borrowing) 和“語法複製” (grammatical replication)：語法借用指的是一個語言（源語）的語法語素（語法性音-義單位，即虛詞及詞綴）遷移到另一個語言（受語）之中，比如很多南方民族語言從漢語借入結構助詞

⁷ 其實跟 M-St-Adj 一樣，南方民族語言裡用“過”義語素作比較標記（如壯語的 *kva*⁵ 和侗水語支的 *ta*⁶）的比較句“Adj-M-St”，也是漢語（粵語）影響的產物（吳福祥 2010）。不過這兩種句法影響所涉及的接觸模式容有不同：部分南方民族語言裡 M-St-Adj 的產生極有可能導源於這些語言跟漢語標準語的接觸，句法擴散的模式語是漢語標準語；另一方面，部分南方民族語言以及東南亞語言裡“過”義語素作比較標記的“Adj-M-St”，則源於這些語言跟粵語（或平話）的接觸，其模式語是粵語（或平話）。

“的”。語法複製則包括“接觸引發的語法化”和“語法結構複製”兩個方面，前者是指一個語言（複製語）對另一個語言（模式語）的語法概念或語法概念演化過程的複製，後者是一個語言（複製語）對另一個語言（模式語）語法結構的複製。

“語法結構複製”可進一步分為兩個小類：

- (i) “語序重組”(reordering) 或“結構重組”(restructuring)，即一個語言（複製語）的使用者依照另一個語言（模式語）的句法和形態模式來重排(rearrange)自己語言裡意義單位的語序，比如印度北部阿薩姆邦境內的阿含語（侗台語）受周邊印度-雅利安語和藏緬語影響，固有的小句語序 SVO、領屬結構式語序 NG 已被重組為 SOV、GN (Enfield 2003, Thomason & Kaufman 1988)。
- (ii) “構式複製”(constructional replication)，即一個語言的使用者依據另一個語言的模式，用自己語言的材料構建出與模式語對等的（形態/句法/話語）結構式，比如很多南方民族語言的使用者利用本族語言的材料複製了漢語正反問結構式“V 不 V”（吳福祥 2008）。吳福祥(2008)將接觸引發的語法演變的機制概括如下：



那麼，前述南方民族語言 M-St-Adj 的產生屬於上述演變機制的哪種呢？既然很多南方民族語言 M-St-Adj 模式中比較標記使用的是漢語借詞“比”，那麼這些語言裡 M-St-Adj 的出現當與“語法借用”有關。不過，這種語序演變也一定涉及“語序重組”，因為邏輯上比較標記的借用跟比較句基準和形容詞語序的變化並沒有必然聯繫，事實上有的民族語言借用了漢語比較標記“比”，但仍然用於 Adj-M-St 格式。比如李錦芳(2007b)報導的布幹語（即布賡語）：

布幹語（李錦芳 2007b:1438-1450）

Adj-M-St:

- (56) a. li⁵⁵ tho³¹ tsai³¹/pi¹³ tau³³. (1447)
 黃牛 大 比 豬
 黃牛比豬大。
- b. mɛ⁵⁵ i³¹ sɛ³¹ tsai³¹/pi¹³ mɛ⁵⁵ mu³¹. (1447)
 孩子 他 乖 比 孩子 你
 他的孩子比你的孩子聰明。

因此我們認為，相對於語法借用而言，語序重組可能是南方民族語言“Adj-M-St > M-St-Adj”演變和變異的主要機制。

綜上所述，導致南方民族語言“Adj-M-St > M-St-Adj”演變和變異的機制是“語法複製（語序重組）”和“語法借用”，更準確地說，是“語序重組”和“語法借用”共同作用的結果。

5. 結語

本文從接觸語言學和語言類型學的角度討論中國南方民族語言（壯侗、苗瑤、南亞及南島）比較句語序的演變或變異。主要結論是：

- (i) 南方民族語言（壯侗、苗瑤、南亞及南島）比較句基準和形容詞的固有語序是 Adj-M-St，部分語言出現的 M-St-Adj 則是語言接觸引發的“Adj-M-St > M-St-Adj”演變和變異的結果。
- (ii) 在這個語序演變過程中漢語是“模式語”，也就是說，漢語比較句基準和形容詞的語序模式 (M-St-Adj) 為這些語言比較句語序的演變和變異提供了複製模式。
- (iii) 演變的主要機制是“語序重組”（語法複製），但同時也涉及“語法借用”。

引用文獻

- Ansaldo, Umberto. 2004. The correlation between surpass comparatives and verby languages. *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistics Society*, ed. by Somsonge, 1-12. Tempe: Program for Southeast Asian Studies, Arizona State University.
- Bo, Wenze (薄文澤). 1997. *Yanghuangyu Yanjiu* 佯僂語研究 [*A Study of the Yanghuang Language*]. Shanghai: Shanghai Far East Publishers.
- Bo, Wenze (薄文澤). 2003. *Mulaoyu Yanjiu* 木佬語研究 [*A Study of the Mulao Language*]. Beijing: Nationalities Press.
- Bo, Wenze (薄文澤). 2007. Caijiahua 蔡家話 [The Caijia language]. *Zhongguo de Yuyan* 中國的語言 [*The Languages of China*], ed. by Hongkai Sun, Zengyi Hu & Xing Huang, 1465-1476. Beijing: The Commercial Press.
- Chen, Guoqing (陳國慶). 2002. *Kemuyu Yanjiu* 克木語研究 [*A Study of the Kemu Language*]. Beijing: Nationalities Press.
- Chen, Guoqing (陳國慶). 2005. *Kemieyu Yanjiu* 克蔑語研究 [*A Study of the Kemie Language*]. Beijing: Nationalities Press.
- Chen, Qiguang (陳其光). 2007. Banayu 巴那語 [The Bana language]. *Zhongguo de Yuyan* 中國的語言 [*The Languages of China*], ed. by Hongkai Sun, Zengyi Hu & Xing Huang, 1602-1610. Beijing: The Commercial Press.
- Chen, Xiangmu (陳相木), Jingliu Wang (王敬榴), and Yongliang Lai (賴永良). 1986. *De'angyu Jianzhi* 德昂語簡志 [*Outline of the De'ang Language*]. Beijing: Nationalities Press.
- Dao, Jie (刀潔). 2007. *Bumangyu Yanjiu* 布芒語研究 [*A Study of the Bumang Language*]. Beijing: Nationalities Press.
- Dryer, Matthew S. 1991. SVO languages and the OV/VO typology. *Journal of Linguistics* 27.2:443-482.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68.1:81-138.
- Dryer, Matthew S. 2007. Word order. *Clause Structure* (2nd edition), ed. by Timothy Shopen, 61-131. Cambridge & New York: Cambridge University Press.
- Dryer, Matthew S. 2009. The branching direction theory revisited. *Universals of Language Today*, ed. by Sergio Scalise, Elisabetta Magni & Antonietta Bisetto, 185-207. Berlin: Springer.
- Enfield, Nick J. 2003. *Linguistic Epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London & New York: RoutledgeCurzon.
- Enfield, Nick J. 2007. *A Grammar of Lao*. Berlin & New York: Mouton de Gruyter.

- Gao, Yongqi (高永奇). 2003. *Mangyu Yanjiu* 莽語研究 [*A Study of the Mang Language*]. Beijing: Nationalities Press.
- Gao, Yongqi (高永奇). 2004. *Buxingyu Yanjiu* 布興語研究 [*A Study of the Buxing Language*]. Beijing: Nationalities Press.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Language*, ed. by Joseph H. Greenberg, 73-113. Cambridge: MIT Press.
- Hawkins, John A. 1983. *Word Order Universals*. New York: Academic Press.
- Heine, Bernd. 2007. Typology and language contact: word order. Spring School on Grammaticalization and Typology, March 21-24, 2007. Rakvere, Estonia.
- Heine, Bernd, and Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27.3:529-572.
- Heine, Bernd, and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge & New York: Cambridge University Press.
- Heine, Bernd, and Tania Kuteva. 2006. *The Changing Languages of Europe*. Oxford: Oxford University Press.
- Heine, Bernd, and Tania Kuteva. 2007. Identifying instances of contact-induced grammatical replication. Paper presented at the Symposium on Language Contact and the Dynamics of Language: Theory and Implications. May 10-13, 2007. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Heine, Bernd, and Tania Kuteva. 2008. Constraints on contact-induced linguistic change. *Journal of Language Contact* – THEMA 2 (2008): 57-90.
- Iwasaki, Shoichi, and Preeya Ingkaphirom. 2005. *A Reference Grammar of Thai*. Cambridge & New York: Cambridge University Press.
- Jacob, Judith M. 1968. *Introduction to Cambodian*. Oxford: Oxford University Press.
- LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 225-254. Oxford: Oxford University Press.
- Lehmann, Winfred P. 1973. A structural principle of language and its implications. *Language* 49.1:47-66.
- Lehmann, Winfred P. 1978. The great underlying ground-plans. *Syntactic Typology*, ed. by Winfred P. Lehmann, 3-55. Austin: University of Texas Press.
- Li, Daoyong (李道勇), Xizhen Nie (聶錫珍), and E'feng Qiu (邱鐸鋒). 1986. *Bulangyu Jianzhi* 布朗語簡志 [*Outline of the Bulang Language*]. Beijing: Nationalities Press.
- Li, Jinfang (李錦芳). 1999. *Buyangyu Yanjiu* 布央語研究 [*A Study of the Buyang Language*]. Beijing: Minzu University of China Press.

- Li, Jinfang (李錦芳). 2001. Chadongyu gaikuang 茶洞語概況 [A survey of the Chadong language]. *Minzu Yuwen* 民族語文 [*Minority Languages of China*] 2001.1:67-80.
- Li, Jinfang (李錦芳). 2007a. Chadongyu 茶洞語 [The Chadong language]. *Zhongguo de Yuyan* 中國的語言 [*The Languages of China*], ed. by Hongkai Sun, Zengyi Hu & Xing Huang, 1324-1337. Beijing: The Commercial Press.
- Li, Jinfang (李錦芳). 2007b. Buganyu 布幹語 [The Bagan language]. *Zhongguo de Yuyan* 中國的語言 [*The Languages of China*], ed. by Hongkai Sun, Zengyi Hu & Xing Huang, 1438-1450. Beijing: The Commercial Press.
- Li, Xulian (李旭練). 1999. *Laiyu Yanjiu* 俛語研究 [*A Study of the Lai Language*]. Beijing: Minzu University of China Press.
- Li, Yunbing (李雲兵). 2000. *Lajiyu Yanjiu* 拉基語研究 [*A Study of the Laji Language*]. Beijing: Minzu University of China Press.
- Li, Yunbing (李雲兵). 2005. *Bugengyu Yanjiu* 布庚語研究 [*A Study of the Bugeng Language*]. Beijing: Nationalities Press.
- Li, Yunbing (李雲兵). 2008. *Zhongguo Nanfang Minzu Yuyan Yuxu Leixing Yanjiu* 中國南方民族語言語序類型研究 [*A Cross-linguistic Typology on Word Order of Minority Languages in Southern China*]. Beijing: Peking University Press.
- Liang, Min (梁敏), and Junru Zhang (張均如). 1997. *Lingaoyu Yanjiu* 臨高語研究 [*A Study of the Lingao Language*]. Shanghai: Shanghai Far East Publishers.
- Liang, Min (梁敏), and Junru Zhang (張均如). 2002. *Biaohua Yanjiu* 標話研究 [*A Study of the Biao Language*]. Beijing: Minzu University of China Press.
- Liang, Min (梁敏), Junru Zhang (張均如), and Yunbing Li (李雲兵). 2007. *Pubiaoyu Yanjiu* 普標語研究 [*A Study of the Pubiao Language*]. Beijing: Nationalities Press.
- Mao, Zongwu (毛宗武), and Yunbing Li (李雲兵). 1997. *Bahengyu Yanjiu* 巴哼語研究 [*A Study of the Baheng Language*]. Shanghai: Shanghai Far East Publishers.
- Mao, Zongwu (毛宗武), and Yunbing Li (李雲兵). 2002. *Jongnaiyu Yanjiu* 炯奈語研究 [*A Study of the Jongnai Language*]. Beijing: Minzu University of China Press.
- Mao, Zongwu (毛宗武), and Yunbing Li (李雲兵). 2007. *Younuoyu Yanjiu* 優諾語研究 [*A Study of the Younuo Language*]. Beijing: Nationalities Press.
- Mao, Zongwu (毛宗武), and Chaoji Meng (蒙朝吉). 1986. *Sheyu Jianzhi* 畚語簡志 [*Outline of the She Language*]. Beijing: Nationalities Press.
- Mao, Zongwu (毛宗武), Chaoji Meng (蒙朝吉), and Zongze Zheng (鄭宗澤). 1982. *Yaozu Yuyan Jianzhi* 瑤族語言簡志 [*Outline of Languages of the Yao Nationality*]. Beijing: Nationalities Press.
- Matthews, Stephen. 2007. Cantonese grammar in areal perspective. *Grammars in Contact: A Cross-linguistic Typology*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 220-236. Oxford: Oxford University Press.

- Miller, Carolyn P. 1964. The substantive phrase in Brôu. *The Mon-Khmer Studies Journal* 1:63-80.
- Morev, Lev N. 1998. Expressing comparison in the Tai languages. *The Mon-Khmer Studies Journal* 28:91-100.
- Ouyang, Jueya (歐陽覺亞). 1998. *Cunyu Yanjiu 村語研究* [*A Study of the Cun Language*]. Shanghai: Shanghai Far East Publishers.
- Ouyang, Jueya (歐陽覺亞), Fang Cheng (程方), and Cuirong Yu (喻翠容). 1984. *Jingyu Jianzhi 京語簡志* [*Outline of the Jing Language*]. Beijing: Nationalities Press.
- Saul, Janice E., and Nancy Freiburger Wilson. 1980. *Nung Grammar*. Arlington: University of Texas.
- Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.
- Thomas, David Dunton. 1967. *Chrau Grammar: A Mon-Khmer Language of Vietnam*. Philadelphia: University of Pennsylvania dissertation.
- Thomason, Sarah Grey, and Terrence Kaufman. 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley & Los Angeles: University of California Press.
- Thompson, Laurence C. 1965. *A Vietnamese Grammar*. Seattle: University of Washington Press.
- Vennemann, Theo. 1974. Analogy in generative grammar: the origin of word order. *Proceedings of the Eleventh International Congress of Linguists* (1972), 79-83. Bologna.
- Vennemann, Theo. 1976. Categorical grammar and the order of meaningful elements. *Linguistic Studies Offered to Joseph Greenberg on the Occasion of his Sixtieth Birthday*, ed. by A. Juillard Saratoga, 615-634. Saratoga: Anna Libri.
- Wang, Fushi (王輔世). 1985. *Miaoyu Jianzhi 苗語簡志* [*Outline of the Miao Language*]. Beijing: Nationalities Press.
- Wang, Jun (王均) et al. 1984. *Zhuang-Dong Yuzu Yuyan Jianzhi 壯侗語族語言簡志* [*Outline of the Zhuang-Dong Languages*]. Beijing: Nationalities Press.
- Wu, Anqi (吳安其). 2008. Shiqian Huanan diqu de yuyan jiechu 史前華南地區的語言接觸 [The language contact in southern China in prehistoric ages]. *Minzu Yuwen 民族語文* [*Minority Languages of China*] 2008.3:21-35.
- Wu, Fuxiang (吳福祥). 2008. Nanfang yuyan zhengfan wenju de lai yuan 南方語言正反問句的來源 [The origin of the A-not-A question sentence in the languages of southern China]. *Minzu Yuwen 民族語文* [*Minority Languages of China*] 2008.1: 3-18.
- Wu, Fuxiang (吳福祥). 2009. Nanfang minzu yuyan li ruogan jiechu yinfa de yufahua guocheng 南方民族語言裡若干接觸引發的語法化過程 [Some grammaticalization processes induced by language contact in the minority languages of southern

- China]. *Yufahua yu Yufa Yanjiu* 語法化與語法研究 [*Studies in Grammaticalization and Chinese Grammar*], Vol. 4, ed. by Fuxiang Wu & Xiliang Cui, 389-444. Beijing: The Commercial Press.
- Wu, Fuxiang (吳福祥). 2010. Yueyu chabishi “X+A+guo+Y” de leixingxue diwei: bijiao fangyanxue he quyu leixingxue de shijiao 粵語差比式 “X+A+過+Y” 的類型學地位：比較方言學和區域類型學的視角 [The typological status of the comparative construction “X+A+guo(過)+Y” in Yue dialects: perspectives of comparative dialectology and areal typology]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2010.3:238-255.
- Yang, Tongyin (楊通銀). 2000. *Moyu Yanjiu* 莫語研究 [*A Study of the Mo Language*]. Beijing: Minzu University of China Press.
- Zhang, Jimin (張濟民). 1993. *Gelaoyu Yanjiu* 仡佬語研究 [*A Study of the Gelao Language*]. Guiyang: Guizhou Nationalities Publishing House.
- Zheng, Yiqing (鄭貽青). 1997. *Huihuiyu Yanjiu* 回輝語研究 [*A Study of the Huihui Language*]. Shanghai: Shanghai Far East Publishers.
- Zhou, Zhizhi (周植志), and Qixiang Yan (顏其香). 1984. *Wayu Jianzhi* 佤語簡志 [*Outline of the Wa Language*]. Beijing: Nationalities Press.

Changes and Variations of Word Order in the Comparative Construction in the Minority Languages of Southern China

Fuxiang Wu

Chinese Academy of Social Sciences

In the minority languages of southern China (Tai-Kadai, Hmong-Mien, Austroasiatic, Austronesian), the word order of Adjective (Adj.) and comparative standard (St.) in comparative construction can be divided into three typologies, namely (a) “Adj-M-St”, (b) “M-St-Adj”, and (c) “Adj-M-St/M-St-Adj”. From the perspective of contact linguistics and linguistic typology, this paper argues that the inherent word order of Adjective and comparative standard in minority languages of southern China should be “Adj-M-St”, and the word order pattern “M-St-Adj” found in some minority languages of southern China is due to the change “Adj-M-St > M-St-Adj” induced by the contact between these languages and Chinese. In this kind of contact-induced grammatical change, Chinese is the model language, and the other Southeast Asian languages are the replica languages. The mechanism of change is primarily reordering.

Key words: comparative construction, word order change, grammatical replication, reordering

Classifiers and Some Typological Considerations^{*}

Dan Xu (徐 丹)¹ and Jingqi Fu (傅京起)²

*INALCO/CRLAO/IUF*¹

*St. Mary's College of Maryland*²

Classifiers, both nominal and verbal, must be taken into consideration as a parameter in typological studies of the numeral classifier languages. Building on previous work, we will show that nominal and verbal classifiers are mutually dependent both chronologically and in terms of their syntactic position. Cross-linguistic data in general, and Chinese in particular, illustrate this dependency: nominal classifiers develop before verbal classifiers and the two kinds of classifiers are in a complementary distribution, occurring before nouns and after verbs.

Key words: classifier, typology, word order, complementary distribution

1. Introduction

Since Greenberg's (1963) works on language typology, discussions and studies of this topic have flourished in general linguistics, and the number of investigated languages has increased considerably. Dryer's (1992, 2003) studies of numerous typological characteristics, for example, are based on 625 language samples. According to Dryer, the Chinese language is considered an atypical VO language. Chinese scholars have also tried to test the assumptions established by Greenberg and Dryer. Among them, Liu Danqing (2003) has analyzed Chinese languages, while Huang Xing (1996) has investigated non-Han languages in China. Their works confirm, for the most part, the ideas put forward by Greenberg (1966) and Dryer (1992). Numerous significant works have been published on different languages. The focus of previous research was predominantly on word order and its plausible implicational relationships with other features. However, in these studies, the presence of classifiers in numeral classifier languages was not a salient feature and was sometimes neglected. Only recent works by Haspelmath et al. (2005) and Comrie (2008) began to take classifiers into account as

^{*} Our gratitude goes to Alain Peyraube who warmly received us when we first came to France in 1982 as graduate students. He has given us enormous support during our studies in Chinese linguistic research. Our works have been supported by the grant *Quantification et Pluralité ANR-06-BLAN0259* awarded by the Research Department of the French Government.

one of the typological parameters in noun phrases. Indeed, in languages with obligatory numeral classifiers, they play an important role in word order typology, at least in Asian languages. Since a significant number of these languages are found in China, belonging to the Sino-Tibetan, Altaic, and Austro-Asian families, it is interesting to observe how classifiers function and how they influence word order in Chinese (contemporary and classical) and in non-Han languages of China. Our work will build on the significant research results found in Chinese publications, such as Sun Hongkai (1988), Dai Qingxia & Jiang Ying (2004), Li Jinfang & Hu Suhua (2005), and others.

Nominal classifiers have been extensively examined and discussed, but verbal classifiers seem to be less studied. In this paper, we try to understand the distribution of nominal classifiers and verbal classifiers and their relationship. We believe that the order Numeral Noun (NumN) has to be studied together with classifiers (CI) in languages with obligatory NumCI. Nominal CI generally develops before verbal CI and the distribution of nominal CI and verbal CI is complementary (Shi Yuzhi 2006). We will speculate on why this is so at the end of this paper.

2. Numeral and noun

The pair NumN and NNum is left unclassified in Dryer (1992:119). NumN order clearly shows a correlation with VO/OV type except in languages of Africa. It is evident in Dryer's statistics that the order NumN strongly correlates with VO order cross-linguistically. Dryer states that "the two orders of numeral and noun are equally common among OV languages" (p.118). Independently, Huang Xing's (1996) statistics based on 35 non-Han languages in China corroborate Dryer's observations. One phenomenon is striking in Huang's (1996:12) table. In the Tibetan-Burman group, which are of the OV language type, and the Altaic group, also of OV type, 17 languages use NNum order while no language uses NumN order. However, many Tibeto-Burman languages possess classifiers; these statistics without classifiers are more or less meaningless. The NumN or NNum order can help us better grasp the word order patterns in languages without an obligatory NumCI system. This is the case in Old Chinese (OC) when classifiers were not fully developed. OC is mainly a VO language (Shen Pei 1992, Peyraube 2000 among others). Does this NumN & VO pattern work for OC? See the following table established by Wu Dan (2005):

Table 1: NumN and NNum in Old Chinese

	NumN	NNum
詩經 <i>Shījīng</i>	25	5
尚書 <i>Shàngshū</i>	22	0
論語 <i>Lúnyǔ</i>	4	1
左傳 <i>Zuǒ zhuàn</i>	54	6
孟子 <i>Mèng Zǐ</i>	35	7
Total	140	19

In this table, the order of NumN and NNum is significant, since at that period classifiers did not exist as a system. It is clear that NumN order is predominant. NNum existed but the proportion of its occurrence is only about one in seven. Consider these examples in OC in which numeral precedes noun:

- (1) 三人行，必有我師焉 (*Lúnyǔ* 論語, Shù ér 述而)
sān rén xíng, bì yǒu wǒ shī yān
three man walk certain have my master final-part
‘Walking with two persons, I must have a teacher among them.’
- (2) 拔二城 (*Zhànguó zònghéngjiā shū* 戰國縱橫家書, 25)
bá èr chéng
conquer two city
‘[Jing Ji] has conquered two cities.’
- (3) 大縣十七 (*Zhànguó zònghéngjiā shū* 戰國縱橫家書, 26)
dà xiàn shíqī
county seventeen
‘seventeen counties’

The order NumN (ex. (1) and (2)) is most frequent, while the NNum order (ex. (3)) is rare in OC (see Table 2). Table 1 also confirms the statistics of Dryer (1992:118): VO languages exhibit a strong tendency to be NumN. How about Middle Chinese (MC), after the Chinese language had undergone a typological change (Xu 2006) starting in the Western Han (3rd century BC)?

3. Nominal classifiers develop before verbal classifiers

According to previous scholars such as Wang Li (1958), Hashimoto (1977), Peyraube (1998) and others, classifiers began to develop during the Han. According to

the commonly accepted view, there are three steps (simplified here) illustrating the use of classifiers in the Chinese noun phrase (the symbol “>” means “evolve toward”):

NumN > NNumCl > NumClN

Some comments are necessary to explain this change. The evolution is not linear (we will discuss this in more detail later). The schema of the first step (NumN) does not imply that NNum did not exist. This phenomenon is parallel to VO vs. OV word order in Chinese; the fact that Old Chinese was mainly a VO language does not exclude some OV order sentences, which existed but were not dominant (Xu 2006). In Table 1, we saw that NumN order coexists with NNum, but the latter does not represent the prevalent word order tendency of that period. At a later time, the dominant order NumN stabilized. Table 2 is a summary of the distribution of NumN, NNum and NNumCl based on an excavated text of the *Zhànguó zònghéngjiā shū* 戰國縱橫家書 (ZZJ ca. 195 BC) [Letters of Strategists in the Warring States period] (hereafter ZZJ).

Table 2: NumN and NNum and NNumCl in the ZZJ (ca. 195 BC)

N and Measure Word	NumN	NNum	NNumCl
Common noun	65	3	0
MW 里 <i>lǐ</i> half kilometer	15	0	0
MW 乘 <i>shèng</i> ‘chariot’	8	0	2
MW 年 <i>nián</i> ‘year’	7	0	0
MW 月 <i>yuè</i> ‘month’	4	0	0
MW 歲 <i>sui</i> ‘year’ (age)	2	0	0
MW 丈 <i>zhàng</i> measure word	2	0	0
MW 仞 <i>rèn</i> measure word	2	0	0
Total	105	3	2

Excluded from this table are proper names such as 三晉 *Sān Jìn*, 三梁 *Sān Liáng* and so forth. The word 乘 *shèng* ‘chariot’ is different from other measure words. It can be used as a noun (8 occurrences) and classifier (twice). The only two examples of NNumCl involve *shèng*:

- (4) 臣以車百五十乘入齊 (ZZJ 戰國縱橫家書, 8)
chén yǐ jū bǎi wǔshí shèng rù Qí
 I take chariot hundred fifty CL enter Qi
 ‘I will take one hundred fifty chariots to go to the State of Qi.’

- (5) 於氏（是）爲長安君約車百乘 (ZZJ 戰國縱橫家書, 18)
yú shì (shì) wèi Cháng'ān jūn yuē jū bǎi shèng
 then for Chang-An prince arrange chariot hundred CL
 'Then [Empress Zhao] arranged one hundred chariots for her son Chang An.'

The statistics in Table 2 are self-evident. NumN order is the most attested at that period in this excavated text. Many linguists such as Peyraube (1998) are right to indicate that measure words were used first, and classifiers showed up later in the same slot. We distinguish measure words from common words in order to better see their development. Measure words like *lǐ* 'half kilometer', *nián* 'year' and *yuè* 'month' are most frequently seen in this text. The order NNumCl is found twice. This is due to the text's style. In the same tomb, other texts about medicine were excavated. According to Zhang Junzhi & Zhang Xiancheng (2002:220), the order NNumCl is attested more than 100 times in the 五十二病方 *Wǔshí'èr bìng fāng* 'fifty-two prescriptions'. It is unfortunate that the authors did not distinguish measure words from classifiers. In fact, more than 60% measure words in their data are included in their statistics. In any case, their studies confirm that the order NNumCl began to develop around the Han period.¹ Observe two examples drawn from this excavated text.

- (6) 烏豕（喙）十四果（顆） (*Wǔshí'èr bìng fāng* 五十二病方)
wū yǐ (huì) shísi guǒ (kē)
 aconitum fourteen CL
 'fourteen aconitum'
- (7) 龍須（鬚）一束 (*Wǔshí'èr bìng fāng* 五十二病方)
lóng xū (xū) yī shù
 rush one CL
 'one bundle of rush'

Recent studies such as Sun Yan (2005), Wu Dan (2005), Wu Fuxiang (2006) among others show that the order NumN without classifier was the dominant order until the Tang Dynasty (618-907). According to Wu Dan (2005) and Wu Fuxiang (2006:556-557), during 5th century AD, NNum or NumN² without a classifier made up 90% of cases in transmitted texts. This situation remained the same until the Yuan Dynasty (1271-1368). It is clear that the situation drastically changed beginning in the Yuan, as

¹ The study of Wei Desheng (2000:128) shows the same tendency, NNumCl order is not rare in the 睡虎地秦墓竹簡 *Shuì hǔ dì Qín mù zhújiǎn*, an excavated text on bamboo slips dated from around the 3rd century BC.

² The authors did not give the statistics on either NumN or NNum.

indicated in Table 3 below. In other words, classifiers became obligatory much later than current studies suggest. Here are two examples from transmitted texts:

- (8) 出一編書 (*Shǐjì* 史記, 55)
chū yī biān shū
 take.out one CL book
 ‘[He] took out one book.’
- (9) 馬三匹 (*Zuǒ zhuàn* 左傳, *Zhuāng gōng* 莊公, 16)
mǎ sān pǐ
 horse three CL
 ‘three horses’

Now let us observe statistics from other scholars. The following table suggests that the order NumN (with some NNum) lasted at least until the Tang, while the order NNumCl began to develop in the Han and coexisted with NumN during the same period. In other words, the development was not linear from the first stage to the second stage, but with overlaps.

Table 3: Proportion of Noun phrase without Cl to NP with Cl
 (by Wu Fuxiang 2006 based on Wu Dan 2005)

Periods	NNum and NumN without Cl	NNum and NumN with Cl
先秦 <i>Xiān Qín</i> around 8 th -3 rd centuries BC	10	1
漢代 <i>Hàndài</i> 206 BC-220 AD	9	1
魏晉南北朝 <i>Wèi-Jìn Nán-Běi Cháo</i> around 3 rd -6 th centuries AD	10	1
唐代 <i>Tángdài</i> 618-907	4	1
宋代 <i>Sòngdài</i> 960-1279	3	1
元代 <i>Yuándài</i> 1271-1368	2	1
明代 <i>Míngdài</i> 1368-1644	1	5
清代 <i>Qīngdài</i> 1644-1911	1	7

The change from the second to the third step (NNumCl > NumClN) does not appear to have taken place earlier than the Tang.³ Studies by Wang Li (1958), Hashimoto (1977), Peyraube & Wiebusch (1993), Sun Yan (2005), Wu Dan (2005) and Wu Fuxiang (2006) converge to the same conclusion. Consider the following table by Wu Fuxiang (2006: 559) based on Wu Dan (2005):

Table 4: Proportion of NumClN and NNumCl

Periods	Num Cl N	N Num Cl
先秦 <i>Xiān Qín</i> around 8 th -3 rd centuries BC	1	5
漢代 <i>Hàndài</i> 206 BC-220 AD	1	3
魏晉南北朝 <i>Wèi-Jìn Nán-Běi Cháo</i> around 3 rd -6 th centuries AD	1	1
唐代 <i>Tángdài</i> 618-907	5	1
宋代 <i>Sòngdài</i> 960-1279	9	1
元明清 <i>Yuán Míng Qīng</i> 1271-1911	15	1

The independent observations of Sun Yan (2005:407) based on a text from the late Tang confirm the above statistics, i.e. NNumCl is attested more than NumClN even though the two orders coexisted. All of these studies show that the Chinese language became a NumCl language after the Tang Dynasty, and the order NumClN won as a dominant order only around the 13th century AD.

Monographs and articles on classifier development are numerous. Scholars note that fully-fledged classifiers are almost all monosyllabic. At the first step, echo classifiers are found on bone inscriptions. Many Tibeto-Burman languages still use echo classifiers. Li Yuming (2000) indicates that this type of classifier is a primitive phase in classifier development. We can say at least that the reverse process, i.e. a classifier becoming an echo classifier, is not reported. When general classifiers are formed or become mature, echo classifiers become restricted and eventually disappear. As scholars have pointed out, measure words were first used with numerals, and by analogy classifiers took the same position later in noun phrases. We will not deal with all these details in this paper.

In numeral classifier languages, verbal classifiers are also attested. In general, verbal Cl develop after nominal Cl and are apparently much less numerous. Now we examine

³ The table suggest Wèi-Jìn Nán-Běi Cháo 魏晉南北朝 as transitional period.

how verbal classifiers evolved. Scholars agree in general that verbal (action) frequency is expressed by numeral or numeral plus classifier in Chinese, which developed in three steps:

$$\text{NumV} > \text{VNum} > \text{VNumCl}$$

Again, each step is not without some exceptions, but the main tendency is illustrated by these steps. According to several different studies, the order NumV was dominant before the Han:

- (10) 吾日三省吾身 (*Lúnyǔ* 論語, *Xué ér* 學而)
wú rì sān xǐng wú shēn
 I day three introspect I body
 ‘I examine myself three times a day.’
- (11) 曷爲三遇齊王而不言事? (*Xúnzǐ* 荀子, *Dàluè* 大略 35)
héwèi sān yù Qí wáng ér bù yán shì?
 why three meet Qi king Conj Neg talk thing/event
 ‘Why did he not mention affairs of state after three audiences with the king of Qi?’

A few instances of the order VNum are also found (cf. Tang Yuming 2002). Here is a table illustrating verbal quantification expressions in OC, from Tang Yuming (2002:203).

Table 5: NumV and VNum in Old Chinese

	尚書	詩經	左傳	論語	孟子	荀子	墨子	莊子	韓非子	戰國策	禮記	Total
	<i>SS</i>	<i>SJ</i>	<i>ZZ</i>	<i>LY</i>	<i>MZ</i>	<i>XZ</i>	<i>MZ</i>	<i>ZZ</i>	<i>HFZ</i>	<i>ZGC</i>	<i>LJ</i>	
NumV	4	1	126	8	13	39	43	38	62	52	26	412
VNum	1	0	9	0	0	0	0	1	2	5	9	27

It is clear that the occurrences of NumV represent the predominant order in OC while few instances of VNum can be seen. In the earliest examples with VNum given by Tang, we notice that in most of them (7 out of 10) the verb is limited to 笞 *chī* ‘flog, beat with bamboo stick’, or 鞭 *biān* ‘flog’. This finding is further confirmed by the statistics of Wei Desheng (2000:127-128) from the *Shuì hǔ dì Qín mù zhújiǎn* 睡虎地秦墓竹簡 (*SHD*), an excavated text on bamboo slips dating from around the 3rd century

BC. Wei noted that there are six instances of VNum order, all with the verb 治(笞) *chī* ‘flog’. The order NumV, in contrast, is used 17 times. Here is an example with VNum order given by Tang in OC:

- (12) 當笞五十 (*Shuì hǔ dì Qín mù zhújiǎn* 睡虎地秦墓竹簡)
dāng chī wǔshí
 must beat.with.bamboo.stick fifty
 ‘This merits fifty beatings with a bamboo stick.’

In example (12), the order VNum is used. But this order is not frequent in OC. Real verbal CIs are found in the Wei-Jin period. Observe example (13):

- (13) 讀書百遍而義自見 (*Sānguó zhì* 三國志)
dúshū bǎi biàn ér yì zì xiàn
 read book hundred CL CONJ meaning self become.visible?
 ‘Reading a book one hundred times, the meaning itself becomes clear.’

In Table 6 we have roughly sketched the evolution of the classifiers in Chinese:

Table 6: Synthesis of the nominal CI and verbal CI development

先秦 <i>Xiān Qín</i>	漢代 <i>Hàndài</i>	魏晉 <i>Wèi-Jìn</i>	唐宋 <i>Táng Sòng</i>	元 <i>Yuán</i>
NumN	NumN	NumN	NumN	NumN
(NNumCI)	NNumCI	NNumCI/ NumCIN	NumCIN	NumCIN
NumV	VNum (VNumCI)	VNumCI	VNumCI	VNumCI

Parentheses mean that the order began to exist. The absence of some orders such as NNum or VNum does not imply that they did not exist at all, but they were not representative at these periods. In comparing classifier development in Chinese, we notice that nominal CIs and verbal CIs began to develop from the same origin but ended up in different places:

NumN > NNumCI > NumCIN
 NumV > VNum > VNumCI

Table 6 shows that:

- (a) At the beginning, Num preceded either N or V;
- (b) The verbal Cl began to develop when Cl was fully-fledged in the NP;
- (c) The order NumCl in NPs became prenominal, while NumCl in a VP remained in the postverbal position.

The development of classifiers in Chinese may lead us to conclude that nominal classifiers develop before verbal classifiers in NumCl-obligatory languages. It is interesting to see that the word order of phrases with nominal Cl and with verbal Cl is complementary. Sun Hongkai (1988:351), Tang Yuming (2002:206) and Shi Yuzhi (2006:192-193) have already noted this phenomenon. We will add some pieces of evidence to contribute to the discussion.

4. Complementary distribution of nominal classifiers and verbal classifiers

In §2, we showed that nominal classifiers and verbal classifiers ended up taking complementary positions in Chinese syntax. One point has to be noted: at the beginning, their distribution was not complementary. A numeral could precede a noun as well as a verb; and NumCl started out at the same side of noun or verb: post nominal and verbal. Now let us observe the distribution of nominal Cl and verbal Cl in non-Han languages and in cross-linguistic data. As we have just mentioned above, scholars noticed the complementary distribution of nominal classifiers and verbal classifiers in Tibetan-Burman languages and in the history of Chinese. Here we provide a larger sample of languages, which goes beyond Tibetan-Burman languages to include Altaic and Austro-Asian languages.

Table 7: Complementary distribution of classifiers in cross-linguistic data

Languages	VO/ OV	NumCIN NNumCl	other orders	VNumCl NumCIV	other orders	Sources
Tibetan-Burman groups of Sino-Tibetan languages						
<i>Tibetan</i>						
瑪曲藏語 <i>Máqǔ</i>	OV	NCINum		CINumV		Zhou 2003
倉洛語 <i>Cāngluò</i>	OV	NCINum	NNum	CINumV		Zhang JC 1986
門巴語 <i>Ménbā</i>	OV	NCINum	NNum	CINumV		Lu 2002
白馬 <i>Báimǎ</i>	OV	NCINum	NNumCl	CINumV		Sun HK et al. 2007
<i>Yi</i>						
彝語 <i>Yíyǔ</i>	OV	NNumCl		NumCIV		ZGDYY. 2007
拉祜 <i>Lāhù</i>	OV	NNumCl		NumCIV		Chang et al. 1986
納西 <i>Nàxī</i>	OV	NNumCl	echo Cl: N1NumN2	NumCIV		He & Jiang 1985
桑孔 <i>Sāngkǒng</i>	OV	NNumCl		NumCIV		Li YS 2002
卡卓 <i>Kǎzhúo</i>	OV	NNumCl		NumCIV		Mu 2003
柔若 <i>Róuruò</i>	OV	NNumCl		NumCIV		Sun HK et al. 2002
白語 <i>Bái</i>	VO	NNumCl	NCl or echo Cl	VNumCl		Xu & Zhao 1984
<i>Jingpo</i>						
蘇龍 <i>Sūlóng</i>	OV	NNumCl NCINum	NNum	CINumV		Li DQ 2004
景頗語 <i>Jǐngpō</i>	OV	NCINum	NNum	NumCIV		Dai & Jiang 2004
格曼 <i>Gémàn</i>	OV	NCINum	NNum	CINumV		Li DQ 2002
阿儂 <i>Ānóng</i>	OV	NNumCl	echo Cl	NumCIV		Sun & Liu 2005
<i>Mian</i>						
浪速 <i>Làngsù</i>	OV	NNumCl		NumCIV		Dai 2005
阿昌 <i>Āchāng</i>	OV	NNumCl		NumCIV		ZGDYY. 2007
載瓦 <i>Zǎiwǎ</i>	OV	NNumCl		NumCIV		ZGDYY. 2007

<i>Qiang</i>						
羌語 <i>Qiāng</i>	OV	NNumCl	NNum when Num=1	NumCIV		Liu GK 1998
普米 <i>Pǔmǐ</i>	OV	NNumCl	NNum	NumCIV		Lu 2001
嘉絨 <i>Jiāróng</i>	OV	NNumCl	NNum		NumV	Jacques 2008
史興 <i>Shǐxīng</i>	OV	NNumCl		NumCIV		Xu Dan 2009
Tai (Dong-Tai) group						
拉珈 <i>Lājiā</i>	VO	NumCIN		VNumCl		Mao et al. 1982
壯語 (武鳴) <i>Zhuàng</i> (<i>Wǔmíng</i>)	VO	NumCIN		VNumCl		Zhang JR et al. 1999
傣語 (德宏) <i>Dǎi</i> (<i>Déhóng</i>)	VO	NNumCl	NCINum when Num=1	VNumCl		Zhou & Luo 2001
莫話 <i>Mòhuà</i>	VO	NumCIN	CINNum when Num=1	VNumCl	VCINum when Num=1	Yang 2000
拉基語 <i>Lājī</i>	VO	NumCIN		VNumCl		Li YB 2000
標話 <i>Biāohuà</i>	VO	NumCIN		VNumCl		Liang & Zhang 2002
木佬 <i>Mùlǎo</i>	VO	NumCIN		VNumCl		Bo 2003
布央 <i>Bùyāng</i>	VO	NumCIN		VNumCl		Li JF 1999
Hmong-Mien (Miao-Yao) group						
苗語 <i>Miào</i>	VO	NumCIN		VNumCl		Wang FS et al. 1985
勉語 <i>Miǎn</i>	VO	NumCIN		VNumCl		Mao 2004
布努 <i>Bùnnǔ</i>	VO	NumCIN		VNumCl		Mao et al. 1982
炯奈語 <i>Jiǒngnài</i>	VO	NumCIN		VNumCl		Mao & Li 2002
Altaic languages						
撒拉 <i>Sālā</i>	OV	NNumCl NumCIN	NNum	NumCIV	NumV	Lin 1985
朝鮮 Korean	OV	NNumCl NumCIN	NNum NumN	NumCIV	NumV	Jin 2005
滿語 Manchu	OV	NNumCl	NumN	NumCIV		Wang QF 2005

Austro-asian languages						
克蔑語 <i>Kèmiè</i>	VO	NumCIN	echo Cl	VNumCl		Chen 2005
莽語 <i>Mǎng</i>	VO	NumCIN	CIN when Num=1	VNumCl	VCINum when Num=1	Gao 2003
京語 <i>Jīng</i>	VO	NumCIN		VNumCl		Ouyang et al. 1984
俛語 <i>Lǎi</i>	VO	NumCIN		VNumCl		Li XL 1999
布賡 <i>Bùgēng</i>	VO	NumCIN NNumCl		VNumCl		Li YB 2005
布興 <i>Bùxīng</i>	VO	NumCIN NNumCl		VNumCl		Gao 2004
克木 <i>Kè mù</i>	VO	NumCIN NNumCl		VNumCl		Chen 2002

The classification of languages in Table 7 is from 中國的語言 *Zhōngguó de yǔyán* (2007). Languages are more or less chosen randomly in one subgroup; recently studied languages and less-known languages have also been added. The list is not exhaustive. In Tibetan-Burman, most groups from Tibetan, Yi, Jingpo and Qiang are represented here. Some languages belong to the Tai (Dong-Tai) and Hmong-Mien (Miao-Yao) groups, as well as the Altaic and Austro-Asian families. We can observe the following based on Table 7:

1. The VO or OV order is pertinent for subgroup languages (correlates with NumCl and N order) except Bai and Dai (Dehong). Two OV patterns are found, in Tibetan-Burman and Altaic. Two VO patterns are also attested, in Tai-Hmong Mien and Austro-Asian. In the Altaic languages, the Mongolic group is absent because it seems that this group does not use Cl (measure words are excluded here); a numeral can directly modify a noun.

2. Nominal Cls and verbal Cls are complementary in most cases (see also Shi Yuzhi 2006). More precisely, if nominal Cls are on the right of nouns in a language, verbal Cls will be on the left of verbs, and vice versa, regardless of the internal order of Num and Cl. According to Sun Hongkai (1988:341), classifiers are not abundant in languages in which Cl precedes Num. Dai Qingxia & Jiang Ying's (2004) article confirms that classifiers in the Jingpo language having a CINum order in NP are at a very primitive stage. In OV languages, NumCl in Yi and Qiang groups and CINum in Tibetan and Jingpo groups tend to be postnominal and preverbal. In VO languages, except the dialect of Dai cited in the table and some Austro-Asian languages, other languages clearly prefer NumCl order. NumCl order is prenominal and postverbal in most of VO languages.

3. Four patterns of the complementary distribution of classifiers are seen in Table 7.

- (a) Tibetan-Jingpo type: NCINum and CINumV.
- (b) Yi-Qiang type: NNumCl and NumCIV.
- (c) Tai (Dong-Tai) and Hmong-Mien (Miao-Yao) type: NumCIN and VNumCl.
- (d) Mixed type in which the complementary distribution of classifiers is less neat. This is displayed in some Austro-Asian and Altaic languages.

OV languages have two types like (a) and (b), while VO languages also have two types like (c) and (d). Contemporary Chinese (Mandarin) belongs to type (c).

4. Some exceptions exist for the complementary distribution of classifiers. The Bai language has NumCl order to the right of both Noun and Verb. In the Dai dialect (Dehong), the distribution is not complementary either. NumCIN and NumCIV coexist in Korean, an OV language, and in Buxing, Kemu and Bugeng, three VO languages in Austro-Asian according to scholars' descriptions. The Buxing or Kemu speakers have close contact with Dai people. This may provide us an explanation why Dai and these languages do not present a complementary distribution of classifiers. In these languages, word order may be continuing to change due to complex contacts between languages.

5. Discussion

In §3 we saw that in some languages, against the general trend, nominal and verbal classifier distributions are not complementary. We propose that if in one language a nominal Cl and a verbal Cl are on the same side with respect to a noun or to a verb, then classifiers in this language should be in a developing stage. This could be the situation in Bai, which is influenced by Chinese and always presents atypical characteristics in the Yi group. In other words, we think that Bai is still undergoing transition in terms of word order change with regard to numeral classifiers. In some Austro-Asian languages, two orders often coexist; this also suggests that word order change is taking place. We hope that further studies of these languages will shed light on this issue.

We have looked at Chinese classifier evolution to try to understand the reason why nominal Cl and verbal Cl tend to occupy opposite positions in the quantifying phrase in most languages. Table 6 shows that the orders NNumCl/NumCIN and VNumCl coexisted during the Wei-Jin, and only from the Tang did nominal NumCl and verbal NumCl became complementary. Is the phenomenon found in contemporary Austro-Asian languages and in some Dai dialects parallel to the evolution of classifiers in Middle Chinese?

In fact, when nominal quantification and verbal quantification take the same form, i.e. at the same side of a noun or verb, confusion is possible:

- (14) 誦《詩》三百 (*Lúnyǔ* 論語, *Zǐ lù* 子路)
sòng «shī» sānbǎi
 recite Odes three hundred
 ‘Recite the three hundred Odes’

This example is cited by Ye Guichen & Luo Zhifeng (2007). The authors argue that verbal CIs formed to distinguish nominal quantification from verbal quantification.⁴ Here the NP “read-poem-three-hundred” refers to the first *Odes Shījīng* which contains 305 poems. But as Ye & Luo pointed out, nothing in this phrase’s syntax would prevent interpreting “three-hundred” as verbal quantification (‘three hundred times’). The following example is used by Tang Yuming (2002):

- (15) 黎明，圍宛城三匝 (*Shǐjì* 史記, *Gāozǔ běnjì* 高祖本紀)
lí míng, wéi Wǎn chéng sān zā
 dawn surround Wan city three circuit
 ‘At the dawn, [the army of Liu Bang] circled the City of Wan three times/three concentric circles.’

This example shows that if nominal CI and verbal CI take the same syntactic position (both at the same side of a noun or of a verb), ambiguities may arise. Tang cited Liu Shiru (1965) who had indicated that the word 匝 *zā* ‘circuit’ is nominal classifier of 城 *chéng* ‘city’ rather than a verbal classifier of the verb 圍 *wéi* ‘surround’. One can suggest that possible confusions between a nominal CI and a verbal CI require readjustment of word order.

Bai language is another language with such a possible but unrealized confusion. In a string of words Verb-Noun-CI-Num, it is possible to interpret CI-Num as part of noun phrase or as verbal modifier.

- (16) $\alpha^{31}ti^{33}$ ηv^{55} $\alpha^{31}mo^{33}$ sua^{44} la^{42} $tõ^{21}$ ka^{33} $tshẽ^{55}$
 Dad O.M.⁵ mom speak Past speech several NCL
 ‘To Dad, Mom said a few sentences.’

⁴ We do not agree with them when they assert that verbal CIs were already being used in the Qin Dynasty.

⁵ O.M. represents object marker.

- (17) $\alpha^{31}ti^{33}$ ηv^{55} $\alpha^{31}mo^{33}$ sud^{44} la^{42} $t\tilde{o}^{21}$ ka^{33} $t\tilde{a}^{21}$
 Dad O.M. mom speak Past speech several VCL
 ‘To Dad, mom spoke several times.’

However, there are no ambiguous classifiers in the language, at least we have not found ambiguous classifier in Bai.

In conclusion, two possibilities explain the non-complementary distribution of nominal CIs and verbal CIs in some languages. (1) The language has yet to complete a word order change during a historical stage; (2) There is no chance of confusion when nominal CI and verbal CI are on the same side of a noun or verb. However, nominal classifiers and verbal classifiers present a complementary distribution in most languages investigated in our paper.

References

- Bo, Wenze. 2003. *Mulaoyu Yanjiu* [*Studies of the Mulao Language*]. Beijing: Nationalities Press.
- Chang, Hong'en et al. 1986. *Lahuyu Jianzhi* [*Outline of the Lahu Grammar*]. Beijing: Nationalities Press.
- Chen, Guoqing. 2002. *Kemuyu Yanjiu* [*Studies of the Kemu Language*]. Beijing: Nationalities Press.
- Chen, Guoqing. 2005. *Kemieyu Yanjiu* [*Studies of the Kemie Language*]. Beijing: Nationalities Press.
- Comrie, Bernard. 2008. The areal typology of Chinese: between north and southeast Asia. *Chinese Linguistics in Leipzig*, ed. by Redouane Djamouri, Barbara Meisterernst & Rint Sybesma, 1-21. Paris: École des Hautes Études en Sciences Sociales.
- Dai, Qingxia. 2005. *Langsuyu Yanjiu* [*Studies of the Langsu*]. Beijing: Nationalities Press.
- Dai, Qingxia, and Ying Jiang. 2004. Mengyaqi liangci de leixingxue tezheng: Jingpoyu liangci de ge'an yanjiu [Typological characteristics of classifiers at a primitive stage-the case of Jingpo]. *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on his Seventieth Birthday*, ed. by Ying-chin Lin, Fang-min Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang & Dah-an Ho, 315-325. Taipei: Institute of Linguistics, Academia Sinica.
- Downing, Pamela. 1996. *Numeral Classifier System: The Case of Japanese*. Amsterdam & Philadelphia: John Benjamins.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68.1:81-138.
- Dryer, Matthew S. 2003. Word order in Sino-Tibetan languages from a typological and geographical perspective. *The Sino-Tibetan Languages*, ed. by Graham Thurgood & Randy J. LaPolla, 43-55. London: Routledge.
- Fang, Mei. 1993. Binyu yu dongliang ciyu de cixu wenti [Word order of object and verbal classifiers]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 1993.1:54-64.
- Fu, Jingqi, and Dan Xu. 2009. SVO yuyan li de binyu qianzhi [Preverbal object in SVO languages]. *Minzu Yuwen* [*Minority Languages of China*] 2009.3:3-15.
- Gao, Yongqi. 2003. *Mangyu Yanjiu* [*Studies of the Mang Language*]. Beijing: Nationalities Press.
- Gao, Yongqi. 2004. *Buxingyu Yanjiu* [*Studies of the Buxing Language*]. Beijing: Nationalities Press.

- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Languages*, ed. by Joseph H. Greenberg, 58-90. Cambridge: MIT Press.
- Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Languages* (2nd edition), ed. by Joseph H. Greenberg, 73-113. Cambridge: MIT Press.
- Hagège, Claude. 1985. Le chinois, l'ordre des mots et l'ordre du pensable. *T'oung Pao* 71:263-274.
- Hashimoto, Mantaro J. 1977. The genealogy and the role of the classifier in Sino-Tibetan. *Computational Analyses of Asian & African Languages* 7:69-78.
- Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernhard Comrie. (eds.) 2005. *The World Atlas of Language Structures*. Oxford & New York: Oxford University Press.
- He, Jiren, and Zhuyi Jiang. 1985. *Naxiyu Jianzhi [Outline of the Naxi Grammar]*. Beijing: Nationalities Press.
- Huang, Shengzhang. 1961. Liang Han shidai de liangci [Classifiers during the Han]. *Zhongguo Yuwen [Studies of the Chinese Language]* 1961.5:21-28.
- Huang, Xing. 1996. Woguo shaoshu minzu yuyan de cixu leixing [The word order typology of minority languages in China]. *Minzu Yuwen [Minority Languages of China]* 1996.1:10-18.
- Jacques, Guillaume. 2008. *Jiarongyu Yanjiu [Studies of the Jiarong (rGyalrong) Language]*. Beijing: Nationalities Press.
- Jin, Haiyue. 2005. Han-Chaoyu shuliang duanyu de xingshi ji yuyi bijiao [Syntactic and semantic comparison of numeral classifiers in Chinese and Korean]. *Han-Zang Yuxi Liangci Yanjiu [Studies of Classifiers in Sino-Tibetan]*, ed. by Jinfang Li & Suhua Hu, 499-505. Beijing: Minzu University of China Press.
- Li, Daqin. 2002. *Gemanyu Yanjiu [Studies of the Geman]*. Beijing: Nationalities Press.
- Li, Daqin. 2004. *Sulongyu Yanjiu [Studies of the Sulong]*. Beijing: Nationalities Press.
- Li, Jinfang. 1999. *Buyangyu Yanjiu [Studies of the Buyang Language]*. Beijing: Minzu University of China Press.
- Li, Jinfang, and Suhua Hu. (eds.) 2005. *Han-Zang Yuxi Liangci Yanjiu [Studies of Classifiers in Sino-Tibetan]*. Beijing: Minzu University of China Press.
- Li, Xulian. 1999. *Laiyu Yanjiu [Studies of the Lai Language]*. Beijing: Minzu University of China Press.
- Li, Yongsui. 2002. *Sangkongyu Yanjiu [Studies of the Sangkong Language]*. Beijing: Minzu University of China Press.
- Li, Yuming. 2000. Kaobeixing liangci ji qi zai Han-Zang yuxi liangci fazhan zhong de diwei [Echo-classifier and its role in classifier development in Sino-Tibetan]. *Zhongguo Yuwen [Studies of the Chinese Language]* 2000.1:27-34.

- Li, Yunbing. 2000. *Lajiyu Yanjiu* [*A Study of the Laji Language*]. Beijing: Minzu University of China Press.
- Li, Yunbing. 2005. *Bugengyu Yanjiu* [*Studies on the Bugeng Language*]. Beijing: Nationalities Press.
- Liang, Min, and Junru Zhang. 2002. *Biaohua Yanjiu* [*Studies of the Biao Language*]. Beijing: Nationalities Press.
- Lin, Lianyun. 1985. *Salayu Jianzhi* [*An Outline of Sala Grammar*]. Beijing: Nationalities Press.
- Liu, Danqing. 2003. *Yuxu Leixingxue yu Jieci Lilun* [*Word Order Typology and a Theory of Adposition*]. Beijing: The Commercial Press.
- Liu, Guangkun. 1998. *Mawo Qiangyu Yanjiu* [*Studies of the Mawo Qiang*]. Chengdu: Sichuan Nationalities Publishing House.
- Liu, Shiru. 1965. *Wei-Jin Nanbeichao Liangci Yanjiu* [*Studies on Classifiers in Wei, Jin, and the Southern and Northern Dynasties*]. Beijing: Zhonghua Book Company.
- Lu, Shaozun. 2001. *Pumiyu Fangyan Yanjiu* [*Studies on Pumi Dialects*]. Beijing: Nationalities Press.
- Lu, Shaozun. 2002. *Menbayu Fangyan Yanjiu* [*Studies on Menba Dialects*]. Beijing: Nationalities Press.
- Mao, Zongwu. 2004. *Yaozu Mianyu Fangyan Yanjiu* [*Studies on the Mian Dialects of the Yao*]. Beijing: Nationalities Press.
- Mao, Zongwu, and Yunbing Li. 2002. *Jiongnaiyu Yanjiu* [*Studies of the Jiongnai Language*]. Beijing: Minzu University of China Press.
- Mao, Zongwu, Chaoji Meng, and Zongze Zheng. 1982. *Yaozu Yuyan Jianzhi* [*Outline of the Yao Languages*]. Beijing: Nationalities Press.
- Mu, Shihua. 2003. *Kazhuoyu Yanjiu* [*Studies of the Kazhuo Language*]. Beijing: Nationalities Press.
- Mu, Shihua. 2005. Lun Naxiyu kaobeixing liangci de yufahua [About grammaticalization of echo-classifiers in Naxi]. *Han-Zang Yuxi Liangci Yanjiu* [*Studies of Classifiers in Sino-Tibetan*], ed. by Jinfang Li & Suhua Hu, 141-165. Beijing: Minzu University of China Press.
- Ouyang, Jueya, Fang Cheng, and Cuirong Yu. 1984. *Jingyu Jianzhi* [*Outline of the Jing Grammar*]. Beijing: Nationalities Press.
- Paris, Marie-Claude. 1981. *Problèmes de syntaxe et de sémantique en linguistique chinoise*. Paris: Collège de France.
- Peyraube, Alain. 2000. Ordre des constituants en chinois archaïque. *Cahiers de linguistique de l'INALCO* 3:99-110.
- Peyraube, Alain, and Thekla Wiebusch. 1993. Le rôle des classificateurs nominaux en chinois et leur évolution historique: un cas de changement cyclique. *Faits de langues* 2:51-61.

- Peyraube, Alain. 1998. Shanggu Zhonggu Hanyu liangci de lishi fazhan [Evolution of classifiers in Old Chinese and Middle Chinese]. *Yuyanxue Luncong [Essays on Linguistics]* 21:99-122. Beijing: The Commercial Press.
- Shen, Pei. 1992. *Yinxu Jiagu Buci Yuxu Yanjiu [Studies on Word Order of Oracular Inscriptions on Tortoise Shells and Bones of the Yin Dynasty Ruins]*. Taipei: Wenjin Press.
- Shi, Yuzhi. 2006. *Yufahua de Dongyin yu Jizhi [Motivation and Mechanism of Grammaticalization]*. Beijing: Peking University Press.
- Sun, Hongkai. 1988. Zang-Mianyu liangci yongfa bijiao [Comparison of classifiers in Tibeto-Burman Languages]. *Zhongguo Yuyan Xuebao [Chinese Linguistics: The Journal of Linguistic Society of China]* 3:339-355.
- Sun, Hongkai, and Guangkun Liu. 2005. *Anongyu Yanjiu [Studies of the Anong]*. Beijing: Nationalities Press.
- Sun, Hongkai, Zengyi Hu, and Xing Huang. (eds.) 2007. *Zhongguo de Yuyan (ZGDYY) [The Languages in China]*. Beijing: The Commercial Press.
- Sun, Hongkai, Chenglong Huang, and Maocao Zhou. 2002. *Rouruoyu Yanjiu [Studies of the Rouruo]*. Beijing: Minzu University of China Press.
- Sun, Hongkai, Katia Chirkova, and Guangkun Liu. 2007. *Baimayu Yanjiu [The Baima Language]*. Beijing: Nationalities Press.
- Sun, Yan. 2005. Ru Tang Qiu Fa Xunli Xingji de liangci [Classifiers seen in the Diary of a Pilgrimage to China in Search of the Law]. *Han-Zang Yuxi Liangci Yanjiu [Studies of Classifiers in Sino-Tibetan]*, ed. by Jinfang Li & Suhua Hu, 387-408. Beijing: Minzu University of China Press.
- Tang, Yuming. 2002. Gu Hanyu dongliang biao shifa tanyuan [Origin of verbal quantification in Old Chinese]. *Zhuming Zhongnian Yuyanxuejia Zixuanji: Tang Yuming [Self-selected Collection of Tang Yuming]*, 201-206. Hefei: Anhui Education Press.
- Wang, Fushi. 1985. *Miaoyu Jianzhi [Outline of the Miao Grammar]*. Beijing: Nationalities Press.
- Wang, Li. 1958. *Hanyu Shi Gao [Outline of the Chinese Language History]*. Beijing: Science Press.
- Wang, Qingfeng. 2005. *Manyu Yanjiu [Studies of the Manchu]*. Beijing: Nationalities Press.
- Wei, Desheng. 2000. *Shuihudi Qinmu Zhujian Yufa Yanjiu [Study on the Grammar of Bamboo Slips found in the Qin Tombs of Shuihudi]*. Beijing: Capital Normal University Press.
- Wu, Anqi. 2005. Fenxixing yuyan de mingliangci [Nominal classifiers in analytic languages]. *Han-Zang Yuxi Liangci Yanjiu [Studies of Classifiers in Sino-Tibetan]*, ed. by Jinfang Li & Suhua Hu, 1-13. Beijing: Minzu University of China Press.

- Wu, Dan. 2005. *Diachronic Investigations on Word Order Change of Numeral Classifiers in Chinese*. Beijing: Chinese Academy of Social Sciences MA thesis.
- Wu, Fuxiang. 2006. Wei-Jin Nanbeichao shiqi Hanyu mingliangci fanchou de yufahua chengdu [Degree of grammaticalization of Chinese classifier category in Wei-Jin Nanbeichao]. *Linguistic Studies in Chinese and Neighboring Languages: Festschrift in Honor of Professor Pang-Hsin Ting on his 70th Birthday*, ed. by Dah-an Ho, H. Samuel Cheung, Wuyun Pan & Fuxiang Wu, 553-571. Taipei: Institute of Linguistics, Academia Sinica.
- Xu, Dan. 2006. *Typological Change in Chinese Syntax*. Oxford & New York: Oxford University Press.
- Xu, Dan. 2009. Xiayou Shixingyu de mouxie tedian [Some characteristics of Shixing at the lower reaches]. *Minzu Yuwen [Minority Languages of China]* 2009.1:25-42.
- Xu, Lin, and Yansun Zhao. 1984. *Baiyu Jianzhi [Outline of the Bai Grammar]*. Beijing: Nationalities Press.
- Xuan, Dewu, Xiangyuan Jin, and Xi Zhao. 1985. *Chaoxianyu Jianzhi [Outline of the Korean Grammar]*. Beijing: Nationalities Press.
- Yang, Tongyin. 2000. *Moyu Yanjiu [Studies of the Mo]*. Beijing: Minzu University of China Press.
- Ye, Guichen, and Zhifeng Luo. 2007. Hanyu dongliangci xingcheng de yuanyin [About the formation of verbal classifiers in Chinese]. *Gu Hanyu Yanjiu [Research in Ancient Chinese Language]* 2007.3:81-85.
- Zhang, Jichuan. 1986. *Cangluo Menbayu Jianzhi [Outline of the Cangluo Menba Grammar]*. Beijing: Nationalities Press.
- Zhang, Junru et al. 1999. *Zhuangyu Fangyan Yanjiu [Studies of the Zhuang Dialects]*. Chengdu: Sichuan Nationalities Publishing House.
- Zhang, Junzhi, and Xiancheng Zhang. 2002. Boshu “wushi’er bingfang” shuliangci yanjiu [Studies on numeral classifiers seen in the Fifty-two Prescriptions on silk]. *Jianbo Yuyan Wenzi Yanjiu [Collected Studies on Grammar and Graphs found on Bamboo Slips and Silks]* 1:191-224. Chengdu: Bashu Shushe.
- Zhou, Maocao. 2003. *Maqu Zangyu Yanjiu [Studies of the Maqu Tibetan]*. Beijing: Nationalities Press.
- Zhou, Yaowen, and Meizhen Luo. 2001. *Daiyu Fangyan Yanjiu: Yuyin, Cihui, Wenzi [Studies on Dai Dialects: Phonetics, Lexicon, Writing]*. Beijing: Nationalities Press.

When Lexicalization Meets Grammaticalization: The Development of ‘*wang*+path’ Adverbials in Northern Chinese^{*}

Christine Lamarre (柯理思)

INALCO-CRLAO

This paper discusses a productive compounding pattern, which combines preposition 往 *WANG* (*wàng* or *wǎng*) ‘towards’, with a path-expressing element, either a monosyllabic localizer, e.g. 往裡 *wǎnglǐ* ‘in’, or a path verb, e.g. 往回 *wǎnghuí* ‘back’ or 往起 *wǎngqǐ* ‘up’. These compounds, like directional complements, express the core path meanings of a spatial motion event (‘up’, ‘down’, ‘in’, ‘out’, ‘across’ etc.), but they appear before the verb, whereas directional complements follow the verb. Only the latter (e.g. 拿進去 *nájìnqu* ‘take in’, 拿回去 *náhuìqu* ‘take back’ or 拿起來 *náqilai* ‘take up’) have a bounding effect on the clause. We argue that one important motivation for this lexicalization process is the need for two complete symmetrical sets of path-marking elements which share the same repertory of path meanings, but have different aspectual implications for the clause. The second type of compounds (*WANG* + directional verb) appears later in history than the first type (*WANG* + localizer), and fills up the gaps existing in the repertory of core path meanings provided by localizers, e.g. *huì* ‘back’, *qǐ* ‘up (source-oriented)’, *guò* ‘over’ etc. We also discuss the role of grammaticalization in this evolution, and conclude on a few typological perspectives.

Key words: path of motion, directionals, localizers, Northern Mandarin, lexicalization, grammaticalization

1. Introduction

We discuss here a productive compounding pattern, which combines preposition 往 *WANG* (*wàng* or *wǎng*) ‘towards’, with a path-expressing element (hereafter path),

^{*} We develop here some sections of a paper read at the 4th Kentridge roundtable on Chinese Linguistics on grammaticalization and lexicalization (National Singapore University, September 2008). I would like to thank here the organizer of the workshop, Pr. Peng Rui, and several other colleagues for their comments. The topic exactly fitted this volume, considering Pr. Alain Peyraube’s outstanding contributions to the study of both locative constructions and grammaticalization.

either a monosyllabic localizer, e.g. 往裡 *wǎnglǐ* ‘in’, or a path verb, e.g. 往回 *wǎnghuí* ‘back’ or 往起 *wǎngqǐ* ‘up’. *WANG*-path compounds, like directional complements, express the core path meanings of a spatial motion event (‘up’, ‘down’, ‘in’, ‘out’, ‘across’ etc.), but they appear before the verb, whereas directional complements follow the verb.

This study is part of a wider project which has the ambition to give another illustration of an important principle of form-meaning pairing in Chinese: the correlation of syntactic position (left or right of the verb) and perfectivity. This topic has been discussed in a number of studies of various theoretical backgrounds, dealing mainly with cases when the same noun phrases appeared before and after the verb, for instance “*zài* + locative noun” and “*gěi* + noun phrase”. In the case discussed here, though, the first component of the *WANG*-PATH adverbials, preposition *WANG* ‘toward’, expresses the direction of the motion and appears only at the left of the verb,¹ contrasting with directionals and with the goal marker 到 *dào* ‘to’ which appear after. The second component is in mainstream Mandarin a localizer, which bears no relation whatsoever to the directional complements that express the path in Verb-Directional compounds. The fact that these two types of path-expressing element belong to separate word classes has made the form-meaning correlation less conspicuous than in other constructions. Through the lexicalization process discussed here, path verbs combine with preposition *WANG* ‘toward’ to form *WANG*-DIR compounds, e.g. 回 *huí* ‘return’ in 往回走 *wǎnghuí zǒu* ‘walk back’ or 起 *qǐ* ‘rise’ in 往起撿 *wǎngqǐ jiǎn* ‘pick up’. As a result, the same path-expressing elements, i.e. directionals, may appear before the verb in an adjunct prepositional phrase (hereafter PP) and after the verb as a resultative, e.g. 走回去 *zǒuhuíqu* ‘walk back’ or 撿起來 *jiǎnqilai* ‘pick up’. This unveils two types of path marking in a ‘mirror image’.

This paper is organized as follows: Section 2 presents an overview of Chinese localizers and directionals, the two path-expressing categories involved here. Section 3 describes both types of *WANG*-PATH compounds and discusses their ‘wordhood’ and their historical development. Section 4 presents data illustrating the use of both types of *WANG*-PATH compounds in standard and less standard Mandarin. Section 5 discusses the motivation for this lexicalization process. The conclusion gives a few typological perspectives.

¹ We showed (Lamarre 2007a, Tang & Lamarre 2007) that verbs followed by *WANG* appear only in written style, and that in the spoken language of the Northern dialects they are not used.

2. The development of *WANG*-PATH adverbials in Mandarin Chinese

2.1 The linguistic expression of the path of motion in Chinese: localizers and directionals

Mandarin Chinese uses various linguistic devices to encode the path of motion in a spatial motion event: besides path verbs (e.g. 進 *jìn* ‘enter’, 出來 *chūlai* ‘come out’), it can use directional complements, prepositions, and localizers, or a combination of these elements. Prepositions and monosyllabic localizers necessarily combine with a locative noun phrase (hereafter locative NP), e.g. source preposition *cóng* ‘from’ and localizer *-li* ‘in’ in (1a), and therefore belong to an *adnominal* type of encoding. Directional complements mainly combine with verbs. Although they may also introduce a locative NP, e.g. *jìn* ‘enter’ in (1b), they often appear without, e.g. (1c), and can thus be considered as an *adverbial* type of path encoding.²

- (1) a. 從 包裡 掏出來
 cóng bāo-lǐ *tāo-chū-lai*
 from bag-inside draw-exit-come
 ‘draw out of the bag’
- b. 放進 包(裡)
 fàng-jìn bāo(-lǐ)
 put-enter bag(-inside)
 ‘put into the bag’
- c. 放進去
 fàng-jìn-qu
 put-enter-go
 ‘put in’

These various types of path encoding combine following complex and numerous patterns which differ according to dialects and style (see Lamarre 2007a, Tang & Lamarre 2007 etc.). The lexicalization pattern discussed here originates in a prepositional phrase where the preposition *WANG* ‘toward’ (*wǎng* or *wàng*, hereafter represented as *WANG*) combines with a locative NP (which may include a localizer) to express the direction of the motion in a motion event. Instead of a whole locative NP, preposition *WANG* ‘toward’ has come to combine with localizers to form an adverbial which expresses the

² The terms of *adverbial* and *adnominal encoding* are borrowed from Berthele (2004). They correspond roughly to the distinction made by Talmy between satellites and prepositions. Satellites modify the verbs and may appear without the reference (or ground) noun phrase.

direction of the motion such as ‘in’ or ‘up’. The focus of this study, *WANG*-PATH adverbials, is divided into two subtypes according to the nature of the path-expressing element: a monosyllabic localizer (LOC) in *WANG*-LOC adverbials, e.g. 往裡 *wǎnglǐ* ‘in’; a directional complement (DIR) in *WANG*-DIR adverbials, e.g. 往裡 *wǎnghuí* 往回 ‘back’ or 往起 *wǎngqǐ* ‘up’. We give below a brief presentation of the two grammatical categories from which these path-expressing elements are taken, i.e. localizers and directional (or path) verbs. In modern Chinese, localizers and directional (or path) verbs form two distinct closed-class categories which have been widely discussed in the literature, but rarely compared one to each other.

2.2 Localizers

Localizers express “the relative position of objects” (Chappell & Peyraube 2008:15). Their main function is to change the ordinary noun they follow into a ‘place word’, which can then function as a locative NP in specific syntactic environments, for instance combine with locative prepositions and verbs like 從 *cóng* ‘from’, 往 *wǎng* ‘toward’, 在 *zài* ‘be at, at’, or 到 *dào* ‘to arrive at, go to’. Chinese localizers follow the evolutionary path proposed by Svorou (1993:101) and by Heine et al. (1991:132), from a noun to an adposition through a genitive construction (see Peyraube 2003). They are considered either as a subcategory of nouns or as postpositions. It is generally admitted that disyllabic localizers function more like nouns and monosyllabic ones more like adpositions: only the former can function as subject or object of a verb, and the latter are bound forms and often undergo some degree of phonetic reduction (tonal neutralization, erosion of the initial consonant etc.). Localizers are also known to combine with prepositions such as 往 *wǎng*, 朝 *cháo*, 向 *xiàng* ‘towards’, or 從 *cóng* ‘from’ etc., like in the *WANG*-PATH compounds examined here.

Chinese localizers, like the *spatial grams* discussed by Svorou (1993:31), designate a spatial region of the locative noun used as the reference of the localization or of the motion (the *landmark* in her terms, or *ground* NP in Talmy’s terms). The repertory of Chinese localizers fits with Svorou’s (1993) core regions used in spatial reference: the INTERIOR/EXTERIOR/TOP/BOTTOM/Front/BACK/SIDE/MEDIAL REGIONS (respectively 裡 *lǐ*, 外 *wài*, 上 *shàng*, 下 *xià*, 前 *qián*, 後 *hòu*, 旁 *páng*, 中 *zhōng* and 間 *jiān* etc., see Peyraube 2003 and Chappell & Peyraube 2008 for more details). If we follow Talmy (2000:53-57) in distinguishing several components of path, localizers express among other things what he calls the *conformation* component of the path, which includes notions such as “the inside of an enclosure” or “the surface of a volume”.

2.3 Directional (or path) verbs and complements

Chinese directional complements grammaticalized from path verbs (or directional verbs *qūxiàng dòngcí* in the traditional terminology; see Ota 1958:210, Peyraube 2006 and Ma 2008), and compound with verbs to form VERB-DIRECTIONAL compounds just as the resultative complements do (see ex. (1)). Several studies of various theoretical backgrounds (Talmy 2000:108-109, Lamarre 2003, 2007b, Shen 2003, Peyraube 2006) have pointed out the similarity of directional complements with Talmy's path satellites (Talmy 2000:101-112). Strictly speaking, only a subclass of Chinese path verbs function as directional complements: their repertory is similar to that of path satellites in typical satellite-framed languages. Slobin (2001), in discussing the role played during language acquisition by the opposition between closed-class categories and open-class categories, noted that although verbs generally form an open-class category, in a language such as Jacaltec (a Mayan language spoken in Central America, as described in Craig 1993), where directionals grammaticalized from verbs, the verbs expressing core path schema actually form a closed-class category within the open category of verbs.³ Table 1 shows the correspondence between Jacaltec path verbs and path directionals (Craig 1993), Standard Mandarin path verbs and path directionals, and Hungarian verb prefixes, to illustrate Slobin's remark about the 'schematic, generalized meanings'. Hungarian verb prefixes do not have any path verbs corresponding to the verb prefixes listed here, and can therefore be considered as a more 'typical' satellite-framed language than Chinese, which uses both path verbs and path satellites (see Lamarre 2003 and 2007b).

³ Slobin (2001:419) noted this about Jacaltec's directionals: "The directionals have all of the defining features of closed-class morphemes: there is a small, phonologically reduced set of bound morphemes, with schematic and generalized meanings. However, each of these suffixes corresponds to a full verb of motion, and such verbs are clearly an open class by standard definitions. [...] In fact, both sets are small and closed, and both have the familiar characteristics of grammaticizable notions. Clearly, the directionals are grammaticized forms of the verbs. And just as clearly — within the 'open class' — these ten motion verbs constitute a small, closed class."

Table 1: The core path verbs and path directionals in Jacaltec, Hungarian and Chinese

Jacaltec VERB > DIRECTIONAL	Hungarian	Mandarin Chinese VERB > DIRECTIONAL
<i>toyí</i> ‘go’ → <i>-toj</i> ‘away from’	VERB <i>megy</i>	去 <i>qù</i> ‘go’ → <i>-qu</i> ‘thither’
<i>tita</i> ‘come’ → <i>-tij</i> ‘toward’	VERB <i>jön</i>	來 <i>lái</i> ‘come’ → <i>-lai</i> ‘hither’
<i>ahi</i> ‘to ascend’ → <i>-(a)h</i> ‘up’	PREFIX <i>fel-</i>	上 <i>shàng</i> ‘ascend’ → <i>-shang</i> ‘up’
<i>ayi</i> ‘to descend’ → <i>-(a)y</i> ‘down’	" <i>le-</i>	下 <i>xià</i> ‘descend’ → <i>-xia</i> ‘down’
<i>oki</i> ‘to enter’ → <i>-(o/e/i)k</i> ‘in’	" <i>be-</i>	進 <i>jìn</i> ‘enter’ → <i>-jin</i> ‘in(to)’
<i>eli</i> ‘to exit’ → <i>-(e/i)l</i> ‘out’	" <i>ki-</i>	出 <i>chū</i> ‘exit’ → <i>-chu</i> ‘out’
<i>ek’i</i> ‘to pass’ → <i>-(e/i)k</i> ‘through’	" <i>át-</i>	過 <i>guò</i> ‘pass, cross’ → <i>-guo</i> ‘over’
<i>paxi</i> ‘to return’ → <i>-pax</i> ‘back/again’	" <i>vissza-</i>	回 <i>huí</i> ‘return’ → <i>-hui</i> ‘back’
<i>kanh</i> ‘to rise’ → <i>-kanh</i> ‘up’	/	起 <i>qǐ</i> ‘rise’ → <i>-qi</i> ‘up’

Although each language has a few idiosyncratic items for which no corresponding form can be found in other languages, a similar set of core path schema is found in many other languages (see François 2003 or Ross 2004 on Oceanic languages). This explains why in order to introduce the Western or the Russian reader to Chinese directionals, Chao (1968:459) and Jaxontov (1958:90) compared Chinese with German and Russian verbal prefixes (about Russian preverbs as path satellites, see Talmy 2000:106).

2.4 The repertoires of localizers and of directionals

We saw that from a semantic point of view, localizers and directional complements express different facets of path or spatial reference. Whereas localizers may also appear in sentences expressing static location or motion, directional complements are devices which (in their spatial meaning at least) encode mainly the path of motion in sentences expressing a motion. In sentences expressing a spatial motion (like examples (1a-c) above), their repertoire displays a partial overlapping shown in Table 2.⁴ Here a distinction must be made for English glosses between ‘back₁’ antonym of ‘front’ (like in ‘step back’) and ‘back₂’ meaning ‘go back to one’s original position’, and between ‘up₁’, expressing an upward motion toward a goal (used for instance in a running motion from the ground-floor to the second floor), and ‘up₂’, which is source-oriented (used for instance when the moving entity stands up from a chair, or when the agent picks up something from the floor). Two items, *shàng* and *xià*, are homomorphic (and polysemous). In *WANG*-PATH adverbials, where *WANG* expresses direction ‘toward’, locatives and directionals come to fill the same slot, and both encode the path of motion.

⁴ This table lists only a few core items. Other items such as 近 *jìn* ‘close’ and 開 *kāi* ‘away’, and the deictic directionals 來 *lái* and 去 *qù* could be added to the repertoire of the directionals, localizers 對(面) *duì(miàn)* ‘facing’, or 中間 *zhōngjiān* ‘middle’ etc. to the repertoire of the localizers.

Table 2: The core members of localizers and directionals: overlapping of path meanings

Loc.	<i>dōng</i> east	<i>páng</i> near	<i>qián</i> front	<i>hòu</i> back ₁	<i>shàng</i> top	<i>xià</i> bottom	<i>lǐ</i> inside	<i>wài</i> outside			
Dir.					up ₁	down	in	out	back ₂	up ₂	over
					<i>shàng</i>	<i>xià</i>	<i>jìn</i>	<i>chū</i>	<i>huí</i>	<i>qǐ</i>	<i>guò</i>

The grey cells at the right of the table show the path meanings for which only directionals are available. *huí* ‘back₂’, *qǐ* up₂ and *guò* ‘over’ are precisely the items which will in some Mandarin dialects combine with *WANG* to form *WANG-DIR* adverbials.

In mainstream Chinese such switches are not usual, which explains why *WANG-DIR* adverbials are rarely discussed. We will see in §4 that the only *WANG-DIR* compound fully integrated in Standard Mandarin is 往回 *wǎnghuí* ‘back (to one’s original position)’, and that other items are more or less marked as dialectal. Fang (2003) described a similar overlapping in the colloquial dialect of Beijing, where *-li* (‘inside’, a localizer) appears after verbs, in the position usually occupied by directional complements:

- (2) 你什麼時候住裡的？
nǐ shì shénme shíhou zhù-lǐ de
 2SG⁵ be which time live-inside NOM
 ‘When did you move in?’

In the case of *WANG-PATH* compounds, the switch of functions occurs the other way round: the directional fills in a slot primarily devised for a localizer. As a result, 往外推 *wàngwài tuī* and 往出推 *wàngchū tuī* both mean ‘push out’ in the Mandarin dialects which use both.

3. More on *WANG-PATH* adverbials

3.1 *WANG-LOC* and *WANG-DIR* adverbials, and their repertory

The *WANG-PATH* compounds discussed here are not usually considered as words by Chinese reference grammars and dictionaries. They are not listed in the authoritative *Dictionary of Contemporary Chinese* (現代漢語辭典 *Xiàndài Hànyǔ Cídiǎn*), but are listed in the recently published dictionary of ‘phrases’ or 短語 *duǎnyǔ* (Li & Xu 2008:

⁵ The following abbreviations are used 1/2/3 (1st 2nd 3rd person), CL (classifier), DUR (durative), NOM (nominalizer), NEG (negation), OM (object marker), PAS (passive marker), PL (plural), PRT (various particles of pragmatic nature), SG (singular).

259). Only the first of the two types discussed here, where *WANG* combines with a monosyllabic localizer, is widely described in the literature. The first descriptions of *WANG-LOC* compounds by Lü (1965[1990]) and Chao (1968:525-526) treat them like phrases consisting of a preposition and a substantive. Although Chao transcribes these compounds without hyphenation, as he does for words, he uses the vague term ‘occurrence after prepositions’, and Lü (1965[1990], 1980) uses the term ‘combination’ (*zǔhé* in Chinese); hence, neither treats them as a word. According to Lü (1980:482), a major reference glossary of grammatical words, PPs formed of the preposition 往 *wàng* ‘toward’ and localizers or locative NPs appear before the verb to express “the direction of the action”. Lü’s examples include both *WANG-LOC* compounds such as in 往前看 *wàngqián kàn* ‘look forward, look ahead’, 往外走 *wàngwài zǒu* ‘walk out’, 往上拉 *wàngshàng lā* ‘pull up’, and PPs where *WANG* combines with a nominal or pronominal element, such as 水往低處流 *shuǐ wàng dīchù liú* ‘water flows toward low places’, or 往我這兒瞧 *wàng wǒ zhèr qiáo* ‘look at me’.

According to Lü’s survey (1965[1990]) on the use of localizers in Mandarin Chinese,⁶ 往 *wàng* ‘toward’ is the preposition which appears the most frequently with the largest repertory of monosyllabic localizers: 上 *shàng* ‘top, up’, 下 *xià* ‘bottom, down’, 前 *qián* ‘front’, 後 *hòu* ‘back’, 裡 *lǐ* ‘inside’, 外 *wài* ‘outside’, 東 *dōng* ‘east’, 南 *nán* ‘south’, 西 *xī* ‘west’, and 北 *běi* ‘north’. Other prepositions expressing direction, 向 *xiàng*, 朝 *cháo* and 衝 *chòng*, also appear in combination with monosyllabic localizers, but less frequently.

Chao (1968:525) also described in detail the possible combinations of various prepositions, including *WANG*⁷ ‘toward’, 向 *xiàng* ‘toward’, 在 *zài* ‘at’, 從 *cóng* ‘from’ etc. with monosyllabic localizers. Both *WANG* and *xiàng* ‘toward’ can combine with most of the monosyllabic localizers, but according to Chao (1968:526), the fact that 內 *nèi* is more literary than its synonym 裡 *lǐ* ‘inside’ accounts for the absence of the combination *望內 **wàngnèi* ‘in’. Earlier studies show some discrepancy regarding the repertory of these compounds for prepositions other than *WANG*. For instance Lü notes that 在 *zài* ‘at’ can usually only combine with polysyllabic localizers, with one exception in his corpus: 在前 *zàiqián* ‘at the front (of it)’. Chao mentions that combinations of 在 *zài* with monosyllabic localizers are subject to various semantic restrictions, which disappear with a polysyllabic localizer (suffixed with *-tou* or *-bianr*). Fang (2003) mentions the use in Beijing colloquial of 在裡 *zàilǐ* ‘in’, considered by Chao as lacking.

⁶ The survey was based on two written corpora amounting to about 10,000 characters, and representing a rather colloquial variety of Mandarin. These include a few chapters of Lao She’s novel *Rickshaw*, also used in the present paper, which reflects a variety of northern Mandarin.

⁷ Chao (1968:525-526) uses character 望 *wàng*, (see the discussion in §3.2). We changed Chao’s original phonetic transcription to pinyin.

As for prepositions expressing the source of the motion, both Chao and Lǚ remark that 從 *cóng* and 由 *yóu* tend to be restricted to symmetrical sentences where both the source and the goal of motion appear, such as ‘from X to(ward) Y’. Lǚ only found one isolate occurrence of 從裡 *cónglǐ* ‘from inside’ in his corpus, a combination given as non-occurring by Chao. This restriction apparently pertains to the semantics of the preposition: Zou (1984) noted that 往裡去 *wǎnglǐ qù* ‘go inside’ or 朝東去 *cháodōng qù* ‘go eastward’ are possible, while 從裡來 *cónglǐ lái* ‘come from inside’ and 自東來 *zìdōng lái* ‘come from the east’ are infelicitous. He attributed this dissymmetry to the semantics of the preposition: only prepositions expressing the direction of the motion, and not those expressing the source, can compound with monosyllabic localizers. The asymmetry disappears with disyllabic localizers, i.e. when the combination operates at the syntactic level.

To conclude, previous studies do not consider *WANG-LOC* compounds as adverbs, but agree that monosyllabic localizers are bound forms, and that they combine freely only with prepositions expressing direction, whereas disyllabic localizers combine freely with any preposition no matter what the meaning.

No mention is made in the studies presented above of the other type of *WANG-PATH* compounds, where the path-expressing elements belong to the category of path (or directional) verbs, and more exactly, to the closed-class category of those path verbs which grammaticalized into directional complements. This may be because most of these compounds are felt as non-standard. Ota (1958:308, §18.8), mentions two *WANG-DIR* items: 往起 *wàngqǐ* ‘up (source-oriented)’ and 往出 *wàngchū* ‘out’ (without giving any example), and remarks that the item which combines with preposition *WANG* is not a substantive, and that these compounds should be analyzed as a kind of adverbial. We follow here this analysis and consider that both *WANG-LOC* (e.g. 往裡 *wànglǐ*) and *WANG-DIR* (e.g. 往起 *wàngqǐ*) are adverbials.

Another peculiar construction where *WANG* combines with verbs or adjectives has been mentioned in the literature (see Lǚ 1980:480). However in such constructions the verb or the adjective is usually followed by the localizer 裡 *lǐ* ‘in’, e.g., 往壞裡改 *wàng huàilǐ gǎi* ‘change for the worse’, 往死裡打 *wàng sǐlǐ dǎ* ‘beat to death’. This is obviously a related issue, especially if we consider that the early occurrences of *wànghuí* ‘back’ were followed by *-li*, but there is insufficient space to discuss it here.

3.2 More on preposition *WANG*: its origin, its phonetic and graphic variants

Preposition *WANG* ‘toward’ comes from the verb 往 *wǎng* ‘to go’ (now obsolete in modern Mandarin), and appeared around the Tang period to introduce the direction of

motion (Ota 1958:253, Jiang & Wu 1997:469-470, Ma 2002:81-85). Another preposition of similar meaning but with a different initial and a different tone (falling tone *wàng*) developed from the verb 望 *wàng* meaning ‘look (into the distance)’. According to Jiang & Wu (1997:469-470), after these verbs independently grammaticalized into prepositions of similar meaning, during the Ming and Qing period, after some time of competition, they eventually merged, and this merging triggered a new reading *wàng* for 往 ‘toward’. Both characters appear in various proportions in historical documents (see also Chao 1968:758).

The pronunciation *wàng* for ‘toward’ (falling tone category) is generally believed to reflect colloquial Northern Chinese. For instance a dictionary of Beijing lexicon (王璞 Wang Pu, 1911, 京音字彙 *Jīngyīn zìhuì*, p.116) mentions a colloquial pronunciation *wàng* for the meaning ‘towards’. Lü Shuxiang (1980:482) transcribes *WANG* as *wàng* when it stands for ‘toward’. This reading *wàng* was however eliminated in 1985 from the list of accepted readings for 往 in the RPC, and now in Standard Chinese only *wǎng* is listed (some dictionaries list *wàng* as an ‘old’ reading). Both tones are observed in various Mandarin areas for preposition ‘toward’: for instance a falling tone in central Shaanxi, and a rising tone in Kaifeng (Henan), which both belong to the Zhongyuan Mandarin subgroup.

In some documents, the Ming novel 金瓶梅 *Jīn Píng Méi* for instance, both 往 and 望 *wàng* appear in *WANG-LOC* compounds. The more recent, northern texts where *WANG-DIR* compounds appear mainly use 往. We draw therefore the conclusion that the complex issues presented above, however important they may be for the history of Chinese prepositions, bear no direct relationship to the focus of this study, i.e. the lexicalization process which leads first to a full set of *WANG-LOC* compounds then to a more or less developed set of *WANG-DIR* compounds. As noted by previous studies, most of the prepositions with the semantics of ‘toward’ (*xiàng*, *cháo*, *chòng* etc.) are likely to compound with monosyllabic localizers. We discuss mainly *WANG* here because it is the most frequent preposition with this semantics in our data.

3.3 *WANG-LOC* and *WANG-DIR* compounds: prepositional phrases or words?

As was pointed in §3.1, neither *WANG-LOC* compounds nor *WANG-DIR* compounds are usually treated as words in Chinese reference grammars and dictionaries. However, they appear in several Chinese-Japanese dictionaries edited in Japan. Some transcribed them as two separate words, and list them as idioms, for instance the *Hakusuisha Chuugokugo Jiten* (白水社中國語辭典, Hakusuisha 2002) which lists 往裡 *wǎng lǐ*, 往外 *wǎng wài*, 往上 *wǎng shàng*, 往下 *wǎng xià*, 往前 *wǎng qián* and 往

後 *wǎng hòu* (all *WANG-LOC* compounds), as well as two *WANG-DIR* adverbials, both given as ‘dialectal’: 往回 *wǎng huí* ‘back’ and 往起 *wǎng qǐ* ‘up (source-oriented)’. The older *Chuunichi Daijiten* compiled by Aichi University (中日大辭典, Taishuukan 1986/89) lists basically the same items but transcribes them as words (*wǎnglǐ*, *wǎngwài*, etc.), adding 往出 *wàngchū* ‘out’, and giving 往回裡 *wànghuǐlǐ* (or *wànghuǐli*) as a synonymous variant of 往回 *wànghuí* ‘back’.

Let’s now return to the issue of the ‘wordhood’ of *WANG-PATH* compounds. Monosyllabic localizers are bound forms (Chao 1968:524, Packard 2000:75-76, etc.), which is a sufficient reason to treat *WANG-PATH* compounds as words. If we follow Packard’s classification of Chinese words, *WANG-LOC* compounds are ‘bound root words’ (formed from a root word, ‘toward’ and a bound root, the localizers). We follow however Chao (1968:359) and call them ‘compounds’, for the sake of simplicity. Packard (2000:95-106) also analyzes VERB-DIRECTIONAL compounds as words (see his answer in note 2 p.81 about his grounds for analyzing 走進 *zǒujìn* ‘walk in’ as a word). To analyze *WANG-PATH* adverbials as words provides a consistent treatment of the compounding process at work on both sides of the verb (left and right). Here are a few other reasons to treat *WANG-PATH* compounds as words (adverbials), and to consider their development as being also a phenomenon of lexicalization.

(a) Some of the dissyllabic *WANG-LOC* compounds undergo semantic change, whereas combinations of prepositions with disyllabic localizers usually keep their spatial meaning (Chao 1968:524~). For instance *wǎnghòu* ‘back’ also has the temporal meaning of ‘from now on, in the future’, and *wǎngxià* ‘down’ is used to mark an ongoing action: 往下說 *wǎngxià shuō* corresponds to 說下去 *shuōxiàqù* ‘go on talking’. We deal here only with the spatial meanings of these forms, but the ‘wordhood’ of these *WANG-LOC* adverbials is the basis for such semantic changes: *WANG-LOC* items are listed in the lexicon (see Packard 2000:298-303), as opposed to phrases where the localizer is polysyllabic (e.g. 往下邊兒 *wǎng xiàbiānr* ‘down’), which function at the syntactic level.

(b) Both *WANG-LOC* and *WANG-DIR* items are unequally distributed in Chinese dialects. Shanghainese, for instance, developed neither *WANG-DIR* nor *WANG-LOC* adverbials, but uses PPs formed from *WANG* and a polysyllabic localizer.⁸ The next sections show that *WANG-PATH* compounds are more frequently used in northern Mandarin than in Standard Mandarin.

⁸ At this point, I would like to thank colleagues and postgraduate students from the University of Tokyo who speak Shanghainese and the Shengzhou dialect, another Wu dialect, for this information, and also Wu Fuxiang for the information that *WANG-D* compounds do exist in his mother tongue (a Mandarin dialect of the Jianghuai subgroup).

3.4 From direction PPs ‘toward NP’ to *WANG*-PATH adverbials

From a preliminary survey of a sample of various historical texts of the last thousand years, we can provide a rough sketch of the development of *WANG*-PATH compounds. *WANG*-LOC adverbials are likely to have originated from expressions where preposition *WANG* combines with a locative NP (followed by a localizer), or with a disyllabic localizer. The first *WANG*-LOC compounds we found appeared in the novel 金瓶梅 *Jin Ping Mei* (where there are not yet any *WANG*-DIR compounds to be found) and in Pu Songling’s 聊齋俚曲 *Liáozhāi lǐqǔ* (late Ming). These appear later, in the novel 紅樓夢 *Hóng Lóu Mèng* (Qing period), as shown in Table 3. In the Qing period, *WANG*-DIR compounds occur only in texts reflecting northern Chinese, so we have not listed any texts reflecting a more southern usage in Table 3.

Table 3: *WANG*-DIR compounds in a few historical texts reflecting northern Chinese

<i>WANG</i> +	<i>huí</i>	<i>huílì/huílai</i>	<i>qǐ</i>	<i>jìn</i>	<i>chū</i>	<i>guò</i>
<i>Lǎo Qǐdà</i> and <i>Piáo Tōngshì</i> (Yuan, Ming and Qing editions), <i>Jīn Píng Méi cíhuà</i> (late 16 th c.)	---	---	---	---	---	---
<i>Hóng Lóu Mèng</i> (18 th cen.) chap. 1-120	---	3/0	---	---	---	---
<i>Èr nǚ yīngxióng zhuàn</i> (mid.19 th cen.)	1	2/2	---	---	---	---
<i>Guānhuà Zhīnán</i> (1881)	---	1/0	---	---	---	---
<i>Xiǎo È</i> (1908)	---	---	1	---	---	---
<i>Luótuo Xiángzi</i> (1937)	4	---	11	---	2	---

The details of this process require a separate investigation: *WANG*-*huí* ‘back’ first appeared not as a disyllabic compound, but as a phrase where *WANG* combined with the directional verb *huí* followed by localizer *lǐ* (thus treated as a noun, like some other verbs or adjectives in modern Chinese) or with the bimorphemic path verb *huílai* ‘come back’. At the present stage we suppose that *WANG*-DIR compounds developed from *WANG*-LOC compounds through some kind of analogy. However, these data point at a switch from an adnominal encoding where preposition *WANG* ‘toward’ combines with a locative NP encoding the spatial reference of the motion, and possibly including a localizer, to an adverbial type of encoding. This shift can be schematically represented as the schema below shows. In STAGE 1, we deal with an average prepositional phrase (hereafter PP) consisting of preposition *WANG* and a locative NP. Many locative NPs consist in an ordinary noun followed by a localizer, either disyllabic (type a) or monosyllabic (type b). When the NP is covert and its interpretation relies on the context or the situation, we go to STAGE 2, where preposition *WANG* combines either with a disyllabic localizer, i.e. a free morpheme of substantive nature, to form a PP (STAGE 2a), or with a monosyllabic localizer, i.e. a bound morpheme, to form a *WANG*-LOC adverbial (STAGE 2b).

In some northern dialects (including Beijing colloquial), there is a STAGE 3, where the path-expressing element is filled in by a directional verb (or path verb) to form *WANG*-DIR compounds. The covert reference noun cannot be recovered without modifying the semantics or the structure of the verb phrase.

STAGE 1

nature of the linkage: syntactic level (PP = *WANG* + locative NP)

type of encoding: adnominal encoding of the DIRECTION of the motion

reference noun: overtly expressed

e.g. (a) 往屋子裡頭擠 *wǎng wūzi lǐtōu jǐ*;

(b) 往屋裡擠 *wǎng wūlǐ jǐ* ‘squeeze into the room’



STAGE 2a: *WANG* + disyllabic localizer

nature of the linkage: PP (phrase, syntactic level)

type of encoding: indeterminate

(the localizer functions like a substantive but the reference noun is covert)

reference noun: covert (recoverable through situation or context)

e.g. 往裡頭擠 *wǎng lǐtōu jǐ* ‘squeeze in(side)’



STAGE 2b: *WANG*-LOC adverbial

nature of the linkage: compound (morphological level)

type of encoding: adverbial encoding of the DIRECTION of the motion

reference noun: covert

e.g. 往裡擠 *wǎnglǐ jǐ* ‘squeeze in(side)’



STAGE 3: *WANG*-DIR adverbials

nature of the linkage: compound (morphological)

type of encoding: adverbial

reference noun: covert or inexistent, overt reference noun impossible to recover

e.g. 往回走 *wǎnghuí zǒu* ‘walk back’, 往起站 *wǎngqǐ zhàn* ‘stand up’

Departing from *WANG*-LOC compounds (STAGE 2b), in STAGE 3 the noun used as reference of the spatial motion cannot be recovered easily for *WANG*-DIR compounds. For instance, in the case of 往回走 *wǎnghuí zǒu* ‘walk back’ we can recover a *WANG*-NP phrase with an overt reference noun 往家裡走(回去) *wǎng jiālǐ zǒu (huíqu)* ‘walk home’, where *huí* disappears or moves to the slot of the directional complement. In the

case of *wǎngqǐ zhàn* ‘stand up’ it is impossible to insert a locative NP if we keep preposition *wǎng* ‘toward’, because *qǐ* is source-oriented (**wǎng chuáng zhàn*). A possible paraphrase will use preposition ‘from’ and express the path ‘up’ as a directional complement: 從床上站起來 *cóng chuángshàng zhànqilai* ‘rise from the bed’. Similarly, 從地上撿(起來) *cóng dìshàng jiǎn(qilai)* ‘pick up from the floor’ will correspond to 往起撿 *wǎngqǐ jiǎn* ‘pick up’. This shows that the link between *WANG*-PATH adverbials and the original PP as the syntactic combination of preposition *WANG* and a nominal locative free morpheme (stage 1), that was still recoverable for *WANG*-LOC adverbials (STAGE 2b), has been cut off for *WANG*-DIR adverbials (stage 3).

4. *WANG*-LOC and *WANG*-DIR adverbials in Northern Mandarin

4.1 Literary corpora reflecting post-war Standard Mandarin and pre-war Beijing colloquial

We mentioned in §3.4 that only texts reflecting northern Mandarin used *WANG*-DIR adverbials. In this section we compare the use of *WANG*-LOC and *WANG*-DIR adverbials in two literary corpora, one representative of prewar Beijing Mandarin (hereafter BM), Lao She’s novels,⁹ the other of the emerging Standard Mandarin of the fifties (hereafter SM), and show that both use *WANG*-LOC adverbials (with a higher frequency in the former), while *WANG*-DIR adverbials are limited to the Beijing Mandarin corpus; the only item appearing in both corpora is 往回 *wǎnghuí* ‘back (to the place of origin)’. Table 4 gives the number of tokens for each type of *WANG*-PATH compound in these two corpora of equivalent length. For the sake of simplicity, only the types of *WANG*-LOC compounds directly relevant to the discussion are taken into account, i.e. those likely to express a path meaning similar to that expressed by directional verbs. Thus *wǎngqián* ‘ahead’, *wǎnghòu* ‘back₁’, *wǎngdōng* ‘eastward’ etc., listed in Table 2, have been omitted in Table 4. Localizers ‘top’ and ‘bottom’ and directional verbs ‘ascend’ and ‘descend’ are homomorphic (*shàng* and *xià*).

⁹ To get a Northern Mandarin corpus of about the same length as the novel chosen for the Standard Mandarin corpus, 青春之歌 *Qīngchūn zhī gē* (*The Song of Youth*) (written in by Yang Mo, 370,000 characters), we added to Lao She’s novel 駱駝祥子 *Luótuo Xiángzi* (*Rickshaw*, 134,000 characters) the first part (chapters 1-34) of another of his novels, 四世同堂 *Sìshì tóngtáng* (*Four generations under one roof*). The representativity of the Standard Mandarin corpus chosen here was controlled with several other texts written in the 1980’s and the 1990’s by various authors. The English translations used here are *Rickshaw*, translated by J. M. James, (University of Hawai’i Press 1979) and *The Song of the Youth*, translated by Nan Ying (Beijing: Foreign Language Press 1964).

Table 4: Frequency of occurrence of *WANG*-LOC and *WANG*-DIR compounds in two types of Mandarin

	in	out		up ₁	up ₂	down	back ₂	over	Total
+ Localizer	<i>lǐ</i>	<i>wài</i>		<i>shàng</i>		<i>xià</i>			
+ Directional verb			<i>chū</i>		<i>qǐ</i>		<i>huí</i>	---	
Beijing Man. corpus	11	75	2	11	21	33	7	---	160
Standard Man. corpus	2	8	---	1	---	2	2	---	15

The variety of the verbs combining with these adverbials is of course much higher in the BM corpus, for instance, *wàngwài* ‘out’ appears only with three verbs (‘walk’, ‘run’ and ‘pull’) in the SM corpus, whereas in the BM corpus it combines with about 30 different verbs. *WANG*-DIR compounds are also very versatile in BM, for instance 往起 *wàngqǐ* ‘up₂’ (source-oriented motion) combines with 11 types of verbs, such as 長 *zhǎng* ‘grow’, 跳 *tiào* ‘jump’, 立 *lì* ‘rise’, 攙 *chān* ‘help by the arm’, 拿 *ná* ‘take’, 裹 *guǒ* ‘tie up’, 收拾 *shōushi* ‘tidy up’ etc., whereas it does not appear in the SM corpus at all. The following examples (3) and (4) are taken respectively from the *Song of Youth* (SM) and *Rickshaw* (BM corpus) and show that *WANG*-LOC adverbials are often translated by verbs particles in English. In ex. (3), students who took part in a demonstration were beaten and ‘dragged away’. In ex. (4), Xiangzi is questioned by the police and does not want to let the detective ‘in’ the house of his master.

- (3) [許多學生頭上流了血，]也有的被警察綁架著往外拉。
yě yǒu de bèi jǐngchá bǎngjià-zhe wǎng-wài lā
 also have NOM PAS police tie.up-DUR toward-out drag
 ‘[Many students were bruised and bleeding from wounds on the head,] some were being tied up and dragged away’. (*Song of Youth* 1 part 1, chap 15, translation p.142)
- (4) 躲他還不行呢，怎能往裡請呢！
duǒ tā hái bù xíng ne, zěn néng wǎng-lǐ qǐng ne!
 avoid 3SG yet NEG go PRT how can toward-inside ask PRT
 ‘I can’t slip away from him but how can I ask him in!’ (*Rickshaw* chap. 11, translation p.103)

4.2 *WANG*-DIR adverbials in Northwest Mandarin dialects: a complete repertory

Let’s now look at the repertory of *WANG*-D adverbials found in the 岐山 Qishan dialect of Shaanxi (west of Xi’an, Guanzhong subgroup of Mandarin dialects, field data

collected by the author in March 2008). We found there the unusual 往過 *wàngguò* ‘over, across’,¹⁰ and the deictic adverbs 往去 *wàngqù* ‘thither’ and 往來 *wànglái* ‘hither’ that we had not seen mentioned before in the literature. Previous studies on northwest Mandarin mentioned 往進 *wàngjìn* and 往出 *wàngchū* (Sun 2007:178), and *wàngguò*, *wàngqǐ*, *wàngchū*, *wàngjìn*, *wànghuí* (Mo 2004). These items also appear on the internet with a huge variety of verbs. In the case of the most conspicuously non-standard items such as *WANG-guò*, *WANG-lái*, *WANG-qù*, *WANG-chū* and *-jìn*, the geographical origin of the internet sources (blogs, local newspapers, novels, or others) when they could be identified were most often located in Shaanxi, Shanxi, Gansu, Inner Mongolia, i.e. in northwestern China.¹¹ We provide below a few examples of frequent collocations for these non-standard *WANG-DIR* adverbials.

- 往過 *WANG-guò* ‘across, over’: 走 *zǒu* ‘walk’, 跑 *pǎo* ‘run’, 開(車) *kāi* ‘drive, move (for a car)’, 拉 *lā* ‘pull’, 推 *tuī* ‘push’, 飛 *fēi* ‘fly’ etc.
- 往起 *WANG-qǐ* ‘up’: 撿 *jiǎn* ‘pick up’, 拿 *ná* ‘take’, 飄 *piāo* ‘float in the air’, 站 *zhàn* ‘stand’, 背 *bēi* ‘carry on one’s back’, 拽 *zhuài* ‘pull, drag’, 翹 *qiào* ‘raise, lift up’, 提 *tí* ‘lift’, 爬 *pá* ‘climb’
- 往進 *WANG-jìn* ‘in’: 衝 *chōng* ‘rush’, 充值 *chōngzhí* ‘charge (a cellular phone)’, 走 *zǒu* ‘walk’, 跟 *gēn* ‘follow’, 拉人 *lā rén* ‘pull someone’, 擠 *jǐ* ‘squeeze’, 吸 *xī* ‘inspire’, 闖 *chuǎng* ‘rush’, 插 *chā* ‘insert’, 爬 *pá* ‘climb’ etc.
- 往出 *WANG-chū* ‘out’: 跑 *pǎo* ‘run’, 倒 *dào* ‘throw away’, 逃 *táo* ‘run away’ etc.

Note that these verbs also frequently combine with the homomorphic directionals. Table 5 shows how these *WANG-DIR* compounds contributed to fill the gaps of the repertory of the path-expressing items for *WANG-LOC* adverbials. The repertory of core path meanings now includes two deictic items for which only demonstratives *zhèr* ‘here’ and *nàr* ‘there’ were available.

¹⁰ *WANG-guò* ‘over’, a *WANG-D* item which does not appear in Lao She’s texts used here as a corpus, is not necessarily an original feature of western Mandarin: several fellow researchers who grew up in Beijing confirm it is used in the Beijing colloquial too (I thank here Fang Mei and Zhao Liyan for their enlightening remarks). We found on the internet site Baidu the following question: “Who could tell me where you go when you 往過走 *wǎng guò zǒu*?” One of the answers provided said: “This is an expression used in Chengde dialect, and means ‘go ahead’” (cf. <http://zhidao.baidu.com/question/74162307.html>). Chengde and Beijing belong to the same dialect subgroup. The term is probably not understood everywhere, even in Beijing (another colleague brought up in Beijing, Prof. Lu Jian, said she did not know its meaning).

¹¹ Another preposition of similar meaning, 朝 *cháo* ‘toward’ forms the same compounds as *WANG* with directional verbs in Xi’an dialect (this information was provided by Prof. Lan Binhuan from Shaanxi Normal University, PC, 2007).

Table 5: *WANG*-PATH adverbials in the Qishan dialect of Shaanxi, and on the internet

	in	out	up ₁	up ₂	down	back ₂	over	hither	thither
+ Demonstratives 'here/there'	---	---	---	---	---	---	---	<i>zhèr</i>	<i>nàr</i>
+ Localizer (<i>WANG</i> -LOC)	<i>lǐ</i>	<i>wài</i>	<i>shàng</i>	---	<i>xià</i>	---	---	---	---
+ Directional (<i>WANG</i> -DIR)	<i>jìn</i>	<i>chū</i>		<i>qǐ</i>		<i>huí</i>	<i>guò</i>	<i>lái</i>	<i>qù</i>
<i>WANG</i> -DIR internet examples	+	+	+	+	+	+	+	+	+

4.3 Two types of adverbial marking of the path of motion

We saw that *WANG*-LOC compounds were often translated by English verb particles (ex. (3) and (4)). Similarly, we found in our corpus '[tears] streamed down [his cheeks]' for 往下流 *wǎngxià liú* (SY), and 'turn back' for 往回跑 *wǎnghuí pǎo* (SY). One consequence of the semantic overlapping of *WANG*-PATH adverbials and directional complements we observed above in §2.4. (Table 2), is that both are rendered by English verb particles 'up', 'down', 'out' etc. For instance the *New Age Chinese-English Dictionary* (新時代漢英大辭典, Commercial Press, 2005) translates 往外走 *wǎngwài zǒu* as 'go out' (p.1589) and 往出走 *wǎngchū zǒu* as 'come out' (p.219), mentioning that the latter is dialectal, and 走出 *zǒuchū* is translated as 'come out' (p.219). To illustrate this similitude in the temporal domain, both 往下說 *wǎngxià shuō* and 說下去 *shuōxiàqù* are translated as 'go on talking/speaking' (p.1589 & 1661, see also §4). More equivalents are given below in (5).

(5)

<i>WANG</i> -LOC compounds +V	VERB-DIRECTIONAL compounds	<i>WANG</i> -DIR compounds
往裡擠 <i>wǎnglǐ jǐ</i>	擠進去 <i>jǐjìnqu</i> 'squeeze in'	
往外租 <i>wǎngwài zū</i>	租出去 <i>zūchuqu</i> 'rent out'	往出租 <i>wǎngchū zū</i>
往上爬 <i>wǎngshàng pá</i>	爬上去 <i>páshangqu</i> 'climb up'	
往下流 <i>wǎngxià liú</i>	流下去 <i>liúxiaqu</i> 'flow down'	
	撿起來 <i>jiǎnqilai</i> 'pick up'	往起撿 <i>wǎngqǐ jiǎn</i>
	走回去 <i>zǒuhuiqu</i> 'walk back'	往回走 <i>wǎnghuí zǒu</i>

WANG-LOC adverbials positioned at the left of the verb function in a mirror image of directional complements, positioned at the right of the verb. In the case of *WANG*-DIR compounds, this mirror image gets even closer, since the path-expressing elements

are taken from the same repertory of directional verbs. However, this similarity is deceiving: *WANG*-PATH adverbials function like PPs, they are adjuncts without any influence on the aspectual features of the VP, whereas directional complements tend to function in northern Mandarin like result complements, and to have a perfectivizing effect on the verb they compound with. Jaxontov (1958:90) noted that Chinese directional complements, like Russian preverbal suffixes, are not aspect markers *per se*, they have an autonomous and concrete (often spatial) meaning, but they also have a perfectivizing effect. Chinese VERB-DIRECTIONAL compounds are not compatible with imperfective suffix *ZHE*, and rarely co-occur with progressive suffix *zài* (see Lamarre 2007a). *WANG*-PATH compounds, like the PPs they evolved from, are adjuncts and do not modify the aspectual features of the clause. They can co-occur with progressive marker *zài* (in Standard Chinese, *zài* is not originally a northern Mandarin marker), as well as with durative verbal suffix *zhe*, for instance in (6) and (7).

- (6) ...把頭髮使勁往上梳著，梳著...
 ...*bǎ tóufa shǐjìn wǎngshàng shū-zhe, shū-zhe...*
 ...OM hair vigorously toward-top com-DUR comb-DUR
 ‘Then she vigorously combed up her hair.’ (*The Song of Youth*, part 1 chap. 29, translation p.247)
- (7) 老車夫的頭慢慢的往下低，低著低著，全身都出溜下去。
 [*Rickshaw* chap. 10]
lǎo chēfū de tóu mǎnmǎn-de wàng-xià dī,
 old rickshaw NOM head slow-ly toward-bottom droop
dī-zhe dī-zhe, quánshēn dōu chūliū-xià-qù.
 droop-DUR droop-DUR whole-body all slide-descend-go
 ‘[Before the tea was ready] the old man’s head sank slowly down and down until he slid onto the floor.’ (*The Song of Youth*, translation p.91)

5. Where lexicalization meets grammaticalization

According to the criteria proposed by Brinton & Traugott (2005:97), *WANG*-PATH adverbials only reflect the first stage L1 on their three-stage cline of lexicality: their meaning is compositional and thus predictable (at least in the case of those *WANG*-PATH compounds which express the path of motion), the morpheme boundaries are still clear, and the compounding process is productive. This is also the case for VERB-DIRECTIONAL compounds which keep their spatial meaning.

On the other hand, as the localizers composing *WANG*-LOC adverbials, as well as the subset of directional verbs used to form *WANG*-DIR compound are closed-class

categories, *WANG*-DIR compounds form a closed-class category too. The development of a closed-class category is a matter of grammaticalization. What was the motivation for the lexicalization of these compounds and the formation of this category of adverbials expressing various directions of the motion?

In §3.1, we noted that only prepositions expressing the direction of the motion, not the source, may occur with monosyllabic localizers. This suggests other factors involved for the lexicalization of *WANG*-PATH compounds than mere prosodic factors: in this case why should a colloquial preposition such as *cóng* ‘from’ be excluded from this process? Why should the *direction* of the motion be the key to this process? The answer is that path-expressing elements in *WANG*-PATH adverbials and directional complements form two closed-class categories with a partly overlapping repertory of path meanings, which contrast by their syntactic position. Thanks to this grammaticalization process, two complete sets of the core schema of motion path ‘in/out/up₂/down/back₂/over’ are now available for an adverbial encoding (i.e. without an overt reference NP) at the left and right of the verb.

The repertory of paths meanings expressed by *WANG*-LOC adverbials showed some gaps: direction ‘back (to one’s original position)’, direction ‘upward’ but in a source-oriented motion, direction ‘across’ or ‘over’, due to the fact that localizers originally encode the location, not the path in a motion (see the grayed part in Table 2 above). The scenario we suggest is that directional verbs, and specifically the set of directional verbs which form VERB-DIRECTIONAL compounds, provided the lexical (or morphological) material for *WANG*-DIR adverbials, and eventually filled these gaps.

From a pan-Sinitic point of view, the correlation between the postverbal position and the perfectivizing function of the resultative predication tends to be more clear-cut in northern dialects (Lamarre 2007a and Tang & Lamarre 2007 show that postverbal locative NPs are restricted to goal NPs in northern dialects), this may explain why *WANG*-DIR adverbials expressing direction (a unbounded path) mainly developed in northern dialects. After the first items developed (probably *wànghuì* ‘back’ then *wàngqǐ* ‘up’ if our preliminary sample survey is accurate), the two sets of path-expressing elements, localizers and directional verbs, were reanalyzed as being available to encode path after *WANG*, and analogy probably accounts for the further development in some Mandarin dialects of *wàng guò* ‘over’, *wàngjìn* ‘in’, *wàngchū* ‘out’, which show some degree of redundancy with *wànglǐ* and *wàngwài*, and of the deictic *wànglái* ‘hither’ and *wàngqù* ‘thither’, which correspond to Standard Mandarin *wàng zhèr* ‘toward here’, and *wàng nàr* ‘toward there’.

6. Conclusion: typological perspectives

To conclude this short study, it is worth noting that Chinese would not be the only satellite-framed language having two symmetrical sets of path-encoding elements, one perfectivizing, the other not. Several studies of Slavic verbal prefixes show how prepositions and prefixes are often related etymologically but opposed aspectually (see for instance Daynovska & Desclés 2004 on Bulgarian; or Dąbrowska 1996 on Polish). To come back to the case of Hungarian, a typical satellite-framed language (but not an Indo-European language), verb prefixes presented in Table 1 have a perfectivizing function besides their specific path meaning, as is shown in example (8) for the preverb *át* ‘across’. When they are moved to a postverbal position, they keep their path meaning but lose their perfectivizing function, and the clause becomes imperfective, as in (9) (examples are taken from Knittel et al. 2002:71):

- (8) Péter *át*-men-t-ø a *híd*-on
 Peter across-go-PAST-3S the bridge-SUPERESSIVE
 ‘Peter crossed the bridge’ [Fr. *a traversé*]
- (9) Péter men-t-ø *át* a *híd*-on
 Peter go-past-3S across the bridge-SUPERESSIVE
 ‘Peter was crossing the bridge’ [Fr. *traversait*]

From this point of view, the motivation for the lexicalization/grammaticalization process discussed is probably not restricted to Chinese. We are also convinced that a systematic comparison of Chinese *WANG*-PATH adverbials and directional complements will contribute to our understanding of the respective role of satellites and prepositions in satellite-framed languages, and to the clarification of the highly controversial category of ‘satellite’. *WANG*-PATH adverbials are not PPs — since they do not include any nominal element — and hold a ‘sister relation to the verb’ — which means that they fall under Talmy’s definition of path satellites (2000:101). We argue here against their treatment as satellites.

References

- Berthele, Raphael. 2004. The typology of motion and posture verbs: a variationist account. *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*, ed. by Bernd Kortmann, 93-126. Berlin & New York: Mouton de Gruyter.
- Brinton, Laurel J., and Elizabeth Closs Traugott. 2005. *Lexicalization and Language Change*. Cambridge & New York: Cambridge University Press.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chappell, Hilary, and Alain Peyraube. 2008. Chinese localizers: diachrony and some typological considerations. *Space in Languages of China*, ed. by Dan Xu, 15-37. Dordrecht: Springer.
- Craig, Colette G. 1993. Jakaltek directionals: their meaning and discourse function. *Languages of the World* 7.2:23-36.
- Dąbrowska, Ewa. 1996. The spatial structuring of events: a study of Polish perfectivizing prefixes. *The Construal of Space in Language and Thought*, ed. by Martin Pütz & René Dirven, 467-490. Berlin & New York: Mouton de Gruyter.
- Daynovska, Denitsa, and Jean-Pierre Desclés. 2004. A cognitive modelling of the semantic relation between prepositions and preverbs in Bulgarian: V and IZ. *Text Processing and Cognitive Technologies* 9:52-61.
- Fang, Mei. 2003. Cong kongjian fanchou dao shijian fanchou [From space to time]. *Yufahua yu Yufa Yanjiu* [Grammaticalization and Studies on Grammar], Vol. 1, ed. by Fuxiang Wu & Bo Hong, 145-165. Beijing: The Commercial Press.
- François, Alexandre. 2003. Of men, hills, and winds: space directionals in Mwoɬlap. *Oceanic Linguistics* 42.2:407-437.
- Heine, Bernd, Ulrike Claudi, and Friederike Hünemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Cambridge & New York: Cambridge University Press.
- Jaxontov, Sergej. 1958. *Hanyu de Dongci Fanchou* [The Verbal Category in the Chinese Language]. Beijing: The Commercial Press. Translated from the Russian edition: Leningrad, 1957.
- Jiang, Jicheng, and Fuxiang Wu. 1997. *Jindai Hanyu Gangyao* [A Compendium of Premodern Chinese]. Changsha: Hunan Education Press.
- Knittel, Marie Laurence, and Gyöngyvér Forintos-Kosten. 2002. Préverbes et aspects en hongrois. *Temps et Aspect: de la Morphologie à L'interprétation*, ed. by Brenda Laca, 49-79. Saint-Denis: Presses Universitaires de Vincennes.
- Lamarre, Christine. 2003. Hanyu kongjian wei yi shijian de yuyan biaoda [The linguistic encoding of motion events in Chinese]. *Xiandai Zhongguo yu Yanjiu* [Contemporary Research in Modern Chinese] 5:1-18.

- Lamarre, Christine. 2007a. Hanyu shumianyu de butong cengci de panding [Spoken and written patterns in Standard Mandarin motion events]. *Chuugoku Gogaku [Bulletin of the Chinese Linguistic Society of Japan]* 254:51-73.
- Lamarre, Christine. 2007b. The linguistic encoding of motion events in Chinese: with reference to cross-dialectal variation. *Typological Studies of the Linguistic Expression of Motion Events*, Vol. 1: *Perspectives from East and Southeast Asia*, ed. by Christine Lamarre & Toshio Ohori, 3-33. Tokyo: Center for Evolutionary Cognitive Sciences, University of Tokyo.
- Lamarre, Christine. 2008. The linguistic categorization of deictic direction in Chinese: with reference to Japanese. *Space in Languages of China*, ed. by Dan Xu, 69-97. Dordrecht: Springer.
- Li, Linding, and Xiaoying Xu. 2008. *Xiandai Hanyu Duanyu Jiexi Cidian [An Analytical Dictionary of Modern Chinese Phrases]*. Beijing: The Commercial Press.
- Lü, Shuxiang. (ed.) 1980. *Xiandai Hanyu Babai Ci [Eight Hundred Words of Modern Chinese]*. Beijing: The Commercial Press.
- Lü, Shuxiang. 1965[1990]. Fangweici shiyong qingkuang de chubu kaocha [A preliminary survey on the usage of localizers]. *Zhongguo Yuwen [Studies of the Chinese Language]* 1965.3:206-210. Also in *Lü Shuxiang Wenji [Collected Works of Lü Shuxiang]*, Vol. 2, 291-300. Beijing: The Commercial Press, 1990.
- Ma, Beijia. 2002. *Jindai Hanyu Jieci [Premodern Chinese Prepositions]*. Beijing: Zhonghua Book Company.
- Ma, Yunxia. 2008. *Hanyu Lujing Dongci de Yanbian yu Weiyi Shijian de Biaoda [The Development of the Path Verbs and the Expression of the Motion Event]*. Beijing: Minzu University of China Press.
- Mo, Chao. 2004. Bailongjiang liuyu Hanyu fangyan de jieci [The prepositions used in the dialect of the Bailongjiang valley]. *Gansu Gaoshi Xuebao [Journal of Gansu Normal Colleges]* 2004.3:33-36.
- Ota, Tatsuo. 1958. *Chuugokugo Rekishi Bunpoo [A Historical Grammar of Modern Chinese]*. Tokyo: Koonan Shoin.
- Packard, Jerome L. 2000. *The Morphology of Chinese*. Cambridge & New York: Cambridge University Press.
- Peyraube, Alain. 1980. *Les constructions locatives en chinois moderne*. Paris: Langages croisés.
- Peyraube, Alain. 2003. On the history of place words and localizers in Chinese: a cognitive approach. *Functional Structure(s), Form and Interpretation*, ed. by Y.-H. Audrey Li & Andrew Simpson, 180-198. London & New York: RoutledgeCurzon.
- Peyraube, Alain. 2006. Motion events in Chinese: a diachronic study of directional complements. *Space in Languages: Linguistic Systems and Cognitive Categories*,

- ed. by Maya Hickmann & Stéphane Robert, 121-135. Amsterdam & Philadelphia: John Benjamins.
- Ross, Malcolm. 2004. The grammaticalization of directional verbs in Oceanic languages. *Complex Predicates in Oceanic Languages*, ed. by Isabelle Bril & Françoise Ozanne-Rivierre, 297-329. Berlin & New York: Mouton de Gruyter.
- Shen, Jiaxuan. 2003. Xiandai Hanyu 'dongbu jieyou' de leixingxue kaocha [The resultative construction in Chinese: a typological perspective]. *Shijie Hanyu Jiaoxue* [Chinese Teaching in the World] 2003.3:17-23.
- Slobin, Dan I. 2001. Form-function relations: how do children find out what they are. *Language Acquisition and Conceptual Development*, ed. by Melissa Bowerman & Stephen C. Levinson, 406-449. Cambridge & New York: Cambridge University Press.
- Sun, Lixin. 2007. *Xi'an Fangyan Yanjiu* [A Study of Xi'an Dialect]. Xi'an: Xi'an Press.
- Svorou, Soteria. 1994. *The Grammar of Space*. Amsterdam & Philadelphia: John Benjamins.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*, Vol. 2: *Typology and Process in Concept Structuring*. Cambridge: MIT Press.
- Tang, Zhengda, and Christine Lamarre. 2007. A contrastive study of the linguistic encoding of motion events in Standard Chinese and in the Guanzhong dialect of Mandarin (Shaanxi). *Bulletin of Chinese Linguistics* 2.1:137-170.
- Zou, Shaohua. 1984. Xiandai Hanyu fangweici de yufa gongneng [The grammatical functions of localizers in Modern Chinese]. *Zhongguo Yuwen* [Studies of the Chinese Language] 1984.3:173-178.

再論“們”的來源

李 藍

中國社會科學院

本文以現代漢語方言為基礎，結合白話文獻來重新探討“們”的來源。我們的結論比較接近呂叔湘先生的意見：“們”應來源於“輩”字。“輩”讀成“們”的原因是發生了三次音變：第一次音變是“輩”的聲母由幫母變成明母，音變的條件是“輩”與古讀鼻音聲母的人稱代詞“我、吾、爾、汝、若”等連用，受這些字的影響而產生感染類化作用，聲母從雙唇清塞音[p-]變讀為雙唇鼻音[m-]，文獻上出現的字形是“弭、彌”等字。第二次也是感染類化音變：因受“我、爾、汝”等上聲字的影響，“弭、彌”等濁平字又變讀為上聲字，文獻上出現的字形是“每”字。第三次音變產生在音節內部：因鼻音聲母而增生了一個鼻音韻尾，文獻上出現的字形有“邁、門、們”等字。最後本文還討論了不同時代的文獻中“們”“每”交替出現及“們”“每”在同一部文獻共現等問題。

關鍵詞：人稱代詞複數標記，感染音變，“們”的來源

1. “們”的來源的四種意見

1.1 先秦漢語的人稱代詞沒有複數標記，到兩漢的文獻中，開始出現了“輩”、“屬”、“等”、“曹”等與指人名詞、人稱代詞連用，表示類屬關係的詞語。唐代文獻裡開始出現了“弭”、“彌”等人稱代詞複數標記，但在主流文獻裡並不通行。在宋及此後的白話文獻中，出現了學術界一致認可的真性人稱代詞複數標記，明清以後的白話文獻裡，“們”字終於成為主要的人稱代詞複數標記，並一直延用到普通話及許多漢語方言中，成為現代漢語中最具優勢地位的人稱代詞複數標記形式。

“們”在文獻裡出現的時代、順序等問題在學術界爭議不大，但“們”的本字，“們”的來源，學術界的意見卻有很大分歧，直到現在，仍沒有一種說法能得到學術界的一致認可。主要有以下四種觀點。

1.2 第一種觀點認為“們”字可能來源於“輩”。呂叔湘先生(1985:54-103)最早、也最系統地研究了“們”在文獻中的書寫形式和來源等問題。呂先生的基本意見是：“們”字應來自“輩”。從“輩”到“們”的演變脈絡是：

輩（時代：漢以前。條件：可用在人稱代詞後表複數）>弭/偉（時代：唐。條件：用在人稱代詞後表複數）>懣/門（時代：北宋。條件：人稱代詞後的複數形式）>們/每（時代：南宋，金。人稱代詞後的複數形式）>們（時代：明以後直到現代，專用的人稱代詞複數形式）。

“們”來源於“輩”的說法近年來因得到馮春田(1997, 1998)的進一步論證而稍占上風，但學術界一般還是認為尚需進一步的論證和研究。詳見蔣紹愚、曹廣順(2005:82-84)。¹

1.3 第二種觀點認為“們”來源於“門”。最早提出這個意見的是日本學者太田辰夫先生(2003:316)。其他持這種看法的人也比較多，俞敏(1989)、張惠英(1995)等人均持此說。太田、俞敏等人只是一個結論性的意見，並未展開討論。張惠英有比較詳細深入的討論，基本的觀點是：“門”字與“家”“裡”等可放在人稱代詞後面表示複數的詞在語義上有相同之外，也有類屬義，宋代文獻中也有“門”字用作複數標記的用例。這個解釋在語義上有明顯的優點：和“家”一樣，也是表人的類屬詞，文獻上也有用例支持。令人不能完全滿足的地方是不能把文獻中從“輩”開始，經“弭、懣、門、每”等再到“們”的整個演變鏈條統合起來作整體解釋。

李豔惠、石毓智(2000)可能沒有看到太田、俞敏及張惠英等人的研究，又重複了太田等人的這個觀點，認為“們”的本字即為“門”字。李、石二人的文章不同於張文之處主要是從語言類型學的角度提了一些理由。

1.4 第三種觀點認為“們”可能來自“每人”的合音。目前只有羅杰瑞先生(Jerry Norman 1999:120-121)一人持這種看法。不過他同時也承認，書面語中未見“每人”用作複數標記的用法。同時我們也要指出，根據我們現在作的調查，現

¹ 梅祖麟先生(1986)基本上也同意“們”字來源於“輩”並演變成“每”，但他根據其他學者的研究成果認為，“每”變讀“們”是受了古“共同阿勒泰語”裡表複數的鼻尾[*-n]的影響，“們”是一個“雙料貨”的詞尾，-n屬於阿勒泰語，“每”的部分屬於漢語，都表示複數。”這個說法雖然解釋了元代的“每”到明代讀成“們”的鼻韻尾問題，但也引起學術界的一些質疑。原因是“們”在北宋剛出現時的字形寫作“懣”、“瞞”、“門”、“們”等字形，從這些字形來看，“們”字本來就是有鼻韻尾的，在這些字形出現的北宋時期，漢語並未受阿勒泰語的影響，到了阿勒泰語對漢語可能有影響的元代，本有鼻韻尾的“們”等字卻讀成了無鼻韻尾的“每”。

代漢語方言裡雖然有六十多個人稱代詞的複數標記形式，但也沒有發現“每人”作人稱代詞複數標記的用例。因此，這種看法既未得到文獻資料，也未得到現代漢語方言的支持。

1.5 第四種觀點認為“們”的本字是“物”。持這種看法的是江藍生先生(1995)。江藍生先生認為，人稱代詞複數標記“們”與“什麼”的“麼”一樣，本字也應該是“物”字。

把“們”的本字定為“物”有兩個困難。一是文獻中未見“物”用在人稱代詞後表複數的例證。二是未能得到現代漢語方言的廣泛支持。雖然在江西安福（雷冬平、胡麗珍 2007，邱斌 2009）裡發現了可作人稱代詞複數標記的“物”字，但這是一個南方的非官話方言。根據呂叔湘先生的意見，“們”應該產生於官話方言。

此外，一些晉語方言（劉勛寧 1994，邢向東 2006）的人稱代詞複數標記也可視為“物”字，但這些讀音形式還可以有其他的解釋和分析，如劉勛寧和邢向東都認為晉語方言的[meʔ]應該是“彌”或“每”舒聲促化的結果，本字不是“們”字。

1.6 以上這四種觀點雖然各不相同，但就研究方法、語言材料等來說卻大同小異。雖然也提及方言，但主要是把方言作為討論時一個備用證據，有時甚至只是把方言因素當作一個無計可施、無法可想時的臨時性理由，迄今為止，還沒有人完全從現代漢語方言出發來研究過“們”的來源和本字問題。本文擬從人稱代詞複數標記的共時分布狀況出發來研究這個問題。具體的做法是：先整理現代漢語方言人稱代詞複數標記的使用情況，然後和文獻對照，最後根據現代方言的情況對文獻中記載的事實進行解釋和分析。

2. 現代漢語方言的人稱代詞複數標記形式

2.1 李藍(2008)曾根據近二十餘種方言調查報告研究過 443 種現代漢語方言人稱代詞的複數表示法和複數標記形式，人稱代詞複數表示法的問題本文不再涉及，下面根據原文相關部分轉述人稱代詞複數標記形式的內容。

2.2 根據 16 種文獻的初步統計，現代漢語方言中共使用了 61 種人稱代詞複數標記。如下〈表 1〉。

〈表 1〉漢語方言人稱代詞複數標記的字形、分布地域及所屬方言²

序號	音節	聲母	字形	分布地域	所屬方言
1	單 音 節	n	ne	湘鄉、安仁	湘語
2		d ^h	d ^h ə	蒲圻、通城	江淮官話
3		l	liɛ ⁶	河源	客贛方言
4		tɕ	teioŋ ¹	大餘	客贛方言
5		tɕ	tɕɛ	孝感	江淮官話
6		ø	肮	婺源	徽語
7		p	輩	莆田	閩語
8			丙	攸縣	客贛方言
9		tɕ	儕	永安	閩語
10		t	代	懷集	粵語
11			的	西安、天水、羅田、運城	中原官話
12			等	翁源	客贛方言
13			底	永濟	中原官話
14			抵	陸川	客贛方言
15			哋	廣州、肇慶、四會、德慶、雲浮、羅定、郁南、清遠、佛岡、英德、韶關、曲江、仁化、樂昌	粵語
16			都	聞喜	中原官話
17			兜	清溪、香港	客贛方言
18			隊	衢州	吳方言
19			多	寧都、弋陽、建寧、邵武	客贛方言
20		ŋ	恩	會同	湘語
21		k	格	永新	客贛方言
22		h	火	大田	閩語
23		tɕ	幾	宿松、汾西、霍州	江淮官話，晉語
24			家	常州、江陰、黃安、麻城、嘉魚、資興、新絳	吳方言，江淮官話，客家話，晉語
25			接	紅安、汝城	江淮官話，客贛方言

² 本文不考本字，〈表 1〉中所用字形或音標均按原文。

26		l	拉	龍游、大冶、藍山	吳方言，江淮官話，湘南土話
27			徠	泰州、南通	江淮官話
28			勒	茶陵、陽新	客贛方言
29			類	連縣	粵語
30			俚	無錫、南通	吳方言
31			哩	連南、醴陵、新余、宜豐、修水、寧鄉、益陽、安化、南縣、湘陰、岳陽、臨湘、平江、醴陵、城步、新寧、武岡、寧遠、嘉禾、耒陽、宜章	客贛方言，湘語
32			良	金華	吳方言
33		m	買	道縣	湘南土話
34			們	全國各省區	各大方言
35			墨	綏德、大同	晉語
36		n	納	茶陵	客贛方言
37			南	祁陽	湘語
38			能	新化、永明	湘語，湘南土話
39			泥	桂東	湘語
40			伲	開化、廣寧、新興	吳方言，粵語
41			農	慶元、黟縣、福鼎、龍岩、宜章	吳方言，徽語，客家話，湘南土話
42		ɿ	人	雲和、績溪、歙縣、屯溪、休寧、三都、桂陽、東安	吳方言，徽語，湘方言
43		tʰ	推	黃岩	吳方言
44		tɕ	已	丹陽	吳方言
45	雙 音 節	m, ɿ	[mɛn ⁵] 人	武平	客贛方言
46		t, tɕ	[tɕɛin]	黃梅	江淮官話
47		p, ɿ	巴人	都昌	客贛方言
48		tɕʰ, m	儕們	長汀	客贛方言
49		t, tɕ	大家	祁門	徽語

50		t, ʂ	兜余	揭西	客贛方言
51		t, z _l	多人	寧化、餘干、南城	客贛方言
52		k, n	各農	福州、古田、寧德、尤溪	閩語
53		h, z _l	夥人	建甌、建陽	閩方言
54		tɕ, k	幾個	蘄春	江淮官話
55		tɕ, Ø	幾位	吉縣	晉語
56		l, n	亂儂	永康	吳方言
57		m, z _l	媽人	吉水	客贛方言
58		m, z _l	乜人	封開	客贛方言
59		tɕ ^h , tɕia	齊家	敘浦	湘語
60		ɕ, n	些農	遂昌	閩語
61		ɕ, n	星農	常山、玉山	吳方言

下面〈表 2〉從方言區的角度來觀察各方言使用人稱代詞複數的情況。表中的“自用標記”指出現在本方言區的人稱代詞複數標記，“合用標記”指同時還出現在其他方言中的複數標記。

〈表 2〉不同方言區人稱代詞複數標記的使用情況³

方言	自用標記	合用標記	合計	複數標記字形例舉
東北官話		1	1	們
北京官話		1	1	們
膠遼官話		1	1	們
冀魯官話		1	1	們
蘭銀官話	6	4	10	們，的，幾，底，曹，哈，家，哈家，tsɿ，teie
西南官話	1	1	2	們，每（陰平）
晉語	3	2	5	們，每，幾位，墨，彌
中原官話		4	4	們，都，的，底
徽語	1	2	3	們，大家，肮
粵語	3		3	哋，代，類
湘南土話	2	2	4	們，能，買，農
江淮官話	5	1	6	們，[te cin]，dʰə，幾個，徠，tse
閩語	6	1	7	們，夥人，火，各農，輩，些農，儕
湘語	5	2	7	們，能，泥，恩，南，ne，齊家
吳語	9	1	10	們，星農，已，推，良，隊，俚，亂儂，拉，佢
客贛方言	15	2	17	們，抵，納，勒，儕們，teioŋ ¹ ，巴人，乜人，lie ⁶ ，媽人，兜余，多，多人，兜，等，[men ⁵]人，格，丙，物

³ 一，〈表 2〉中的數字與現代漢語方言人稱代詞複數標記的實際使用情況可能會有一些出入，肯定會少一些。這主要是目前還沒有全國性的完整調查材料，另外，我們也沒有把單篇論文中討論的其他複數標記形式全部收集起來。從表中看起來，客贛方言的情況顯得最為複雜，其主要原因是未把這兩個方言分開，如果各算一半，就沒有吳語複雜。客贛方言人稱代詞複數標記顯得比較複雜的另一個原因是對分布在農村山鄉的方言調查得比較深入細緻。但總體上說來，這個表還是能基本反映現代漢語各大方言在人稱代詞複數標記複雜程度上的大致差別。

二，蘭銀官話的情況據張成材、莫超 (2005:30-35) 補。

三，贛語的“物”據雷冬平、胡麗珍 (2007) 補。

3. 人稱代詞複數標記的古今對比

3.1 下面我們對比一下文獻記載情況與現代漢語方言在人稱代詞複數表示法及人稱代詞複數標記在類型、複數標記的使用數量、複數標記詞源類型、複數標記的讀音類型等幾個方面的差別。

(i) 人稱代詞複數表示法的類型：

文獻使用了同用型、換用型和複用型三種，複數表示法則用了同字法、增標法兩種，計三類 4 種。現代漢語方言使用了四類 24 種。

(ii) 複數標記的數量：

白話文獻中共使用了“偉”、“彌”、“滿”、“懣”、“瞞”“門”、“每”、“們”等約 10 個左右，現代漢語方言使用了 61 個。

(iii) 複數標記的詞源類型：

近代漢語中的“偉”、“彌”、“懣”、“門”、“每”、“們”等複數標記的詞源難以確認而需討論，但現代漢語方言裡的一些複數標記則有明確的詞源，大致可以分成以下六種類型：

- (1) 源於“人”義：如“儂”、“農”、“俚”、“人”、“多人”、“巴人”、“乜人”、“夥人”等；
- (2) 源於指人的“多數”義：如“幾個”、“幾位”、“夥”、“火”等；
- (3) 源於“家庭”義：如“家”、“大家”、“儕家”等；
- (4) 源於第一人稱的反身代詞：如“已”、“家自”等；
- (5) 源於古漢語：如“輩”、“曹”、“儕”、“彌”、“等”等。
- (6) 有的可能是兩個詞語連讀而成的合音形式：如“恩”、“良”等。

在上列第 (5) 類中，曾出現在古代文獻裡的“輩”、“曹”、“等”、“等輩”、“儕”等詞與人稱代詞或指人名詞連用時主要是表“類屬”義，不是真正的人稱代詞複數標記。但在現代漢語方言裡，這些詞語都變成真正的人稱代詞複數標記。於此可見，從連類複數義詞語轉變成人稱代詞複數標記，是現代漢語方言人稱代詞複數標記的一個重要來源，一種演變方式。

當然，現代漢語方言裡也有許多人稱代詞複數標記的本字不明。

(iv) 複數標記的讀音類型：

文獻中出現的人稱代詞複數標記除了“偉”字可疑外，其他全都是唇音聲母字，屬古“幫”、“明”（微）兩母。

方言的情況則非常複雜，包括零聲母在內，共使用了唇音、舌尖前、舌尖中、舌尖後、舌面、舌根及喉等發音部位，計 18 個聲母：

p m, ts, t d^h n l, tɕ tɕ^h ɕ, tʂ tʂ^h ʂ ʐ, k ŋ h, Ø

3.2 現代漢語方言人稱代詞複數標記與文獻記載的關係。

(i) 先看人稱代詞複數表示法。

文獻中是三類 2 種，現代漢語方言中有四類 24 種，文獻的人稱代詞複數表示法類型與現代漢語方言的人稱代詞複數類型是包含與被包含的關係，現代漢語方言的人稱代詞複數類型完全覆蓋文獻中出現的類型。

從文獻與方言的對應關係看：從白話文獻中出現複數標記後，文獻中複數標記的用字類型與東北、北京等地的方言完全相同，與多數的官話方言基本一致，與多數南方漢語方言不一致。

(ii) 其次看人稱代詞複數標記的用字情況。

從複數標記的用字情況看，文獻中共出現了 6 個字形，方言裡有 61 個字形。

(iii) 第三看讀音。

從讀音情況看，文獻裡只有單字節的複數標記，方言裡不但有單音節複數標記，也有雙音節的複數標記。文獻裡只有唇音聲母，方言裡有從唇音聲母到牙喉音在內的各種聲母。

(iv) 第四看詞源類型。

從複數標記的詞源類型看，現代方言裡可以明確詞源的主要來源於“指人”類、“家庭”義、人稱代詞（或反身代詞）類及“多數”義詞語。還有一些南方漢語方言的複數標記則來自古漢語。

總的說來，對比文獻和現代方言，漢語人稱代詞複數從表示法類型到複數標記的數量都是涵蓋與被涵蓋的關係。而且，官話方言的複數標記形式與白話文獻的關係更為密切，一些南方漢語的人稱代詞複數標記則可能直接導源於秦漢時期表類屬關係的詞語。

4. 從現代漢語方言看“們”的來源

4.1 現在我們重新回到本文的開頭。比較上列各家之說後可以看出，各家的觀點各有優缺點。如果從“對文獻和方言都可以作出統一解釋”的角度來看，呂叔湘先生的看法應該還是最適切的解釋。

下面我們把文獻和方言結合起來觀察現代漢語方言中使用的人稱代詞複數標記與文獻的對應關係。結合本文第 3 節的討論，暫不涉及白話文獻中未出現的複數標記形式，可以做成〈表 3〉。

〈表 3〉現代漢語方言中一些人稱代詞複數標記之文獻溯源⁴

複數標記	分布地域及方言	文獻溯源	
		秦漢	唐宋
(1) 們	東北、華北、西南、西北等地的官話方言，及東南地區的漢語方言	—	荷公們遠來，亦欲有所補助（朱子語類） 盡他們劣心腸，偏有你（金谷遺音）
(2) 每	貴州黔西北的西南官話、陝北晉語	—	他每孤恩，適來倒埋怨人（董西廂）
(3) 彌	陝北晉語	—	我彌當家沒處得盧皮遐來（因話錄 4.10）
(4) 輩	閩語	情之所鍾，正在我輩（世說新語）	+
(5) 曹	甘肅省漳縣、通渭、秦安、會寧、定西等地的蘭銀官話	我曹言願自殺（漢書霍光傳）	+
(6) 等	客贛方言	如彼等者，無足與計天下事（史記黥布列傳）	+
(7) 儕	湘語	吾儕小人，食而聽事猶懼不給命（左傳襄公十三年）	+

⁴ 表中的白話文獻用例均據呂叔湘(1995:1-37)，秦漢文獻用例均據王力(1980:270-271)。

根據朱慶之 (1993)、龍國富 (2008) 等人的研究，在東漢時期的佛經中，與梵文相應的人稱代詞複數形式主要是用“人稱代詞+曹/等/輩”的方式去對譯。這說明，本來是用在指人名詞表類屬義的“曹、等、輩”等詞由於和人稱代詞長期固定搭配使用，已具備了轉化成真性人稱代詞複數標記的句法條件。而在秦漢時期屬於漢語核心區域的甘肅方言中發現了“曹”用在人稱代詞後表示複數的用法，則為文獻研究中發現的演變線索指明了此後的演變方向和演變結果。

4.2 從〈表 3〉中可以明顯看出，從“輩”到“彌、每、們”，文獻證據的連續性，語音上的高關聯度，都很難讓人拒絕“們”來源於“輩”的想法。呂先生 (1995:1-37) 曾指出，古幫、明兩母字有通轉之例，如“陌”（明母字）從“百”（幫母字）得聲之類。但上古漢語的諧聲關係，音變條件不明確的“通轉說”等，雖然在一定程度上也可用來解釋從“輩”到“們”的變化，但始終不能讓人完全放心。這也可能是至今學術界大都在引用呂先生觀點的同時又有所保留的原因。

其中的關鍵是如何解釋從“輩”到“們”的語音演變。

我們認為，從漢代的“輩”到宋的“們”，主要的原因是發生了三次音變，其結果是，字隨音變，“輩”字最後變成了“們”字。

“輩”的第一次音變（讀“彌”）是聲母由幫母（雙唇清塞音）變成明母（雙唇鼻音）。發生這個音變的語言條件是兩個：

一是“輩”作為後字與人稱代詞連用。二是古漢語中最常用的四個人稱代詞“吾（古疑母）、我（古疑母）、爾（古日母）、汝（古日母）”的聲母都是鼻音。

在這樣的語言條件下，“輩”字因受讀鼻音聲母的前字的影響，從幫母字變讀為明母字。這個音變過程可以構擬如下：

(i) 第一次音變：吾 *ŋ-/我 *ŋ-/爾 *n-/汝 *n-+輩 *p- → *m- 彌

這個音變是在唐代完成的。

從這個音變條件還可推測到濁平的“彌”為什麼會變讀成上聲的“每”：這可能是因為“我、爾、汝”等主要的第二、三人稱代詞都是上聲字，所以，“彌”字又一次被同類感染，變讀為上聲的“每”。

音變過程可構擬如下：

(ii) 第二次音變：我_{古上聲}/爾_{古上聲}/汝_{古上聲}＋彌_{古濁平}→每_{古上聲}

“們”讀“每”音已見於大量文獻，不用舉例。在現代漢語官話方言中也可以得到一些語言事實的支援。如筆者的母語貴州大方話（西南官話）就是這樣。在大方話中，“們”字有三個讀音：最文的讀音是讀陽平的[mən²¹]，讀陰平調的[mən⁵⁵]是最普遍的讀音，現在大多數人都用這個音。[mei⁵⁵]是最口語化的讀音，多見於城裡的老派或鄉下。從大方話文白異讀的歷史層次關係看，“每”是早於“們”的形式。

由“每”變成“們”是“輩”的第三次音變。這一次是開尾韻變成了鼻韻尾。這次音變不是由外部條件引起的，而是一次在音節內部發生的音變，音變的過程是因受鼻音聲母[m-]的影響而增生了一個鼻音韻尾。

音變過程可構擬如下：

(iii) 第三次音變：每 m+i → n 們

在現代漢語方言中，[m-n-ŋ-]等鼻音聲母影響圓唇母音或高母音韻母使其變成鼻音韻尾不是一種特別少見的音變現象。比如湖南、湖北等地的漢語方言把“木”字讀成[ɱmoŋ]（趙元任等 1948，楊時逢 1974），貴州和四川的一些方言把“質”說成[moŋ²]（李藍調查），山東牟平（羅福騰 1992:37）把“昧”讀成[men²]，山東壽光（張樹錚 1995:32）把“每天”的“每”讀成[ɱmẽ]，湖南城步儒林鎮話把“牛”說成[ɱnioŋ]（李藍調查），等等。

以上是現代漢語方言的情況。從歷史上看也有同類現象，如源於“物”的“麼”也在一些方言中讀成了“們”。這個音變應該也是發生在音節內部，受鼻音聲母影響後產生的鼻音韻尾。詳說見江藍生(1995)，本文不贅。

4.3 以上所構擬的三次音變只是大致的音變順序，無論從歷史還是從現代漢語方言看，實際情況可能都要更複雜一些。

比如說，從歷史看，文獻上還有“懣”、“門”等字形，這種情況顯然與第二次音變無關。讀去聲的“懣”應該是保留了“輩”字的原聲調，讀“門”則應該是繼承了“彌”的濁平調。

從現代方言看，西南官話中的“每”就沒有經歷第三次音變。晉語中今讀入聲的“墨”等字形則應如劉勛寧(1994)所言，是從“彌”字“舒聲促化”而來，

音變路向與其他北方官話方言截然不同。至於閩語中保留的“輩”，則只能解釋為直接承接了秦漢漢語的“輩”字。這種情況與西北方言中保留了“曹”字相當。

此外，對於幾乎所有的東南方言來說，這些地方讀鼻音韻尾的“們”可能都是經過白話文獻、普通話或其他官話方言的途徑擴散進去的，這些地方的方言另有自己的人稱代詞複數標記，從“輩”到“們”的音變過程與這些方言沒有關係。

5. “每”、“們”的反覆出現及共現問題

5.1 如果按上列音變順序來觀察這些字形在文獻中出現的情況，有兩種情況常令人困惑不已。

5.2 宋代文獻本來已出現了有鼻音韻尾的“懣、門、們”等字形，但元代人稱代詞複數標記的主要用字卻是無鼻音韻尾的“每”，到明代及此後“們”字又逐漸成為主流用法。這樣就出現了“們”→“每”→“們”的反覆變化。學術界對這種情況有不同的解釋，詳見黎新第(2001)。我們認為，結合形、音、義來看，“們”其實是一個失收於韻書的俗字，因此，從音韻發展源流的角度來討論這個字意義不大。我們的意見仍接近呂叔湘先生的看法，但這種反覆變化反映的可能是南北兩種官話方言的交替，更可能是通語與方言，方言與方言的交替。

我們認為，古代中國的國都及其周圍地區的“雅言”大概也會像今天的北京話一樣，對全國的方言，對書面語都會有影響。但是，當政治中心轉換變更後，舊雅言會逐漸式微，最後淪落為普通方言，以新國都為核心的新雅言會很快興起並對其他方言及書面語發生影響。當這種因國都及其方言背景的轉換映射到不同時代的文獻中時，一些呈現在不同時代的文獻上的“歷時演變”實際上是不同時代、不同地域的方言背景轉換，而非同一個方言的歷時演變。因此，我們在文獻上看見的“們”“每”“們”的來回變化，很有可能是“輩”字各個階段的不同讀音在不同時代，以不同的方言背景進入通語的結果。比如說，宋代文獻的“們”類字是甲方言中的讀音類型，如華北地區的官話方言。元代的“每”是乙方言的音變類型，如晉語或西南官話之類的方言。明清的“們”反映的有可能是甲方言，也有可能是發生了同類音變的丙方言在文獻中的書寫形式。

從“們”、“每”等複數標記所存在的話語環境來看這個問題，則主要是在小說、戲曲、佛經故事等貼近生活，反映時代和地域特點的通俗文學作品中，方

言口語必然要進入文獻，反映自己的存在。而深入研究這些特殊的語言現象，可能恰好是研究了漢語的真正變化和發展。

5.3 “每”、“們”在音變順序上本來是分先後的，先有“每”後有“們”，從語言符號的經濟性來說，應該是一次有一個就夠用了，不必同時使用兩個。但文獻調查的結果表明，“每”和“們”是可以在同一部文獻同時出現的。

“每”、“們”共現的原因比較複雜，我們認為可以從以下幾個方面來考慮這個問題。

一，如江藍生(1995)所指出的，這種情形可以用現代漢語方言裡常見的文白異讀來理解。在早期文獻裡，“們”與“每”共現可能反映了人稱代詞複數標記形成初期的文白競爭情形。這種情況應該是和一些現代漢語方言一樣，“每”和“們”都同時出現在一個方言裡，但“每”是早期的讀音，白讀音；“們”是晚期的讀音，是文讀音。“每”“們”在同一部文獻出現實際上是語法範疇的歷史層次現象。

二，在戲曲、小說等文體中，由於作者要形象逼真地描摹說話者的口氣，往往會使用符合說話者身分背景和地域特徵的詞語，這樣也會使得在同一部文獻裡同時出現不同的人稱代詞複數標記形式。

三，白話文獻在詞語使用上有一定的傳承性，後世作品往往沿用前代作品中使用過的一些詞語。這可能也是“每”、“們”會在一些白話文獻裡同時出現的原因。

引用文獻

- Chao, Yuen Ren (趙元任). 1956. *Xiandai Wuyu de Yanjiu* 現代吳語的研究 [*Studies in the Modern Wu Dialects*]. Beijing: Science Press.
- Chao, Yuen Ren (趙元任) et al. 1948. *Hubei Fangyan Diaocha Baogao* 湖北方言調查報告 [*Report on a Survey of the Dialects of Hupeh*]. Shanghai: The Commercial Press.
- Chen, Zhiwen (陳治文). 1988. Yuandai you zhiwu mingci jia ‘mei’ de shuofa 元代有指物名詞加“每”的說法 [Nouns referring to objects plus *mei*’ construction in Yuan Dynasty]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1988.1:71-72.
- Chinggeltei (清格爾泰). 1991. *Mengguyu Yufa* 蒙古語語法 [*A Grammar of the Mongolian Language*]. Huhehot: Inner Mongolia People’s Publishing House.
- Chou, Fa-kao (周法高). 1959. *Zhongguo Gudai Yufa: Chengdai Pian* [*The Grammar of Ancient Chinese: Reference and Pronouns*]. Hong Kong: The Chinese University of Hong Kong Press.
- Comrie, Bernard. 1981. *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: University of Chicago Press.
[Chinese version: *Yuyan Gongxing he Yuyan Leixing* 語言共性和語言類型 [*Language Universals and Linguistic Typology: Syntax and Morphology*], translated by Jiaxuan Shen. Beijing: Huaxia Publishing House, 1989.]
- Croft, William. 2000. *Typology and Universals*. Beijing: Foreign Language Teaching and Research Press.
- Deng, Xingfeng (鄧興鋒). 1993. Yuandai fushuxingwei ‘mei’ de duyin: jian lun Hanyu fushuxingwei de lai yuan ji qita 元代複數形尾“每”的讀音：兼論漢語複數形尾的來源及其他 [The pronunciation of the plural suffix *mei* in Yuan Dynasty]. *Journal of Nanjing University (Philosophy, Humanities and Social Sciences)* 1993.4:103-112.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68.1:81-138.
- Feng, Chuntian (馮春田). 1997. Hanyu fushu ciwei ‘men’ lai yuan shi shuo 漢語複數詞尾“們”來源試說 [The exploration to the source of the plural suffix *men* in Chinese]. *Suyuyan Yanjiu* 俗語言研究 [*Study of the Folklore Language*] 4:76-85.
- Feng, Chuntian (馮春田). 1998. Hanyu fushu ciwei ‘men’ lai yuan yu ‘bei’ zi shuo xudiao 漢語複數詞尾“們”來源於“輩”字說續貂 [Another exploration of the plural suffix *men* originated from *bei*]. *Yuyuan Xieying: Qingzhu Tang Zuofan Jiaoshou Qishi Shouchen Xueshu Lunwenji* 語苑擷英：慶祝唐作藩教授七十壽

- 辰學術論文集 [*Festschrift in Honor of Professor Zuofan Tang on his 70th Birthday*], 302-309. Beijing: Language and Culture University Press.
- Gao, Guofan (高國藩). 1988. Qu Nuo fengsu he Dunhuang minjian geyao Erlangwei 驅儼風俗和敦煌民間歌謠兒郎偉 [The driving Nuo customs and the folk songs Erlangwei in Dunhuang]. *Wenshi* 文史 [*Literature and History*] 1988.29:287-299.
- Greenberg, Joseph H. 1963. Some universals of grammar, with particular reference to the order of meaningful elements. *Universals of Language*, ed. by Joseph H. Greenberg, 73-113. Cambridge: MIT Press.
- [Chinese version: Mouxie zhuyao gen yuxu youguan de yufa pupian xianxiang 某些主要跟語序有關的語法普遍現象 [Some universals of grammar, with particular reference to the order of meaningful elements], translated by Bingfu Lu & Zhiji Lu. *Guowai Yuyanxue* 國外語言學 [*Contemporary Linguistics*] 1984.2:45-60.]
- Hopper, Paul J., and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge & New York: Cambridge University Press.
- Hu, Mingyang (胡明揚). 1987. Haiyan fangyan de rencheng daici 海鹽方言的人稱代詞 [The personal pronouns in the Haiyan dialect]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 1987.1:82-84.
- Huang, Borong (黃伯榮). (ed.) 1996. *Hanyu Fangyan Yufa Leibian* 漢語方言語法類編 [*Concordance of Chinese Dialect Grammar*]. Qingdao: Qingdao Press.
- Huang, Xiaoshan (黃笑山). 2001. 'Erlangwei' he 'Xitansong' de hesheng “兒郎偉”和“悉曇頌”的和聲 [The harmony of Erlangwei and Xitansong]. *Journal of Henan Radio & TV University* 2001.3:3-5.
- Huang, Zheng (黃征). 1992. Dunhuang yuanwen 'Erlangwei' kaolun 敦煌願文“兒郎偉”考論 [A discussion on Erlangwei in Dunhuang praying texts]. *Wenxue Luncong* 文學論叢, 150-166. Hangzhou: Hangzhou University Press.
- Ji, Xianlin (季羨林). 1993. Lun 'Erlangwei' 論“兒郎偉” [On Erlangwei]. *Qingzhu Rao Zhongyi Jiaoshou Qishiwu Sui Lunwenji* 慶祝饒宗頤教授七十五歲論文集 [*A Festschrift in Honour of Professor Jao Tsung-I on the Occasion of his Seventy-fifth Anniversary*]. Hong Kong: The Chinese University of Hong Kong.
- Jiang, Lansheng (江藍生). 1995. Shuo 'me' yu 'men' tongyuan 說“麼”與“們”同源 [On the homology of me and men]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1995.3:180-190.
- Jiang, Shaoyu (蔣紹愚). 1994. *Jindai Hanyu Yanjiu Gaikuang* 近代漢語研究概況 [*An Outline of the Study of Early Modern Chinese*]. Beijing: Peking University Press.
- Jiang, Shaoyu (蔣紹愚), and Guangshun Cao (曹廣順). (eds.) 2005. *Jindai Hanyu Yufashi Yanjiu Zongshu* 近代漢語語法史研究綜述 [*An Overview of Studies about the Early Chinese Grammar History*]. Beijing: The Commercial Press.

- Kortmann, Bernd. 1999. Typology and dialectology. *Proceedings of the 16th International Congress of Linguists, Paris 1997*. CD-ROM. Amsterdam: Elsevier Science.
[Chinese version: Leixingxue yu fangyanxue 類型學與方言學 [Typology and dialectology], translated by Haiyan Liu. *Fangyan* 方言 [Dialect] 2004.2:148-157.]
- Kortmann, Bernd. 2002. New prospects for the study of English dialect syntax: impetus from syntactic theory and language typology. *Syntactic Microvariation*, ed. by Sjef Barbiers, Leonie Cornips & Susanne van der Kleij, 185-213. Amsterdam: SAND.
- Kortmann, Bernd. 2003. Comparative English dialect grammar: a typological approach. *Fifty Years of English Studies in Spain (1952:2002): A Commemorative Volume*, ed. by Ignacio M. Palacios, María José López Couso, Patricia Fra & Elena Seoane, 65-83. Santiago de Compostela: University of Santiago.
- Labov, William. 2001. *Studies in Sociolinguistics: Selected Papers by William Labov*. Beijing: Beijing Language and Culture University Press.
- Lei, Dongping (雷冬平), and Lizhen Hu (胡麗珍). 2007. Jiangxi Anfu fangyan biao fushu de ‘wu’ 江西安福方言表複數的“物” [The plural form *wu* in the Anfu dialect spoken in Jiangxi Province]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 2007.3:255-257.
- Li, Lan (李藍). 1999. Hunan Chengbu Qingyimiao Tuhua de rencheng daici he zhidaici 湖南城步青衣苗土話的人稱代詞和指代詞 [The personal and demonstrative pronouns in the Chengbu Qingyimiao Tuhua in Hunan Province]. *Minzu Yuwen* 民族語文 [Minority Languages of China] 1999.6:35-38.
- Li, Lan (李藍). 2002. Hunan fangyan de rencheng daici 湖南方言的人稱代詞 [The personal pronouns of the Chinese dialects in Hunan Province]. *Studies in Chinese Linguistics* 2002.2:63-74.
- Li, Lan (李藍). 2008. Hanyu de rencheng daici fushu biaoshifa 漢語的人稱代詞複數表示法 [The ways to express the plural forms of personal pronouns in Chinese]. *Fangyan* 方言 [Dialect] 2008.3:224-243.
- Li, Rulong (李如龍), and Song-Hing Chang (張雙慶). 1999. *Daici* 代詞 [Pronouns]. Guangzhou: Ji’nan University Press.
- Li, Xindi (黎新第). 2001. Ye shuo ‘men–mei–men’ de fanfu bianhua 也說“們—每一們”的反覆變化 [Another study on the bi-directional change between *men* and *mei*]. *Journal of Chongqing Normal University* 2001.1:89-99.
- Li, Y.-H. Audrey (李豔惠), and Yuzhi Shi (石毓智). 2000. Hanyu liangci xitong de jianli yu fushu biaoji ‘men’ de fazhan 漢語量詞系統的建立與複數標記“們”的發展 [The plural marker *men* in Chinese]. *Dangdai Yuyanxue* 當代語言學 [Contemporary Linguistics] 2000.1:27-36.

- Li, Yongsui (李永遂). 1983. Yi, Mian, Jingpo sange yuzhi diyi, er rencheng daici bijiao 彝、緬、景頗三個語支第一、二人稱代詞比較 [The contrast between first and second pronouns in Yi, Mian and Jingpo]. *Yuyan Yanjiu 語言研究 [Studies in Language and Linguistics]* 1983.1:179-191.
- Liao, Qiuzhong (廖秋忠). 1984. Yuyan Gongxing yu Leixing shuping 《語言共性與類型》述評 [A review of *Language Universals and Linguistic Typology: Syntax and Morphology*]. *Guowai Yuyanxue 國外語言學 [Contemporary Linguistics]* 1984.4:1-15.
- Liu, Danqing (劉丹青). 2003. *Yuyan Leixingxue yu Jieci Lilun 語言類型學與介詞理論 [Language Typology and Adposition Theory]*. Beijing: The Commercial Press.
- Liu, Xunning (劉勛寧). 1994. Shanbei Qingjianhua rencheng daici he zhiren mingci yuwei [-mi] tanyuan 陝北清澗話人稱代詞和指人名詞語尾 [-mi] 探源 [The origin of the personal pronoun and human noun suffix [-mi] in the Qingjian dialect of Shaanhsi]. *Chinese Languages and Linguistics*, Vol. 2: *Historical Linguistics*, ed. by Paul Jen-kuei Li, Chu-Ren Huang & Chih-Chen Jane Tang, 251-261. Taipei: Institute of History and Philology, Academia Sinica.
- Liu, Yingjiu (柳應九). 1993. Laoqida zhong de 'zhemen' 'namen' yu 'zheban' 'naban' 《老乞大》中的“這們”“那門”與“這般”“那般” [On the *zhemen*, *namen* and *zheban*, *naban* in *Laoqida*]. *Yuyan Yanjiu 語言研究 [Studies in Language and Linguistics]* 1993.2:140-143.
- Long, Guofu (龍國富). 2008. Cong Fan-Han duikan kan zaoqi fanyi dui yijing rencheng daici shu de yingxiang 從梵漢對勘看早期翻譯對譯經人稱代詞數的影響 [Effects of the early translation of Buddhist Scriptures on the numbers of the personal pronoun in Chinese Buddhist Scriptures]. *Waiyu Jiaoxue yu Yanjiu 外語教學與研究 [Foreign Language Teaching and Research]* 2008.3:218-223.
- Lu, Xiaoqun (盧小群). 2004. *Xiangnan Tuhua Daici Yanjiu 湘南土話代詞研究 [A Study of the Pronouns in the Sub-dialects in Southern Hunan Province]*. Beijing: China Social Sciences Press.
- Lü, Shuxiang (呂叔湘). 1985. *Jindai Hanyu Zhidaici 近代漢語指代詞 [Demonstratives in Early Chinese]*, supplemented by Lansheng Jiang. Shanghai: Xuelin Press.
- Lü, Shuxiang (呂叔湘). 1995. *Lü Shuxiang Wenji 呂叔湘文集 [The Collected Works of Lü Shuxiang]*, Vol. 2. Beijing: The Commercial Press.
- Lü, Shuxiang (呂叔湘). 1999. *Hanyu Yufa Lunwenji 漢語語法論文集 [Collection on Chinese Grammar]*. Beijing: The Commercial Press.
- Luo, Futeng (羅福騰). 1992. *Muping Fangyan Zhi 牟平方言志 [A Survey of the Muping Dialect]*. Beijing: Language and Culture Press.
- Luo, Pengnian (雒鵬年). 1997. Gansu fangyan jilei shici zhong cunzai de yixie yufa xianxiang 甘肅方言幾類實詞中存在的一些語法現象 [Some grammatical

- phenomena in several types of content words in the dialects in Gansu Province]. *Journal of Northwest Normal University* 1997.1:57-60.
- Mei, Tsu-Lin (梅祖麟). 1986. Guanyu jindai Hanyu zhidaici: du Lü zhu Jindai Hanyu Zhidaici 關於近代漢語指代詞：讀呂著《近代漢語指代詞》[A review to *Demonstratives in Early Chinese* written by Lü Shuxiang]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1986.6:401-412.
- Mei, Tsu-Lin (梅祖麟). 2004. Suzhouhua de ‘wudu’ (nimen) yu Handai de ‘ruoshu’ 蘇州話的“唔篤”（你們）與漢代的“若屬”[An etymology for [n to?] ‘you (plural)’ in the Suzhou dialect]. *Fangyan* 方言 [Dialect] 2004.3:247-255.
- Norman, Jerry (羅杰瑞). 1999. *Chinese*. Cambridge & New York: Cambridge University Press.
- Ota, Tatsuo (太田辰夫). 1991. *Hanyushi Tongkao* 漢語史通考 [Comprehensive Studies of the History of the Chinese Language], translated by Lansheng Jiang & Weiguo Bai. Chongqing: Chongqing Press.
- Ota, Tatsuo (太田辰夫). 2003. *Zhongguoyu Lishi Wenfa* 中國語歷史文法 [A Historical Grammar of Modern Chinese] (2nd edition), translated by Shaoyu Jiang & Changhua Xu. Beijing: Peking University Press.
- Peking University, Department of Chinese Language and Literature (北京大學中文系). 1995. *Hanyu Fangyan Cihui* 漢語方言詞匯 [Vocabularies of Chinese Dialects] (2nd edition). Beijing: Language and Culture Press.
- Qiu, Bin (邱斌). 2009. ‘Jiangxi Anfu fangyan biao fushu de “wu” jiyin xianyi 江西安福方言表複數的“物”〉記音獻疑 [Some questions about paper ‘The plural form *wu* in the Anfu dialect spoken in Jiangxi Province’]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 2009.1:89.
- Shen, Jiaxuan (沈家煊). 1999. *Buduicheng he Biaojilun* 不對稱和標記論 [Asymmetry and Markedness]. Nanchang: Jiangxi Education Press.
- Shen, Jiaxuan (沈家煊). 2000. *Yuyan Leixingxue yu Pubian Yufa Tezheng daodu* 《語言類型學與普遍語法特徵》導讀 [Introduction to W. Croft’s *Typology and Universals*]. Beijing: Foreign Language Teaching and Research Press.
- Sun, Xixin (孫錫信). 1990. Yuandai zhiwu mingci hou jia ‘men (mei)’ de youlai 元代指物名詞後加“們（每）”的由來 [The origin of which nouns referring to objects are suffixed by *men (mei)* in Yuan Dynasty]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1990.4:302-303.
- Wang, Li (王力). 1980. *Hanyu Shi Gao* 漢語史稿 [An Outline of Chinese Language History]. Beijing: Zhonghua Book Company.
- Wei, Pei-chuan (魏培泉). 1990. *Han Wei Liuchao Chengdaici Yanjiu* 漢魏六朝稱代詞研究 [A Study to Pronouns in Han, Wei, and Six Dynasties]. Taipei: National Taiwan University dissertation.

- Wu, Yunji (伍雲姬). 2000. *Hunan Fangyan de Daici* 湖南方言的代詞 [*Pronouns in Hunan Dialects*]. Changsha: Hunan Normal University Press.
- Xing, Xiangdong (邢向東). 2006. *Shanbei Jinyu Yufa Bijiao Yanjiu* 陝北晉語語法比較研究 [*The Contrastive Study to the Jin Dialects in Northern Shanxi*]. Beijing: The Commercial Press.
- Yang, Shin-Feng (楊時逢). 1974. *Hunan Fangyan Diaocha Baogao* 湖南方言調查報告 [*Report on a Survey of the Dialects of Hunan*]. Taipei: Institute of History and Philology, Academia Sinica.
- You, Rujie (游汝杰). 1999. Wuyu li de rencheng daici 吳語裡的人稱代詞 [Personal pronouns in Wu dialects]. *You Rujie Zixuanji* 游汝杰自選集 [*You Rujie's Anthology*]. Guilin: Guangxi Normal University Press.
- Yu, Min (俞敏). 1989. Gu Hanyu de rencheng daici 古漢語的人稱代詞 [Personal pronouns in Ancient Chinese]. *Yu Min Yuyanxue Lunwenji* 俞敏語言學論文集 [*Yu Min's Collection of Linguistic Papers*]. Harbin: Heilongjiang People's Publishing House.
- Zhan, Bohui (詹伯慧), and Yat-shing Cheung (張日昇). 1988. *Zhujiang Sanjiaozhou fangyan cihui duizhao* 珠江三角洲方言詞彙對照 [*Contrastive Vocabulary between Dialects in the Pearl River Delta Region*]. Hong Kong: New Century Press.
- Zhang, Chengcai (張成材), and Chao Mo (莫超). 2005. Gansu fangyan tezheng juyao 甘肅方言特徵舉要 [Some grammatical features in Gansu dialects]. *Gansu Gaoshi Xuebao* 甘肅高師學報 [*Journal of Gansu Normal Colleges*] 2005.3:30-35.
- Zhang, Huiying (張惠英). 1995. Fushu rencheng daici ciwei 'jia', 'men', 'li' 複數人稱代詞詞尾“家”、“們”、“徕” [The plural markers *jia*, *men*, and *li* suffixed to personal pronouns]. *Zhongguo Yuyan Xuebao* 中國語言學報 [*Chinese Linguistics: The Journal of Linguistic Society of China*] 5:28-34.
- Zhang, Huiying (張惠英). 1997. Hanyu fangyan daici yanjiu 漢語方言代詞研究 [Studies of pronouns of Chinese dialect]. *Fangyan* 方言 [*Dialect*] 1997.2:88-96.
- Zhang, Min (張敏). 1998. *Renzhi Yuyanxue he Hanyu Mingci Duanyu* 認知語言學與漢語名詞短語 [*Cognitive Linguistics and Nominal Phrases in Chinese*]. Beijing: China Social Sciences Press.
- Zhang, Shuzheng (張樹錚). 1995. *Shouguang Fangyan Zhi* 壽光方言志 [*A Survey of the Shouguang Dialect*]. Beijing: Language and Culture Press.
- Zhou, Shengya (周生亞). 1980. Lun Shanggu Hanyu rencheng daici fanfu de yuanyin 論上古漢語人稱代詞繁複的原因 [The explanation to the complicated system of personal pronouns in Archaic Chinese]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1980.2:127-136.

- Zhu, Qingzhi (朱慶之). 1993. Hanyi fodian yuwen zhong de yuandian yingxiang chutan 漢譯佛典語文中的原典影響初探 [The primary investigation to the influence of the original edition to the translated edition of Buddhist Scriptures]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1993.5:379-385.
- Zu, Shengli (祖生利). 2002. Yuandai baihua beiwen zhong ciwei 'mei' de teshu yongfa 元代白話碑文中詞尾“每”的特殊用法 [The special usage of the plural suffix *mei* in the vernacular inscriptions of Yuan Dynasty]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 2002.4:72-80.

A New View about the Origin of *men*

Lan Li

Chinese Academy of Social Sciences

There are four kinds of views regarding the origin of *men* which is a plural marker of personal pronouns. According to the data drawn on Chinese dialects and modern Chinese, the paper points out that Professor Lǚ Shuxiang' view about the change from *bei* to *men* may be the closest to the truth. We point out that the *bei* changed to *men* after three sound change stages. In the first sound change stage, the bilabial initial of *bei* changed to the nasal initial *mei*; in the second sound change stage, the falling tone *mei4* became a rising tone *mei3*; in the third sound change stage, the vowel end of *mei* became the nasal end of *men*. These three types of sound changes are infection of sound change. The paper also discusses other questions such as why *men* and *mei* appeared in the same data and how *men* and *mei* alternate in some data of different periods.

Key words: plural marker of personal pronoun, infection of sound change, origin of *men*

Shape Classifiers in Earlier Southern Min Texts^{*}

Chinfa Lien (連金發) and Liching Livy Chiu (邱力璟)

National Tsing Hua University

This paper tackles shape classifiers in earlier Southern Min texts. It focuses on three kinds of shape classifiers, which collocate with noun phrases denoting specific sets of objects bearing common features. We flesh out the distributional pattern of each type of classifier and furnish a perceptual structure of shape classifiers based on a feature analysis. We also try to account for how less-straightforward shape classifiers function in terms of their etymological origin.

Key words: shape classifier, Southern Min, perceptual structure, feature

1. Introduction

Examination of the linguistic nature of classifiers will shed light on human cognition as well as categorization of objects. Allan (1977) proposes that there are four types of classifier languages in the world. Mandarin Chinese (MC for short) is one of the typical numeral classifier languages. There are basically two approaches to research on classifiers. One is to tackle classifiers in terms of the semantic nature of their collocating nouns (Tai & Wang 1990,¹ Tai & Chao 1994, and Tai 1999). The other approach is to sort out

^{*} The paper is based on a research project partially supported by NSC 95-2411-H-013-MY3 and Studies on Constructions in Taiwanese Southern Min, National Tsing Hua University. Thanks are due to anonymous reviewers for insightful comments and to Edison Chang and May Wang for timely assistance. We are grateful to Caitlin Keenan for editorial assistance in the penultimate version of the paper.

¹ Note that the tones for all Mandarin examples are indicated by superscript numbers after each syllable. Three explicit linguistic criteria are proposed in Tai & Wang (1990) to distinguish between classifiers and measure words: (1) *De*⁰ 的 cannot appear between a classifier and the following noun phrase, as in **yi*⁴ *zhang*¹ *de*⁰ *zhuo*¹ *zi*⁰ 一張的桌子 ‘a desk’. But it is possible for *de*⁰ 的 to occur between a measure word and the following noun phrase, as in *shi*² *bang*⁴ *de*⁰ *rou*⁴ 十磅的肉 ‘ten pounds of meat’; (2) Classifiers can be replaced by *ge*⁰ 個 without changing their semantic meaning, as in *yi*⁴ *zhang*¹ / *ge*⁰ *zhuo*¹ *zi*⁰ 一張/個桌子 ‘a desk’. In the case of measure words, however, *ge*-replacement results in a different semantic interpretation, as in *shi*² *bang*⁴ / **ge*⁰ *rou*⁴. 十磅/*個肉 ‘a pound of meat’; (3) Classifiers vary from dialect to dialect, but measure words are rather consistent between dialects. Such criteria are not without dispute

the semantic interpretations and historical changes of classifiers. Lee (2005) proposes that metonymy and metaphor play a role in this respect and can be used to gain insight into the use of classifiers in the Hakka dialect. This study intends to explore shape classifiers as attested in earlier Southern Min texts. We adopt Allan's seven categories of classification (Allan 1977) to re-examine our data and further suggest that dimensionality is not the only salient aspect for classifiers of the shape category. Several minor features of the object are also crucial for the linguistic output of classifiers. The interaction of these parameters is, therefore, important, and worthy of careful consideration.

The first thing that must be done when researching classifiers is to distinguish real classifiers from measure words. The syntactic distributions of classifiers and measure words actually overlap in MC. However, the nature of these two types of elements can be teased apart by focusing on the semantic relationship between the element (classifier or measure word) and the following NP. Classifiers refer to persistent and salient characteristics of objects, while measure words refer to the temporary quantity of objects. The paper is organized as follows: we will first review existing literature on the theoretical models of shape classifiers and previous studies of Chinese shape classifiers. Drawing data from earlier Southern Min texts, we will then explore three kinds of shape classifiers, viz., one-dimensional, two-dimensional and three-dimensional classifiers. For each type of shape classifier we will furnish a range of objects that occur with them and give relevant explanations. We will then discuss the correlated distributions of the classifiers and provide a perceptual structure of the classifiers backed up by feature analysis. Finally, we will briefly explore the etymological origins of some less straightforward classifiers in an attempt to give an adequate explanation of shape classifiers in operation.

2. Literature review

2.1 Allan's categorization of classifiers

Allan (1977) identifies seven categories of classification: (1) "material", (2) "shape", (3) "consistency", (4) "size", (5) "location", (6) "arrangement", and (7) "quanta". We adopt his model and take each classificational category as a parameter within which features are specified. The category "shape" can be dimensional or non-dimensional. Non-dimensional shapes include curved exterior, hollow, and angular forms. The dimensional classifiers, on the other hand, are used for long, flat, and round objects; these are the basic target of our study.

—see Tang (2005) for more explorations. In identifying the classifiers in the texts in this article, we basically follow Tai & Wang (1990).

For Allan's seven types of classifiers, each can be further subdivided into more precise subcategories. "Material" is defined as "the essence of the entities referred to by nouns" and includes three subcategories: [animacy], [actions],² and [inanimacy]. The parameter "consistency" refers to the flexibility or rigidity of an object. [Flexible], [hard], [rigid], and [discrete] are the relevant features. The parameter "size" is straightforward, as it indicates the perceptual features [big] and [small]. "Location", on the other hand, is about the constraints of the environment in which the nominal expression is situated. The typical features of "arrangement" are [pleat], [fold], and [twisted-off], all of which actively interact with other parameters to indicate an entity. Finally, "quanta" concerns the (grammatical) number of the objects.

2.2 Tai's model

Tai & Wang (1990) propose that each classifier has a salient perceptual property which serves as the typical condition for categorization. For example, the one-dimensional extension is relevant for *tiao*² 條, and the three-dimensionality of a long, rigid object is salient for *gen*¹ 根. Later, Tai & Chao (1994) adopt the major parameters mentioned by Allan (1977) in conjunction with the model of prototype theory and divide the referents into three kinds of members: central members, natural extension members, and metaphorical extension members. The members of these semantic domains can be illustrated in Table 1, distilled from Tai & Wang (1990:47).

Table 1: Distribution of kinds of objects associated with *tiao*²

	Nominal Origin	Central Members	Natural Extension	Metaphorical Extension
條 <i>tiao</i> ²	條子 <i>tiao</i> ² <i>zi</i> ⁰	魚 <i>yu</i> ²	路 <i>lu</i> ⁴	新聞 <i>xin</i> ¹ <i>wen</i> ²
	柳條兒 <i>liu</i> ³ <i>tiao</i> ² <i>r</i>	黃瓜 <i>huang</i> ² <i>gua</i> ¹	街 <i>jie</i> ¹	意見 <i>yi</i> ⁴ <i>jian</i> ⁴
	麵條兒 <i>mian</i> ⁴ <i>tiao</i> ² <i>r</i>	褲子 <i>ku</i> ⁴ <i>zi</i> ⁰	河 <i>he</i> ²	消息 <i>xiao</i> ¹ <i>xi</i> ⁰
	木條兒 <i>mu</i> ⁴ <i>tiao</i> ² <i>r</i>	被單 <i>bei</i> ⁴ <i>dan</i> ¹	走廊 <i>zou</i> ³ <i>lang</i> ²	理由 <i>li</i> ³ <i>you</i> ²
		凳子 <i>deng</i> ⁴ <i>zi</i> ⁰		

The core members and extensions represent the domain structure in a synchronic state resulting from a historical development.³

² Although Allan (1977) takes [abstract and verbal nouns] as the second subcategory of material, he does not actually touch on abstract nouns. We therefore replace this category with the label [actions].

³ For the dating of various domains of *tiao*² in a diachronic perspective, see Erbaugh (1986) and Tai & Wang (1990).

Tai (1999) further concludes that the prototype theory of categorization has an explanatory value in the Chinese classifier system. This model is rather straightforward in that it captures the choice of classifiers in terms of a simple semantic mechanism.

2.3 Wu's analysis

Modeling his analysis after Allan (1977) and Pinker (1989), Wu (1998) specifies the referents of a classifier in terms of the dimensionality of an object and its saliency. An object can have at most three dimensions, one or two of which can be saliently focused. Dimensionality of an object has nothing to do with its size and shape, but rather with the ratio of the three dimensional axis. Based on this mechanism supplemented by some other features, Wu's research is claimed to be descriptively adequate.

2.4 Experimental analysis of Lien & Wang (1999)

Lien & Wang (1999) conducted a psycho-linguistic experiment on speakers of Taiwanese and Mandarin dialects. Their study uncovered the patterns of interaction of shape classifiers (two-dimensional and three-dimensional). In particular, it proved that the prototypical effect does apply. There is also a dialect-specific general tendency in the use of classifiers. Here are some points worth mentioning: (1) Prototypical effects on different classifiers are observed; the prototypical effect of *li⁴* 粒, for example, is greater than that of *ke¹* 顆, which means that *ke¹* 顆 is a more general classifier; (2) The experiment helps us tease out the central member associated with a classifier so that we can examine the core features that differentiate one classifier from another. For instance, [precious] is the main feature distinguishing *ke¹* 顆 from *li⁴* 粒; (3) Dialectal comparisons and related phenomena are accounted for. For example, Group A participants (Mandarin native speakers) used fewer items associated with shape than Group B participants (Min native speakers) did. In other words, Taiwanese speakers tend to use shape classifiers such as *ke¹* 顆 and *li⁴* 粒 instead of the neutral *ge⁴* 個.

3. Shape classifiers in earlier Southern Min texts

Allan (1977) prompted many of the studies on shape-oriented classification touched on above. Inspired by this line of enquiry, the present paper surveys a range of shape-related classifiers in the earlier Southern Min text “The Legend of the Litchi Mirror”.⁴ We also distinguish shape classifiers into three main categories in accordance

⁴ There are four texts of “The Legend of the Litchi Mirror” in our corpus. JJ, WL, SZ, and GX

with their salient dimensionality, and compare the early Southern Min data with that of other research.

3.1 One-dimensional shape classifiers

3.1.1 The referents of one-dimensional classifiers

We have assumed that the salient features for each of the one-dimensional classifiers —*tiau*⁵ 條, *ki*¹ 枝, *chiah*⁴ 隻, and *ber*² 尾—are all subsumed under the main [shape] category and the sub-feature [one dimensional].⁵ This is a hierarchical categorization. The structure of classifiers is amenable to further analysis from different perspectives. Given in (1) are the four one-dimensional shape classifiers *tiau*⁵ 條, *ki*¹ 枝, *chiah*⁴ 隻 and *ber*² 尾, gleaned from earlier Southern Min texts.⁶

- (1) a. *tiau*⁵ *chhiu*²-*phe*³ 手帕 ‘handkerchief’₃ *soa*ⁿ³ 線 ‘line’₁
 *peh*⁴-*au*⁷-*loo*⁵ 白後羅 ‘embroidery’₁ *be*²-*pi*ⁿ¹ 馬鞭 ‘horsewhip’₁
 *oe*⁵-*toe*²-*soa*ⁿ³ 鞋底線 ‘shoelace for the boot’₂
 b. *ki*¹ *chhiu*⁷-*ki*¹ 樹枝 ‘branch, twig’₁ *hoe*¹ 花 ‘flower’₅
 *kim*¹-*thoe*¹ 金釵 ‘gilt hairpin’₃ *chher*⁵ 簍 ‘rod, poker’₂⁷
 *khi*²-*thok*⁴ 齒托 ‘tooth-pick’₁ *chha*⁵ 柴 ‘fire wood’₂

are the respective abbreviations for the *Jia*¹-*Jing*⁴ 嘉靖 (AD 1522-1566) edition, the *Wan*⁴-*Li*⁴ 萬曆 (AD 1573-1619) edition, the *Shun*⁴-*Zhi*⁴ 順治 (AD 1644-1661) edition, and the *Guang*¹-*Xu*⁴ 光緒 (AD 1875-1908) edition. See Wu (2001a, b, c, d) for the annotated editions of the four play scripts. The language represented in the texts reflects mainly colloquial forms of Southern Min developed from a hybridization of the Quanzhou 泉州 and Chaozhou 潮州 varieties, as well as some earlier Mandarin and classical elements. The numbers in the parentheses stand for the date of each edition. The modified Church Romanization system is adopted with numerical superscripts for the tonal categories in Southern Min.

⁵ The spelling of Southern Min in this paper is based on the Church Romanization given in Douglas (1873). Some modifications have been made. In particular, the diacritic tone marks have been replaced by numerical superscripts. 1, 2, 3, 4, 5, 7 and 8 stand for the tone categories yin ping, yin shang, yin qu, yin ru, yang pang, yang qu and yang ru respectively. No distinction is made between *ch* and *ts* or *chh* and *tsh* as they do not involve phonemic contrast. Open /ɔ/ and closed /o/ are rewritten as *oo* and *o*, as in *too*⁵ 圖 ‘drawing’ and *to*⁵ 逃 ‘escape’. /er/ and /ir/ stand for /ɤ/ and /u/ respectively.

⁶ The subscripts on the right of the Chinese expressions indicate the number of tokens for each item.

⁷ There are two occurrences of *chher*⁵ 簍. The one which occurs as a free form means a long thin rod made of wood or bamboo. The other occurs in the compound *her*²-*chher*⁵ 火簍 fire-rod ‘poker’ (Douglas 1873:67).

- | | | |
|-----------------------------|--|--|
| | <i>tek⁴-poe¹</i> 竹杯 ‘bamboo rod’ ₁ ⁸ | <i>chiak⁴</i> 燭 ‘candle’ ₃ |
| | <i>kim¹-chiam¹</i> 金針 ‘hairpin’ ₂ | <i>sau³-chhiu²</i> 掃帚 ‘broom’ ₁ |
| c. <i>chiah⁴</i> | <i>se²-chun⁵</i> 小缸 ‘boat’ ₂ | <i>be²</i> 馬 ‘horse’ ₃ |
| | <i>lang²-a²</i> 籠子 ‘cage’ ₁ | <i>kau²</i> 狗 ‘dog’ ₃ |
| | <i>kim¹-chiam¹</i> 金針 ‘golden hairpin’ ₁ | <i>tir¹-ko¹</i> 豬 ‘sex maniac’ ₁ ⁹ |
| d. <i>be²</i> | <i>bo⁵-bat⁸-le⁵</i> 無目鸚 ‘blind fish’ ₁ | <i>choa⁵</i> 蛇 ‘snake’ ₂ |

As for *tiao²* 條 and *zhi¹* 枝 in Mandarin (Tai 1994), the feature of rigidity can be evoked to distinguish *tiau⁵* 條 and *ki¹* 枝 in earlier Southern Min texts, as shown in (1a) and (1b). Thus the secondary parameter for *tiau⁵* 條 is [–rigid], whereas that for *ki¹* 枝 is [+rigid]. However, *ki¹* 枝 could be further narrowed down in its [material] specification. The data show that all the referents are made of wood or related to wooden products or metal. If we specify something as [+wood] or [+metal], it naturally implies the feature of [+rigid].

In the analysis of Tai & Wang (1990), animals with long shapes take *zhi¹* 隻 as their classifier in Mandarin, and this is taken as the default marker for the category of animacy.¹⁰ However, this feature is not at work in the objects that involve the classifier *chiah⁴* 隻 in the Southern Min texts, as in (1c). Rather, what is involved is the feature of moveability [+move], specifically the potential of changing location.¹¹

As shown in Tai (1999), the classifier *ber²* 尾 collocates with nouns like *hir⁵* 魚 ‘fish’ and *choa⁵* 蛇 ‘snake’ in modern Southern Min. Scant data gleaned from earlier Southern Min texts also show that the classifier *ber²* 尾, as in (1d), applies to reptiles and fish with the salient feature ‘tails’, and rarely (in only one instance) applies to birds with the same salient feature.¹² We therefore sort [+coda], in other words ‘with a tail as a part of

⁸ *Tek⁴-poe¹* 竹杯 is glossed as ‘a flat bamboo rod for beating criminals or scholars’ in Douglas (1873:382).

⁹ This usually means ‘boar’, as in *khan¹ tir¹-ko¹* 牽豬哥 lead-pig male ‘to lead a boar to a sow’ (Douglas 1873:228). But in the example adduced in the text (3.019 GX, see Wu 2001d), it takes on the metaphoric sense as indicated in the list.

¹⁰ As an anonymous reviewer perceptively remarked, ‘In Mandarin, animals with long shape take *tiao²* 條 (fish, snake, crocodile); *zhi¹* 隻 is also used for other objects (the characteristic feature usually being specified as “one of a pair”). But the latter is used for any kind of quadruped or bird, even roundish ones, in its more general use’.

¹¹ Moveability includes auto-motion and passive moveability. This feature is only a necessary rather than sufficient condition for the use of *chiah⁴* 隻. For example, *chhia¹* 車 ‘car’ can collocate with both *chiah⁴* 隻 and *tai⁵* 台. The same object may be amenable to more than one means of interpretation due to different construals in our cognitive process.

¹² *Nng⁷ be² bo⁵-bat⁸-le⁵* 兩尾無目鸚 two CL not-have eye oriole ‘two eyeless orioles’ is found in GX (1875-1908). But the character 鸚 is rendered as ‘魚遷’ in SZ (1644-1661), an even earlier text. Since 鸚 appears later than ‘魚遷’, with a time gap of about 200 years, it is

the body’, into the arrangement category and place Allan’s (1977) [+animate] feature into the material category.¹³

3.1.2 The distribution of one-dimensional shape classifiers

There are four typical one-dimensional classifiers in earlier Southern Min texts, namely *tiau*⁵ 條, *chiah*⁴ 隻, *ber*² 尾, and *ki*¹ 枝, as shown in Table 2.¹⁴

Table 2: The distribution of one-dimensional shape classifiers

	<i>tiau</i> ⁵ 條	<i>chiah</i> ⁴ 隻	<i>ber</i> ² 尾	<i>ki</i> ¹ 枝	Total	Word count	Percentage
JJ	1	3	0	3	7	7331	0.00095
WL	0	2	0	3	5	4361	0.00115
SZ	4	3	2	9	18	8208	0.00219
GX	3	4	1	7	15	9147	0.00164
Total	8	12	3	22	45		

Some studies, such as Tai & Wang (1990:37), suggest that *tiao*² 條 is one of the most frequently-used classifiers. The concept of ‘extension in length’ underlying *tiao*² 條 is also very common among other classifier languages. Erbaugh (1986) shows that *tiao*² 條 is the most frequently used and extended classifier in child Mandarin. However, in earlier Southern Min texts, *tiau*⁵ 條 is no more frequent than other classifiers. In fact, among the four classifiers, *ki*¹ 枝 is the most frequently used in the earlier Southern Min texts. *Ber*² 尾, on the other hand, appears the least frequently in the texts. It is difficult to determine whether there is a chronological increase in the use of one-dimensional shape classifiers based on the ratio of the sum total of tokens of classifiers to the total number of words in each text because there is a tangible gap between JJ and WL.

3.1.3 The perceptual structure of one-dimensional classifiers

In the previous section, each of the one-dimensional classifiers was further defined

reasonable to surmise that barring a slip of the scriptwriter, what is involved is a categorical extension of fish and reptiles to birds for the classifier *ber*² 尾. Such a semantic extension is motivated by the salient feature of the coda that they share. Other kinds of birds with the salient coda feature, such as *oan*¹-*iu*ⁿ¹ 鴛鴦 ‘mandarin duck’, *too*⁷-*koan*¹ 杜鵑 / *chu*²-*kui*¹ 子規 ‘cuckoo’, and *i*ⁿ²-*a*² 燕仔 ‘swallow’, take on the general classifier *e*⁵ 个 rather than *be*² 尾.

¹³ Allan (1977) sub-divides the material category into animacy, inanimacy, and abstract and verbal nouns.

¹⁴ The numbers indicate the token count of each classifier.

in terms of one or more prominent perceptual features. We propose that those features must be constrained within Allan's (1977) seven parameters. Only when the possible categories for the perceptual features are not random can our analysis be said to reach a generalization. Example (2) summarizes the analysis of one-dimensional classifiers given above.

- (2) a. *tiau*⁵ [-rigid] (interacts with 'consistency')
 b. *ki*¹ [+rigid] [+wood] [+metal] (interacts with 'consistency')
 c. *chiah*⁴ [+move] (interacts with 'location')
 d. *ber*² [+coda] [+animate] (the part of an animal in relation to 'arrangement')

The features adopted here are all related to Allan's parameters for classifier languages. The whole organization of the mechanism is represented in Figure 1. Each classifier is on a tier of the 'shape' category, and is related to other perceptually salient features via linking.

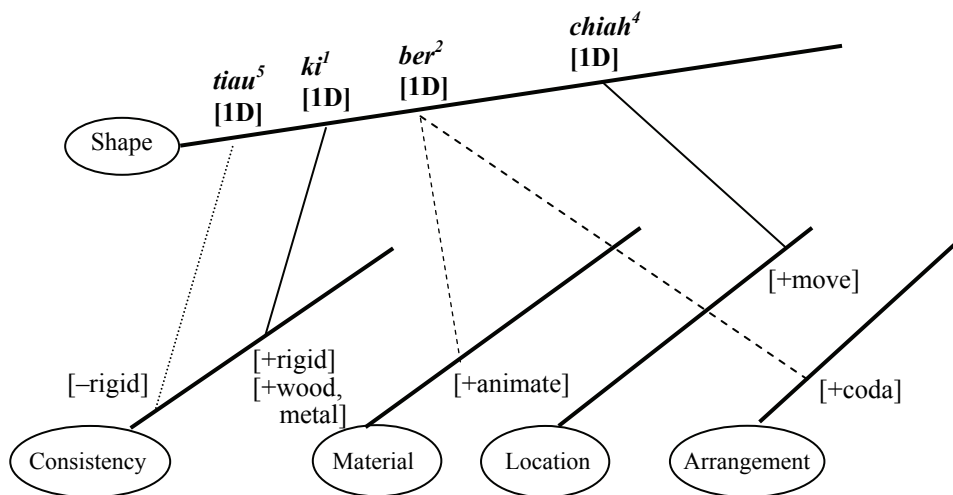


Figure 1: The perceptual structure of one-dimensional classifiers

One noteworthy point is that an object may be associated with more than one classifier. Tai (1994) acknowledges this point for MC, pointing out the overlap between *tiao*² 條 and *gen*¹ 根, on the one hand, and between *gen*¹ 根 and *zhi*¹ 枝, on the other. Similar phenomena are found in earlier Southern Min texts, as exemplified by the coexistence of *chit*⁸ *ki*¹ *kim*¹-*chiam*¹ 一枝金針 one CL golden-needle 'a hairpin' and *chit*⁸ *chiah*⁴ *kim*¹-*chiam*¹ 一隻金針 one CL golden-needle 'a hairpin' in the same text, particularly in the GX edition. Each classifier teases out a different perspective of the same object. Each of the expressions must be used in a different setting. For example,

*chit*⁸ *ki*¹ *kim*¹-*chiam*¹ 一枝金針 is used in the gift-giving context,¹⁵ as in (3), whereas *chit*⁸ *chiah*⁴ *kim*¹-*chiam*¹ 一隻金針 occurs in the setting of conveyance of an object reflecting the semantic property of *chiah*⁴ 隻, that is to say moveability, as in (4).¹⁶

- (3) *Chit*⁸ *ki*¹ *kim*¹-*chiam*¹ *sia*⁷ *lu*²
 一 枝 金針 謝 汝 (7.071, GX)
 one CL golden-needle thank you
 ‘Here is a golden hairpin for you as a token of thanks.’
- (4) *Kang*⁷ *a*¹-*ma*² *teh*⁸ *chit*⁸ *chiah*⁴ *kim*¹-*chiam*¹ *lai*⁵
 共 亞媽 提 一 隻 金針 來 (7.069, GX)
 for grandma fetch one CL golden-needle come
 ‘Fetch me a golden hairpin.’

3.2 Two-dimensional shape classifiers

3.2.1 The referents of two-dimensional classifiers

On a par with our account of the one-dimensional classifiers, we display the collocating objects for two-dimensional classifiers *tiu*ⁿ¹ 張, *hong*¹ 封, *pak*⁴ 幅, *pun*² 本, *pan*² 板, *bin*⁷ 面, and *phi*ⁿ³ 片 in (5).

- (5) a. *tiu*ⁿ¹ *chong*⁷ 狀 ‘written petition’₁ *choa*²-*ji*⁷ 紙字 ‘written paper,’
*pai*⁵-*phio*³ 牌票 ‘warrant’₁ letter’₃
 b. *hong*¹ *sir*¹ 書 ‘letter’₃₇ *sir*¹-*sin*³ 書信 ‘letter’₆
*thian*¹-*chir*²-*chiau*³ 天子詔 ‘imperial decree’₁
 c. *pak*⁴ *oe*⁷ 畫 ‘painting’₁
 d. *pun*² *nui*⁷-*chi*¹ *ki*³ 荔枝記 ‘The Legend of the Litchi Mirror’₁
 e. *pan*² *phah*⁴ ...*toa*⁷ *pan*² 打...大板 ‘give ... a flogging’₉
 f. *bin*⁷ *ki*ⁿ³ ...*chit*⁸-*bin*⁷ 見...一面 ‘meet’₉
 g. *phi*ⁿ³ *hoo*⁵-*soat*⁴ 胡說 ‘talk nonsense’₁
*giok*⁸-*kiat*⁸-*peng*¹-*chhing*¹ 玉潔冰清 ‘purity; virginity’₁

*Tiu*ⁿ¹ 張, *hong*¹ 封, *pak*⁴ 幅 and *pun*² 本 are basically classifiers for objects featuring a salient two-dimensionality. As listed in (5a), a piece of paper, a two-dimensional object, takes on a special function when co-occurring with the classifier *tiu*ⁿ¹ 張. A written

¹⁵ Gift-giving could, but does not necessarily, involve movement of the object from one person to another. It may only imply the shift of ownership without movement in space.

¹⁶ The numbers in parentheses in the following two examples stand for the scene and the input item respectively.

complaint is needed in a lawsuit; a written note is used to inform; a warrant is good for arrest. Despite its manifold functions in collocation with a range of noun phrases, the common feature in the object associated with the classifier *tiuⁿ¹* 張 is [spread], which is subsumable under the category of consistency. By [spread] is meant that something can be spread out. *Hong¹* 封, as shown in (5b), emphasizes the sealing and delivery of the letter. This element may be implicated in the origin of the lexical items *tiuⁿ¹* 張 and *hong¹* 封 in Chinese. The two words are verbs and mean ‘to stretch a bow’ and ‘to seal a letter’, respectively. Lee (2005) interprets this as a process of metaphorical extension. Here the feature [+move] is proposed on a par with the mechanism of one-dimensional objects in order to highlight the process of delivery. The classifiers *pak⁴* 幅, as in (5c), and *pun²* 本, as in (5d), only favor paintings and books respectively.

Unlike the nominal classifiers just touched on, *pan²* 板, *bin⁷* 面 and *phiⁿ³* 片, as given in (5e-g), are verbal classifiers. They function as classifiers for actions or events. They therefore take on the feature [+event]. Although they are both two-dimensional objects in their original meanings, they are related to the counting of events. *Pan²* 板, meaning ‘wooden board’, is an instrument used in the action of flogging, and when collocated with a numeral denotes a single or plural event. *Bin⁷* 面, with an original meaning of ‘human face’, assumes the function of a counter for meeting events. It is often but not always occurs in constructions with the verb *kiⁿ³* 見 ‘meet’. *Phiⁿ³* 片 may be used as a counter of gibberish utterances or more abstract entity such as ‘purity or virginity’.

3.2.2 The distribution of two-dimensional shape classifiers

Given in Table 3 is the distribution and frequency of the two-dimensional classifiers *tiuⁿ¹* 張, *hong¹* 封, *pak⁴* 幅, *pun²* 本, *pan²* 板, *bin⁷* 面, and *phiⁿ³* 片.

Table 3: The distribution of two-dimensional shape classifiers

	<i>tiuⁿ¹</i> 張	<i>hong¹</i> 封	<i>pak⁴</i> 幅	<i>pun²</i> 本	<i>pan²</i> 板	<i>bin⁷</i> 面	<i>phiⁿ³</i> 片
JJ	1	15	0	0	1	3	1
WL	0	10	1	1	2	2	1
SZ	2	9	0	0	4	3	0
GX	3	10	0	0	2	1	1

3.2.3 The perceptual structure of two-dimensional classifiers

Given in (6) are the feature specifications of the two-dimensional classifiers:

- (6) a. *tiu*ⁿ¹ [+spread] [-sealed]
 b. *hong*¹ [+move] [+sealed]
 c. *pak*⁴ [+material] [+painting]
 d. *pun*² [+material] [+print]
 e. *pan*² [±single]
 f. *bin*⁷ [±single]
 g. *phi*ⁿ³ [+single]

For a proper treatment of two-dimensional classifiers, in the last section we proposed to include the feature [+spread] in the category “consistency” to account for the spreadability of the object denoted by the noun phrase in collocation with the classifier *tiu*ⁿ¹ 張. We also postulated the feature [+sealed] in the category “arrangement” to distinguish *hong*¹ 封 from *tiu*ⁿ¹ 張 in earlier Southern Min texts. In particular, *hong*¹ 封 is used to categorize sealed letters, whereas *tiu*ⁿ¹ 張 is the classifier for papers not sealed in envelopes. *Pak*⁴ 幅 is used specifically for paintings, so we locate its selection restriction in the category “material”. Compared to other classifiers for different kinds of paper, *pun*² 本 denotes a volume comprising written pages or prints. We specify this classifier in terms of the feature [+print] subsumed under the category “material”.

*Pan*² 板, *bin*⁷ 面 and *phi*ⁿ³ 片 function as verbal classifiers denoting actions and events. Thus, according to Allan’s analysis (1977), they are specified in terms of the feature [+action], subsumed under the parameter of material. Figure 2 encapsulates the perceptual structure of two-dimensional classifiers.

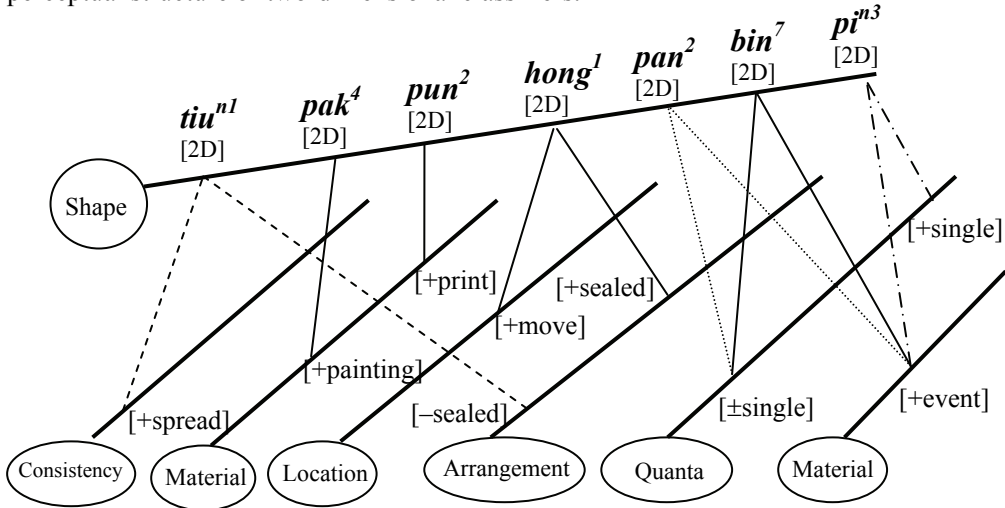


Figure 2: The perceptual structure of two-dimensional classifiers

3.3 Three-dimensional shape classifiers

3.3.1 The objects collocating with three-dimensional classifiers

Three-dimensional classifiers to be examined, such as *ter*³ 塊, *liap*⁸ 粒, *thng*⁵/*oan*⁵ 團, *lun*⁵ 輪 and *lui*² 蕊, are listed in (7):

- (7) a. *ter*³ *kiam*⁵-*chhai*³ 鹹菜 ‘salted vegetable’₁ *gun*⁵ 銀 ‘silver ingot’₂
 *chioh*⁸-*thau*⁵ 石頭 ‘stone’₂ *kia*ⁿ³ 鏡 ‘mirror’₃
 *tiam*²-*sim*¹ 點心 ‘snack’₁ *ker*² 粿 ‘rice-cake’₂
 *choa*² 紙 ‘paper’₂ *pang*¹ 版 ‘board’₁
 (*bin*⁵-*chhng*⁵) *pang*¹ (眠床) 枋 ‘(bed) board’₃
- b. *liap*⁸ *bi*² 米 ‘rice’₂
- c. *thng*⁵/*oan*⁵ *tia*ⁿ²-*phi*² 鼎丕 ‘rice crust’₁¹⁷ *leng*²-*png*⁷ 冷飯 ‘cold rice’₁
- d. *lun*⁵ *kng*¹-*gerh*⁸ 光月 ‘full moon’₇ *ang*⁵-*jit*⁴ 紅日 ‘sun’₂
 *gerh*⁸-*chiau*³ 月照 ‘moonbeam’₁
- e. *lui*² *hoe*¹ 花 ‘flower’₄

*Ter*³ 塊 in (7a), denoting a lumpish object, functions as a classifier for chunks of stone, bits of ink, etc. (Douglas 1873:482). It is a productive classifier for three-dimensional objects, as exemplified by *kiam*⁵-*chhai*³ 鹹菜 ‘salted vegetables’, *kia*ⁿ³ 鏡 ‘mirror’, *chioh*⁸-*thau*⁵ 石頭 ‘stone’, *ker*² 粿 ‘rice cake’, *gun*⁵ 銀 ‘silver (ingot)’, *pan*² 版 ‘pattern’, (*bin*⁵-*chhng*⁵) *pang*¹ (眠床) 枋 ‘(bed) plank’, *tiam*²-*sim*¹ 點心 ‘snack’, and *choa*² 紙 ‘lump of paper’.

*Liap*⁸ 粒 in (7b) denotes a tiny and globular mass of an object such as a grain of rice. The feature [+small] can be posited in the ‘size’ parameter to capture the meaning of this classifier.¹⁸ Whereas *liap*⁸ 粒 must denote an individual object, the classifier *thng*⁵/*oan*⁵ 團 in (7c), which also denotes a globular mass, can be used to refer to an aggregate of objects such as a cluster of rice crusts. It is therefore bigger in size and takes on the features of [–small] [+collective] in our specification. The situation is much simpler for *lun*⁵ 輪 in (7d) and *lui*² 蕊 (7e): they are simply used to refer to disc-like heavenly bodies — such as the sun and the moon — and flowers, respectively.

¹⁷ *Tia*ⁿ²-*phi*² 鼎丕 is glossed as ‘incrustation on a rice-boiler from boiling rice or other things’ in Douglas (1873:394). It corresponds to *kuo*¹-*pa*¹ 鍋巴 in Mandarin. See also Wu (2001d:345).

¹⁸ The feature [+small] does not figure in the classifier *liap*⁸ 粒 in modern Southern Min, as it is acceptable to use this classifier in reference to a watermelon (Lien & Wang 1999).

3.3.2 The distribution of three-dimensional shape classifiers

Except for *ter*³ 塊, three-dimensional classifiers are hard to come by. The distribution and frequency of three-dimensional classifiers are given in Table 4.

Table 4: The distribution of three-dimensional shape classifiers

	<i>ter</i> ³ 塊	<i>liap</i> ⁸ 粒	<i>oan</i> ⁵ 團	<i>lun</i> ⁵ 輪	<i>lui</i> ² 蕊
JJ	6	1	0	2	0
WL	6	0	1	2	0
SZ	14	1	1	3	2
GX	3	1	0	3	2

As expected, the three-dimensional classifier *ter*³ 塊 is the most productive, and we will see in the next section that its referents are also the most diverse. But the distribution of *lun*⁵ 輪 shows that the content of the legend plays a role in the frequency of occurrence as well. The number of tokens of this classifier is high, but the variety of objects it applies to is small.

3.3.3 The perceptual structure of three-dimensional classifiers

The feature specifications of the three-dimensional classifiers are given in (8):

- (8) a. *ter*³ [+lumpish]
 b. *oan*⁵ [–small] [+collective]
 c. *liap*⁸ [+small] [–collective]
 d. *lun*⁵ [+heavenly body] (in interaction with the category ‘material’)
 e. *lui*² [+flower] (in interaction with the category ‘material’)

Sub-parameters can be furnished to better specify the three-dimensional classifiers: [smallness] and [collectiveness]. One interesting point is that, theoretically, the classifiers shown in (8) do not exhaust the possible realizations of the feature combinations for [smallness] and [collectiveness]. For instance, we can say that the case of [–small] and [–collective] is realized as *ter*³ 塊 but what about the opposite: a classifier with the features [+small] and [+collective]? It is logically possible for a small object to comprise a set of even smaller elements. However, the data is not sufficient for us to make any further claims. We leave this question open for further research. The perceptual structure of three-dimensional classifiers is given in Figure 3.

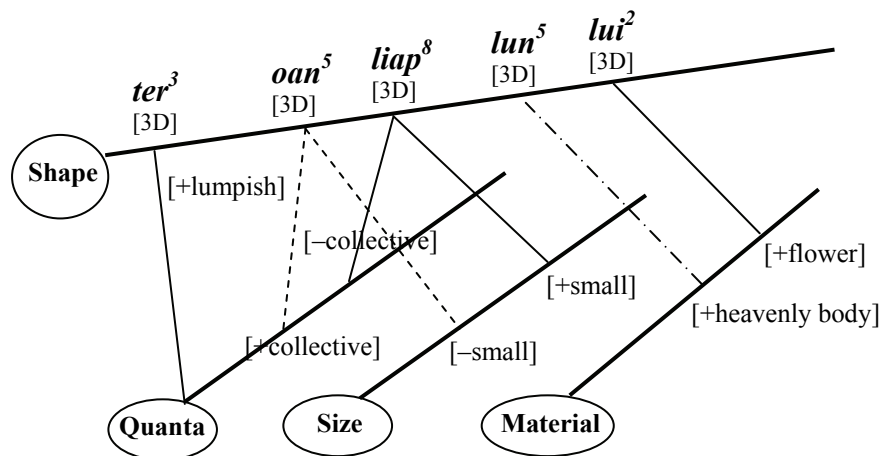


Figure 3: The perceptual structure of three-dimensional classifiers

4. The semantic interpretations of classifiers

In this section, we look into the etymological origins of the classifiers. As we noted in the first section, the meanings of the lexical items themselves are worth exploring. Some of the historical changes undergone by these lexical items provide reasonable explanations for the linguistic uses of the classifiers. Since our study focuses on earlier Southern Min texts, and since a basic account of many dimensional shape classifiers has been given in the previous sections, we will only focus on semantic interpretations of the classifiers that we have not done full justice to.

For one-dimensional classifiers we tackle the earlier and more basic senses of *tiau*⁵ 條, *ki*¹ 枝 and *chiah*⁴ 隻, but not of *ber*² 尾, which is pretty straightforward. While *tiau*⁵ 條 originally referred to the long and thin branch of a tree (Wang et al. 2000:483), it is no longer confined to the branch of a tree but can indicate anything long and slender (Douglas 1873:498). It has shed its arboreal property while keeping its configurational character. The default value of this classifier is the feature [-rigid], which denotes flexible objects; however, this feature does not always apply, as exemplified by the compatibility of *tiau*⁵ 條 with objects such as *kim*¹-*tiau*⁵ 金條 ‘a bar of gold’ and *han*¹-*chir*⁵ 蕃薯 ‘sweet potatoes’ (Douglas 1873:498). The original meaning of *ki*¹ 枝 was a tree branch or twig (Wang et al. 2000:464) but, as a classifier, it takes on the sense of something long, straight, and characterizable by the feature [+rigid] (Douglas 1873:210). Such a feature distinguishes *ki*¹ 枝 from *tiau*⁵ 條. *Ki*¹ 枝 is quite common in modern Southern Min especially in its extended use, as it can co-occur with *ien*⁵-*pit*⁴ 鉛筆 ‘pencil’, *ki*⁵-*koain*¹ 旗桿 ‘pole’, *ong*⁷-*lai*⁵ 旺梨 ‘pineapple’, etc. (Douglas 1873:210). On the basis of its graphic make-up, we can surmise that the earliest sense of *chiah*⁴ 隻

was the getting or holding of birds; in the modern language, however, its sense extends to single as opposed to double, as in *shuang*¹ 雙 (Ogawa et al. 1994:156). This lexeme is etymologically related to 獲, which indicates the setting of dogs to catch birds and beasts (passim 648). Such an etymological origin can account for the confinement of the *zhi*¹ 隻 to animals in Mandarin.¹⁹ Perhaps there is a semantic link between catching animals and identifying the animals to be caught. In Southern Min, this classifier has undergone further semantic extension by shedding the feature of animacy while retaining the feature of moveability.²⁰

For two-dimensional classifiers let's consider *tiu*ⁿ¹ 張 and *pun*² 本. *Tiu*ⁿ¹ 張 originally meant 'stretching a bow', but such a particular action has extended first to stretching anything and then to anything stretched. This then leads to two-dimensional entities. *Yi*⁴ *zhang*¹ *wang*³ 一張網, 'a net' in MC, could be taken as a telling example reflecting such a propensity. However, notwithstanding the historical link, if all noun phrases that the classifier *zhang*¹ 張 covers in modern MC are taken into consideration, we can see the prototype effect, since synchronically it is difficult to pin down the semantic common features that all the noun phrases involved share. *Pun*² 本 is commonly known to denote the root of a tree in its earlier sense.²¹ In the modern language, it can also denote the plant itself. This extension is probably a result of the operation of metonymy where part (root) stands for whole (tree), as in *toa*⁷ *pun*² 大本 'large plant' in modern Southern Min. Although the precise development which caused *pun*² 本 to become a classifier for books remains a mystery, we deem it quite plausible that the sense of book as a 'basis or foundation', could lay the grounds for developing a new function for this classifier.

5. Closing remarks

In this paper we have examined one-dimensional, two-dimensional and three-dimensional classifiers in earlier Southern Min texts, viz., "The Legend of Litchi Mirror", dating back to as early as the sixteenth century. For each kind of classifier, examples have been furnished and supported by relevant explanations, and their distribution and

¹⁹ As mentioned here, however, there has been another route of development for this classifier as well, with *zhi*¹ 隻 denoting one of a pair, such as *yi*¹ *zhi*¹ *yan*³ *jing*¹ 一隻眼睛 'one eye', *yi*¹ *zhi*¹ *wa*⁴ *zi*⁰ 一隻襪子 'a sock' and *yi*¹ *zhi*¹ *xiu*⁴ *zi*⁰ 一隻袖子 'a sleeve'. This is a development that Mandarin but not Southern Min has experienced.

²⁰ It would be immensely interesting to examine the discrepancy between Mandarin and Southern Min regarding the development of *zhi*¹ 隻 from a diachronic perspective in line with Erbaugh (1986) and Peyraube (1998).

²¹ See Liu (1965:96-97).

frequency of occurrence have been given. Finally, figures featuring the perceptual structures of the classifiers have been presented on the basis of detailed feature analysis. We have concluded the paper with an attempt to flesh out the etymological origins of some less straightforward classifiers and have hopefully provided a viable account for the behavior of the classifiers in question.

References

- Allan, Keith. 1977. Classifiers. *Language* 53.2:285-311.
- Douglas, Rev. Carstairs. 1873. *Chinese-English Dictionary of the Vernacular or Spoken Language of Amoy with the Principal Variations of the Chang-chew and Chin-chew Dialects*. London: Trübner and Co.
- Erbaugh, Mary S. 1986. Taking stock: the development of Chinese noun classifiers historically and in young children. *Noun Classes and Categorization*, ed. by Colette G. Craig, 399-436. Amsterdam & Philadelphia: John Benjamins.
- Lee, Wei-che. 2005. *Hai-Lu Hakka Classifiers: An Interactional View*. Hsinchu: National Tsing Hua University MA thesis.
- Lien, Chinfa, and Pen-ying Wang. 1999. Shape classifiers in Mandarin and Taiwanese —A psycholinguistic perspective. *The Biological Bases of Language*, ed. by Ovid J. L. Tzeng, 189-221. Berkeley: Project on Linguistic Analysis.
- Liu, Shiru. 1965. *Wei Jin Nanbei Chao Liangci Yanjiu* [Studies of Classifiers in the Wei, Jin and Nanbei Chao Period]. Beijing: Zhonghua Book Company.
- Ogawa, Tamaki, Nichida Taichiroo, and Akatsuka Kiyoshi. (eds.) 1994. *Kadokawa Shin Jigen* [Kadokawa New Dictionary of Etymology]. Tokyo: Kadokawa Shoten.
- Peyraube, Alain. 1998. On the history of classifiers in archaic and medieval Chinese. *Studia Linguistica Serica*, ed. by Benjamin K. Tsou, 131-145. Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong.
- Pinker, Steven. 1989. *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge: MIT Press.
- Tai, James H-Y. 1994. Chinese classifier systems and human categorization. *In Honor of William S-Y. Wang: Interdisciplinary Studies of Language and Language Change*, ed. by Matthew Chen & Ovid J. L. Tzeng, 479-494. Taipei: Pyramid.
- Tai, James H-Y. 1999. A note on the classifier *bue*⁵³ 尾 in Southern Min. *Contemporary Studies on the Min Dialects*, ed. by Pang-Hsin Ting, 225-228. Berkeley: Project on Linguistic Analysis.
- Tai, James H-Y., and Fang-yi Chao. 1994. A semantic study of the classifier *zhang*. *Journal of Chinese Language Teachers Association* 29.3:67-78.

- Tai, James H-Y., and Lianqing Wang. 1990. A semantic study of the classifier *tiao*. *Journal of Chinese Language Teachers Association* 25.1:35-56.
- Tang, Chih-Chen Jane. 2005. Nouns or classifiers: a non-movement analysis of classifiers in Chinese. *Language and Linguistics* 6.3:431-472.
- Wang, Li, Zuofan Tang, Xiliang Guo, Xianzhao Cao, Jiuying He, Shaoyu Jiang, and Shuangdi Zhang. (eds.) 2000. *Wangli Gu Hanyu Zidian [Wang Li's Old Chinese Dictionary]*. Beijing: Zhonghua Book Company.
- Wu, Shouli. 2001a. *Ming Jiajing Kan Li Jing Ji Xiwen Jiaoli [Annotated Texts of the Romance of Li Jing Ji of Ming Jiajing Edition (JJ)]*. Taipei: Ts'ung-I Workshop.
- Wu, Shouli. 2001b. *Ming Wanli Kan Li Zhi Ji Xiwen Jiaoli [Annotated Texts of the Romance of Li Zhi Ji of Ming Wanli Edition (WL)]*. Taipei: Ts'ung-I Workshop.
- Wu, Shouli. 2001c. *Qing Shunzhi Kan Li Zhi Ji Xiwen Jiaoli [Annotated Texts of the Romance of Li Zhi Ji of Qing Shunzhi Edition (SZ)]*. Taipei: Ts'ung-I Workshop.
- Wu, Shouli. 2001d. *Qing Guangxu Kan Li Zhi Ji Xiwen Jiaoli [Annotated Texts of the Romance of Li Zhi Ji of Qing Guangxu Edition (GX)]*. Taipei: Ts'ung-I Workshop.
- Wu, Suemei. 1999. A semantic study of the classifiers *tiau*⁵, *tiu*ⁿ¹ and *liap*⁸ in the Southern Min dialect. *Selected Papers from the Fifth International Conference on Chinese Linguistics*, ed. by H. Samuel Wang, Feng-fu Tsao & Chinfa Lien, 369-386. Taipei: Crane.
- Wu, Yaolin. 1998. *The Study of Shape-Based Classifiers in Chinese*. Hsinchu: National Tsing Hua University MA thesis.

粵語“先”、“添”虛實兩用的跨域投射解釋*

劉丹青

中國社會科學院

廣州/香港粵語中 VP 後的“先”和“添”有後置副詞狀語和語氣詞兩種用法。本文基於行、知、言三域跨域投射學說，統一闡述後狀語用法和語氣詞用法的密切關係。“先”和“添”的語氣詞用法，是其副詞用法（行域）投射到言域或知域的產物。副詞“先”表示所修飾的 VP 的行為先行發生，而語氣詞“先”則表示“我要說/問”這類隱性母句先行發生，即表示“我先要說/問……”。副詞“添”表示添加一次 VP 的行為，語氣詞“添”則表示添加一次“我知道”的隱性母句，即“我還知道……”。功能詞的跨域用法可能造成詞義的分化，並經規約化凝固成新的義項甚至新的詞性（句法屬性），分化後的不同用法甚至還能在同一句中疊用。本文對“添”在引申過程中的一些複雜情況也做了解釋。

關鍵詞：粵語，後置副詞，語氣詞，行知言三域，跨域投射

1. 引言

粵語廣州話（包括香港粵語）中用在動詞短語 VP 之後的“先”和“添”是很常用的兩個動詞後置成分，已有諸多學者做過研究。這兩個詞都有偏實（後置副詞）和偏虛（語氣詞）兩大類用法。現有研究成果有一個有趣現象。對它們的副詞用法，諸家釋義比較接近，都能跟普通話中相應的前置副詞狀語直接對應，因而通常被認定為後置的副詞狀語（參看詹伯慧 1958），即使不採用這一定性，如施其生 (1995)，但釋義仍很一致。後狀語“先”表示普通話“先”的意思，如“我走先”等於“我先走”。副詞“添”表示“再、又、還”一類意思，如“飲

* 本文寫作和修改中得到同事麥耘教授在語料和分析方面的重要幫助，深致謝意。初稿曾在中國社科院語言研究所高研報告會（2008 年 10 月）和第 13 屆國際粵方言研討會（2008 年 12 月，香港城市大學）報告，與會者多所是正。一併致謝，尤其感謝粵語專家詹伯慧教授及李寶倫、單韻鳴等年輕粵語研究者在會上會後的鼓勵和指正。文中尚存偏誤盡歸筆者。謹以此文向尊敬的貝羅貝教授賀壽！

杯酒添”（再喝一杯酒）。而對於它們的句末語氣詞用法，雖然各家定性比較一致，但釋義差別頗大，也難以在普通話中找到恰當的對應詞語。如“先”，鄧思穎 (2006b) 分出三個義項，並說不區分後置副詞和語氣助詞，但後文又明確說“先₂、先₃”是語氣助詞。“先₁”（我走先）是比較公認的後置狀語，“先₂”用“暫且先做”、“先……，（然後）”來釋義，如“準備好先”，“先₃”用語氣詞“到底”來釋義。而詹伯慧主編 (2002:95) 將相當於鄧思穎“先₂”、“先₃”的用例都放在“狀語後置”的“副詞”中，如“佢要點先？”（他到底要什麼？），語氣詞中有“啫”無“先”。“添”的釋義分歧更大。有的文獻釋為“甚至”，這意味著它是表量級極點的算子，有的釋為表示強調或誇張（詹伯慧 1958，李新魁等 1995:504），與“甚至”略微接近，也有的釋為“而且”，這意味著它起表遞進的關聯作用，有的認為其表示遺憾（寫作“啫”，見詹伯慧 2002:88，方小燕 2003:40），有學者認為是表示意外，“在預設之外”，用“竟然”來對譯（見鄧思穎 2006b 及所引文獻）。李寶倫和潘海華 (Lee & Pan 2006) 是探討“添”字最詳盡的文獻，該文提出用一個解讀來覆蓋“添”的所有虛實用法，釋其為表示量級增量 (scale increment) 的句末助詞，即“添”表示本句所述內容在該量級中居於比預設中的對比命題更高的位置，因而“添”字句排斥程度或數量降低的句子。

以上對語氣詞“先”，特別是對“添”的諸家釋義，雖然不無交匯之處，但彼此距離也很明顯，所對應的普通話詞語出入很大，如“竟然”和“而且”，“程度增加”和“表意外”，“誇張”和“遺憾”，都難以說是相同相近的意義。這些看法分歧至少說明，“先”、“添”的語氣詞語義不像其後狀語語義那麼顯豁、直接和單純，既難找到直接對應的普通話詞語，也缺乏現成的語法範疇可供參照。要用單一的現成概念來概括其用法殊為不易，而且現有釋義也沒能清楚解釋副詞用法和語氣詞用法的聯繫。本文擬另闢蹊徑，從跨域投射這一主觀化機制的角度，著重考察後狀語用法和語氣詞用法的相互關係，以此獲得對“先”、“添”的語義解釋的新看法，並據此審視兩種用法的句法屬性。

2. 跨域投射與語義分化及詞性分化

2.1 “先”、“添”從後狀語用法引申出語氣詞用法，是一種由跨域投射引起的主觀化。所跨之域，即與人類的言語行為有關的三種概念域：

- 1) 行域——以直接表達的方式完成言語行為，包括陳述、祈使、疑問等。
- 2) 知域——表達說話人的主觀認定。
- 3) 言域——表達說話人的語言本身。

本文借鑒的三域概念是沈家煊 (2003) 分析一些複句連詞時所採用的框架，但該文本身只就具體複句類型舉例分析，沒有對三域下適用面更廣的定義，以上定義是我們根據沈文內容所作的概括。下面以沈文所舉的“如果……那麼”為例來顯示三域的含義：

- (1) 如果明天下雨，比賽就取消。 [行域]
- (2) 如果比賽取消了，昨天就下雨來著。 [知域]
- (3) 如果比賽真的取消，太陽就從西邊出來了。 [言域]

“如果”一詞引出的充分條件在三個域中的性質可以這樣來表述：

- (1') 行域：p 的發生是 q 發生的充分條件。（如果 p，那麼 q）
- (2') 知域：知道 p 是我得出結論 q 的充分條件。（如果知道 p，那麼我推斷 q）
- (3') 言域：狀態 p 是我聲稱 q 的充分條件。（如果 p，那麼我聲稱 q）

對比 (1)-(3) 和 (1')-(3')，可以看出，行域 (1) 是最直接、最客觀的表達方式。而知域和行域都不是字面上的直接表達，而是以隱性形式存在一個以說話人為主語的母句（高層命題）“我知道/我認為/我推斷……”和“我聲稱”，表層出現的句子其實是這個母句的賓語從句。這兒的關鍵是“隱性形式”。假如“我知道”、“我推斷”、“我聲稱”之類形式出現在表層，句子就是直接表達，屬於行域。而在句子中以隱性形式表達的內容是以第一人稱為主語即“言者主語”的命題，這是典型的主觀性 (subjectivity)。此時“如果……就”的關聯對象不再指向這兩個分句的命題本身，而指向言者主語的母句。至於這種跨域使用所憑藉的機制，沈家煊 (2003) 指出，“知域”和“言域”的語義關係是從“行域”的語義關係通過隱喻“投射”形成的。跨域投射使句子有了主觀性，是一種典型的主觀化 (subjectivization) 操作。

回到粵語，我們發現，“先”和“添”的語氣助詞用法，正是其副詞用法投射到言域或知域的產物。具體地說，副詞“先”表示所修飾的 VP 先行發生，而

語氣詞“先”則表示“我要說/問”這一隱性母句先行發生；副詞“添”表示添加一次 VP 的行爲，語氣詞“添”則表示添加一次“我知道”的隱性母句。

“先”和“添”的修飾對象從主句所表達的事件轉向隱藏的言者主語的知情狀態。以“先”爲例：

(4) 阿良飲啲酒先。(阿良先喝點兒酒)

(5) A：你叫阿良飲多的酒。(你叫阿良多喝點兒酒)

B：阿良飲唔飲酒先？(阿良到底喝不喝酒？≈(我先要問)阿良喝不喝酒！)

(4)的“先”修飾“飲啲酒”(喝些酒)，是標準的時間狀語的用法。而從(5)所示大致翻譯看，語氣詞“先”語義上不修飾句子謂語所表示的喝酒這一事件，而是修飾括號中隱性的“我要問”，而“問”是一種言語行爲，所以應當屬於言域的表達。再看“添”：

(6) 阿良飲咗三杯酒添。(阿良又喝了三杯酒)

(7) 阿強飲咗兩杯酒，阿良飲咗三杯酒添。((我知道) 阿強喝了兩杯酒，我還知道，阿良喝了三杯呢)

(6)的“添”修飾“飲咗三碗酒”，表示“又(用於未然事件時則相當於“再”)”，意味著“阿良”之前已經喝過酒了，總共肯定不止三杯。而從(7)的翻譯可知，語氣詞“添”不再指向謂語動詞，即不表示阿良之前已經喝酒，而指向隱性的言者主語的知情狀態，表示說話人除了知道上句信息之外，還知道下一句的信息——“阿良喝了三杯酒”(總共三杯)，因而是知域的表達。

“先”的另一種用法，根據鄧思穎(2006b)，是表示“暫且……，(然後)”。我們注意到，這個“先”總是用在祈使句、意願句中，表達說話人的提議(往往是不同於預設行爲次序的提議)。假如它僅僅表示時間性的“暫且”，不應局限於祈使意願句(如陳述句：他現在暫且(先)寄住在親戚家)。因此，這種用法的祈使義意願義，實際上來源於“先”的另一種言域投射，涉及的隱性言語動詞是“提議”一類詞，如：

(8) 搞掂呢件事先。(暫且先解決這件事！/我先提議解決這件事)

2.2 功能詞的跨域使用可能造成詞義的分化，“先”、“添”因此形成有別於副詞義的語氣詞義項。

例(1)-(3)另有一點需要注意，同一對關聯詞“如果……就”，由於用於不同的概念域，實際表達的邏輯關係頗為不同。(1)句表達的是真正的假設條件關係，(2)句表達的不是字面上兩個分句間的假設條件關係，而是推斷關係，因為事件“比賽取消了”不是事件“昨天下雨來著”的條件，而是“我推斷昨天就下雨來著”的理由。同樣，(3)句表達的也不是字面上兩個分句間的假設條件關係，因為事件“比賽取消了”更不是事件“太陽就從西邊出來了”的假設條件，而是“我宣稱太陽就從西邊出來了”的條件，它表達的是歸謬推理關係（該句的確切推導解讀過程需要進一步的會話分析，本文不贅）。換言之，功能詞語的跨域使用，會導致詞語的實際功能和語義出現偏差和分化，但它的基本語義仍會得到保留，如上例中“如果……就”的假設條件作用。其他的語義分化可以視為這個基本語義的概念域變體，這些變體會受具體語境的影響而表現出更多的差異，如上面(3')中就有歸謬的意味。

“先”、“添”一類副詞，在語氣詞用法中仍保留其基本語義，同時因跨域投射而出現了一些概念域變體，並且因為攙雜了語境因素而形成更多的語義變異，這些變異既沒有現成的普通話詞語可以對譯，也難以概括為一個現成範疇。以往對它們的解釋，尚未充分認識其表面之異和背後之同，導致隨文釋義、解讀蜂起。當然，假如某項概念域變體因為高頻使用而規約化，則凝固下來成為新的義項，並可以導致有關虛詞以主觀化為主要表現的語法化。

在上引沈文所舉的連詞例子中，跨域投射主要是通過語用推理實現的，其語義變異只是其臨時變體，不能視為新的義項，也沒見辭典為這類變體新設義項。下面我們來看一些跟本文論題有關的由跨域投射引起義項分化的例子。普通話的“又”表示已然行為重複以往發生過的行為（未然行為則用“再”），屬於累加性修飾，與粵語“添”部分相同。這一基本義項是“又”的行域用法。此外，“又”還有知域用法。比較：

- (9) a. 行域：他又喝了一杯酒。（他之前已喝了至少一杯酒。總數超過一杯）
- b. 知域：醫生讓他戒煙戒酒，他出院當天就抽起了煙，第二天，他又喝了一杯酒。（……除了知道他抽煙之事，我又知道，第二天，他喝了酒。總數為一杯）

(9a) 句副詞“又”是客觀敘述一件事情的累加性發生，是直接表達，意義實在，重讀，屬於行域用法。(9b) 的“又”不表示同一事件（喝酒）本身的累加性發生，而是強調“他”背離醫囑一類行為的累加。“又”之所以有主觀強調作用，是因為存在一個隱性的言者主語，“又”由累加副詞投射到知域，表示說話人對此類事知情狀態的累加，這個說話人就是主觀性的來源。此時，“又”不再重讀。這種投射用法因常用而規約化，凝固為連貫性類同行為的關聯副詞，已收為辭書義項。如《現代漢語八百詞》（呂叔湘 1980）將 (9a) 類用法歸為義項 1：“1. 表示一個動作（狀態）重複發生……”，而將 (9b) 類用法歸入義項 2：“2. 表示幾個動作、狀態情況累積在一起”。當然，將 (9b) 分析為累加性副詞的知域用法是本文的看法，我們認為正因為它比行域用法有所虛化和主觀化，所以不再重讀。

實際上，普通話中“也、就、才、再”等副詞都可用跨域投射分析其義項分化和詞義主觀化，跨域投射是形成常用副詞多義現象的重要機制之一。

與“又”相似，粵語“先”、“添”的語氣詞用法也經歷了跨域投射的規約化。就“先”來說，它的疑問語氣詞用法主要用來引出一個問題或提出一個提議，這就是一種規約化。因為，投射到言域，理論上既可以表示疑問、祈使，也可以表示陳述，只要求其指向語言活動，不要求其專門指向疑問句和祈使句。現在，“先”的語氣詞用法只見於疑問句和祈使句，表明這裡起作用的不但有普遍存在的跨域投射機制，還有粵語特定的選擇——只投射到言域中的疑問句域和祈使句域，排除了陳述句域。如：

(10) A： 今日我哋去睇電影得唔得？‘今天我們去看部電影行不行？’

B₁： 放緊邊部電影先？（我先問一下在放什麼電影？）

B₂： 我哋去飲杯咖啡先。（我先要提議大家去喝杯咖啡，（然後再……））

B₃： *今日有好睇嘅電影先。（我先要告訴你們今天沒好看的電影。）

B₁ “先” 投射到言域中的疑問句，句子成立。B₂ “先” 投射到言域中的祈使句，句子也成立。B₃ “先” 投射到言域中的陳述句，理論上與 B₁、B₂ 相對應，有可能成立，事實上在粵語中該句不成立。可見，這一個“先”已規約化為一個增強疑問語氣或表示祈使提議的句末語氣詞。鄧思穎將其釋義為“到底”、“暫且”，是試圖用普通話詞語來對譯，事實上很難做到詞對詞的翻譯，雖然這兩個意思有點接近，但無法體現“我先要問”、“我先提議”這樣的隱含義。“添”

的意義更加多樣化，其中有些也可以認為有新的義項，詳第 3 節的討論。

2.3 跨域投射引起的主觀化用法，有時基本不改變詞類屬性，有時會引起句法行為的調整，導致詞性改變和虛化。“先”、“添”都發生了副詞向句末語氣詞的轉類。

沈家煊 (2003) 所分析的“如果……就”的三域用法，雖然所在句子的語義關係不同，但相關功能詞的句法行為並沒有明顯改變，“如果”沒有失去其複句連詞的詞性，“就”也沒有失去其關聯副詞的詞性。上文分析的普通話副詞“又”，雖然引起了義項的分化，但句法上，它們仍是典型的副詞，整體詞性尚未改變，但在副詞的小類之間有所波動。“又”的行域用法是意義偏實的累加性副詞（或張誼生 2000:21 所說的“重複副詞”），知域用法則進入了意義偏虛的小類（關聯副詞），是一種語法化現象。不過，跨域投射也可能再進一步，引起詞性的改變。粵語“先”、“添”就屬於這種情況。下面我們參考鄧思穎 (2006a, 2006b) 的句法分析來討論它們的詞性改變。

鄧思穎的上引論文，通過句法測試確定這些詞在線性序列中的位置。其中 (2006a) 文著重分析副詞“先”和疑問語氣詞“先”的不同句法位置，(2006b) 文則全面分析動詞後多個功能詞的相互位置和句法地位。對於“先”，鄧文的結論是，副詞“先₁”在表示時間和焦點類語義的語氣詞之前（＝之內），而疑問語氣詞“先₃”（在 2006a 文中記作“先₂”，但 2006b 文則用“先₂”記祈使語氣詞，疑問語氣詞記作“先₃”，本文從 b 文）在表示時間和焦點的語氣詞之後（＝之外）。我們參照鄧思穎 (2006b) 的表示法公式化如下：

VP 後：先₁ > 時間、焦點類語氣詞 > 先₃

鄧思穎 (2006b) 全面探討了多個動詞後功能詞的相互位置和句法地位。得出以下序列：

添₁ > 先₁ > 咁滯/住/嚟 > 添₂/番/話/先₂/先₃

僅從以上序列，還不足以看出“先”、“添”的轉類，因為無論是副詞狀語（例如在動詞前的狀語或英語中動詞後的狀語），還是句末語氣詞，都並不排斥同類依次疊用。鄧思穎爲了提煉上述序列，進行了細緻的句法測試。對於本文關注的“先”、“添”這兩個詞而言，引起我們注意的是他所指出的下述現象。

1) “先₁”和“先₂”，“先₁”和“先₃”，都能以不同的功能疊加連用。例如：

- (11) 食一碗添先₁先₂。(鄧釋義：暫且先多吃一碗。/ 劉分析：我先要提議先再吃一碗。)
- (12) 咪講先₁住先₂。(鄧釋義：暫且先別講。/ 劉分析：我先提議先別講。)
- (13) 邊個去先₁先₃? (鄧釋義：到底誰先去? / 劉分析：我先要問誰先去呀?)
- (14) 佢食先₁未先₃? (他到底先吃沒有? / 劉分析：我先問一下他先吃了沒有?)

(11)-(12)句都是“先₁”和“先₂”連用，(13)-(14)句則是“先₁”和“先₃”疊加連用。劉丹青(2001)曾經指出，同一個詞以不同句法身分疊加的現象可以用作判斷語法化程度的一杆標尺：它意味著雙方的語法化等級已顯著不同，這就為語法化的漸進連續統插入了一個難得的顯性離散點。“先₁”的意義和功能與普通話副詞“先”相同，區別只在後置還是前置於VP，而狀語在動詞之後是VO-前置詞類核心居前優勢的語言中更加和諧的現象。因此“先₁”是典型的副詞狀語。而且，“先₁”之前要麼直接緊跟動詞或VP，要麼只能插入“添₁”之類意義較實的副詞狀語，無法插入語氣詞一類成分，這也說明“先₁”沒有只能分析為語氣詞的情況。而“先₂”、“先₃”以不同的語義和功能與“先₁”連用，證明其句法身分已經與副詞相當不同，其位置在句末，分析為語氣詞是很恰當的，也符合其比較虛化的語義。

2) “添₁”與“先₁”連用時位於比“先₁”更加靠左貼近VP的位置，副詞狀語的身分更加無疑，其意義上也與普通話副詞狀語“再、又”和英語VP後的副詞狀語again等對應；而“添₂”與“先₂”、“先₃”佔據同樣的句法位置，同屬句末語氣詞。鄧文未舉出“添₁”與“添₂”疊加連用的例句，但麥耘先生為本文提供了顯示兩者可以連用的廣州話例句，更清楚地顯示了“添₂”的語氣詞屬性：

- (15) A：唔係今晚值班，我而家重飲得半斤添啊！‘不是今晚值班，我現在可以再喝半斤。’
B：噉話噉，若果聽日唔上班，我而家飲得三斤添添！‘照這麼說，要是明天不上班，我現在還能再喝三斤呢。’

3) 在“添₁”、“先₁”和“添₂”、“先₂”、“先₃”之間，可以出現一批與時間、強調等有關的語氣詞，例如表示近過去及確認判斷語氣的“嚟”（“來”的不規則讀音），與這個“嚟”功能相近的還有普通話的“來著”、官話方言及近代漢語中表近過去或已然“來”（參看麥耘 2006 及所引相關文獻），都是典型的語氣詞。這進一步說明“添”、“先”由語法化導致了顯著的句法分化，原有的副詞身分和重新分析後的身分差距大於它們各自與“嚟”等語氣詞的差距。“添₁”、“先₁”是比“嚟”更靠近動詞的副詞狀語，而“添₂”、“先₂”、“先₃”是比“嚟”更邊緣的語氣詞。粵語學者多在詞類上將它們分開處理，這是有句法和語義依據的。也有些學者用不同的方式對待“先”和“添”。如方小燕 (2003) 在其廣州話句末語氣詞專著中將“添₂”收作句末語氣詞助詞，並為之專造了“啖”字，而未提與副詞“添”的關係，反映了在廣州話母語人語感中“添”的語氣詞用法已經很虛了；另一方面，方著沒有收錄“先”作為句末語氣助詞。

4) “添₁”、“先₁”和“添₂”、“先₂”、“先₃”之間的詞性分化，還有語音上的表徵。承麥耘教授提示，同樣出現在句末，“添₁”、“先₁”念它們的本調（陰平為高平 55），而“添₂”、“先₂”、“先₃”念語氣詞專用的語調（陰平語氣詞一律變高降 51。參看麥耘 1998）。如“點樣先₁”（先怎麼樣）之“先”念 55，而“點樣先₂”（先說說怎麼樣）之“先”念 51。

3. 跨域投射語義的規約化與“先”、“添”既有解讀的評述

上文說明，“先”和“添”原來分別是表先行和表累加的後置副詞，它們從行域到言域或知域的跨域投射在語法系統中規約化為新的義項和新的詞性，而原來的語義在新用法中仍然有所保存，只是從限制 VP 變成限制言者主語的母句。至於規約化的表現也有多種。對“先”來說，它的言域用法所修飾的隱性母句不限於單純的言語動詞，而有特有的選擇，即表示的不是簡單的“我說”，而是“我要問”、“我提議”等，於是形成“我先要問”、“我先提議”這樣的規約性含義。“添”的規約性語義則主要由它的語用含義凝固而成，表現得更加複雜多樣。“先”和“添”的句末語氣詞用法的詞性分化和語義變異均可以從跨域用法得到比較統一的解釋，而以往文獻中對這些用法的解讀分歧也可以得到更合理的解釋。

關於“先₂”，鄧思穎 (2006b) 認為是表示“暫且先做……（然後）”。梁仲森（2005，引自鄧思穎 2006b）認為這個“先”“強調說話那一刻”。這裡有兩點值得注意。第一，所謂“強調說話那一刻”，體現其有直指（deictic，或稱現場指）作用。而“先₁”只表示相對次序在前，用於過去現在將來各種情況。第二，“先₂”只用於表示祈使或意願的句子，而祈使和意願都是第一人稱即說話人發出的以言行事行為；“先₁”則不限於祈使句，可以用於陳述或疑問。“先₂”比“先₁”多出來的直指義就源於第一人稱的言語行為，因為第一人稱的以言行事行為必然是現場性的，“先₂”的限定對象不再是 VP 這一行為，而是現場的言語行為，即隱性母句“我提議”，於是得到“我先提議”這一解讀。而“先₂”比“先₁”多出的祈使義也來自“我提議”這個母句。行域向言域（此處實現為言域中的祈使言語行為）投射的機制使這些含義得到統一而合理的解釋。

關於“先₃”，鄧思穎將其解釋為表示疑問語氣的“到底”。“到底”是語氣副詞，表示說話人迫切追問的語氣。這一解釋雖然比較貼近，但仍不足以反映該詞對說話人“優先回答問題”的要求，也無法充分反映“先₃”和“先₁”的聯繫：“先₁”表時間次序，“先₃”表疑問語氣，兩義聯繫何在？只有聯繫跨域投射，找出其中隱性的言者主語及所帶的言語行為動詞“我要問”，才能解釋疑問語氣的由來。而疑問語氣的迫切性，來自“先”對隱性母句的修飾所產生的“優先”回答的要求。本文由此對其疑問語氣與“先₁”的關係給予了統一的解釋。

“添₂”的用法更加複雜，文獻中的解釋分歧也更大，其實也都與知域用法有關。

副詞用法“添₁”表示類同行為的累加性重複。一些學者將其解釋為表示賓語數量的增加（如詹伯慧 1958），因為他們認為這種用法的句子中賓語前總帶數量限制，如“食一碗飯添”（再吃一碗飯）。不過，“添₁”其實也可以用於無數量限制的賓語，如：

- (16) 等埋小陳添喇，急乜嘢啊？（（除了等了某人之外）再等等小陳吧，急什麼呢？）（李新魁等 1995:504）

總體上，它是表示說話人的祈使或願望。黎美鳳 (2003) 注意到這一用法的“添”只以第一第二人稱為主語（例外情況見黎文注 5。這實際上已是語氣詞“添₂”了）。這表明了它與祈使願望這類言語行為的高度相關性。根據語用學的言語行為理論，不帶言語行為動詞的祈使句都可以看作是一個隱性的言語行為動詞所支配的補足語從句（參看 Levinson 1983[2001:244]），即“你快走”相當於

“我要你快走”。但是，與“添₂”限制言語類動詞不同，“添₁”的句子即使補出言語行為動詞，也在“添₁”所管不著的上層母句中，“添₁”修飾的是祈使言語動詞所支配的從句內的動詞短語，局部地簡化圖解如下：

(17) (我要你)[[食一碗]添] (圓括號裡是母句成分)

“添”字祈使功能的隱性言者主語，為它發展出真正指向知域母句的功能創造了便捷的條件，“添”的知域用法，只需要將“添”的修飾層次往外擴展一下。如“出便落緊雨添”（外面下著雨）可以圖解為：

(18) [[[我知道][出便落緊雨]]添]

上述分析為我們深入剖析“添”的知域用法和相關語義變體及其規約化奠定了基礎，也使我們能為諸家的不同釋義提供一個統一的解釋。

知域用法的本質就是在字面句子之上有“我知道（或發現）”這種表明說話人認知狀態的隱性母句，相關的功能詞不再指向原來的命題，而是指向這個母句。“添”作為副詞表示類同行為的累加性，近似於“再”、“又”、“還”。限制母句“我知道”後成為“我又知道、我還知道”（粵語表達為“我重知道”）的隱性母句。這是各種“添₂”句的基礎意義。

以“出便落緊雨”為例，加上“添”後，便產生下列效應：

(19) 出便落緊雨+添 → 出便落緊雨添 = 我重發現：出便落緊雨添（我又發現：外面下著雨呢）

“添₂”作為語氣詞的各種語義變體，都來自這一基本語義與語境的互動，其中有的變體已經規約化。“添”的主觀語氣義，與知域本身所含的言者主語有關。

在正常情況下，說外面下著雨這樣的命題，沒有必要在前面加上“我又知道”、“我又發現”這樣的母句，因為任何陳述性的命題都是言者所知道的內容，多數情況下不必添加。一旦這麼說，便是一種違背適量原則的超量表達，必然要產生會話涵義。本句通過“添”的宣示，首先可以催生預設的存在，例如正要出門郊遊，忽然有人來電話告知要來拜訪，此時說話人說“出邊落緊雨添”，是要指明另一件同樣干擾郊遊的事情——“下雨”，朝同一方向累加了“今天郊遊困難真多”的信息。這可以視為一種尚未規約化的會話含義。

關於這類句子的預設，李寶倫、潘海華 (Lee & Pan 2003) 有一個部分合理的分析。他們認為“添”字句都有一個存在預設，即存在一個與本句構成對比的相關命題（本文修改為“類同”而不是“構成對比”，詳下）。他們認為這是由“添”的基本義——量級增量所決定的，預設的命題是作為本句（斷言句）表增量的襯托。我們認為，預設確實存在，但“添”的基本義不是量級增量，而是類同行為的累加（“飲杯酒添”預設的是已飲過酒，都是飲酒行為），投射到知域之後是認知行為的累加（我又知道、我還發現）。“添”所修飾的層次不同，兩者的構成條件也不同。“添₁”指向本句 VP 的行為，因而預設要求前後的行為性質上類同。“添₂”指向上層的隱性母句，表示“知道、發現”的累加，並不要求其子句所述的行為與預設行為完全類同，不過仍要求有共同的意義傾向，而不是李、潘文所說的對比關係。此外。如李、潘文所言，這個預設可以真實存在於話語的上文中，也可以存在於言談雙方的共識中。不過，根據我們對語料的分析，在某些已經規約化的語義變體中，這種預設並不必然存在（詳下）。

下面我們據此來進一步分析“添₂”的語義變體及其規約化。

“添₂”被多位作者提到的一種解釋是表示預期（或預設）之外。鄧思穎用“竟然”來翻譯“添₂”，就屬於這種解讀。黎美鳳 (2003) 也常用“驚訝”來解釋“添₂”之義。這種語義可以從“添₂”句知域用法的隱性母句“我還知道”得到解釋。“我還知道/發現”這一命題（不管表現為顯性還是隱性）含有兩面性。一方面，說明類同的信息已在上文或預設中存在，而類同的命題理論上信息量不高，說的必要性不強。另一方面，說話人明確聲明“我又知道/我還發現”，實際上是在明確宣示說此句的必要性，不是廢話，這就同時帶上了預示高信息量的話語功能。因此，這個相當於“我又知道/我還發現”一類母句的“添”帶上了標記高信息量的重要功能。信息論上，信息量與出現頻率成反比。越是預期之外的內容，出現率越低，則信息量越高；越是可預期，則信息量越低。因此，標記高信息量的“添₂”本性上就與反預期信息更加親和，更容易匹配。試比較下列語境下的句子：

- (20) （語境：大家知道很多人可能不來，炳叔也很可能不來。）
〈普〉a.（很多人沒來，）??我又知道，炳叔也沒來。
〈粵〉b.（好多人冇來，）*炳叔冇嚟添。
- (21) （語境：大家知道很多人可能不來，炳叔則不可能不來。）
〈普〉a.（很多人沒來，）我又知道，炳叔也沒來！
〈粵〉b.（好多人冇來，）炳叔冇嚟添。

從 (20)、(21) 可見，“添”和反預期信息匹配、對可預期信息排斥，這跟“我又知道”這樣的顯性母句的表現是一致的。

當然，“添”和反預期信息的匹配，不意味著它本身就表示反預期義，但經常與反預期信息共同出現，就會受其浸潤而逐漸吸收反預期含義，使之最終成為“添₂”的固有語義，這就是語用含義的規約化。再進一步，“添”字句可以脫離上下文或預設而在單句中獨立表達這種規約化的含義，從而使反預期進一步凝固為“添₂”的主要語義之一。例如：

- (22) a. 炳叔有嚟添！（炳叔竟然沒來啊！）
 b. 我重估你唔番嚟添！（我還以為你不回來了呢？）（李新魁等 1995:504）
 c. 落雨添。（下雨了麼！）（c, d 引自黎美鳳 2003，普通話譯文為引者所加）
 d. 部影印機無紙添。（這台複印機沒紙了麼？）

(22) 各句都表示與預期不符的情況，它們都可以作為單句出現，不需要有類同的行為事件作為預設。到這個階段，“添₂”的語義與“添₁”相比，已有顯著的變異，不再有累加性的修飾義，不管是行為事件的累加還是命題的累加。黎美鳳認為這類句子有與字面義相反的預設，如 (22c) 的預設是“沒有下雨”（也許是“不會下雨”——引者），(22d) 的預設是“影印機是有紙的”（也許是“複印機應當有紙”——引者）。按此說法，“添”促發的預設也發生了變化，從預設類同的事件，到預設相反的情況。這更證明其主要詞義可以轉向反預期。

其實與“添”對應的普通話“又”也有類似的語義引申，只是前置性的副詞不容易發展到語法上的語氣詞這種階段。“又”的基本義是表示同類甚至同一行為的累加性重複（已然行為，未然的累加則用“再”），是一種客觀的陳述。如前文 (9a) “（他喝過一些酒了），他又喝了一杯酒”。這個“又”一般重讀。而“又”的一個引申義是表示相關聯的一些行為的累加，如前文 (9b) “（醫生讓他戒煙戒酒，）他出院當天就抽起了煙，第二天，他又喝了一杯酒。”我們前面已指出，這種用法實際上是第一種“又”的知域用法，即表示“我又知道”，是“我知道”這一母句的累加性重複，而不是行為本身的重複。所以不要求事件本身相同，但要求信息的傾向相同（都是醫生勸止的行為），這種“又”有所虛化和主觀化，不再重讀。此外，“又”還有更加主觀化的語氣用法，如“我又不喝酒，你管我什麼？”，這類“又”已經不需要上文和預設命題，也不重讀。顯

然，這種用法離“又”的本義更遠，是從知域用法進一步發展來的。由此可以推知，作為單句表反預期的“添”，也是從知域用法發展而來的。有時同一個句子可以有不同主觀化階段的解讀。如“出便落雨添”，正如前文所分析的，此句可以是在已經存在不利於出門的因素時的知域用法，表示說話人所知道的類同性命題的累加，也可以脫離預設獨立使用，表示言者的主觀意外語氣。從上面與“又”的比較看，單句用法顯然是從關聯用法進一步發展而來。

有些文獻將“添₂”的語義概括為表“遺憾”。這一解讀能得到部分語料的支持，如(22c)可以用來表示對下雨的後果（如取消郊遊之類）的遺憾，但“遺憾”不是對這類用法的合理解釋。首先，遺憾只適合於無法挽回的不如意事件，而有些不如意事情是可以挽回的，不存在遺憾，如複印機無紙固然不如意，但紙張可以添加，談不到遺憾。用“不如意”可以更好地覆蓋這部分“添₂”句的語氣。其次，即使是“不如意”，也不是“添₂”的固定語義，而是反預期義的語境變體。我們注意到，所有表不如意的“添₂”句都同時有反預期的語義和語氣，但不是所有表反預期的“添₂”句都表不如意，如上面(22b)就只表反預期，不表不如意或遺憾。可見不如意還不是“添₂”的一個固定義項。當然，不如意的語氣在表反預期的“添₂”句中佔優勢。這是由語用原則造成的。人類的思維、認知和語句都存在“求好準則”，即主觀上朝好的方向預期和表達。反過來，表達為反預期的情況，自然就傾向於不如意，“居然、不料”等反預期副詞都有不如意的傾向。假如有一天，表不如意的句子可以不帶反預期的含義，則可以認定“添”獨立獲得了固定的表不如意的意義。目前“不如意”還是可以看作反預期的一種具體表現或一種可選性的傾向。

詹伯慧(1958)最早將“添₂”解釋為表強調誇張，並認為很難看出它與後狀語“添₁”的淵源關係。方小燕(2003)、李新魁等(1995)也持強調誇張說，不過李新魁等仍然承認它與“添₁”的聯繫，只是認為這種聯繫比較“隱晦，變成一種強調、誇張的語氣”。而從他們所舉的例句看，這些句子實際上主要體現了“添₂”的反預期義。例如上面被用作強調誇張之例的(22b)，以及詹伯慧(1958)所舉的“我以為你唔返來添”。“強調、誇張”總體上就指增強語句的信息強度。不過，關於強調誇張的描寫，只有落實到怎樣強調、強調哪個成分、屬於什麼焦點等等，才是有效的語言學描寫。否則，它只是一個理論前(pretheoretic)的籠統感覺。相比而言，“反預期”是一個可以測試的定性。值得注意的是，反預期信息本身的信息量很高，而專門用來凸顯反預期信息的標記天然帶有強調誇張的作用（如：竟然、居然、連、甚至、比……還）。我們看到上述文獻中所謂強調誇張的“添”字句，基本都有反預期屬性，有的同時帶有反預期所伴隨的不如意語

氣，而黎美鳳 (2003) 也常將與強調誇張有關的例句釋義為“驚訝、強調”、“驚訝、誇張”，可見“添”所表示的強調、誇張主要是反預期語義的附帶特徵，不是一種獨立義項。此外，黎文提到，表示強調誇張的“添”都與前面的“重”（俗作“仲”，相當於普通話副詞“還”）同現——當然事實上這只是一個傾向，黎文本身就有不用“重”而表強調的例句。但至少說明，表強調誇張經常是由“還”和“添”共同完成的。這個問題涉及“重”的作用和“重”“添”兩字關係的問題，本文不擬展開，只想提一點簡單的看法。普通話“還”作為副詞與粵語“添”有語義上重疊的部分（但不是等同，副詞“添”也對應於“再”或“又”），“還”有反預期兼強調誇張用法（這個木木的小夥子還是個研究生！），而且“還”的強調誇張用法正是“還”的言域用法，就是沈家煊 (2001) 所分析的“還”的“元語”用法，即指向說話人所說的語言本身的用法，這符合沈家煊後來 (2003) 所說的言域用法。可見“還”（粵語“重”）和“添”有重疊的副詞語義和類似的跨域投射（知域和言域都是主觀化的投射，都會滋生表語氣功能），顯示了相似的引申機制，其強調誇張用法也都與反預期用法合為一體。

“重”和“添”在句中的作用是一種表達上的近義強化，“重……添”也可以像鄧思穎 (2006b) 那樣分析為一個框式虛詞，不過由於句法位置不同，功能側重點略有不同，“添”作為句末語氣詞處在更加外層，更多偏向說話人主觀語氣。

根據上面的分析，我們認為“添”的所謂強調誇張用法，基本上是反預期用法的自然的表達效應，總是伴隨著它的反預期用法，不必分析為“添”的一個獨立的語義變體。

很多人注意到“添”字句常常有進層（遞進）、程度強化、數量（比上句或預設）增加這類現象。李寶倫、潘海華 (Lee & Pan 2003) 集中在這類語義的探討。他們的結論是，“添”字句表示量級增量 (scalar increment)，即比上文或預設中一個對比性的命題在程度或數量的量級上佔據更高位置。他們的分析比起以往籠統談論遞進、增加，有可取之處：1) 注重預設的存在。“添₁”的用法，包括某些“添₂”的用法，確實存在這種預設。2) “量級增量”是個更具有操作性的可測試標準，他們的測試也證明很多“添”字句符合這一概括。但是，他們的結論還無法覆蓋所有的語言事實。問題如下。

1) 他們沒有區分本文所提出的三個概念域，使他們的結論雖能很好概括“添”的行域用法，卻未必能概括投射到知域之後的用法。如他們所舉的下例：

(23) 喺北京嗰陣，我食過蠍子，阿美食過 gaatzaa 添。（原文 (14) 例）

他們認為後一分句用“添”意味著“阿美”吃過的東西必須多於上一分句的“我”所吃的，所以“阿美”除了吃了句中明言的 *gaatzaa*（蟑螂）外，肯定還吃了“我”所吃過的蠍子。此句確實可以表達這一意義，這是它的行域用法的解讀（添₁）。但是，在知域（添₂），只要滿足“我又知道”所適合的內容要求，即信息的累加及反預期強度的增加，命題內容上允許不增量。根據我們與麥耘教授的討論，如果“阿美”只吃過蟑螂，也符合此句的真值條件，因為吃蟑螂的反預期性更強。而且，即使脫離上一分句及相關的預設，“阿美食過 *gaatzaa* 添”作為單句也成立，並且只涉及她吃蟑螂的事，句子內容很符合“添₂”的反預期要求，可見比預設的內容增量似乎不是完美的概括。他們觀察到這裡涉及到量級並要求量級增量是有合理性的，但在知域層面，它是信息量的大小或反預期的強度的量級增量。命題內容的增量（比預設遞進）可以導致信息增量，但這不是使用“添”的要求。此外，量級增量不適合“添”用於無預設單句表反預期的情況。

2) 他們忽略本文所強調的“添”要求的類同行為或類同命題的限制，即本句表達內容的傾向要與上文或預設中存在的同類命題一致。我們注意到，凡是單純因後分句量級不超過前分句而不合增量要求的句子，在李、潘文中都只打問號不打星號，這說明這一概括不是句法限制，至多是一種語用傾向。他們打星號的句子，其實有其他的不合格原因。如：

(24) *阿美好佻，佢就唔咁佻添。

（原 (36a) 例。阿美很聰明，他則沒這麼聰明“添”。）

此句不成立，不是因為後分句沒有增量反而減量，而是因為兩句表達的傾向相反，一句說人聰明，一句說人不聰明，不是類同性命題，並有“就”（相當於普通話“則”）凸顯其對比性。對比性的命題其實恰恰是阻礙“添₂”合格性的因素。我們發現，雖然程度沒有增加，但只要信息有反預期且不如意的屬性，就符合“添₂”的要求。例如，假設阿強向來是喝酒的，而阿良一向是不喝酒的，那麼下句完全成立：

(25) 今晚阿強飲咗三杯酒，阿良都飲咗添！‘今晚阿強喝了三杯酒，連阿良都喝了酒。’

兩句都表示飲酒的事實，內容傾向一致，都表示喝酒氣氛濃烈。兩句在程度或數量的量級上不可比，此句並不要求阿良喝得比阿強多，但不喝酒的阿強喝了

酒，使後分句有更強的反預期性，雖然不是增量而是減量，根據李、潘文的預測該句應當是不合格的，但句子很自然。

4. 結語

粵語中的“先”、“添”詞都有後狀語和語氣詞的用法。參照行、知、言三域的框架（沈家煊 2003），語氣詞的用法是由這些副詞從行域用法（直接指向該句常規的言語行為）投射到言域（指向支配該句的隱性言語行為母句）或知域（指向支配該句的隱性認知行為的母句）的產物，即原來的副詞義不再修飾原句謂語 VP，而修飾隱性的以言者為主語的“我說、我問、我提議、我知道、我發現”等母句，構成“我先要問、我先提議（先₂、先₃）、我又知道、我還發現（添₂）”等基本語義。這種跨域投射使相關功能詞暗藏言者主語，帶來表達的主觀性，因而又屬於主觀化操作。這些語義同特定語境的結合，便產生各種主觀性的語義變體，其中有些凝固下來，便成為新的句末語氣詞的義項。而在考察它們的句法表現（例如鄧思穎 2006a, 2006b 的連用測試和語序描寫）後，可以確定這些跨域投射的用法確實引起了詞性的轉變，使“先、添”在副詞之外還兼有了句末語氣詞的詞性。

“先”作為副詞表示行為的先時，在投射至言域時只指向疑問和祈使兩種言語行為，分別形成“我先要問”（先₃）和“我先提議”（先₂）兩個規約化的語氣詞義項，不能指向陳述言語行為，沒有形成“我先說”的義項。

“添₁”作為副詞表示類同行為的累加，預設類同行為的存在，並以表示祈使願望為基本功能。在投射至知域時，指向認知行為，形成“我又知道/我還發現”（添₂）的基本語氣詞義項，預設有類同命題的存在，同時負載多種伴隨語義，有的規約化為新的義項。“添”的“我又知道/我還發現”之義宣示了所在句子的強信息量，使“添₂”與反預期信息更加匹配，並吸收此義，使反預期成為“添₂”的規約化語義要素之一，多附帶不如意的語氣，但“不如意”或“遺憾”不是“添₂”的獨立義項。反預期義進一步規約化為獨立的義項，使“添₂”可以脫離預設的類同命題而作為單句的語氣詞存在，主要表達反預期語氣。“添₂”的反預期義意味著“添₂”句信息量很大，文獻中所說的強調誇張語氣其實都是“添₂”句天然存在的大信息量的反映，不是獨立的義項。文獻中常提到“添₂”有遞進、增量一類語義。實際上這是因為“添₂”句常配大信息量的命題，但遞進、增量一類語義並不是“添₂”的獨立義項和使用條件，只要符合類同命題的要求而且有其他增大信息量的條件，如反預期內容，內容上並不遞進或增量的“添₂”句同樣成立。

跨域投射在後狀語中不是只對“先、添”有效，粵語中還有些後狀語的語義分化和詞性分化由跨域投射機制所造成，所以本文的視角可能對其他後狀語的語氣詞用法的研究有一定幫助。

引用文獻

- Fang, Xiaoyan (方小燕). 2003. *Guangzhou Fangyan Jumo Yuqi Zhuci* 廣州方言句末語氣助詞 [*Sentence-final Particles in Cantonese*]. Guangzhou: Ji'nan University Press.
- Lai, Mei Fung (黎美鳳). 2003. *Yueyu 'tim' de Yixie Yuyan Tedian* 粵語“添”的一些語言特點 [*Some Linguistic Properties of tim in Cantonese*]. Hong Kong: The Hong Kong Polytechnic University MA thesis.
- Lee, Peppina Po Lun, and Haihua Pan. 2006. Cantonese sentence final particle *tim* ‘add’ as a scalar increment particle: a comparison with additive particles and English even. Manuscript. Hong Kong: City University of Hong Kong.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge & New York: Cambridge University Press. Cited here from 2001's Beijing edition: Foreign Language Teaching and Research Press.
- Leung, Chung-sum (梁仲森). 2005. *Dangdai Xianggang Yueyu Yuzhuci de Yanjiu* 當代香港粵語語助詞的研究 [*A Study of the Utterance Particles in Cantonese as Spoken in Hong Kong*]. Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong.
- Li, Xinkui (李新魁), Jiajiao Huang (黃家教), Qisheng Shi (施其生), Yun Mai (麥耘), and Dingfang Chen (陳定方). 1995. *Guangzhou Fangyan Yanjiu* 廣州方言研究 [*Studies in Cantonese*]. Guangzhou: Guangdong People's Publishing House.
- Liu, Danqing (劉丹青). 2001. Yufahua zhong de gengxin, qianghua yu diejia 語法化中的更新、強化與疊加 [Renovation, reinforcement and superposition in grammaticalization]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 2001.2:71-81.
- Lü, Shuxiang (呂叔湘). 1980. *Xiandai Hanyu Babai Ci* 現代漢語八百詞 [*Eight Hundred Words in Modern Chinese*]. Beijing: The Commercial Press.
- Mai, Yun (麥耘). 1998. Guangzhouhua yudiao shuolue 廣州話語調說略 [A brief study on intonation in Cantonese]. *Guangzhouhua Yanjiu yu Jiaoxue* 廣州話研究與教學 [*Cantonese Teaching and Researching*], Vol. 3, ed. by Ting-Au Cheng & Jianhua Cai, 269-280. Guangzhou: Sun Yat-sen University Press.

- Mai, Yun (麥耘). 2006. Guangzhouhua de jumo yuqici 'lai' 廣州話的句末語氣詞“來” [The sentence final particle *lai* in Cantonese]. Paper presented at the Third International Conference on Chinese Dialectal Grammar, December 2-3, 2006. Guangzhou: Ji'nan University.
- Shen, Jiaxuan (沈家煊). 2001. Gen fuci 'hai' youguan de liangge jushi 跟副詞“還”有關的兩個句式 [Two constructions with the *hai* in Mandarin Chinese]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2001.6:483-493.
- Shen, Jiaxuan (沈家煊). 2003. Fujin san yu: xing, zhi yan 複句三域“行、知、言” [Compound sentences in three conceptual domains: acting, knowing, and uttering]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2003.3:195-204.
- Shi, Qisheng (施其生). 1995. Lun Guangzhou fangyan xu chengfen de fenlei 論廣州方言虛成分的分類 [On the classification of function elements in Cantonese]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 1995.1:114-123.
- Tang, Sze-Wing (鄧思穎). 2006a. Yueyu yiwenju mo 'sin' zi de jufa tedian 粵語疑問句末“先”字的句法特點 [Syntactic properties of *sin* in Cantonese interrogatives] *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2006.3:225-232.
- Tang, Sze-Wing (鄧思穎). 2006b. Houzhi chengfen he Hanyu fangyan yufa de bijiao 後置成分和漢語方言語法的比較 [Postverbal elements and the comparative study of Chinese dialectal grammar]. Paper presented at the Third International Conference on Chinese Dialectal Grammar, December 2-3, 2006. Guangzhou: Ji'nan University.
- Zhan, Bohui (詹伯慧). 1958. Yuefangyan zhong de xuci 'ceon, zyu, faan, maai, tim' 粵方言中的虛詞“親、住、翻、埋、添” [The function words *ceon, zyu, faan, maai, tim* in Cantonese]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1958.3:119-122.
- Zhan, Bohui (詹伯慧). (ed.) 2002. *Guangdong Yue fangyan Gaiyao* 廣東粵方言概要 [*An Outline of Yue Dialects in Guangdong*]. Guangzhou: Ji'nan University Press.
- Zhang, Yisheng (張誼生). 2000. *Xiandai Hanyu Fuci Yanjiu* 現代漢語副詞研究 [*Studies in Adverbs in Mandarin Chinese*]. Shanghai: Xuelin Press.

A Cross-domain Projection Account for the Content and Function Usages of *sin* and *tim* in Cantonese

Danqing Liu

Chinese Academy of Social Sciences

The Cantonese words *sin* (先) and *tim* (添) each have two usages, one as a post-verbal adverb and the other as a sentence-final particle (SFP henceforth). This paper proposes a unified account for the connection between adverbial and SFP usages based on the cross-domain projection among the three conceptual domains, i.e. acting, knowing and uttering. The SFP usage is the result of the projection of their adverbial usage from the acting domain onto the knowing and speaking domains. The adverb *sin* indicates that the VP event occurs first while the SFP *sin* indicates that a covert matrix sentence like 'I will say/ask VP' occurs first, the latter of which means 'I will say or ask VP first'. The adverb *tim* indicates the increment of the VP event while the SFP *tim* indicates the increment of a covert matrix sentence like 'I know VP', the latter of which means 'I also know VP'. The cross-domain usages of function words might lead to the divergence of lexical meanings of the words in question and even new syntactic status in the word class system due to the conventionalization of these usages. Different usages of the same word can even co-occur in one sentence. The paper also offers an account for some complicated phenomena in the extension of these words.

Key words: Cantonese, post-verbal adverbs, sentence final particle, three domains of acting, knowing and speaking, cross-domain projection

Cantonese *sin* 先 and the Question of Microvariation and Macrovariation

Rint Sybesma (司馬翎)
Leiden University

This paper investigates the question of whether or not the atypical postverbal occurrence of the adverbial *sin* 先 in Cantonese is due to language contact, as has been suggested. After teasing apart several different *sins*, which differ, to varying degrees, from one another in semantics, pronunciation and distribution, and after confirming that the postverbal *sin* is indeed an adverbial, we look at adverb placement in the Tai language Zhuang, observing that in this language (as is the case with other Tai languages) many more adverbials are found in postverbal position than is the case in Sinitic languages. We may therefore conclude that the positioning of *sin* behind the verb is indeed due to language contact. The question that comes up then, however, is why *sin* is the only postverbal adverbial in Cantonese. The speculative answer is that there were more, but that they have been reanalysed into verbal and sentence final particles, which may explain why Cantonese has many more such particles than Mandarin.

Key words: Cantonese, Tai, adverb placement, *sin/xian*

1. Introduction

To celebrate Charles Darwin's 200th birthday, the *NewScientist* issue of 24 January 2009 headlined: "Darwin was wrong". The cover story (Lawton 2009) gives a detailed account of how it became necessary to uproot the tree of life. The tree of life, picturing how one species developed into another, was central to Darwin's thinking; he used it as an organizing principle to explain the evolutionary relationships between different species. The discovery of DNA, however, has upset this orderly picture of the tree and its ever-smaller branches and twigs. "Darwin assumed that descent was exclusively 'vertical', with organisms passing traits down to their offspring. But what if species also routinely swapped genetic material with other species, or hybridised with them? Then the neat branching pattern would quickly degenerate into an impenetrable thicket of interrelatedness, with species being closely related in some respects, but not others" (Lawton 2009: 36). And, as Lawton points out, this is exactly how the biological reality turns out to be: the interrelationships between species pattern like a web rather than like a branching tree.

In linguistics, we have long grown used to the idea that family trees, as a means to organize the relationships between languages, are too beautiful to be true. We know that language contact can have serious consequences for the nature of a language, setting it apart typologically, lexically or otherwise from the other members of what is supposed to be one family. To be sure, languages can be affected by language contact to different degrees, with the “mixed languages” documented in Bakker & Mous (1994) being on one extreme end. More towards the other end of the spectrum, we have language families like Romance and Sinitic, the members of which are all definitely recognizable as being related in many respects (not just lexically), the differences being attributable to the different substrate languages which the different member languages are supposed to have (at least according to Norman 1988).

An interesting question in this context relates to the difference that may or may not exist between macrocomparative and microcomparative differences. Defining these notions in relative terms, Kayne (2005:6) writes: “Work on a more closely related set of languages or dialects is more microcomparative than work on a less closely related set”, adding that “[a]s a first approximation, we can take degree of historical relatedness as an informal guideline for degree of syntactic ‘closeness’” (*loc. cit.*). He goes on to state that microcomparative syntax “is the closest we can come... to a controlled experiment in comparative syntax” (2005:8), explaining that such experiments, were they possible, would involve taking a particular language, let’s say, Italian, in which pronominal clitics follow infinitives, “and alter it minimally, for example, by giving it a ‘twist’ in such a way as to change the position of its clitics relative to infinitives. We would then carefully look at this new language... to see if any other syntactic properties have changed as an automatic result of our experimental twist... These experiments would dramatically increase our knowledge of what clusters of syntactic properties are linguistically significant” (2005:8). Since we cannot do such experiments, all we can do is examine sets of very closely related languages, “languages that differ from one another in only a relatively small number of syntactic ways” and hope to “achieve something of the same effect” (2005:8). With respect to the difference between microcomparison and macrocomparison, Kayne’s idea is that when we study closely related languages, we get closer to a “controlled experiment” than when we study unrelated or less closely related languages.

Kayne is right, of course, that we cannot do such experiments, but would it be possible that they could actually be observed to take place in language contact situations? As I mentioned above, languages change in contact situations, and some do so more radically than others. In cases in which there is little syntactic change, we may try to find out whether or not the changes that exist are interconnected.

In this paper, I will investigate one point in which Cantonese (Yuè) is often reported to differ from Mandarin, namely the postverbal occurrence of adverbials, and see whether

this difference can be taken as a Kaynian “twist”, since it seems to have arisen as a result of contact with Tai languages; Tai is supposed to be a substrate language of Cantonese (see references mentioned below). Our point of departure is the distribution of the element *sin* in Cantonese.

2. *Sin*: the facts

Only considering the form, we seem to need to distinguish two different “*sin*” elements: one with a high level tone, [55] in the tradition set by Y. R. Chao, which we will represent as “*sin1*”, and one with a high falling tone, [53] (or lower than three), which we will represent as “*sin*”. Taking the meaning and other factors into consideration as well (while at the same time excluding adjectival *sin1* as well as *sin1* in fixed expressions from the discussion right away), we seem to have to distinguish minimally four different *sins*: one *sin1*, meaning ‘first’, and three *sins*: *sin*₁, occurring preverbally, meaning ‘only’ (among other things); *sin*₂, occurring postverbally, meaning something like ‘and then we’ll see’; and *sin*₃, a sentence final particle (henceforth, “SFP”) which is found in interrogative sentences to add emphasis (Tang 2006).¹

¹ For this overview, we have consulted the following works: Cheung (1972), Gao (1980), Lucas & Xie (1994), Mai (1993), Matthews & Yip (1994), Yuan (1983) and Zeng (1991). Although Cheung (1972) and Matthews & Yip (1994:355) observe that, tone-wise, there are two different “*sin*1s”, the distributional characterization they give is different from ours. Gao (1980) marks the tones in a way that is consistent with our marking method, but does not make any explicit comments about it. Lucas & Xie (1994) and Mai (1993) provide good and quite comprehensive descriptions, but are silent about tonal differences: the former has “*sin1*” consistently, the latter, written entirely in characters, indicates the pronunciation once in the beginning of the article, marking the tone as [53] (p.64). Despite this, Lucas & Xie (1994) also distinguish four different *sins*, which are more or less the same as the ones we distinguish in this paper.

All data have been checked with native speakers. I would like to thank Lisa Cheng, Tak Cheung, Joanna Sio and Leo Wong for their help in this regard.

The source of the examples is indicated by a capital letter (or two) and a number. The number is a page number, the letters refer to some of the works just mentioned, as follows: “G” stands for Gao (1980), “LX” for Lucas & Xie (1994), “M” for Mai (1993), “MY” for Matthews & Yip (1994), and “Y” for Yuan (1983). The glosses and the translations may not be same as in the original (if they had them). The same applies to the characters used to write the sentences here. All Cantonese examples are presented in Jyutping, the romanization system designed by the Linguistic Society of Hong Kong, even if the original source used a different system. In case the original did not mark tones, they have been added.

Key to the abbreviations used in the glosses: CLASSifier; DEMonstrative; EXPeriential; Sentence Final Particle; SUCCessive; PERFective; PROGressive; 1/2/3 SG/PL: first/second/third person singular/plural.

2.1 *Sin1* ‘first’

This paper will mainly be concerned with *sin1*, with a high level tone, an adverbial meaning ‘first of all, before anything else’. Here are some examples, with *sin1* occurring in preverbal position as well as in postverbal position.^{2,3} Note that, in postverbal position, when the verb is followed by a complement such as an object, *sin1* follows that complement.

- (1) a. 啲女仔先唱歌。(LX196, adapted)
 di1 lei5-zai2 sin1 coeng3-go1
 CL^{PL} girls first sing-song
 ‘The girls sing first.’
- b. 我哋先睇下點再諗辦法。(MY192)
 ngo5-dei6 sin1 tai2 ha6 dim2 zoi3 lam2 baan6-faat3
 1PL first look a.bit how again think solution
 ‘We’ll see how it goes first, and then think of the solution.’
- (2) a. 唔該你斟杯茶俾我先。(Y228)
 m4-go1 nei5 zam1 bui1 ca4 bei2 ngo5 sin1
 please 2SG pour cup tea give 1SG first
 ‘Please pour me a cup of tea first.’
- b. 喺 McDonald’s 食嘢要俾錢先。(LX198)
 hai2 McDonald’s sik6-je5 jiu3 bei2 cin2 sin1
 at McDonald’s eat-things will give money first
 ‘When you eat at McDonald’s you have to pay first (in advance).’
- c. 你哋去工地先，唔係聽下趕唔撤㗎喇。(G137)
 lei5-dei6 heoi3 hung1-dei6 sin1, m4 hai6 ting1 ha1
 2PL go workplace first not be wait a.bit

² Matthews & Yip (1994:192) observe that *sin1* ‘first’ is often accompanied by the delimitative marker *ha6* following the verb; the examples in other sources generally do not contain this marker.

³ It has also been reported as meaning ‘just now, earlier’:

- (i) 你先唔喺度，有人嚟搵你。
 lei5 sin1 m4 hai2 dou6, jau5 jan4 lei4 wan2 lei5 (M64)
 2SG SIN not be there have person come look.for 2SG
 ‘Just now you were not there; someone was looking for you.’

Lucas & Xie (1994:195) report that their informants reject this sentence and prefer to replace *sin1* by 頭先 *tau4sin1* ‘just now, earlier’; our informants agree with theirs.

gon2-m4-zit3 gaa3 le4
haste-not-SUCC SFP SFP
‘You guys go to work first, otherwise, if you wait, you won’t make it on time.’

- (3) a. 你先行啦。(M64)
lei5 sin1 haang4 laa3
2SG first leave SFP
- b. 你行先啦。(M66)
lei5 haang4 sin1 laa3
2SG walk first SFP
BOTH: ‘You go first.’

With respect to the positioning of *sin1* in preverbal or postverbal position, according to our informants, there is not much difference in meaning between members of minimal pairs such as the one in (3). Or the minimal pair formed by (2b) and (4):

- (4) 喺 McDonald’s 食嘢要先俾錢。
hai2 McDonald’s sik6-je5 jiu3 sin1 bei2 cin2
at McDonald’s eat-things will first give money
‘When you eat at McDonald’s, you have to pay first (in advance).’

Some of the informants consulted have the feeling that if there is any difference at all, it is that if *sin1* is preverbal, it is more likely that we will be talking about different events conducted by the same person, whereas with postverbal *sin1* it is more a sequence of events independent of the agent. Not all informants feel equally strong about this. Also, with respect to minimal pairs, some informants seem to have a preference for the sentence with preverbal *sin1*, a preference which diminishes if postverbal *sin1* is followed by other material, like an SFP, as in (3b).

With respect to preverbal *sin1* there is no doubt that it is an adverb, semantically or distributionally, as it seems to occupy the same position as preverbal adverbs such as *ji4-gal* ‘now’:

- (5) a. 我而家同你傾。
ngo5 ji4gal tung4 nei5 king1
1SG now with 2SG talk
‘I’ll talk with you now.’

- b. 我先同你傾。
 ngo5 sin1 tung4 nei5 king1
 1SG first with 2SG talk
 ‘I’ll talk with you first.’

I would like to establish beyond any doubt that in the postverbal position, *sin1* ‘first’ is also an adverbial, and that it is the same *sin1* as in the preverbal position. In order to do that we need to distinguish it from the other *sins* first. This is not necessarily easy, despite the differences in meaning and pronunciation. The differences in meaning are not always easy to hold up and with respect to the pronunciation, the difference is between a high level and a high falling tone, the latter of which is no longer recognized as a separate tone phonologically; at least, Matthews & Yip (1994:21) say that many speakers use the high level and high falling tones interchangeably, but they mention explicitly that the high falling tone is “distinctively heard” on certain elements, *sin* being one of them. We will come back to this question below.

2.2 Preverbal *sin1*

With respect to preverbal *sin1* ‘only’, let us look at some examples illustrating its temporal ‘only, only then’ semantics.⁴

- (6) a. 佢過咗十二點鐘先去瞓覺。(M65)
 keoi5 gwo3-zo2 sap6-ji6-dim2[,] sin heoi3 fan3-gaaui3
 3SG past 12 o’clock only go sleep
 ‘He only goes to bed after 12 o’clock.’
- b. 佢試咗半個鐘頭先放棄。(LX196)
 keoi5 si3-zo2 bun3-go3 zung1-tau4[,] sin fong3-hei3
 3SG try-PERF half-CL hour only give.up
 ‘He only gave up, after having tried for half an hour.’
- c. 係佢嚟咗，我先走嘅。(M65)
 hai6 keoi5 lei4-zo2[,] ngo5 sin zau2 ge3
 be 3SG come-PERF 1SG only leave SFP
 ‘I’ll only leave after he has arrived.’

⁴ Note that, in accordance with the pronunciation of the informants, I have consistently written “sin”, even if the original sources don’t indicate a falling tone. Also, I have added pauses ‘[,]’ where my informants inserted them.

- d. 我翻屋企先同你傾。(Y228)
 ngo5 faan1 uk1-kei2[,] sin tung4 nei5 king1
 1SG return home only with 2SG talk
 'I'll go home, and only then we'll talk about it.'

In expressing the notion of 'only', *sin*_I is not exclusively temporal. Very close to its use in sentences with consecutive events (as (6b-d)), and often hard to distinguish from it, is its use as a marker of the conditional; Matthews & Yip (1994:305) call this type of sentence "implicit conditionals".

- (7) a. 你做咗功課先睇電視。(LX197)
 nei5 zou6-zo2 gung1-fo3[,] sin tai2 din6-si6
 2SG do-PERF homework only watch television
 'You can only watch TV after you finish your homework.'
- b. 你要出大價錢，先有人肯制。(M65)
 nei5 jiu3 coet1 daai3-gaa3-cin4, sin jau6 jan4 hang2-zai3
 2SG if present big-amount-money only have people willing
 'Only if you come up with a big amount of money, will there be people willing.'
- c. 我借到錢，先可以買樓。(cf. MY306)
 ngo5 ze3-dou2 cin2[,] sin ho2-ji6 maai6 lau2
 1SG borrow-SUCC money only can buy place.to.live
 'Only if I manage to borrow money, can I buy an apartment.'

Besides this use, *sin*_I can also express 'only' in a quantificational sense of 'no more/after than'.

- (8) a. 今日先禮拜三。(LX195)
 gam1jat6 sin lei5-baa3-saam1
 today only Wednesday
 'Today is only Wednesday.'
- b. 一個月先輪到一次。(M66)
 jat1-go3-jyut6 sin loen4-dou2 jat1-ci3
 one-CL-month only turn-SUCC one-time
 'It turns only once a month.'

- c. 佢哋總共先得八個人。(M66)
 keoi5-dei6 zung2-gung6 sin dak1 bat3-go3-jan6
 3PL altogether only have eight-CL-person
 ‘They altogether only have eight people.’

Finally, preverbal *sin*_I is used for emphasis, especially, as Matthews & Yip (1994: 192) note, in contrastive contexts; the last example indicates the semantic closeness of this usage to ‘only’:

- (9) a. 嗰場波先精彩啊！(M65)
 go2-coeng4-bo1 sin zing1-coi2 aa3!
 that-CL-game only exciting SFP
 ‘That game was exciting!’
- b. 你先係，我先唔係啊！(M65)
 lei5 sin hai6, ngo5 sin m4 hai6 aa3!
 2SG sin be 1SG only not be SFP
 ‘It’s you! Not me!’
- c. 佢有耳窿㗎？係女仔先有㗎嗎！？(LX195)
 keoi5 jau5 ji5-lung1 gaa4? hai6 leoi5-zai2 sin jau5 gaa4 maa1!?
 3SG have ear-hole SFP be girls only have SFP SFP
 ‘He has a pierced ear? Isn’t it that only *girls* have that?!’

With respect to sentences like (9c), Matthews & Yip (1994:192) note that outside of contrastive contexts, *sin*_I can be seen as emphasizing the preceding constituent; this can be a time adverbial or something else, as the following examples show:

- (10) a. 我聽日先俾電話你啦！(cf. MY192)
 ngo5 ting1-jat6 sin bei2 din6-waa2 lei5 laa1
 1SG tomorrow only give telephone 2SG SFP
 ‘I won’t call you until tomorrow.’
- b. 熱辣辣先好食㗎嘛。(cf. MY192)
 jit6-laai6-laai6 sin hou2-sik6 gaa2 maa3
 very hot only delicious SFP SFP
 ‘It only tastes good, when it’s hot, you know that.’

Two more comments are in order with respect to the pronunciation and status of *sin* in these sentences. First, with a high-level tone, i.e. with *sin*_I, the sentences in (6) through (10) would be ungrammatical (or would have a different meaning). In fact, one

of our informants clearly has the feeling that not only do we have a falling tone here, but the coda *n* also has to be extended (“as if an extra syllable is added”). This automatically brings us to the second comment. Although our informants find all sentences given above more or less acceptable, they strongly prefer the use of *sin1-zi3* (先至) instead of *sin*. Significantly, in the texts commenting on the examples illustrating the use of *sin1*, it is suggested that *sin1*, *sin1-zi3* and *zi3* are equivalent and are interchangeable in the sentences, but in the actual examples they give, we exclusively find *sin1-zi3*! This means that we may have to look at *sin1* as equivalent to *sin1-zi3*, in which case it is sufficiently different, in form and meaning/function, from *sin1* ‘first’ for us to safely exclude it from further discussion in this paper.

2.3 Postverbal *sin2* ‘and then we’ll see’

Postverbal *sin2* is not as easily teased apart from *sin1*. The meaning it expresses seems to be something like ‘first’, but colored by an extra meaning aspect, paraphrasable as ‘and then we’ll talk again’ or ‘and then we’ll see’, as if it is a truncated conditional. Here are some examples (again, with “*sin*” even if the original source did not mark a falling tone):

- (11) a. 呢件事聽老王下晝返嚟先。(M66)
 li1-gin6-si6 ting1 lou5-wong2 ha6-zau3 fan1-lei4 sin
 this-CL-thing wait old-Wong afternoon return SIN₂
 ‘As to this matter, wait till old Wong comes back this afternoon.’
- b. 咪理佢嚟唔嚟，你通知佢先。(M66)
 mai5 lei5 keoi5 lei4-m4-lei4, lei5 tung1-zi1 keoi5 sin
 don’t bother 3SG come-not-come 2SG inform 3SG SIN₂
 ‘Don’t pay attention to whether he comes or not, you just inform him first (and then we’ll know).’
- c. 而家問題係佢個度得唔得先。(M67)
 ji4-ga1 man6-tai4 hai6 keoi5-go2-dou6 dak1-m4-dak1 sin
 now question be at his place OK-not-OK SIN₂
 ‘The immediate question now is whether his place is fine or not.’
- d. 擦咗牙先!
 cat3-zo2 nga4 sin!
 brush-PERF tooth SIN₂
 ‘Only after you have brushed your teeth!’

- e. 聽日先！
 ting1-jat6 sin!
 tomorrow SIN₂
 ‘Wait until it’s tomorrow and then we’ll see!’

The meaning of *sin*₂ ‘only then...’ is very close to that of *sin*₁ ‘first’. In many cases we can actually translate it with ‘first’. However, when we try to replace the *sins* in these sentences with *sin*₁ (that is, with the level, rather than the falling tone), we see that in some of the cases it leads to ungrammaticality, such as in (11c) and (11e). In other cases it is possible, but a meaning change ensues. Consider the following minimal pair (cf. (11b)).

- (12) a. 通知佢先。
 tung1-zi1 keoi5 sin1
 inform 3SG first
 ‘(Before you do anything else) inform him first.’
 b. 通知佢先。
 tung1-zi1 keoi5 sin
 inform 3SG SIN₂
 ‘Inform him first (and then we’ll know/see what happens/talk again).’

Two comments are in order here. First, this difference is very close to what we reported above with respect to the difference observed by some native speakers between *sin*₁ in the preverbal and postverbal positions, that the placement had to do with the identity of the agent of the sequential events, the same for all (preverbal) or not necessarily the same for all (postverbal). Secondly, also in line with what we said above, (12a) is actually not considered very good by most informants consulted. It can be salvaged by adding an SFP, such as *laa*₃.

- c. 通知佢先啦。
 tung1-zi1 keoi5 sin1 laa3
 inform 3SG first SFP
 ‘(Before you do anything else) inform him first.’

In (12b) there is no such problem: *sin*₂ ‘and then...’ does not have to be followed by anything to obtain a well-formed/acceptable sentence. We will return to the differences and similarities between *sin*₁ ‘first’ and *sin*₂ ‘and then...’ below.

2.4 Emphatic *sin*₃

Finally, there is *sin*₃, which, as we will see later on, can safely be called an SFP. It is an emphatic marker, introducing subjectivity, used in questions only (Lucas & Xie 1994, Tang 2006). Matthews & Yip (1994:356) remark that “[i]n questions, [*sin*₃] may express reservations about a possible course of action”. It often expresses a kind of urgency on the part of the speaker. Tang (2006) remarks that *sin*₃ is pronounced slightly elongated.

- (13) a. 噉嘅天，去邊先？(LX201)
 gam2-ge3 tin1, heoi3 bin1 sin?
 such day go where SIN₃
 ‘With such weather, where’s the best place to go?’
- b. 同你講過㗎啦，係唔係先？(LX201)
 tung4 nei5 gong2-gwo3 gaa3 laa4, hai6-m4-hai6 sin?
 with 2SG speak-EXP SFP SFP be-not-be SIN₃
 ‘Didn’t I talk to you about it? I did, didn’t I?’
- c. 我有乜嘢着數先？(MY356)
 ngo5 jau6 mat1-je5 zoek6-sou3 sin?
 1SG have what advantage SIN₃
 ‘In the end what’s in it for me?’

2.5 Teasing apart postverbal *sin*₁, *sin*₂ and *sin*₃

Let us now turn to the question of what grammatical status these postverbal elements have. To start with *sin*₃, Lucas & Xie (1994) and Tang (2006) agree that it is an SFP, and I think that this view is correct. First of all, as Tang points out, what it has in common with many SFPs is that its distribution is sensitive to the type of sentence but not to the nature of the predicate in the sentence; this is illustrated in the following sentences (based on Tang 2006). As we saw, distributionally, *sin*₃ is limited to interrogative sentences; hence the ungrammaticality of (14c).

- (14) a. 邊個最靚先？
 bin1-go3 zeoi3 leng3 sin?
 who most beautiful SIN₃
 ‘So, who is the most beautiful then?’

- b. 邊個去先?
 bin1-go3 heoi3 sin?
 who go SIN₃
 ‘Who will go first, then?’
- c. *佢最靚先?
 *keoi5 zeoi3 leng3 sin
 3SG most beautiful SIN₃
 [not an interrogative sentence]

Furthermore, like all SFPs, we only find it in matrix sentences. As the following examples (adapted from examples in Tang 2006:227) show, we do not find *sin*₃ in subordinate and relative clauses:

- (15) a. 因為邊個做咗老闆(*先)你好開心?
 jan1-wai3 bin1-go3 zou6-zo2 lou2-baan2 (*sin), lei5 hou2 hoi1-sam1?
 because who do-PERF boss SIN₃ 2SG very happy
 ‘(*In the end,) you were very happy because who became the boss?’
- b. 邊個寫(*先)嘅書最好睇?
 bin1-go3 se2 (*sin) ge3 syu1 zeoi3 hou2-tai2?
 who write SIN₃ SUB book most good
 ‘(*In the end,) the book that who wrote is the best?’

This is in clear contrast, as Tang continues to point out, to *sin1* ‘first’, even in the postverbal position, which is found only in sentences in which the predicate allows for temporal sequencing (contrasting (16a) and (16b)), but which is insensitive to sentence-type. We have seen examples of the postverbal *sin1* ‘first’ in declaratives above; in (16b) and (16c) we see it in two different types of interrogatives.

- (16) a. *佢最靚先? (Tang 2006:229)
 *keoi5 zeoi3 leng3 sin1
 3SG most beautiful first
 INTENDED: ‘Who was the most beautiful first?’
- b. 邊個去先啊?
 bin1-go3 heoi3 sin1 aa3?
 who go first SFP
 ‘Who goes first?’

- c. 佢去先咩? (Tang 2006:229)

keoi5 heoi3 sin1 me1?

3SG go first SFP

‘Will he go first?’

The postverbal *sin1* has the same co-occurrence restrictions as the preverbal *sin1*. This is a good argument for the claim that it is an adverb. Also, *sin1* is not limited to matrix sentences. We find it in relative clauses and in subordinate clauses such as the following (again from Tang 2006:230):

- (17) a. 佢去先冇問題。

keoi5 heoi3 sin1 mou5 man6-tai4

3SG go first not.have problem

‘That he goes first is not a problem.’

- b. 做功課先嘅細路仔最乖。

zou6 gung1-fo3 sin1 ge3 sai3-lou6-zai2 zeoi3 gwaai1

do homework first SUB children most well-behaved

‘Children who do their homework first are very sweet.’

Finally, as it has also been pointed out by others, in sentences with an SFP, *sin1* always precedes these particles. It even co-occurs with, or below, particles that are thought to occupy a low position in the structure in most accounts (see Tang 1998, Sybesma & Li 2007). For example, *laa3* (the head of “Deik(tic)P”, which, according to Sybesma & Li, is a kind of TenseP, see (3b)), *zaa3* (supposedly heading FocusP) and *ge3* (FinitenessP, the lowest functional projection headed by an SFP in the structure):

- (18) a. 我飲水先咋。(Tang 2006:229)

ngo5 jam2 seoi2 sin1 zaa3

1SG drink water first SFP

‘I only drank some water first.’

- b. 佢係嚟先嘅。

keoi5 hai6 lei4 sin1 ge3

3SG be come first SFP

‘He was the one to arrive first.’

- c. 唔使擔心，佢會俾錢你先嘅。

m4-sai2 daam1-sam1, keoi5 wui5 bei2 cin2 lei5 sin1 ge3

no-need worry 3SG will give money 2SG first SFP

‘Don’t worry, he will certainly give you the money first.’

Tang (2006:230) has examples with *sin1* and *sin3* in one sentence; here is one of them, which also shows, incidentally, that *sin3* follows some of the other SFPs (Tang has more examples to show this), thus confirming its own status as one.

- (19) 佢食先未先 ?
 keoi5 sik6 sin1 mei6 sin?
 3SG eat first Q-SFP SIN₃
 ‘Now, is he going to eat first or not?’

On the basis of all this evidence, we can safely conclude that *sin3* is an SFP and that *sin1* is not. *Sin1* is best analysed as an adverb. As we noted in the beginning, preverbally it behaves like an adverb both functionally and distributionally, and on the basis of the distributional properties we reviewed in this section, the conclusion that it is an adverb in the postverbal position as well seems quite inescapable. Like other adverbs, and like preverbal *sin1*, it is sensitive to the nature of the predicate. Also, semantically, both *sin1*s behave like adverbs as well in that they modify the predicate, and in this respect there does not seem to be a difference between the two. Furthermore, postverbal *sin1* occurs postverbally, but is not an SFP, and since it is also not: (a) a complement; (b) a verbal particle (which we will discuss below) since it follows the object; nor (c) the head of any other functional projection, it must be an adverbial.

The remaining question to be answered in this section is: what is the status of *sin2* ‘and then...’? We pointed out above that *sin1* ‘first’ and *sin2* ‘and then...’ are semantically very close to one another. We also pointed out that the high level and high falling tones are not different phonologically in modern Cantonese. We can add to this that *sin1* ‘first’ and *sin2* ‘and then...’ cannot co-occur sequentially (see (20a)), as also reported by Lucas & Xie (1994), although *sin2* ‘and then...’, as shown in (20b), can co-occur when *sin1* ‘first’ is preverbal (see also Cheung 1972:190, Mai 1993:69):

- (20) a. *通知佢先先 / 先先。
 *tung1-zil keoi5 sin1 sin / sin sin1
 inform 3SG sin sin sin sin
 INTENDED: ‘Before you do anything else inform him and then we’ll see.’
 b. 先通知佢先。
 sin1 tung1-zil keoi5 sin
 sin inform 3SG sin
 ‘Before you do anything else inform him and then we’ll see.’

On the other hand, in connection with (12), we noticed that there is a difference between postverbal *sin1* ‘first’ and *sin2* ‘and then...’, namely, that the latter does not need to be followed by any other material to get a grammatical sentence, whereas the former does.

Now, the fact that *sin2* ‘and then...’ does not have to be followed by, say, an SFP, could be taken as an indication that it is an SFP itself. The falling tone strengthens this idea, if we realize that question SFP *sin3* also has a falling tone (as well as *tim1*, which we will mention briefly below). Work with informants has given me the impression that *sin2* ‘and then...’, besides having a lot in common with *sin1* ‘first’, also seems closely related to the emphatic question SFP *sin3*. *Sin3* can be interpreted as involving some kind of conditional interpretation too: “tell me quickly, otherwise...”. It could be the case that they are actually the interrogative and non-interrogative counterparts of each other and that we are actually dealing with one single element. As we will suggest below, both may be the SFP version of *sin1*.

2.6 Conclusion

The four *sins* we have distinguished are indeed different, if not semantically and formally, at least distributionally. Yet, at the same time, we need to acknowledge that the three *sins* we encounter in the postverbal position also have a lot in common. What we set out to do was to determine the status of *sin1* in the postverbal position. We succeeded and found that all empirical evidence points in the direction of it being an adverb. *Sin3* is definitely an SFP, as may be *sin2*.

3. An adverb in the postverbal position

We concluded above, that *sin1* ‘first’ is an adverbial, which can follow and precede the verb. Its postverbal occurrence is interesting because adverbials in Chinese languages generally precede the verb, so why is *sin1* different? We will look at this question from several angles. We will consider other adverbials in Cantonese, as well as verbal particles and SFPs, and we will compare Cantonese with Mandarin and the Tai language Zhuang.

3.1 Adverbs in Cantonese and Mandarin

With respect to adverb placement in Mandarin, the general rule is that locative, temporal and other adverbials (such as manner adverbs) precede the verb/VP and duratives and frequentatives follow the verb. With some intransitives, locatives can follow the verb, in which case they are best analysed as (part of) the complement (Mulder &

Sybesma 1992). Furthermore, we find so-called “descriptive phrases” in the postverbal position, connected to the verb with *de*, and although it is not clear how such phrases must be analyzed structurally, they sometimes can be interpreted like manner adverbs (which does not mean that they *are* manner adverbs, in the structural sense). In short, duratives and frequentatives aside, adverbials precede the verb; crucially, temporal adverbs never follow the verb.⁵

Generally, Cantonese is like Mandarin in all these respects. However, it is reported that there are some exceptions. In the literature we can find short inventories of supposedly adverbial elements which may follow the verb in Cantonese (Gao 1980:136, Yuan 1983:228-229, and Zeng 1991:247). *Sin1* ‘first’ is one of these elements, but it is not clear that all elements in these inventories are adverbials, like *sin1*. In fact, when studied carefully, we see that most of them are not; either they are verbal particles (aspectual or otherwise: *faan1* ‘again’, *saai3* ‘all’, *do1* ‘much’), which always follow the verb directly and never follow the object, if there is one, or they are SFPS, such as *tim1* ‘too’ (or degree modifiers as in fn.5). Here are some examples:

- (21) a. 佢着翻啲鞋。
 keoi5 zeok6-faan1 di1-hai4
 3SG put.on-‘re’ CL^{PL}-shoe
 ‘He is putting on his shoes again.’ (‘re-putting on his shoes’)
- b. 佢睇曬我啲書。
 keoi5 tai2-saai3 ngo5-di1-syu1
 3SG read-all 1SG-CL^{PL}-book
 ‘He read all my books completely.’
- c. 佢執翻好曬啲書。(Wong 2008)
 keoi5 zap1-faan1-hou2-saai3 di1-syu1
 3SG tidy-‘re’-done-all CL^{PL}-book
 ‘He tidied up the books again completely.’
- (22) 重可以游水添！(M&Y 356)
 zung6 ho2-ji5 jau4-seoi2 tim1!
 still can swim SFP
 ‘And you can swim too!’

What is interesting about the elements in (21) is that they have a meaning which in English has to be rendered by using an adverb, which can be interpreted as an indicator

⁵ There is the rare case in Mandarin of a degree modifier directly following the adjective it modifies: *jile* as in *hǎo-jile* /good-*jile*/ ‘very good’.

of their lexical, rather than functional, status. On the other hand, it is clear that syntactically, they are one of the verbal particles, which Cantonese has a large number of (in any case, many more than Mandarin, a fact we will return to below).⁶ This point can be illustrated with respect to *do1* ‘much, many’ and *siu2* ‘little, few’ as well. Consider sentences like the following:

- (23) a. 食少一塊。
 sik6 siu2 jat1-fai3
 eat few one-piece
 ‘Eat one piece less.’
 b. 佢日日想睇多兩本書。
 keoi5 jat6-jat6 seung2 tai2 do1 leung5-bun2-syu1
 3SG day-day want read much two-CL-book
 ‘He would like to read two books more every day.’

In Cantonese, this is the only possible order. In contrast, in Mandarin, *duō* ‘many, much’ and *shǎo* ‘little, few’, the counterparts of Cantonese *do1* and *siu2*, only occur in preverbal position:

- (24) 少/多吃一塊。(cf. (23a))
 shǎo/duō chī yī-kuài
 little/much eat one-piece
 ‘Eat one piece less/more.’

What is interesting about *do1* ‘much’ and *siu2* ‘little’ is that despite the lexical meaning, they, too, behave like verbal particles syntactically. In any case, just like elements such as *jyun4* ‘finished’, which is an aspectual verbal particle by all possible criteria, they precede perfective marker *zo2*. *Zo2* always follows the verb directly and can only be separated from it by verbal particles, never by objects.

- (25) a. 佢食完咗啲飯。
 keoi5 sik-jyun4-zo2 di1-faan6
 3SG eat-finished-PERF CL-rice
 ‘He finished the rice.’

⁶ For discussion of these elements, see Liu (2006:133-151).

- b. 佢食多啲啲飯。
 keoi5 sik-do1-zo2 di1-faan6
 3SG eat-much-PERF CL-rice
 ‘He ate some more rice.’

Turning, finally, to *tim1* in (22) — which, by the way, has a falling tone, like *sin*₂ and *sin*₃ — we note that it is a much discussed element in Cantonese grammar. It is often described as a focus element meaning ‘too, in addition’. According to the criteria used to determine the status of *sin*₃ above, we can say that, distributionally, *tim1* is an SFP.

In short, when we look at the overall results, we observe that there is actually one element in the postverbal position which can be analysed as an adverb, and that is *sin1* ‘first’. All other elements are verbal particles or SFPs.

Lucas & Xie (1994:200) point out that postverbal *sin1* ‘first’ did not exist in classical or medieval Chinese, and they mention the possibility that it has entered the language through contact with non-Chinese languages spoken in the south of what is now China. They also point out that cognates of *sin1* can be found in other southern dialects as well, such as Wu, Xiang and Min. This is confirmed by You (2004), who claims that, although in current Wenzhou (Wu), both orders are possible, the one with the counterpart of *sin1* ‘first’ in the postverbal position is older than any other order. He also mentions the possibility of non-Chinese influence. This ties in well with the well-known thesis put forth by Mantaro Hashimoto in a series of papers in the 1970s, saying that “Chinese” is divided into a northern and a southern branch, the former being influenced by Altaic, the latter by Tai (for the references, see Chappell 2001). In the following subsection, we will look at Zhuang, a Tai language, currently spoken in and around Guangxi.

3.2 Adverbs in Zhuang

Although in the Zhuang noun phrase modifiers to the noun consistently follow the noun they modify, at the sentential level, Zhuang basically displays the same pattern as Mandarin and Cantonese in the sense that verbal modifiers generally precede their verb. However, we find a number of exceptions to this basic pattern.

First, there are a small number of adverb-like elements which can only follow the verb; with respect to these elements, the literature gives the same confusing picture as those we got for Cantonese, with many elements included in the inventories which are not easily analysed as adverbs. Wei (1985) includes a number of aspectual elements in his list, such as *gvaq* ‘PAST’ and *dwk* ‘PROGR’, though we have to realize that these aspectual particles, including *liux* ‘PERF’, have a freer distribution than their Chinese

counterparts do, in the sense that they do not always have to follow the verb directly; in some cases they can, for instance, be separated from the verb by an object. Another element included is *dem*, the Zhuang counterpart of Cantonese *tim1* (and according to Huang 1997 its origin), which in Zhuang is much harder to pin down as an SFP instead of an adverb than is the case for *tim1* in Cantonese. Here is an example (Wei 1985:96):

- (26) song duz bya neix lij lix dem hw!
 two CL fish DEM still live TIM SFP
 ‘These two fish are still alive!’

For discussion, and many more examples, see Qin (1995:159-162).

The two elements in these inventories which seem to be most like adverbials are *dahraix* ‘really’ and *gonq* ‘first’. Here are two examples (Wei 1985:96):

- (27) a. lwgnyez cungj aeu doeg saw gonq
 child all want read book first
 ‘The children all want to read first.’
 b. de doeg saw dahraix bw!
 3SG read book really SFP
 ‘He is really reading!’

The literature provides a lot of examples of *gonq* with the meaning ‘first, before anything else’. Here are some more examples (from Luo et al. 2005:521-522, see also Huang 1997:71, and Zhang & Qin 1993:172-173):

- (28) a. gou beij de daengz gonq, vihmaz mbouj hawj gou gonq?
 1SG than 3SG arrive first why not give me first
 ‘I got here before he did, why not give it to me first?’
 b. sou gwn gonq, mbouj yungh caj gou lo
 2SG eat first not need wait 1SG SFP
 ‘You go ahead and eat first; don’t wait for me.’

Note that, whereas we find Cantonese *sin1* ‘first’ in the preverbal position as well, *gonq* is found exclusively in postverbal position (Qin 1995:195). Qin (1995:163ff) presents a number of examples in which *gonq* ‘first’ is more like an SFP, very similar to Cantonese *sin2* ‘and then...’ (not like question particle *sin3*).

Secondly, a sizeable number of temporal adverbs can occur on both sides of the verb (Qin 1995, and Luo 2008). According to Qin (1995:192) the canonical position of

these time adverbials is preverbal, but they can be moved to a postverbal position “for emphasis”. Here are some examples, from Qin (1995:193-194):

- (29) a. de ndwenneix miz
3SG this.month have
b. de miz ndwenneix
3SG have this.month
BOTH: ‘She will have a baby this month.’
- (30) a. haemhneix mbanj raeuz miz denqyingj
tonight village 1PL have film
b. mbanj raeuz miz denqyingj haemhneix
village 1PL have film tonight
BOTH: ‘There will be a film in our village tonight.’
- (31) a. haxbaenh miz vunz ra mwngz
just.now have person seek 2SG
b. miz vunz ra mwngz haxbaenh
have person seek 2SG just.now
BOTH: ‘Somebody was looking for you just now.’

Qin emphasizes that the positioning of the temporal adverb before or after the VP makes no difference to the meaning, except that placement in postverbal position emphasizes the adverb more. He also emphasizes, that these temporal cases are different from *gonq* ‘first’ because *gonq* can only occur in the postverbal position.

Finally, we find some other adverbial-like elements in the postverbal position.⁷ Some of these are manner adverbs, others seem to indicate the reason why or the circumstances under which certain things happen (Qin 1995:198-199, and Zhang & Qin 1993:175). Here are some examples (from Qin 1995:199):

- (32) a. duz-vaiz de dai nit lo
CL-buffalo DEM die cold SFP
‘That buffalo died of cold.’

⁷ There are also some degree adverbs following what they modify; see, e.g., Qin (1995:104); cf. fn.5.

- b. aen-maenz neix gyuk, mbouj ndaej gwn ndip
 CL-yam DEM dirty not fit eat raw
 ‘This yam is dirty, you cannot eat it raw.’
- c. de ninz gutgungq
 3SG sleep back.bent
 ‘He sleeps curled up.’

If we were to translate these sentences into Chinese (Cantonese or Mandarin), the modifying elements would all precede the verb.

What we can conclude from this short overview is that although Zhuang displays the same basic adjunct-verb order as we saw in Chinese, there are quite a number of circumstances in which we find adverbial elements in the postverbal position. One of the elements that is mentioned in all overviews as an example of this more exceptional order, is *gongq* ‘first’, the Zhuang counterpart of *sinl* ‘first’. In fact, this is the picture that is displayed by most (Kam-)Tai languages spoken in China: modifiers to the verb tend to precede the verb, but there is quite some flexibility, especially with time adverbials, degree modifiers and some manner adverbs, and there are a small number of elements which always follow the verb, one of these being the Zhuang cognate *gongq* ‘first’.⁸ On the other hand, it must be noted that Tai languages not spoken in China may show different patterns. In any case, Lao seems to have many more adverbial expressions in the postverbal position (Enfield 2007) and so does Thai (Smyth 2002, Higbie & Thinsan 2002).

4. Discussion and conclusion

In view of the data from the different Sinitic and non-Sinitic languages reviewed above, it seems reasonable to conclude that *sinl* ‘first’ occurs in the postverbal position in Cantonese under influence of Tai languages such as Zhuang (note: Zhuang *gongq* and Cantonese *sinl* incidentally are not phonologically related). On the face of it, we are dealing with a syntactic change; we find an adverb in a position where we do not find it in some related Sinitic languages. Nonetheless, going back to our point of departure, it seems that our experiment has not succeeded—the change resulting from language contact has a very limited effect, it only affects one single element. Why would that be the case?

⁸ This is at least the impression one gets from reading through the series of *jiǎn-zhì* ‘simple records’ of the Kam-Tai (Zhuang-Dong) languages collected in Wang Jun et al. (1984). It includes descriptions of the following languages: Zhuang, Buyi, Dai, Dong, Mulao, Shui, Maonan and Li.

An often observed difference between Cantonese and Mandarin is that the former has many more SFPs as well as many more verbal particles than the latter, and that particles of both types are often quite lexical in their meaning: they mean ‘only’, ‘again’ or ‘all’. What appears to have happened is that Cantonese was influenced by Tai languages in admitting adverbials in the postverbal position, but that it subsequently reanalysed them into either of the two categories it already had in the postverbal position: verbal particles and SFPs. The only element still holding out as an adverb is *sin1* ‘first’ but even *sin1* already has its reanalyzed offspring in the form of *sin2* ‘and then...’ and focus SFP *sin3*. If this is correct, it may explain why Cantonese has many more of these particles than Mandarin.

The question of why such reanalyses took place instead of adverbials just being accepted in the postverbal position is hard to answer. The reason could be socio-linguistic; it is possible that “Chinese” just had too strong a presence. But it could also be that in order to accommodate adverbials in the postverbal position, the language would have had to change too much. Asking for the accommodation of postverbal adverbials would have had to result in a “macro” change. If this is the reason, then it could still be possible that Kayne’s experiments can be observed to happen in language contact situations, but only if the contact concerns languages that are closely related.

References

- Bakker, Peter, and Maarten Mous. (eds.) 1994. *Mixed Languages: 15 Case Studies in Language Intertwining*. Amsterdam: IFOTT.
- Chappell, Hilary. 2001. Language contact and areal diffusion in Sinitic languages. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 328-357. Oxford & New York: Oxford University Press.
- Cheung, Samuel Hung-nin. 1972. *Xianggang Yueyu Yufa de Yanjiu* [Studies on Cantonese as Spoken in Hong Kong]. Hong Kong: The Chinese University of Hong Kong.
- Enfield, Nicholas J. 2007. *A Grammar of Lao*. Berlin & New York: Mouton de Gruyter.
- Gao, Huanian. 1980. *Guangzhou Fangyan Yanjiu* [Investigations into the Dialect of Guangzhou]. Hong Kong: The Commercial Press.
- Higbie, James, and Snea Thinsan. 2002. *Thai Reference Grammar: The Structure of Spoken Thai*. Bangkok: Orchid Press.
- Huang, Yuanwei. 1997. The interaction between Zhuang and the Yue (Cantonese) dialects. *Comparative Kadai: The Tai Branch*, ed. by Jerold A. Edmondson & David B. Solnit, 57-76. Dallas: SIL; Arlington: University of Texas.
- Kayne, Richard. 2005. Some notes on comparative syntax, with special reference to English and French. *The Oxford Handbook of Comparative Syntax*, ed. by Guglielmo Cinque & Richard Kayne, 3-69. Oxford & New York: Oxford University Press.
- Lawton, Graham. 2009. Uprooting Darwin's tree. *New Scientist*, 24 January 2009, 34-39.
- Liu, Shuxin. 2006. *Yueyu Zhuang-Daiyu Wenti* [The Question of Yue and Tai]. Beijing: The Commercial Press.
- Lucas, Alain, and Honghua Xie. 1994. "There is practically one universal Chinese grammar": a propos de *xian1* en mandarin et *sin1* en cantonais. *Cahiers de Linguistique – Asie Orientale* 23.1-2:189-206.
- Luo, Liming, Yaowu Qin, Zhenyu Lu, and Fulong Chen. (eds.) 2005. *Cuengh Gun Yingh Swzdenj* [Zhuang-Chinese-English Dictionary]. Beijing: Nationalities Press.
- Luo, Yongxian. 2008. Zhuang. *The Tai-Kadai Languages*, ed. by Anthony V. N. Diller, Jerold A. Edmondson & Yongxian Luo, 317-377. London & New York: Routledge.
- Mai, Yun. 1993. Guangzhouhua 'sin1' zai fenxi [Analyzing Cantonese *sin1* once more]. *Guangzhouhua Yanjiu yu Jiaoxue* [Research and Teaching of Cantonese], ed. by Ting-Au Cheng & Xiaobing Zhou, 63-73. Guangzhou: Sun Yat-Sen University Press.

- Matthews, Stephen, and Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. London & New York: Routledge.
- Mulder, René, and Rint Sybesma. 1992. Chinese is a VO language. *Natural Language and Linguistic Theory* 10.3:439-476.
- Norman, Jerry. 1988. *Chinese*. Cambridge & New York: Cambridge University Press.
- Qin, Xiaohang. 1995. *Zhuangyu Teshu Yufa Xianxiang Yanjiu* [A Study of the Special Grammatical Phenomena of Zhuang]. Beijing: Nationalities Press.
- Smyth, David. 2002. *Thai: An Essential Grammar*. London & New York: Routledge.
- Sybesma, Rint, and Boya Li. 2007. The dissection and structural mapping of Cantonese sentence final particles. *Lingua* 117.10:1739-1783.
- Tang, Sze-Wing. 1998. *Parametrization of Features in Syntax*. Irvine: University of California dissertation.
- Tang, Sze-Wing. 2006. Yueyu yiwenju mo 'sin1' zi de jufa tedian [The syntactic properties of sentence final *sin1* in interrogative sentences in Cantonese]. *Zhongguo Yuwen* [Studies of the Chinese Language] 2006.3:225-232.
- Wang, Jun et al. (eds.) 1984. *Zhuang-Dong Yuzu Yuyan Jianzhi* [Simple Records of Kam-Tai Languages] Beijing: Nationalities Press.
- Wei, Qingwen. 1985. *Zhuangyu Yufa Yanjiu* [Investigations into the Grammar of Zhuang]. Nanning: Guangxi Nationalities Publishing House.
- Wong, Nai-Fai. 2008. How verbal particles structure events together: a work in progress report. Paper presented at Taalkunde in Nederland-dag (TiN-dag) 2008, February 2, 2008. Netherlands: Utrecht.
- You, Rujie. 2004. Wuyu yufa de lishi cengci [On the historical stratification in the grammar of Wu]. Paper presented at Leiden University, October 2004.
- Yuan, Jiahua. 1983. *Hanyu Fangyan Gaiyao* [An Outline of Chinese Dialects]. Beijing: Language Reform Press.
- Zeng, Zifan. 1991. *Colloquial Cantonese and Putonghua Equivalents* (3rd edition), translated by S. K. Lai. Hong Kong: Joint Publishing.
- Zhang, Yuansheng, and Xiaohang Qin. 1993. *Xiandai Zhuang-Hanyu Bijiao Yufa* [The Comparative Grammar of Contemporary Zhuang Language-Chinese]. Beijing: Minzu University of China Press.

湖南南部處置和被動同標記現象的再思考*

伍雲姬

墨爾本大學

我在 1998 年對湖南 100 個方言點的處置和被動標記進行了普查 (Wu 1999)，發現在湖南南部的 11 個方言點裡，處置和被動都採用同一的、來源於表給予動詞的標記。我在 2005 年的文章中對這種同標記的現象作了一些初步的探討。我從語義和句法的角度討論了為什麼表給予的動詞既可以充當被動標記也可以充當處置標記，討論了當被動/處置採用同一標記時，人們如何消除“張三給（被/把）李四打了一頓”的歧義等問題。這篇文章從另一個角度對上述現象進行了研究，我們利用湖南南部 14 個方言點的材料提出這樣一個假設：在漢語史上，被動和處置的語法化途徑至少有兩條：一條是像書面語和大多數北方方言那樣，被動處置採用不同標記；一條是像湖南南部方言一樣，被動處置採用同一標記。表給予的動詞演化成被動或處置的標記可能在中古就開始了，這兩條途徑在歷史上，至少在口語裡，曾有過並立甚至爭雄的階段。

關鍵詞：湖南方言，處置式，被動式，給字句

我在 1998 年對湖南 100 個方言點的處置和被動標記進行了普查 (Wu 1999)，發現在這 100 個點中，處置標記來源於表給予動詞的共有 74 個方言點，這些表給予的動詞多為“把”，也有“給”和“得”；被動標記來源於表給予動詞的共有 38 個，這些表給予的動詞多為“給”和“得”。而在湖南南部的 11 個方言點裡，如雙牌（湘方言）、常寧（贛方言）以及江永（湘南土話），處置和被動都採用同一的、來源於表給予動詞的標記。我在 2005 年的文章中對這種同標記的現象作了一些初步的探討。我從語義和句法的角度討論了為什麼表給予的動詞既可以充當被動標記也可以充當處置標記，討論了當被動/處置採用統一標記時，

* 本文初稿曾在 2007 年歐洲語言學會的年會上發表，本稿作了較大幅度的修改。謹向對此文初稿提出建議和意見的同仁表示謝忱。筆者還要特別感謝本稿審查人所提出的意見，他們指出了文中好幾處不可原諒的失誤。

人們如何消除“張三給（被/把）李四打了一頓”的歧義等問題。本文將從另一個角度對上述現象進行研究，我們試著用湖南南部的材料提出這樣一個假設：在漢語史上，被動和處置的語法化途徑至少有兩條：一條是像書面語和大多數北方方言那樣，被動處置採用不同標記，為方便論述，我們暫且稱之為“把被分立”；一條是像湖南南部方言一樣，被動處置採用同一標記，我們暫且稱之為“把被同一”。我們認為表給予的動詞演化成被動或處置的標記可能在中古就開始了，這兩條途徑在歷史上，至少在口語裡，曾有過並立甚至爭雄的階段。

1. “把被同一”現象及其地域分布

迄今為止，共發現湖南南部 14 個點有“把被同一”的現象（伍 1999 年的文章裡僅列舉了 11 個點）。在我們掌握的這些有限的材料中，有 8 個點中“把被同一”的語料比較豐富，因此我們利用這些材料作為本文的主要研究對象。其中寧遠官話、嘉禾廣發以及宜章赤石的材料是 1999 年和 2005 年的文章中不曾提及的。

〈表 1〉湖南南部方言中的處置和被動標記

方言點以及材料來源	處置標記	被動標記
1. 常寧（吳啓生 1998）	得 [te ³³]	得 [te ³³]
2. 嘉禾（縣城）(Tsuji 1987)	□ [ɔŋ ⁴⁵]	□ [ɔŋ ⁴⁵]
3. 嘉禾（廣發）（盧小群 2002）	□ [mi ²⁴]	□ [mi ²⁴]
4. 江永（黃雪貞 1993）	分 [pai ⁴⁴]	分 [pai ⁴⁴]
5. 臨武（李永明 1988）	挨 [a ³³]	挨 [a ³³]
6. 寧遠官話（張曉勤 1998）	給 [kə ⁴⁵]	給 [kə ⁴⁵]
7. 寧遠土話（張曉勤 1999）	與 [ie ⁴⁵]	與 [ie ⁴⁵]
8. 宜章（赤石）（盧小群 2004）	拿 [no ⁴⁴]	拿 [no ⁴⁴]
9. 郴州（官話） ¹	給 [kɛ3]	給 [kɛ3]
10. 藍山（土話）	給 [kæ3]	給 [kæ3]
11. 雙牌	給 [kei3]	給 [kei3]
12. 汝城（官話）	拿 [no2]	拿 [no2]
13. 冷水江	拿 [la2]	拿 [la2]
14. 宜章（土話）	□ [tou5]	□ [tou5]

¹ 下面六個點的材料來自作者 1998 年的田野調查。當時採用的是調類記錄法：1 代表中古陰平，2 代表陽平，3 代表濁上，4 代表陰上和陰去，5 代表濁去，6 代表入聲。詳情請參看 Wu (1999)。

〈表 1〉給我們提供了下面一些資訊：

(i) 這八個點的地理位置很接近，都集中在湖南南部。而這麼鄰近的八個點，竟然有八個不同的詞形式。不同的詞形式按同一語法方向發展，可見這種語法化現象不是孤立的。

(ii) 這八個點的處置/被動標記雖都可以用作動詞，表給予，但其來源各不相同。除了嘉禾兩個點的詞源不明，臨武的“挨”的本字尚待考證以外，“得、分、拿、給、予”五個動詞都有一個共同的特點，那就是這些動詞在這些方言裡都至少牽涉到兩個物件：“得”牽涉“獲得者”和“給予者”；“分”牽涉到共用的雙方，“拿”牽涉到“取方”和“被取方”；“給”和“予”牽涉到“付出方”和“接收方”。“得”可能是從“使對方得到”義中滋生出“給予”義，“分”可能是從共同受益義中滋生出“給予”義，“拿”可能是從“使對方取得”義中滋生“給予”義的。

(iii) 這八個點的表給予的動詞都具有多重功能，除了能充當被動/處置標誌以外，還可以用作與格、受益格以及工具格的標誌。請看〈表 2〉：

〈表 2〉“給予”動詞在湖南南部方言中的用法

	常寧	嘉禾 (縣城)	嘉禾 (廣發)	江永	臨武	寧遠 官話	寧遠 土話	宜章 (赤石)
用做動詞表“給予”	+	+	+	+	+	+	+	+
用做工具格標記	-	-	+	-	-	+	+	+
用做受益格標記	+	-	+	+	+	+	-	-
用做與格標記	+	-	+	+	+	+	+	+
用做被動標記	+	+	+	+	+	+	+	+
用做處置標記	+	+	+	+	+	+	+	+

值得說明的是，表中的“-”表示的是在所採用材料中沒有發現該用法，不表示該方言中就一定沒有這種用法。下面是臨武方言的例子。臨武方言表給予的動詞為“挨”[a³³]，[a³³]除了作表“給予”的動詞以外，還可以作與格、受益格、工具格、被動、處置標記。李永明(1988)的書中共有 15 個例句。下面圓括號裡的數字是例句數：

1.1 用作動詞 (2)

- (1) 挨書挨我。(p.378)

a³³ ɕy³³ a³³ ŋo³³.

把書給我。

1.2 用作與格標誌 (1)

- (2) 挨書挨我。(p.378)

a³³ ɕy³³ a³³ ŋo³³.

把書給我。

1.3 用作受益格標誌 (1)

- (3) 挨我寶寶送冷飯。(p.428)

a³³ uai⁵⁵ pau⁵⁵ pau⁵⁵ soŋ¹¹ leŋ⁵⁵ faŋ¹¹.

給我孩子送飯。

1.4 用作被動標記 (2)

- (4) 我挨狗咬㗎。(p.378)

ŋo³³ a³³ kai⁵⁵ ŋau⁵⁵ tie.

我被狗咬了。

- (5) 我挨碗打爛㗎腦殼。(p.379)

ŋo³³ a³³ ua⁵⁵ to⁵⁵ na¹¹ tie lau⁵⁵ k^ho⁵³.

我被碗打破了頭。

1.5 用作處置標記 (9)

- (6) 我挨碗打爛㗎。(p.378)

ŋo³³ a³³ ua⁵⁵ to⁵⁵ na¹¹ tie.

我把碗打破了。

值得強調的是，雖說從現在已有的材料來看，“把被同一”的方言點並不是太多，但也絕對不是僅限於湖南南部方言。請看〈表 3〉：

〈表 3〉漢語中“把被同一”的方言點

方言點及資料來源	被動標記	處置標記
北京（徐丹 1992:54）	給 [kei ²¹⁴]	給 [kei ²¹⁴]
福建福清（馮愛珍 1993:211）	乞 [k ^h ɿ ²²]	乞 [k ^h ɿ ²²]
廣東北鄉（張雙慶 2000:349）	□ [a ²¹²]	□ [a ²¹²]
廣東桂塘（張雙慶 2000:349）	□ [a ¹³]	□ [a ¹³]
湖北鄂東（黃伯榮 1996:660-661）	把	把
河南葉縣（張雪平 2005:301）	叫 [tɕiau ³¹²]	叫 [tɕiau ³¹²]
山西交城（黃伯榮 1996:660-661）	給 [kuɿ ¹³]	給 [kuɿ ¹³]
湖南湘鄉（黃伯榮 1996:669）	鈇	鈇
江西上饒等 7 個點（胡松柏、葛新 2003:244）	把	把
	拿	拿
	拿/擺	拿/擺
	背/撥/拎/端	背/撥/拎/端
	益/端	益/端
	益/拎	益/拎
	端	端
江西龍南、定南、全南（劉綸鑫 1999:745）	□ [ia ²¹]	□ [ia ²¹]
江西萍鄉（劉綸鑫 1999:745）	把	把

〈表 3〉告訴我們，除了湖南南部的 14 個方言點（見〈表 1〉）以外，漢語方言裡至少還有 17 個點有“把被同一”的現象（我相信還可以找到一些）。雖說〈表 3〉提供的材料裡大多數只簡單的列出了處置和被動的標記，沒有對它們的詞源和用法展開討論，因而我們不知道它們是否都和表“給予”的動詞有關，是否都具有多重功能。但是，這些材料至少透漏了一個資訊，既然“把被同一”並不只限於南方，北方（如北京、山西、河南等地）也有，那麼“把被同一”是否可能在漢語史上曾與“把被分立”並立呢？寧遠方言的材料印證了我們的推測，在寧遠的方言裡，“把被同一”和“把被分立”兩套系統是同時並用的。

2. 寧遠方言裡的“把被同一”和“把被分立”

寧遠處於西南官話和土話並用的雙語區，不少人既能說官話，也能說土話。

有意思的是，不管是在寧遠官話還是寧遠土話裡，都有兩套系統，“把被同一”和“把被分立”。

2.1 寧遠官話裡的“把被同一”

在寧遠官話裡，表處置和被動的標記均為“給”。我們在張曉勤(1998)的文章裡共找到了 22 個例句。“給”除了作表“給予”的動詞以外，還可以作與格、受益格、工具格、被動、處置標記。

2.1.1 用作動詞表“給予”(2)

- (7) 給本子給我。(p.80)
kə⁴⁵ pən⁴⁵ tsi⁴⁵ kə⁴⁵ ŋo⁴⁵.
把本子給我。

2.1.2 用作與格標記(2)

- (8) 給碗給他。(p.80)
kə⁴⁵ van⁴⁵ kə⁴⁵ t^ha³³.
把碗給他。

2.1.3 用作受益格標記(3)

- (9) 有好東西從不給他吃。(p.95)
iəu⁴⁵ xau⁴⁵ tɔŋ³³ ɕi²¹ ts^hoŋ²¹ pu²¹ kə⁴⁵ t^ha³³ tɕ^hi²¹.
有好東西從不給他吃。

2.1.4 用作工具格標記(8)

- (10) 沒的碗了，你就給口杯吃飯。(p.81)
mə²¹³ tə²¹ van⁴⁵ liau⁴⁵, li⁴⁵ tɕiəu²¹³ kə⁴⁵ k^həu⁴⁵ pəi³³ tɕ^hi²¹ fan²¹³.
沒碗了，你就用杯子吃飯吧。

- (11) 這座屋是給水磚起起的。(p.82)

tɕi²¹³ tso²¹³ vu²¹ si²¹³ kə⁴⁵ suəi⁴⁵ tɕyn³³ tɕh⁴⁵ tɕh⁴⁵ ti³³.

這間屋子是用水磚建成的。

2.1.5 用作被動標記 (2)

- (12) 我給狗嚇怕了。(p.78)

ŋo⁴⁵ kə⁴⁵ kəu⁴⁵ xə²¹ p^ha²¹³ liau⁴⁵.

我被狗嚇怕了。

- (13) 那本書給一個同學借走了。(p.78)

la²¹ pən⁴⁵ ɕy³³ kə⁴⁵ i²¹ ko²¹³ t^hoŋ²¹ ɕio²¹ tɕie²¹³ tsəu⁴⁵ liau⁴⁵.

那本書被一個同學借走了。

2.1.6 用作處置標記 (5)

- (14) 給那本書拿過來。(p.79)

kə⁴⁵ la²¹³ pən⁴⁵ ɕy³³ la²¹ ko²¹³ læi²¹.

把那本書拿過來。

- (15) 你莫給腳搭到桌子高頭。(p.79)

li⁴⁵ mo²¹ kə⁴⁵ tɕio²¹ ta²¹ tau²¹³ tso²¹ tsɿ⁴⁵ kau³³ t^hau²¹.

別把腳擱在桌子上。

2.2 寧遠土話的“把被同一”

寧遠土話裡表被動和處置的標記均為“與”。“與”的用法與官話中的“給”很相似，也可以作受益格、工具格、被動、處置標記（只是沒有找到做與格標記的例句）。在張曉勤(1999)的著作裡，我們共找到了 21 個例句。

2.2.1 用作動詞表“給予” (7)

- (16) □與我的 50 文花錢還勿得用呱。(p.233)

ie³³ ie⁵³ io⁵³ kə²¹ ŋ⁵³ sl²¹ mian⁵³ xo⁴³⁵ ts^haŋ²¹³ xa²¹³ mə²¹ tə²¹³ iaŋ²¹ kua.

媽媽給我的 50 元錢還沒用完。

- (17) □□清起，爹爹與呱我五文花錢。(p.235)
tɕi²¹³ kaŋ⁴³⁵ tɕ^hio⁴³⁵ ɕi³³ tia³³ tia³³ ie⁵³ kua io⁵³ ŋ⁵³ miaŋ⁵³ xo⁴³⁵ ts^haŋ²¹³.
今天早上父親給了我五塊錢。

2.2.2 用作受益格標記 (1)

- (18) 要逮起豆子秧子與我盯。(p.298)
iaŋ²¹ tə³³ ɕi³³ t^həu²¹ lə iŋ⁴³⁵ lə ie⁵³ io⁵³ tio²¹³.
你應該拿豆秧給我看。

2.2.3 用作工具格標記 (2)

- (19) 你與火照倒那條魚。(p.229)
ŋ⁵³ ie⁵³ fu³³ tsɿ⁵³ tie la⁵³ t^hi²¹³ ŋo²¹³.
你用火照著那條魚。
- (20) 與兩粒衩口載起。(p.297)
ie⁵³ liŋ⁵³ læ²¹³ ts^ho⁴³⁵ k^həu³³ tɕia⁵³ ɕi³³.
用兩個口袋裝著（那些東西）。

2.2.4 用作被動標記 (2)

- (21) 手與斑毛草割開呱了。(p.267)
səu³³ ie⁵³ pie⁴³⁵ mie²¹³ tɕ^hie³³ kə²¹³ xə⁴³⁵ kua lie.
手被茅草割開了。

2.2.5 用作處置式標記 (9)

- (22) 與那本書逮過來。(p.264)
ie⁵³ la⁵³ pian³³ ɕie⁴³⁵ tə³³ ku⁵³ la²¹³.
把那本書拿過來。
- (23) 你哪怎勿與伊點事告訴我呢？(p.265)
ŋ⁵³ lie²¹ tsəŋ⁵³ mə²¹ ie⁵³ a²¹ kuəi⁵³ sl²¹ kə⁵³ sə²¹ io⁵³ lie⁵³.
你怎麼不把這件事告訴我呢？

從上面的例子可以看出，寧遠官話和土話中的“與”和“給”的用法幾乎是一致的。這種一致是不謀而合還是相互影響的結果我們暫且不能下結論。可以下結論的是這兩個動詞“與”和“給”在語法化的道路上都走得相當遠了，也就是說，它們的虛化程度已經相當高了。關於這個問題，我們在下節中還會討論。

2.3 寧遠官話中的“把被分立”

根據張曉勤(1998)的描述，寧遠官話中除了“給”可以表處置和被動以外，“拿倒”(意為“拿”)也可以作處置標記；“兜倒”可以作被動標記。我認為“兜”的本字可能為“逗”，“招惹之意”，讀平聲，在湖南不少方言裡都能找到。“拿倒”和“給”用作處置式可以互換，但是“兜倒”和“給”用於被動時有條件，“給”不能用於不如意的事情，“給”後面一定要有名詞，而“兜倒”沒有這樣的要求。在下面的例句中，(24)、(25)中的“拿到”可以用“給”代替，而(26)中的“給”不能換成“兜倒”，(27)中的“兜到”不能換成“給”。

- (24) 風拿到門吹開呱了。(p.77)

xoŋ³³ la²¹ tau⁴⁵ mən²¹ ts^huei³³ k^hæ³³ kua²¹ liau⁴⁵.

風把門吹開了。

- (25) 領導拿到他臭罵了一頓。(p.77)

lin²¹³ tau²¹³ la²¹ tau⁴⁵ t^ha³³ tɕ^hiəu²¹³ ma²¹³ liau⁴⁵ i²¹ tən²¹³.

領導把他狠狠地罵了一頓。

- (26) 他給領導表揚了一通。(p.83)

t^ha³³ kə⁴⁵ lin⁴⁵ tau²¹³ pia⁴⁵ iaŋ²¹ liau⁴⁵ i²¹ t^hoŋ³³.

他受到了領導的表揚。

- (27) 那本書兜倒老師沒收了。(p.82)

la²¹³ pən⁴⁵ ɕy³³ təu³³ tau⁴⁵ lau⁴⁵ si³³ mə²¹ ɕiəu³³ liau⁴⁵.

那本書被老師沒收了。

從“給”和“兜倒”的這一區別可以看出，它們各自有著自己的發展軌道。“兜倒”遵循漢語被動句的發展方向，只用於不如意的事情，“兜倒”後可以不帶賓語，而“給”並不遵循這一發展軌道。

2.4 寧遠土話的“把被分立”

寧遠土話中除了“與”可以表處置和被動以外，“逮倒”也可以表處置，“兜倒”可以表被動。“逮倒”可以作動詞，義為“捉”（在吳方言裡，“捉”可以作處置標記，參看蔣冀騁 2003）。請看張曉勤 (1999) 給的例句。“兜倒”是否可以作動詞尚不得知。

- (28) 賊佬兜倒我逮倒呱了。(p.266)

tsə²¹³ lɔ³³ təu⁴³⁵ tie io⁵³ tɔ³³ tie kua lie.

賊被我抓住了。

- (29) 那本書兜倒先生繳呱去了。(p.266)

la⁵³ pian³³ ɕie⁴³⁵ təu⁴³⁵ tie saŋ⁴³⁵ ɕie⁴³⁵ tɕiəu³³ kua xo⁵⁴ lie.

那本書被老師沒收了。

- (30) 爹爹逮倒我斥了一場。(p.263)

tia³³ tia tə³³ tie io⁵³ tso⁵³ lie i²¹³ ts^həŋ²¹³.

爺爺把我罵了一頓。

- (31) 你才歸來，逮倒你奶奶急死呱了。(p.263)

ŋ⁵³ tɕ^hia²¹³ kuæ⁴³⁵ la²¹³, tə³³ tie ŋ⁵³ lian⁵³ lian⁵³ tɕi²¹ sæ³³ kua lie.

你才回來，把你奶奶急死了。

“與”和“逮倒”“兜倒”一般可以互換，但有時也不可以，但到底區別在哪裡，還有待研究。綜上所述，寧遠官話和土話都有兩套系統，見下表：

〈表 4〉寧遠話裡的“把被同一”和“把被分立”

方言	把被同一	把被分立	
		被動	處置
寧遠官話	給 [kə ⁴⁵]	兜倒 [təu ³³ tau ⁴⁵]	拿倒 [la ²¹ tau ⁴⁵]
寧遠土話	與 [ie ⁵³]	兜倒 [təu ⁴³⁵ tie]	逮倒 [tə ³³ tie]

這兩套系統在口語裡面的實際運用情況如何呢？很遺憾，我們的材料不夠。從張曉勤所記載的兩個土語講述的民間故事和一個用官話講述的民間故事中只找到了八個例句。例句太少，不能說明什麼問題，但我們還是統計如下，以供參考。

〈表 5〉寧遠話裡的“把被同一”和“把被分立”在口語中的運用

	被動	處置	被動	處置
寧遠官話	給 [kə ⁴⁵] (0)	給 [kə ⁴⁵] (1)	兜倒 [təu ³³ tau ⁴⁵] (1)	拿倒 [la ²¹ tau ⁴⁵] (1)
寧遠土話	與 [ie ⁵³] (0)	與 [ie ⁵³] (1)	兜倒 [təu ⁴³⁵ tie] (1)	逮倒 [tə ³³ tie] (3)

要特別指出的是，當“張三給（把/被）李四打了一頓”的歧義句產生時，人們會使用“把被分立”來解決歧義。

3. “把被同一”與“把被分立”在漢語中的地位

從上文的討論來看，漢語被動和處置標誌的語法化至少有兩個不同的途徑。一個是“把被分立”的途徑，被動和處置標記的來源不同，產生的先後也不同，就像書面語中的“被”和“把”。另一個是“把被合一”的途徑，被動和處置標記來源相同，均來自同一個動詞，就像湘南方言裡的“給”類動詞。湖南南部方言中的“與”“得”等演變為與格、受益格、工具格、被動、處置等標記的歷程可能早在中古就開始了，它和“把被分立”的系統在口語裡可能有過並立甚至爭雄的階段。理由如下：

3.1 “把被同一”不是孤立的現象

“把被同一”的方言點目前已找到的雖不多，但分布卻很廣。湘、贛、閩、粵、西南官話、北京話裡都可以找到。這種分布提醒我們，“把被同一”不是孤立或偶發的現象。

3.2 “把被同一”中的標記已經高度語法化了

“把被同一”的標記可以用於表與格、受益格、工具格、處置以及被動。學者們對給予動詞語法化的順序有過討論，有的認為應是動詞→處置標記→被動標記（如 Bennett 1981）；有的認為應是動詞→受益格→處置標記→被動標記（如李宇明、陳前瑞 2005）。不管“給予”動詞是如何演變成不同標記的，有一點可

以肯定，那就是它們語法化的過程是相當長的。如果說與格、受益格的標記中還有動詞“給予”的影子，那麼工具格、表處置、表被動的標記與動詞“給予”的意義就相去甚遠了。

3.3 “把被同一”和“把被分立”兩套系統可共見於同一個方言點中

在寧遠土話和官話裡，“把被同一”用的標記是單音節的“與/給”，而“把被分立”的標記是雙音節的“兜倒/拿到；兜倒/逮倒”；在寧遠土話和官話裡，“把被同一”用的標記可以有多種作用，而“把被分立”用的標記只有單一的用途。這似乎在提醒我們，“把被同一”在寧遠方言裡的歷史可能比“把被分立”的更長。

3.4 “把被同一”的標記有來源於“與”的

在湖南南部的方言裡，“把被同一”的標記有來源於“給”這個 18 世紀才出現在書面語中的給予動詞，也有來源於“與”等在漢語史上出現得很早的表給予的動詞。根據貝羅貝 (1986) 對書面語言的研究，在中古時，動詞“與”在雙賓語的結構中就取代了其他表示“給予”的動詞，在近代漢語的初期就已經虛化成介詞了。到 18 世紀時，“與”讓位給了“給”。而口語的情況就不太一樣了。

“與”（“與”爲本字，見張曉勤 1999:105）一直活躍在寧遠土話裡，而一批從別的詞義發展到“給予”義的動詞，如“得”“拿”“分”等，有著跟“與”同樣的語法歷程。

4. 尚待研究的問題

“把被同一”的現象雖說 Bennett 在 1981 年已有專文論述，但是到近年才給予較多的關注（如徐丹 1992，蔣紹愚 2001, 2002，石毓智 2004，李宇明、陳前瑞 2005 等）。本文利用湖南南部八個方言點的材料提出了對“把被同一”現象的一些思考。要真正弄清楚“把被同一”和“把被分立”兩套系統的關係，弄清楚表給予的動詞發展爲處置/被動標誌的歷程，還有許多問題值得探討。比如，“把被同一”的方言點在全國範圍內到底有多少？像寧遠方言那樣兩套系統並存的方言點有多少（北京方言應該可以算在此列。參看徐丹 1992）？這兩套系統在實際

口語的運用中的情況如何？這兩套系統之間是什麼關係？寧遠官話和北京話中“給”字的運用如此相似，而這兩個方言在地理上相隔甚遠，這又說明了什麼？如此等等，都是值得進一步研究的。

根據 Heine & Kuteva (2002:1) 對世界上 500 種語言的普查（普通話亦在普查之列），處置式標記 *patient*² 來源於表給予動詞的很常見；但在他們普查的 500 種語言裡，卻沒有一例是表被動的標記來源於表給予的動詞的。由此可見，漢語方言裡的“把被同一”現象無疑能給世界語言的研究提供新的資料。

引用文獻

- Bennett, Paul A. 1981. The evolution of passive and disposal sentences. *Journal of Chinese Linguistics* 9.1:61-89.
- Chang, Song-Hing (張雙慶). (ed.) 2000. *Lechang Tuhua Yanjiu* 樂昌土話研究 [*A Study on the Lechang Dialect*]. Xiamen: Xiamen University Press.
- Feng, Aizhen (馮愛珍). 1993. *Fuqing Fangyan Yanjiu* 福清方言研究 [*A Study on the Fuqing Dialect*]. Beijing: Social Sciences Academic Press.
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge & New York: Cambridge University Press.
- Hu, Songbo (胡松柏), and Xin Ge (葛新). 2003. Gan dongbei Gan, Wu, Huiyu jieyuan didai fangyan de chuzhi jieci yu beidong jieci 贛東北贛、吳、徽語接緣地帶方言的處置介詞與被動介詞 [The disposal and passive markers in northeastern Jiangxi where Gan, Wu and Hui dialects co-exist]. *Hanyu Fangyan Yufa Yanjiu he Tansuo: Shou Jie Guoji Hanyu Fangyan Yufa Xueshu Yantaohui Lunwenji* 漢語方言語法研究和探索：首屆國際漢語方言語法學術研討會論文集 [*The First International Conference on Chinese Dialect Grammar*], ed. by Zhaoming Dai, 241-245. Harbin: Heilongjiang People's Publishing House.
- Huang, Borong (黃伯榮). (ed.) 1996. *Hanyu Fangyan Yufa Leibian* 漢語方言語法類編 [*Concordance of Chinese Dialect Grammar*]. Qingdao: Qingdao Press.
- Huang, Xuezhen (黃雪貞). 1993. *Jiangyong Fangyan Yanjiu* 江永方言研究 [*A Study on the Jiangyong Dialect*]. Beijing: Social Sciences Academic Press.
- Jiang, Jicheng (蔣冀騁). 2003. Lun Mingdai Wu fangyan de jieci 'zhuo' 論明代吳方言的介詞“捉” [The preposition *zhuo* in Wu dialect of the Ming Dynasty]. *Gu Hanyu Yanjiu* 古漢語研究 [*Research in Ancient Chinese Language*] 2003.3:36-40.

² 在書中，作者把普通話的處置標記“把”列入 *patient*（見 289 頁）。

- Jiang, Shaoyu (蔣紹愚). 2001. *Jindai Hanyu Yanjiu Gaikuang* 近代漢語研究概況 [*An Outline of the Study of Early Modern Chinese*] (2nd edition). Beijing: Peking University Press.
- Jiang, Shaoyu (蔣紹愚). 2002. ‘Gei’ ziju, ‘jiao’ ziju biao beidong de lai yuan: jian tan yufahua, leitui he gongneng kuozhan ‘給’字句、‘教’字句表被動的來源：兼談語法化、類推和功能擴展 [The origin of passive sentence with *gei* and *jiao*: on grammaticalization, analogy and expansion of grammatical function]. *Yufahua yu Yufa Yanjiu* 語法化與語法研究 [*Studies in Grammar and Grammaticalization*], Vol. 1, ed. by Fuxiang Wu & Bo Hong, 202-223. Beijing: The Commercial Press.
- Li, Yongming (李永明). 1988. *Linwu Fangyan: Tuhua yu Guanhua de Bijiao Yanjiu* 臨武方言：土話與官話的比較研究 [*Linwu Dialect: A Comparison between Mandarin and the Local Dialect*]. Changsha: Hunan People's Publishing House.
- Li, Yuming (李宇明), and Qianrui Chen (陳前瑞). 2005. Beijinghua ‘gei’ zi beidongju de diwei ji qi fazhan 北京話“給”字被動句的地位及其發展 [The status and development of the passive *gei* construction in Beijing dialect]. *Fangyan* 方言 [*Dialect*] 2005.4:289-297.
- Liu, Lunxin (劉綸鑫). (ed.) 1999. *Ke Gan Fangyan Bijiao Yanjiu* 客贛方言比較研究 [*A Comparative Research between the Gan and Hakka*]. Beijing: China Social Sciences Press.
- Lu, Xiaoqun (盧小群). 2002. *Jiahe Tuhua Yanjiu* 嘉禾土話研究 [*A Study on the Jiahe Local Dialects*]. Changsha: Central South University Press.
- Lu, Xiaoqun (盧小群). 2004. Xiangnan Tuhua zhong biao shi ‘gei’ de zi 湘南土話中表示“給”的字 [The words meaning ‘to give’ in the local dialects in southern Hunan]. *Xiangnan Tuhua Luncong* 湘南土話論叢 [*Essays on Xiangnan Tuhua*], 247-254. Changsha: Hunan Normal University Press.
- Peyraube, Alain (貝羅貝). 1986. Shuang binyu jiegou cong Handai zhi Tangdai de lishi fazhan 雙賓語結構從漢代至唐代的歷史發展 [Double-object construction: historical development from Han Dynasty to Tang Dynasty]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1986.3:204-216.
- Shi, Yuzhi (石毓智). 2004. Jian biao beidong he chuzhi de ‘gei’ de yufahua 兼表被動和處置的“給”的語法化 [*Gěi*: Its double functions as passive and disposal markers]. *Shijie Hanyu Jiaoxue* 世界漢語教學 [*Chinese Teaching in the World*] 2004.3:15-26.
- Tsuji, Nobuhisa. 1987. *Lexicon of a Chinese Dialect in South Hunan, China*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Wu, Qizhu (吳啟主). 1998. *Changning Fangyan Yanjiu* 常寧方言研究 [*A Study on the Changning Dialect*]. Changsha: Hunan Education Press.

- Wu, Yunji. 1999. An etymological study of disposal and passive markers in Hunan dialects. *Journal of Chinese Linguistics* 27.2:90-123.
- Wu, Yunji. 2005. *A Synchronic and Diachronic Study of the Grammar of the Chinese Xiang Dialects*. Berlin & New York: Mouton de Gruyter.
- Xu, Dan (徐丹). 1992. Beijinghua zhong de yufa biaoji 'gei' 北京話中的語法標記詞“給” [*Gei* in the Beijing dialect]. *Fangyan* 方言 [*Dialect*] 1992.1:54-60.
- Zhang, Xiaoqin (張曉勤). 1998. Ningyuan fangyan de jieci 寧遠方言的介詞 [Coverbs in the Ningyuan dialect]. *Hunan Fangyan de Jieci* 湖南方言的介詞 [*Coverbs in the Hunan Dialects*], ed. by Yunji Wu, 74-97. Changsha: Hunan Normal University Press.
- Zhang, Xiaoqin (張曉勤). 1999. *Ningyuan Pinghua Yanjiu* 寧遠平話研究 [*A Study on the Ningyuan Dialect*]. Changsha: Hunan Education Press.
- Zhang, Xueping (張雪平). 2005. Henan Yexianhua de 'jiao' ziju 河南葉縣話的“叫”字句 [The passive and disposal construction with particle *jiao* in Yexian dialect, Henan Province]. *Fangyan* 方言 [*Dialect*] 2005.4:301-305.

Some Further Thoughts on the Identical Disposal and Passive Markers in Southern Hunan

Yunji Wu

The University of Melbourne

In 1998 I conducted a survey of the disposal and passive markers in 100 localities in Hunan, Wu (1999). The findings showed that in eleven localities in southern Hunan, the disposal and passive markers are identical, both derived from words meaning ‘to give’. A brief analysis of this phenomenon was published in Wu (2005). In that article I discussed from both a semantic and a syntactic point of view why the word meaning ‘to give’ can act as both a disposal and a passive marker. I also discussed how people distinguish meanings such as *Zhangsan ba Lisi da le yidun* [Zhangsan gave Lisi a beating] and *Zhangsan bei Lisi da le yidun* [Zhangsan was given a beating by Lisi] in those dialects. The current paper will discuss the same phenomenon but from different perspectives. I will use the materials found in 14 localities in southern Hunan to propose the following hypothesis: In the history of the Chinese language there have been at least two modes of grammaticalization of disposal and passive markers: one in which the disposal marker and the passive marker are not identical, as in the written language and in most dialects in northern China; the other in which the disposal marker and the passive marker are identical, as in the dialects in southern Hunan. The evolution of the words meaning ‘to give’ into a disposal or a passive marker may have begun in Middle Chinese. Historically there was a stage in which these two modes co-existed and competed, at least in spoken Chinese.

Key words: Hunan dialects, disposal construction, passive construction, *gei* construction

La marque du redoublement graphique en chinois archaïque*

Chrystelle Maréchal (麥里筱)

CNRS-EHESS-CRLAO

Au cours des années 1930, les propos de Guo Moruo sur l'apparition de la marque chinoise du redoublement graphique (二) dans les inscriptions Shang sur os et carapaces ont provoqué de nombreuses réactions de la part d'autres spécialistes. L'objectif de cet article est de résumer les différents points de vue sur le sujet tout en proposant une interprétation inédite de l'emploi initial de ce signe.

Mots-clés: Shang, Zhou, *jiaguwen*, *jinwen*, marque de redoublement

1. Introduction

Dans les plus anciens documents écrits chinois, les inscriptions gravées en majorité sur des os de bovins et des carapaces de tortues (en chinois *jiaguwen* 甲骨文, désormais abrégé JGW) de l'époque de la fin des Shang (XIII^e - XI^e s. av. J.-C.), trois procédés de redoublement graphique sont attestés:

Le premier consiste simplement à écrire deux fois de suite une même graphie comme dans l'exemple suivant:¹

(1) 王占曰其雨佳壬午允雨 (Heji 3. 902 verso)

wang zhan yue qi yu wei ren ren-wu

roi pronostiquer dire FUT pleuvoir être ren ren-wu

yun yu

effectivement pleuvoir

«Le roi fait le pronostic suivant: 'il pleuvra un jour *ren*'. Le jour *ren wu*, il plut effectivement.»

* Une première version plus succincte de cet article est parue en chinois en 2008 dans la revue *Guwenzi yanjiu* 古文字研究 27:171-177. Nous exprimons nos sincères remerciements à Geneviève Barman, Yau Shun-chiu ainsi qu'aux deux relecteurs de cet article pour leurs remarques et commentaires.

¹ Des erreurs graphiques sont toujours possibles comme dans le cas de *xu xu* 戌戌 (*Jiaguwen heji*, désormais abrégé *Heji*, 12. 37 651) où le premier caractère erroné aurait dû être écrit *wu* 戊, de manière à former le binôme *wu-xu* 戊戌, le 35^e jour du calendrier sexagésimal chinois, voir Yao et al. (1988:7).

Le second consiste à faire double emploi d'une graphie comme dans le cas suivant où le contexte permet de rétablir la graphie *jia* 甲 manquante:²

- (2) 乙亥卜又十牢十伐大甲申 (Heji 10. 32 201)
yi-hai bu you shi lao shi fa da-jia (jia) shen
yi-hai divination sacrifier dix bœuf dix victime Da Jia (jia)-shen
 «Divination du jour *yi-hai*, on sacrifie dix bœufs sacrificiels et dix victimes humaines décapitées à l'ancêtre Da Jia le jour *jia-shen*.»

Enfin, le troisième procédé, sujet de cet article, consiste à ajouter deux petits traits placés sur le côté inférieur d'une graphie devant être redoublée comme dans l'exemple suivant:

- (3) 貞卯（劉）二牢王受𠄎 (Heji 10. 31 110)
zhen liu er lao wang shou you (you)
 soumettre sectionner deux bœuf roi recevoir avoir (aide)
 «On soumet à l'oracle: '(Si) on sectionne deux bœufs, le roi recevra de l'aide providentielle.'»

2. États des faits et hypothèse

Luo Zhenyu (1915) est le premier à relever, dans les JGW de l'époque des Shang, l'existence de la graphie 𠄎 qu'il considère être une variante graphique de *you* 𠄎 (又), pictogramme d'une main emprunté dans ce cas pour écrire l'homonyme *you* 𠄎 (祐) «aide providentielle».³ Au premier abord, l'hypothèse que 𠄎 puisse être une variante graphique de 𠄎 est séduisante en raison de l'attestation de *wang shou you* 王受又 𠄎 (王受又, roi-recevoir-aide) «Le roi reçoit de l'aide providentielle» au côté de 王受又 𠄎 (cf. Heji 10. 31 029 et 31 108). L'observation des inscriptions JGW démontre cependant que c'est une erreur. En effet, sachant que *you* 𠄎 a cinq acceptions dans les JGW, à savoir *you* 有 «avoir», *you* 右 «droite», *you* 又 «encore; et», *you* 祐 «aide providentielle»⁴ et *you* 佑 «nom d'un rite», nous devrions nous attendre à pouvoir trouver 𠄎 dans n'importe laquelle de ces cinq acceptions parmi les très nombreuses

² Cet usage concerne essentiellement les graphies servant à désigner les dix troncs célestes et les douze rameaux terrestres du cycle sexagésimal, qui servent à l'appellation des ancêtres chinois, voir Qiu (1980).

³ Pour les différentes explications de 𠄎, on pourra se référer à Yu (1996: t. I, n°905).

⁴ Notons que certains spécialistes interprètent 𠄎 tantôt *you* 祐 «aide providentielle», tantôt *you* 佑 «aide», ce qui est sans réelle conséquence sémantique.

occurrences de *you* 𠄎, si nous étions réellement en présence de variantes. Or il n'en est rien, puisque les occurrences de 𠄎 apparaissent uniquement dans l'environnement immédiat du verbe *shou* 受 «recevoir», donnant le plus souvent les phrases *wang shou you you* 王受𠄎 (roi-recevoir-𠄎) «le roi reçoit de l'aide providentielle», *wo shou you you* 我受𠄎 (nous- recevoir-𠄎) «nous recevons de l'aide providentielle» ou encore *yu shou you you* 余受𠄎 (Je (royal)-recevoir-𠄎) «Je reçois de l'aide providentielle». C'est sans doute ce qui, dans les années 30, a conduit Guo Moruo à innover en défendant l'idée que les deux petits traits sur le côté de 𠄎 étaient la preuve de l'existence de la marque du redoublement dans les JGW.⁵

D'après Guo (1933[1983], 1937), l'apparition de la graphie *you* 𠄎 est attestée au plus tôt dans les inscriptions datées de la III^e période selon la périodisation des inscriptions JGW de Tung Tso-pin (1933).⁶ Guo interprète la graphie 𠄎 comme *you you* 𠄎𠄎 (𠄎𠄎), qu'il lit *you you* 有祐 (avoir - aide providentielle) en s'appuyant sur les attestations de *you you* 𠄎𠄎 dans une inscription gravée sur une carapace de tortue datée de la III^e période (*Heji* 9. 28 011) et de *you you* 𠄎𠄎 (有祐) dans diverses inscriptions datées de la I^{re} période.⁷

Le *Heji*, réunissant l'essentiel des JGW, constitue un corpus suffisamment conséquent pour nous donner une idée de la fréquence des occurrences des différentes

⁵ Si ce point de vue visionnaire est aujourd'hui accepté par le plus grand nombre, il n'en a pas toujours été ainsi. Jusqu'à la fin de sa vie, le célèbre épigraphiste Yu Xingwu (1896-1984) n'a, par exemple, pu se résoudre à cette idée. D'après un de ses disciples, Yu a fait part de ses doutes sur l'existence de la marque de redoublement dans les JGW à l'occasion de cours de formation professorale qu'il donnait en fin de carrière, voir Chen (1989:18).

⁶ Tung a divisé l'ensemble des JGW en cinq périodes, correspondant aux règnes des différents rois Shang, qui s'échelonnent comme suit (les dates indicatives sont celles fournies par *Xia Shang Zhou duandai gongcheng zhuanjia*, 2000:88): I^{re} période: règne du roi Wu Ding 武丁 (1250-1192); II^e période: règnes des rois Zu Geng 祖庚 et Zu Jia 祖甲; III^e période: règnes des rois Lin Xin 廩辛 et Kang Ding 康丁 (ces quatre règnes s'étendent de 1191 à 1148); IV^e période: règnes des rois Wu Yi 武乙 (1147-1113) et Wen Ding 文丁 (1112-1102); V^e période: règnes des rois Di Yi 帝乙 (1101-1076) et Di Xin 帝辛 (1075-1046). Dans le présent article, nous nous satisferons de cette périodisation malgré ses imperfections, voir note 9. Signalons que nous avons relevé, dans les inscriptions datées de la II^e période, une occurrence de 𠄎𠄎 (*Heji* 13. 40 982) et de 𠄎 isolée (*Heji* 8. 22 619). Si la datation de ces inscriptions n'est pas remise en cause, il conviendrait donc de faire remonter l'apparition de la marque de redoublement à la II^e période.

⁷ La structure *shou you you* 𠄎𠄎𠄎 à côté de 𠄎𠄎 est à mettre en parallèle avec celles de *shou you nian* 𠄎𠄎𠄎 (受有年, recevoir - avoir - récolte) et de *shou nian* 𠄎𠄎 (受年 recevoir - récolte) «recevoir une (bonne) récolte» (voir respectivement *Heji* 4. 787 et 801). Pour plus de détails sur cet emploi de *you* 有, voir Djamouri (1992).

possibilités d'écrire «recevoir de l'aide» à l'époque des Shang.⁸ On constate que les syntagmes verbaux les plus anciens, employés initialement à peu près à la même fréquence, n'ont pas connu la même évolution. Tandis que *shou* (X) *you* 𠄎 (X) 𠄎 «recevoir de l'aide (de quelqu'un)» a su s'imposer jusqu'à la V^e période, *shou you you* 𠄎 𠄎 𠄎, qui n'a pas survécu à la I^{re} période, a été supplanté par *shou you you* 𠄎 𠄎 𠄎 comme le montre le tableau suivant:

Tableau 1: fréquence des syntagmes pour «recevoir de l'aide» dans le *Heji*

	I ^{re} période 57,13% du <i>Heji</i>	II ^e période 11,28% du <i>Heji</i>	III ^e période 12,49% du <i>Heji</i>	IV ^e période 8,62% du <i>Heji</i>	V ^e période 10,48% du <i>Heji</i>
𠄎 (X) 𠄎	110		311	49	19
𠄎 𠄎 𠄎	125				
𠄎 𠄎 𠄎			2		
𠄎 𠄎		1	99		23

L'emploi de la marque de redoublement à l'époque des Shang ne semble pas avoir été systématique: si elle est fréquente dans les inscriptions JGW des III^e et V^e périodes, aucune occurrence n'est attestée dans celles de la IV^e période — les scribes de cette époque se contentant du syntagme verbal *shou* (X) *you* 𠄎 (X) 𠄎 (受 (X) 又).⁹ Par

⁸ Par commodité, cet article se réfère aux transcriptions des quelques 41 956 inscriptions du *Heji* données dans Yao et al. (1988) et aux corrections de Bai (2004).

Si l'orientation du pictogramme d'une main est décisive quand il s'agit de noter les mots «droite» 𠄎 (*you* 右) et «gauche» 𠄎 (*zuo* 左), force est de constater qu'il n'y a aucune différence sémantique entre 𠄎 et 𠄎. Ces deux formes ont été prises en compte en notant toutefois une nette préférence pour 𠄎 en termes de fréquence d'emploi, puisque seules 18 occurrences de 𠄎, datant toutes de la III^e période, ont été relevées. Enfin, signalons que les deux petits traits de 𠄎 peuvent être plus ou moins verticaux (cf. 𠄎 *Heji* 9. 29 627, 𠄎 *Heji* 9. 27 079, 𠄎 *Heji* 9. 27 136, 𠄎 *Heji* 9. 27 423; 𠄎 *Heji* 10. 30 388).

⁹ Les problèmes de périodisation dépassent le cadre de cet article, notons toutefois que l'absence d'occurrence de 𠄎 dans les inscriptions de la IV^e période de Tung constitue un argument en faveur de l'hypothèse de Li (1977), confortée par Qiu (1981), considérant préférable de faire remonter leur datation de la fin du règne du roi Wu Ding à celui de Zu Geng. Nous remercions Olivier Venture d'avoir souligné ce point lors d'un exposé de cette étude au séminaire mensuel du Centre de recherches linguistiques sur l'Asie orientale le 26 mars 2008 à l'EHESS, Paris. Si les inscriptions de la IV^e période sont finalement datées de la I^{re} et du début de la II^e périodes de Tung, cela implique, d'une part, que la IV^e période n'aurait plus de raison d'être et, d'autre part, que la marque de redoublement ne saurait être apparue avant le règne de Zu Jia (voir note 6).

ailleurs, à côté de la bonne centaine d'occurrences de la graphie 𠂔 dans les JGW des Shang, nous ne disposons que d'un seul autre cas de graphie redoublée, *fāng* 方 «pays frontalier», dont il sera question à la 6^e section. L'inscription d'un bronze Shang atteste également que les graphies 𠂔 et 𠂔, représentant respectivement un enfant et la lune, ont, elles aussi, pu être dotées de la marque de redoublement comme nous le verrons plus loin.

Pour tenter de cerner les raisons de l'irrégularité de l'emploi de la marque de redoublement à ses débuts, nous n'avons pas d'autre choix que de les déduire des seules données graphiques que les scribes Shang nous ont laissées. La motivation de ceux qui ont créé cette marque était sans doute d'ordre pratique, son emploi permettant d'éviter la réécriture d'une même graphie.¹⁰ Il y a peu de chance pour que les conditions d'emploi de cette marque de redoublement aient été rigoureusement consignées par écrit au moment de son élaboration. Cela étant, on peut supposer que l'usage courant du syntagme *shou you you* 𠂔 𠂔 𠂔 (受有祐) ait pu inciter un scribe à l'abréger pour gagner du temps; la particularité de correspondre à des emplois distincts (en chinois *tong xíng yì yòng* 同形異用) de la graphie *you* 𠂔 lui a sans doute suggéré l'idée de placer sur le côté inférieur un signe diacritique, possible miniaturisation de la graphie *er* 二 «deux», pour indiquer qu'elle devait être lue deux fois. Par la suite, certains scribes ont pu être amenés à penser, consciemment ou non, que le recours à cette marque distinctive ne pouvait se faire que dans le cas d'homographes. C'est en tout cas l'hypothèse que nous postulons.







3. Les deux cas de redoublement dans les inscriptions sur bronzes des Shang




Pour l'époque des Shang, l'emploi de la marque de redoublement sur un bronze n'est attestée que sur un vase: le *Xiaozi Feng you* 小子𠂔𠂔 (Yin Zhou jinwen jicheng, désormais abrégé *Jicheng*, 10. 5417)¹¹ dont l'inscription semble présenter deux graphies dotées d'une marque du redoublement, à savoir 𠂔 et 𠂔.


En complément du *Heji*, relevons l'absence de 𠂔 dans les inscriptions découvertes en 1991 à l'est du village de Huayuan. Les rédacteurs du *Yinxu Huayuanzhuang dongdi jiagu* (2003: vol. 6, 1696) soupçonnent cependant que la graphie *you* 友 (𠂔) dans l'inscription H3: 1042 est l'équivalent de 𠂔 interprété comme *you* 祐 «aide providentielle». Ils considèrent donc que 𠂔 est une simple variante graphique de 𠂔, point de vue dont on a démontré la non recevabilité.

¹⁰ Chen (2006:33) partage, par exemple, cet avis.


¹¹ Ce bronze, de provenance inconnue, est exposé à la Galerie des Beaux-Arts Hakutsuru 白鶴美術館 à Kobe au Japon; son authenticité ne fait aucun doute (communication personnelle de Li Xueqin, août 2004).


L'exemple de  indique que  est à lire deux fois de suite; le contexte permettant de comprendre que  est, dans un premier temps, à interpréter comme le signe cyclique *si* 巳 s'associant à la graphie *yi* 乙 qui le précède pour former le binôme *yi si* 乙巳, le 42^e jour du calendrier sexagésimal, tandis que  doit, dans un second temps, se comprendre comme *zi* 子 «Grand ministre», sujet du verbe qui suit, *ling*  «ordonner». Si tous les spécialistes sont d'accord sur la lecture de ce premier cas de redoublement,¹² il n'en va pas de même pour le second, en l'occurrence .

C'est dans la dernière colonne de l'inscription du bronze que la graphie  apparaît, précédée des graphies  (zai shi 在十) et suivie de  (wei 唯). Il est donc en partie question d'une date: *zai shi yue* 在十月二, toutefois l'interprétation des deux petits traits a suscité au moins trois explications différentes.

Yen Yi-p'ing (1980:162), Ma Chengyuan (1988: t.3, p.3) ou Zhang Yujin (1994: 213) considèrent que les deux petits traits dans  doivent être compris comme la graphie *er* 二 «deux», le tout se lisant *shi yue er* 十月二 «12^e mois».

Hong Jiayi (1988:9-10) et Zhang Yachu (2001:107) lisent simplement *shi yue* 十月 «10^e mois».¹³ Qiu Xigui (1983) fait de même en expliquant que «les deux caractères *shi* 十 et *yue* 月 de l'inscription n'occupent qu'une seule place, ainsi considérer les deux petits traits sur le côté droit comme une 'marque d'association graphique' (*hewen* 合文), semble très logique».¹⁴

La troisième lecture de  est celle de Yu Xingwu (1932[1998:221]), Rong Geng (1941: Vol. I, 82), Chen Mengjia (1956:304), Li Xueqin (1958:69) ou encore des rédacteurs du '*Yin Zhou jinwen jicheng*' *shiwen* (2001:6. 161), qui considèrent les deux petits traits comme une marque de redoublement. Tout récemment, Jao Tsung-i (2007: 4) a également opté pour cette solution. Bien qu'aucun de ces spécialistes ne se soit avancé sur l'explication de cette lecture, nous partageons en partie leur point de vue et nous nous en expliquerons plus loin.

Les lectures «10^e mois» et «12^e mois» permettent de contourner le problème grammatical que pose la graphie *yue* 月 suivie de *wei* 唯. En effet, si le signe diacritique sur le côté droit de la graphie  est vraiment une marque de redoublement, cela engendre la structure inhabituelle «...*yue yue wei*...月月唯...». La graphie *wei* 唯 est une particule grammaticale fréquemment employée, entre autres, pour introduire un syntagme temporel comme dans l'inscription qui nous occupe:

¹² Voir Yu (1932[1998]), Rong (1941), Yang (1959[1997]), Qiu (1983) ou encore Ma (1988).

¹³ Selon Hong, les deux petits traits sous *yue* 月 seraient une marque superflue.

¹⁴ Cette citation est extraite d'une lettre que Qiu Xigui a eu la gentillesse de nous écrire, le 27/08/07, en réponse à un courrier que nous lui avons adressé au cours de la préparation de cet article; nous le remercions vivement de nous avoir répondu de manière si détaillée.

(4) 唯子曰令璽人方

wei zi yue ling wang Ren fang

être ministre dire ordonner inspecter Ren pays

«C'était quand le Grand ministre ordonna d'inspecter le pays de Ren».¹⁵

Ainsi, *yue yue wei* 月月唯... semble vraisemblablement devoir se lire «... *yue yue. Wei ...*», ce qui, à première vue, rend l'inscription difficile à comprendre. C'est sans doute la raison pour laquelle certains évitent le redoublement, tandis que ceux qui optent pour cette dernière solution ponctuent *yue, yue wei* 月, 月唯...

4. La marque de redoublement dans les JGW des Zhou

La découverte de JGW à Zhouyuan a permis de mettre au jour une autre occurrence de 𠄎 apparaissant dans un contexte semblable à celui de l'inscription du *Xiaozi Feng you*. Les archéologues responsables des fouilles de Zhouyuan (Shaanxi, 1979:41) ont, en effet, transcrit l'inscription (H11: 2) comme suit: 自三月至于三月月唯五月𠄎(惟)尙. Ils soulignent le redoublement de la graphie *yue* 月 «mois» en mentionnant l'inscription du *Xiaozi Feng you*, mais n'expliquent pas comment il faut le comprendre. Plus récemment, Peng et al. (2005: Vol. 6, 2005) transcrivent et ponctuent cette inscription comme suit: 自三月至于三月; 月唯五月; 𠄎尙.¹⁶ En l'absence de point et de virgule dans les JGW comme dans les inscriptions sur bronzes, rien n'indique que les deux graphies *yue wei* 月唯 soient liées. En fait, la structure *wei* 唯 suivie d'une date est bien connue. Dès les JGW des Shang, par exemple, on trouve *wei wang er si* 唯王二祀 «à la seconde année royale», tandis que la formule *wei si yue* 唯𠄎月 «au 4^e mois» est attestée dans les JGW de Zhouyuan (H11: 40).¹⁷ Li Xueqin et Wang Yuxin (1980) soupçonnent, eux, que le deuxième *san* 三 «trois» dans 三月至于三月 est une erreur pour *si* 𠄎 «quatre», ce qui n'empêche pas la césure inhabituelle de nos deux phrases.¹⁸

À notre connaissance, ni Qiu Xigui ni Ma Chengyuan ne se sont prononcés sur la lecture de cette inscription des Zhou. Les deux petits traits ne peuvent ici être interprétés ni comme la graphie *er* 二 «deux», *san yue er* 三月二 étant impossible, ni comme la marque d'une association graphique (*hewen* 合文), car *san yue zhi yu san yue* 三月至于三月 «du 3^e mois au 3^e mois» ne fait guère de sens.

¹⁵ Voir Djamouri (1988:308).

¹⁶ Cao (2002:2) opte pour la même transcription, mais il ne ponctue pas.

¹⁷ Voir aussi l'inscription H11: 55 dans Cao (2002).

¹⁸ En ne mentionnant plus son article co-écrit avec Li Xueqin et en insistant sur les articles défendant l'idée du mois intercalaire (voir plus loin), Wang (1984) donne l'impression qu'il soutient à présent cette dernière hypothèse.

Pour Xu Xitai (1979, 1991), le redoublement de la graphie *yue* 月 indique qu'il s'agit d'un mois intercalaire, et qu'il convient de comprendre 三₂ (三月月) comme «3^e mois intercalaire», point de vue que partagent, entre autres, Yen Yi-p'ing (1980:162), Chen Quanfang (1982:343) ou encore Chu Ki-cheung (1997:5).¹⁹ Chen Quanfang, dans un ouvrage collectif, défend désormais l'idée qu'il s'agit du 3^e mois de l'année suivante, sans plus d'explication.²⁰

L'interprétation de la marque de redoublement dans 三₂ comme l'indicateur d'un mois intercalaire est très discutable, car la marque du redoublement indique que seule la graphie qui en est dotée est à redoubler, rien de plus. 三月₂ ne devrait donc pouvoir se lire que 三月月. Lire 三月₂ comme «3^e mois intercalaire» impliquerait que les deux petits traits soient, tout au plus, une marque de lecture spécifique indiquant un mois intercalaire;²¹ par ailleurs, il faut prendre en considération la graphie précédant *yue* 月 pour déterminer de quel mois intercalaire il s'agit. Le traitement *ad hoc* de cette marque ne répondant pas à la définition de la marque de redoublement, elle n'a par conséquent rien à voir avec elle.²²

L'éventail des propositions pour expliquer les deux petits traits sous la graphie *yue* 月 dans l'inscription du bronze de l'époque des Shang comme dans celle des JGW de Zhouyuan met en relief tout l'embarras que suscite son interprétation. Si la graphie *yue* 月 est bien dotée d'une marque de redoublement, les spécialistes ont le plus souvent fourni une interprétation *ad hoc* ne répondant pas à la définition de cette marque. C'est la raison pour laquelle il nous semble nécessaire de traiter ce problème en nous conformant à la définition de la marque de redoublement, plutôt que d'en faire fi afin de rendre lisible, en apparence, une phrase énigmatique.

¹⁹ Soulignons qu'à l'époque des Shang, à l'exception de la notation possible d'un 13^e mois, voire même d'un 14^e mois, l'ajout de mois intercalaire se faisait sans indication particulière. C'est par déduction que les spécialistes sont parvenus à démontrer qu'un mois pouvait être ajouté à n'importe quel moment de l'année sans aucun signallement graphique particulier. Voir Chang (1998).

²⁰ Voir Chen et al. (2003:98). Cao (2000) avait déjà argumenté en faveur de cette hypothèse. Nous remercions Cao Dingyun de nous avoir signalé dans un courrier daté du 8/12/08 son article traitant précisément de ce problème. Toutefois, son explication ne tenant pas compte de la règle d'emploi de la marque de redoublement, nous ne pouvons y adhérer.

²¹ Zhang (1998) opte, lui, pour l'idée d'une «marque de redondance» (*chong yi fu hao* 重益符號) et propose de lire 三月₂ comme «3^e mois intercalaire»; Wei (2004:84) partage son avis. Cette marque de redondance est finalement une marque de lecture revenant, à peu de chose près, à l'explication de Xu Xitai.

²² Dans le même ordre d'idée, Lian (1986:168) a proposé d'interpréter les deux petits traits comme une marque distinctive servant à noter le «jour de la pleine lune» (en l'occurrence du 3^e mois), tandis que Gao (1987[1996:330]) pense qu'elle permet d'indiquer qu'il est question de «trois mois», en l'occurrence, selon lui, des 3^e, 4^e et 5^e mois.

5. Nouvelle interprétation de la graphie *yue* 月 = à la fin des Shang

Dans les JGW, la graphie 𠄎 est attestée tantôt au sens de *yue* 月 «mois», tantôt au sens de *xi* 夕 «crépuscule; nuit». Dans leur préface Yao et al. (1989:7-8) font remarquer que «les graphies 𠄎 et 𠄎 étaient déjà bien distinguées ... mais en raison de différences temporelles et scripturales, les notions que représentent 𠄎 et 𠄎 ont toujours été fluctuantes. Cependant, d'un point de vue purement comparatif, les deux graphies ne sauraient changer sans règle fixe. En fait, quelle que soit la situation, quand 𠄎 sert pour «mois» 月, on peut être assuré que 𠄎 l'est pour «crépuscule» 夕; de même que lorsque 𠄎 est utilisé pour «mois» 月, c'est 𠄎 qui l'est pour «crépuscule» 夕. À travers la comparaison des exemples de JGW, les notions que représentent 𠄎 et 𠄎 peuvent être distinguées.»²³ Ce phénomène se rencontre également sur bronzes où la graphie *yue* 月 normalement écrite 𠄎 est, par exemple, attestée au sens de *xi* 夕 (cf. *Li fang ding* 曆方鼎, in *Jicheng* 5. 2614). Nous émettons donc l'hypothèse que la graphie 𠄎 (𠄎𠄎) du bronze *Xiaozi Feng you* des Shang et de l'inscription JGW de Zhouyuan note les deux homographes *yue* 月 et *xi* 夕, le sens de 𠄎 devant être compris comme *yue xi* 月夕 «fin du mois».²⁴

Dans la littérature transmise, *xi* 夕 au sens de «fin» est bien attesté comme dans le *Shijing* 詩經 «Le livre des Odes», au chant *Junzi yu yi* 君子于役 du livre *Wang feng* 王風, où l'on peut lire:

(5) 日之夕矣，羊牛下來

ri zhi xi yi, yang niu xialai

jour DET fin PF mouton bœuf retourner

«... vers la fin du jour, les brebis et les bœufs reviennent dans leurs étables»

(trad. Couvreur 1896[1992:77])

ou encore dans le *Shangshu dazhuan* 尚書大傳 où l'on peut lire:

²³ Seul le contexte permettra de trancher en cas de doute, voir par exemple *Heji* 12. 35 498.

²⁴ Dans la perspective de notre hypothèse, il conviendrait de repenser la compréhension d'inscriptions JGW telles que *jin ri xi feng* 今日夕風 comprise habituellement comme «cette nuit, il y a du vent».

À titre d'information, relevons la correction de lecture proposée par Bai (2004:109) de l'inscription JGW du *Heji* 5. 12 396, datée de la I^{re} période. Si cette correction s'avérait exacte, nous aurions une occurrence de deux graphies 𠄎 écrites deux fois de suite.

(6) 歲之夕，月之夕

sui zhi xi, yue zhi xi
année DET fin mois DET fin
«fin de l'année, fin du mois».

Sous l'entrée de Cao Shu 曹攄, dans la «Chronique des bons fonctionnaires» du *Jinshu* 晉書, on trouve *sui xi* 歲夕 «fin de l'année (voire le dernier soir de l'année)»; extrait de la phrase:

(7) “歲夕攄行獄” (《晉書·良吏傳·曹攄》) ²⁵

sui xi shu xing yu
année fin Shu aller prison
«À la fin de l'année, Shu se rendit à la prison»

Pour conforter notre lecture de *yue xi* 月夕 «fin du mois», tournons-nous vers le *Xunzi* 荀子 (III^e s. av. J.-C.), au chapitre *Li lun* 禮論 «Des rites», où il est écrit:

(8) 然後月朝卜日，月夕卜宅，然後葬也。

ranhou yue zhao bu ri, yue xi bu zhai,
puis mois début divination jour mois fin divination lieu
ranhou zang ye
puis enterrer PF

«Puis on procède, en début de mois [月朝], aux opérations divinatoires qui permettront de choisir la date des funérailles et en fin de mois [月夕] à celles qui permettront d'en établir le lieu, après quoi l'on y procède.» (trad. Kamenarović, 1987:231).²⁶

Ces occurrences de *xi* 夕 au sens de «fin» sont relativement tardives, cependant on ne peut exclure que cette évolution soit liée à l'emploi de *xi* 夕 dans les JGW.

6. La graphie dans les JGW de *Xiaotun nandi*

Pour l'époque des Shang, Qiu Xigui (1992:147) a relevé la présence d'un cas

²⁵ D'après une note de Zheng Xuan: «du 9^e au 12^e mois, c'est la fin de l'année ... la dernière décade, c'est la fin du mois.», voir *Hanyu da cidian* 漢語大詞典 (1994:3.1146).

²⁶ Dans une note, Yang Liang 楊倞 des Tang (IX^e s.) précise que *yue xi* 月夕 a le sens de «fin du mois», voir Yang (1985:529).

supplémentaire de graphie dotée de la marque de redoublement dans les JGW découverts au sud de Xiaotun près d'Anyang.²⁷ Il s'agit de la graphie *fang* 方 dans l'inscription N° 2 651, datée de la III^e période, 戊辰〔卜〕戊執正斁方=不往 que nous traduisons, selon la lecture de Qiu, comme suit:

(9) 戊辰〔卜〕：戊執正斁方，方不往。

wu chen (bu) shu zhi zheng Gui fang

wu-chen (divination) garde brandir.armes attaquer Gui pays

fang bu wang

Fang NEG aller

«(Divination du jour) *wu-chen*: 'Le garde-frontière brandit les armes et attaque le pays de Gui, les Fang n'iront pas.'»

Yao et al. (1985:106) nous apprennent que «dans les documents connus, Gui 斁 est le nom d'un terrain de chasse... (et) Gui fang 斁方 n'avait encore jamais été attesté. Les Yin n'ont cessé d'annexer les pays voisins pour en faire leur terrain de chasse; ce genre de phénomène est très courant dans les JGW». Si Yao et al. reproduisent les deux petits traits sous la graphie *fang* 方, ils se contentent de transcrire une seule graphie 方. Avant que Qiu Xigui ne rectifie cette erreur de transcription, les spécialistes faisaient abstraction des deux petits traits.²⁸ Il faut dire que le sens de cette inscription est obscur. Qiu (1992) ne dit rien du rapport entre les deux graphies *fang* 方. Si comme Yang (1946, 1954[2006:63]) et Chen (1956[1988:270]) le soutiennent, Fang est le nom d'un pays dans les JGW, il pourrait être raisonnable de considérer que le premier *fang* s'associe avec *gui* 斁 pour former le pays de Gui (斁方), tandis que le deuxième *fang* renvoie à un autre pays, celui de Fang. Dans l'inscription suivante, l'emploi conjoint de *fang* 方 et X *fang* 𠂔方 confirme que, dans une même inscription, *fang* 方 ne renvoie pas forcément à une même entité:

(10) 貞：方不至于𠂔方？ (Heji 4. 8626)

zhen fang bu zhi yu X fang

soumettre Fang NEG atteindre à 𠂔 pays

«on soumet à l'oracle: Fang n'atteindra-t-il pas le pays de 𠂔?»

²⁷ Voir *Xiaotun nandi jiagu* (1983: 2^e vol., I^{re} partie, p.1034). L'auteur remercie Cao Dingyun de lui avoir fait parvenir des photos du recto et du verso de la pièce, gracieusement concédées par l'Institut d'archéologie de Beijing; elles offrent une meilleure appréciation de l'estampage qui conforte l'identification de Qiu.

²⁸ Cf. Zhan (1985:61) ou encore Jung (1989:237).

Si les deux *fang* sont bien différents, on peut donc dire que la graphie *fang* remplit la condition de cette époque pour pouvoir être dotée de la marque de redoublement, à savoir qu'il s'agit de la notation de deux homographes.

7. L'attestation 𠄎 𠄎 (𠄎 𠄎) dans les JGW de la V^e période

Dans les JGW de la V^e période apparaissent à plusieurs reprises deux caractères 𠄎 qui se suivent comme dans:

- (11) 丙戌卜貞武 𠄎 𠄎 其牢茲用 (Heji 12. 35 828)
bing-xu bu zhen Wu Ding ding qi lao
bing-xu divination soumettre Wu Ding sacrifice FUT bœuf
zi yong
 DEM utiliser
 «Divination (du jour) *bing xu*, on soumet à l'oracle: sacrifice *ding* à Wu Ding, on sacrifiera un bœuf. Cette (divination) a été utilisée.»

Comme la plupart des spécialistes chinois, nous considérerons que la seconde graphie 𠄎 rend compte d'un verbe lié à une pratique rituelle.²⁹ Les deux 𠄎 correspondant à une seule et même graphie ayant des emplois différents, on aurait pu s'attendre à voir la seconde remplacée par la marque de redoublement. Il est donc légitime de se demander pourquoi cela n'a pas été le cas?

Depuis les inscriptions de la I^{re} période subsiste l'habitude traditionnelle de redoubler un caractère en le gravant deux fois de suite. Malgré l'apparition de la marque de redoublement, un scribe a toujours eu la possibilité de graver deux fois de suite un même caractère. C'est la raison pour laquelle le cas qui nous occupe ne constitue en rien une exception; il s'agit juste d'un choix scriptural.

8. Conclusion

Il apparaît que l'emploi de la marque de redoublement, initié par des scribes Shang, était également attesté dans les JGW des Zhou. La datation de ces derniers documents n'a pas toujours pu être clairement établie. Ainsi, aucun indice n'a permis aux spécialistes de dater précisément la pièce sur laquelle la graphie 月𠄎 apparaît, en revanche deux des

²⁹ La tendance actuelle est de suivre Yang (1946, 1954[2006:43]) qui identifie le second 𠄎 à *beng* 𠄎. Graphiquement parlant, il semble plus juste d'opter pour l'identification de *ding* 𠄎 en considérant qu'il s'agit d'un emprunt rendant compte d'un nouveau rituel.

trois pièces où apparaissent 𠄎 (H11: 82 et 84) sont datées du roi Wen 文 des Zhou, le père du roi Wu 武.³⁰ Les plus anciennes inscriptions de Zhouyuan sont donc contemporaines de celles de la fin des Shang, tandis que les autres s'inscrivent dans leur continuité.³¹ Une habitude ou un style scriptural est souvent hérité d'une tradition plus ancienne. À titre d'illustration, citons Wang Guowei expliquant que l'écriture de Qin correspond à la période comprise entre les Zhou et les Qin: «l'écriture du *Shi Zhou pian* est celle de Qin, elle correspond à l'écriture des terres de l'Ouest au cours de la période comprise entre les Zhou et les Qin ... les écrits du mur de Confucius... correspondent à l'écriture des pays de l'Est au cours de la période comprise entre les Zhou et les Qin.»³² L'habitude d'écrire du début des Zhou s'inscrit dans le prolongement de la période précédant la victoire des Zhou sur les Shang; les JGW des Shang et des Zhou ont vraisemblablement une origine commune, ce qui n'empêche pas des différences notoires entre elles. Par ailleurs, il est fort probable que les Zhou aient recruté des scribes Shang à leur service, d'où le transfert de certaines habitudes scripturales héritées des Shang dans les JGW des Zhou. En tant que phénomène linguistique, l'écriture doit refléter certaines règles; une fois hors de ses frontières, ou du lieu de son développement, elle peut évoluer différemment et, loin de son lieu d'origine, il arrive qu'elle soit plus conservatrice du point de vue du style.

L'attestation de 〇 〇 dans les JGW Shang de la V^e période et de celle de 𠄎 dans les JGW des Zhou, représentant deux habitudes scripturales différentes, se situe à la veille de la généralisation de la marque de redoublement. Les inscriptions sur bronzes et les inscriptions JGW sont deux choses distinctes tant du point de vue calligraphique que technique, les premières demandant plus de rigueur, car le processus de travail est plus complexe. Les inscriptions sur bronzes devenant de plus en plus longues à partir du début des Zhou, les scribes ont eu plus souvent recours à la marque de redoublement.³³ À l'évidence, les contraintes d'emploi de cette marque sont devenues moins strictes dès le début des Zhou, puisque désormais les graphies redoublées ont le même sens, comme l'attestent, entre autres, l'apparition de la formule 子孫 (qu'il faut lire *zi zi sun sun* 子

³⁰ Aucune date n'a pu être donnée pour la pièce H11: 130 où figure la troisième occurrence de 𠄎. Selon Chen (1988:148) «la datation de ces inscriptions JGW remonte pour les plus anciennes au roi Wen 文 des Zhou et pour les plus récentes au roi Mu 穆 des Zhou; à l'appui de leur contenu et de leur style graphique, on peut conclure *grosso modo* qu'elles s'échelonnent entre la veille et le début des Zhou occidentaux.»

³¹ Rappelons que la graphie 月 𠄎 est attestée sur un bronze de la fin des Shang. Pour les connaissances des Zhou héritées des Shang, voir Liu (1989:515).

³² Voir *Shi Zhou pian shu zheng* 史籀篇疏證 dans Wang Guowei (1983: vol. 6).

³³ L'inscription d'un vase *gui* 簋 découvert en 1983 au Shandong, datant de l'époque des Zhou, prouve qu'un scribe pouvait toujours écrire deux fois de suite une même graphie, mais c'est très rare (cf. *Wenwu* 文物, 2004.8:79-80).

子孫孫) pour exprimer l'abondance de «descendants». Dans l'inscription du *Zeling fang yi* 𠄎令方彝 (*Jicheng* 16. 9901), datée du début des Zhou, les graphies *hou* 侯 et *yue* 月 sont dotées chacune d'une marque de redoublement; le contexte démontre que le sens de chacune de ces graphies est le même. Ainsi, à partir du début des Zhou, il apparaît clairement que la marque de redoublement s'applique à toute graphie sans condition particulière.³⁴ Un assouplissement des conditions d'emploi de cette marque de redoublement était de fait indispensable dans le cadre de sa généralisation.

Références

- Bai, Yulan (白于藍). 2004. 'Yinxu Jiagu Keci Moshi Zongji' Jiaoding 殷墟甲骨刻辭摹釋總集校訂 [*Révisions du Yinxu jiagu keci moshi zongji*]. Fuzhou: Fujian People's Publishing House.
- Cao, Dingyun (曹定雲). 2000. Zhouyuan jiagu 'chong sanyue' bushi 'run sanyue' 周原甲骨“重三月”不是“閏三月” [Le '3^e mois redoublé' dans les os et carapaces de Zhouyuan n'est pas un '3^e mois intercalaire']. *Kaogu* 考古 [*Archaeology*] 2000.8:65-69.
- Cao, Wei (曹瑋). 2002. *Zhouyuan Jiaguwén* 周原甲骨文 [*Les inscriptions sur os et carapaces de Zhouyuan*]. Beijing: World Publishing Corporation.
- Chang, Yuzhi (常玉芝). 1998. *Yin Shang Lifa Yanjiu* 殷商曆法研究 [*Étude du calendrier à l'époque de la fin des Shang*]. Changchun: Jilin Literature and History Press.
- Chen, Chusheng (陳初生). 2006. Tantan heshu, chongwen, zhuanming fuhao wenti 談談合書、重文、專名符號問題 [À propos du problème des marques d'association graphique, de redoublement et de nom propre]. *Kang Le Ji: Zeng Xiantong Jiaoshou Qishi Shouqing Wenji* 康樂集：曾憲通教授七十壽慶文集 [*Recueil Kangle: articles célébrant le soixante-dixième anniversaire du Professeur Zeng Xiantong*], 27-34. Guangzhou: Sun Yat-sen University Press.
- Chen, Mengjia (陳夢家). 1956[1988]. *Yinxu Buci Zongshu* 殷虛卜辭綜述 [*Présentation générale des inscriptions oraculaires de Yinxu*]. Beijing: Zhonghua Book Company.
- Chen, Quanfang (陳全方). 1982. Shaanxi Qishan Fengchucun Xi Zhou jiaguwén gailun 陝西岐山鳳雛村西周甲骨文概論 [Introduction aux inscriptions sur os et carapaces des Zhou occidentaux du village de Fengchu dans le district de Qishan au Shaanxi]. *Guwenzi Yanjiu Lunwenji* 古文字研究論文集 [*Recueil d'articles concernant les recherches sur l'écriture archaïque*], 305-424. Chengdu: Sichuan People's Publishing House.

³⁴ Cette marque a également évolué vers d'autres fonctions qui feront l'objet d'une prochaine recherche.

- Chen, Quanfang (陳全方). 1988. *Zhouyuan yu Zhou Wenhua* 周原與周文化 [*Zhouyuan et la culture des Zhou*]. Shanghai: Shanghai People's Publishing House.
- Chen, Quanfang (陳全方), Zhiyi Hou (侯志義), et Min Chen (陳敏). 2003. *Xi Zhou Jiawen Zhu* 西周甲文注 [*Notes sur les inscriptions sur os et carapaces des Zhou occidentaux*]. Shanghai: Xuelin Press.
- Chen, Shihui (陳世輝). 1989. Huainian Yu Xingwu xiansheng 懷念于省吾先生 [À la mémoire de M. Yu Xingwu]. *Guwenzi Yanjiu* 古文字研究 16:17-20. Beijing: Zhonghua Book Company.
- Chu, Ki-cheung (Zhu, Qixiang 朱歧祥). 1997. *Zhouyuan Jiagu Yanjiu* 周原甲骨研究 [*Étude des os et carapaces de Zhouyuan*]. Taipei: Student Book.
- Couvreur, Séraphin. 1896[1992]. *Cheu King*. Taipei: Institut Ricci-Kuangchi Press.
- Djamouri, Redouane. 1988. *Étude des formes syntaxiques dans les écrits oraculaires gravés sur os et écaïlle de tortue*. Paris: Thèse de doctorat de l'EHESS.
- Djamouri, Redouane. 1992. Un emploi particulier de YOU en chinois archaïque. *Cahiers de Linguistique – Asie Orientale* 21.2:231-289.
- Gao, Ming (高明). 1987[1996]. *Zhongguo Guwenzixue Tonglun* 中國古文字學通論 [*Introduction à l'écriture archaïque chinoise*]. Beijing: Peking University Press.
- Guo, Moruo (郭沫若). 1933[1983]. *Buci Tongzuan* 卜辭通纂 [*Compilation d'inscriptions oraculaires*]. Beijing: Science Press.
- Guo, Moruo (郭沫若). 1937[1965]. *Yinqi Cuibian* 殷契粹編 [*Collection d'inscriptions gravées sur os et carapaces des Yin*]. Beijing: Science Press.
- Guo, Moruo (郭沫若) et al. (eds.) 1982. *Jiaguwen Heji* 甲骨文合集 [*Compilation d'inscriptions sur os et carapaces*], 13 vols. Beijing: Zhonghua Book Company.
- Hong, Jiayi (洪家義). 1988. *Jinwen Xuan Zhushi* 金文選注釋 [*Explications annotées d'un choix d'inscriptions sur bronzes*]. Nanjing: Jiangsu Education Press.
- Jao, Tsung-i (Rao, Zongyi 饒宗頤). 2007. Yin Zhou jinwen buci suo jian Yifang Xibei dili kao 殷周金文卜辭所見夷方西北地理考 [Étude géographique du pays Yi dans le nord-ouest fondée sur les inscriptions sur bronzes des Yin et des Zhou]. *Journal of Chinese Studies* 47:1-11.
- Jung, Bor-sheng (Zhong, Bosheng 鍾柏生). 1989. *Yin Shang Buci Dili Luncong* 殷商卜辭地理論叢 [*Essais sur la géographie dans les inscriptions oraculaires de la fin des Shang*]. Taipei: Yee Wen Publishing Company.
- Kamenarović, Ivan P. 1987. *Xun Zi* (Siun Tseu). Paris: Éditions du Cerf.
- Li, Xueqin (李學勤). 1958. Di Yi shidai de fei wang buci 帝乙時代的非王卜辭 [Les inscriptions oraculaires non royales de l'époque de Di Yi]. *Kaogu Xuebao* 考古學報 [*Acta Archaeologica Sinica*] 1958.1:43-74.
- Li, Xueqin (李學勤). 1977. Lun 'Fu Hao' mu de niandai ji youguan wenti 論婦好墓的年代及有關問題 [À propos de la datation de la tombe de Fuhao et de quelques problèmes s'y rapportant]. *Wenwu* 文物 [*Cultural Relics*] 1977.11:32-37.

- Li, Xueqin (李學勤), et Yuxin Wang (王宇信). 1980. Zhouyuan buci xuanshi 周原卜辭選釋 [Sélection d'inscriptions oraculaires de Zhouyuan]. *Guwenzi Yanjiu* 古文字研究 4:245-257. Beijing: Zhonghua Book Company.
- Lian, Shaoming (連劭名). 1986. Du Zhouyuan chutu de jiagu keci 讀周原出土的甲骨刻辭 [Lectures d'inscriptions gravées sur os et carapaces excavées à Zhouyuan]. *Guwenzi Yanjiu* 古文字研究 13:161-173. Beijing: Zhonghua Book Company.
- Liu, Yu (劉雨). 1989. Xi Zhou jinwen zhong de jizu li 西周金文中的祭祖禮 [Les rites sacrificiels aux ancêtres dans les inscriptions sur bronzes des Zhou occidentaux]. *Kaogu Xuebao* 考古學報 [*Acta Archaeologica Sinica*] 1989.4:495-522.
- Luo, Zhenyu (羅振玉). 1915. *Yinxu Shuqi Kaoshi* 殷虛書契考釋 [*Études d'inscriptions gravées provenant des ruines de Yin*]. Réédité dans *Yinxu Shuqi Kaoshi San Zhong* 殷虛書契考釋三種 [*Trois études sur les inscriptions gravées provenant des ruines de Yin*], 2 vols. Beijing: Zhonghua Book Company, 2006.
- Ma, Chengyuan (馬承源). 1988. *Shang Zhou Qingtongqi Mingwen Xuan* 商周青銅器銘文選 [Sélection d'inscriptions sur bronzes des Shang et des Zhou], 3 vols. Beijing: Cultural Relics Press.
- Peng, Bangjiong (彭邦炯), Ji Xie (謝濟), et Jifan Ma (馬季凡). 2005. 'Jiaguwen Heji' *Bubian* 《甲骨文合集》補編 [*Compilation supplémentaire au Jiaguwen heji*], Vol. 6. Beijing: Language and Culture Press.
- Qiu, Xigui (裘錫圭). 1980. Jiagu wenzi kaoshi 甲骨文字考釋 [Études de graphies sur os et carapaces]. *Guwenzi Yanjiu* 古文字研究 4:153-175. Beijing: Zhonghua Book Company.
- Qiu, Xigui (裘錫圭). 1981. Lun 'Lizu buci' de shidai 論“歷組卜辭”的時代 [À propos de l'époque des 'inscriptions oraculaires du groupe Li']. *Guwenzi Yanjiu* 古文字研究 6:263-321. Beijing: Zhonghua Book Company.
- Qiu, Xigui (裘錫圭). 1983. Guanyu Shangdai de zongzu zuzhi yu guizu he pingmin liangge jieji de chubu yanjiu 關於商代的宗族組織與貴族和平民兩個階級的初步研究 [Recherche préliminaire sur l'organisation clanique et les deux classes sociales aristocratique et populaire de l'époque des Shang]. *Wenshi* 文史 17:14-17. Repris dans *Gudai Wenshi Yanjiu Xintan* 古代文史研究新探 [*Nouvelles explorations sur l'histoire et la civilisation antiques*], 296-342. Nanjing: Jiangsu Ancient Book Publishing House, 1992.
- Qiu, Xigui (裘錫圭). 1992. Zai tan jiaguwen zhong chongwen de shenglue 再談甲骨文中重文的省略 [Reparlons de l'abréviation de la marque de redoublement dans les inscriptions sur os et carapaces]. *Guwenzi Lunji* 古文字論集 [*Recueil d'articles sur l'écriture archaïque*], 147-150. Beijing: Zhonghua Book Company.
- Rong, Geng (容庚). 1941. *Shang Zhou Yiqi Tongkao* 商周彝器通考 [*The Bronzes of Shang and Zhou*], 2 vols. Peiping: Harvard-Yenching Institute.

- Shaanxi Zhouyuan Kaogudui (陝西周原考古隊). 1979. Shaanxi Qishan Fengchucun faxian Zhouchu jiaguwen 陝西岐山鳳雛村發現周初甲骨文 [Inscriptions sur os et carapaces du début des Zhou découvertes dans le village de Fengchu à Qishan au Shaanxi]. *Wenwu* 文物 [Cultural Relics] 1979.10:38-43.
- Tung, Tso-pin (Dong, Zuobin 董作賓). 1933. *Jiaguwen Duandai Yanjiuli* 甲骨文斷代研究例 [Exemples d'étude de datation des inscriptions sur os et carapaces]. Réimprimé dans *Tung Tso-pin Xiansheng Quanjì* 董作賓先生全集 [Œuvres complètes de M. Tung Tso-pin], Vol. 2, 323-424. Taipei: Yee Wen Publishing Company, 1977.
- Wang, Guowei (王國維). 1983. *Wang Guowei Yishu* 王國維遺書 [Œuvres posthumes de Wang Guowei], 16 vols. Shanghai: Shanghai Classics Bookstore.
- Wang, Yuxin (王宇信). 1984. *Xi Zhou Jiagu Tanlun* 西周甲骨探論 [Sur les os et carapaces des Zhou occidentaux]. Beijing: China Social Sciences Press.
- Wei, Tz'u-te (Wei, Cide 魏慈德). 2004. Du 'Zhouyuan Jiaguwen' 讀《周原甲骨文》 [Lecture du Zhouyuan jiaguwen]. 2004 Nian Anyang Yin Shang Wenming Guoji Xueshu Yantaohui Lunwenji 2004 年安陽殷商文明國際學術研討會論文集 [Recueil d'articles du colloque international sur la civilisation de la fin des Shang de 2004 à Anyang], ed. by Yuxin Wang (王宇信), Zhenhao Song (宋鎮豪) et Xianwu Meng (孟憲武), 83-86. Beijing: Social Sciences Academic Press.
- Xia Shang Zhou Duandai Gongcheng Zhuanjia (夏商周斷代工程專家). 2000. *Xia Shang Zhou Duandai Gongcheng 1996-2000 Nian Jieduan Chengguo Baogao* 夏商周斷代工程 1996-2000 年階段成果報告 [Rapport des résultats des années 1996-2000 du groupe de travail de datation des époques Xia, Shang et Zhou]. Beijing: World Publishing Corporation.
- Xu, Xitai (徐錫台). 1979. Tanta Zhouyuan jiaguwen zhong youguan Zhouchu de lifa wenti 探討周原甲骨文中有關周初的曆法問題 [Discussion approfondie des problèmes calendaires se rapportant au début des Zhou dans les inscriptions sur os et carapaces de Zhouyuan]. *Guwenzi Yanjiu* 古文字研究 1:203-207. Beijing: Zhonghua Book Company.
- Xu, Xitai (徐錫台). 1991. *Zhouyuan Jiaguwen Zongshu* 周原甲骨文綜述 [A Summary of Inscriptions on Bones of Zhouyuan]. Xi'an: San Qin Press.
- Yang, Liuqiao (楊柳橋). 1985. *Xunzi Gushi* 荀子詁釋 [Commentaires du Xunzi]. Ji'nan: Qilu Press.
- Yang, Shuda (楊樹達). 1954[2006]. *Jiweiju Jiawen Shuo* 積微居甲文說 [À propos des graphies sur carapaces du Studio Jiwei]. Shanghai: Shanghai Ancient Book Publishing House.
- Yang, Shuda (楊樹達). 1959[1997]. *Jiweiju Jinwen Shuo* 積微居金文說 [À propos des graphies sur bronzes du Studio Jiwei]. Beijing: Zhonghua Book Company.

- Yao, Xiaosui (姚孝遂), et Ding Xiao (肖丁). 1985. *Xiaotun Nandi Jiagu Kaoshi* 小屯南地甲骨考釋 [*Étude des os et carapaces du sud de Xiaotun*]. Beijing: Zhonghua Book Company.
- Yao, Xiaosui (姚孝遂), et Ding Xiao (肖丁). 1988. *Yinxu Jiagu Keci Moshi Zongji* 殷墟甲骨刻辭摹釋總集 [*Recueil de copies manuscrites et de transcriptions des inscriptions gravées sur os et carapaces des ruines de Yin*]. Beijing: Zhonghua Book Company.
- Yao, Xiaosui (姚孝遂), et Ding Xiao (肖丁). 1989. *Yinxu Jiagu Keci Leizuan* 殷墟甲骨刻辭類纂 [*Compilation des inscriptions gravées sur os et carapaces des ruines de Yin*]. Beijing: Zhonghua Book Company.
- Yen, Yi-p'ing (Yan, Yiping 嚴一萍). 1980. Zhouyuan jiagu 周原甲骨 [Les os et carapaces de Zhouyuan]. *Zhongguo Wenzi* 中國文字 New Series 1:159-165.
- Yu, Xingwu (于省吾). 1932[1998]. *Shuangjianchi Jijin Wenxuan* 雙劍謠吉金文選 [*Sélection d'inscriptions sur bronzes du Studio Shuangjian*]. Beijing: Zhonghua Book Company.
- Yu, Xingwu (于省吾). 1996. *Jiagu Wenzi Gulin* 甲骨文字詁林 [*La forêt des commentaires des graphies sur os et carapaces*], 4 vols. Beijing: Zhonghua Book Company.
- Zhan, Qinxin (詹勤鑫). 1985. Du 'Xiaotun Nandi Jiagu' zhaji 讀《小屯南地甲骨》札記 [Notes de lecture du *Xiaotun nandi*]. *Kaogu yu Wenwu* 考古與文物 [*Archaeology and Cultural Relics*] 1985.6:61-65.
- Zhang, Guiguang (張桂光). 1998. Jia jin wen zhong de chongyi fuhao yu Shang Zhou de runyue wenti 甲金文中的重益符號與商周的閏月問題 [La marque de redondance dans les inscriptions sur carapaces et bronzes et le problème du mois intercalaire à l'époque des Shang et des Zhou]. *Rong Geng Xiansheng Bainian Danchen Jinian Wenji* (Guwenzi Yanjiu Zhuanhao) 容庚先生百年誕辰紀念文集 (古文字研究專號) [*Recueil d'articles en hommage à Rong Geng célébrant le centenaire de sa naissance*] (numéro spécial de la revue *Guwenzi Yanjiu*), 252-257. Guangzhou: Guangdong People's Publishing House.
- Zhang, Yachu (張亞初). 2001. 'Yin Zhou Jinwen Jicheng' Yinde 《殷周金文集成》引得 [*Index au Yin Zhou jinwen jicheng*]. Beijing: Zhonghua Book Company.
- Zhang, Yujin (張玉金). 1994. *Jiaguwen Xuci Cidian* 甲骨文虛詞詞典 [*Dictionnaire des mots vides dans les inscriptions sur os et carapaces*]. Beijing: Zhonghua Book Company.
- Zhongguo Shehui Kexueyuan Kaogu Yanjiusuo (中國社會科學院考古研究所). 1980-1983. *Xiaotun Nandi Jiagu* 小屯南地甲骨 [Les os et carapaces retrouvés au sud de Xiaotun]. Beijing: Zhonghua Book Company.
- Zhongguo Shehui Kexueyuan Kaogu Yanjiusuo (中國社會科學院考古研究所). 1984-1994. *Yin Zhou Jinwen Jicheng* 殷周金文集成 [*Compilation d'inscriptions sur*

- bronzes des Yin et des Zhou*], 18 vols. Beijing: Zhonghua Book Company.
- Zhongguo Shehui Kexueyuan Kaogu Yanjiusuo (中國社會科學院考古研究所). 2001. 'Yin Zhou Jinwen Jicheng' Shiwen 《殷周金文集成》釋文 [*Transcriptions des inscriptions du Yin Zhou jinwen jicheng*], 6 vols. Hong Kong: Institute of Chinese Studies, The Chinese University of Hong Kong.
- Zhongguo Shehui Kexueyuan Kaogu Yanjiusuo (中國社會科學院考古研究所). 2003. *Yinxu Huayuanzhuang Dongdi Jiagu* 殷墟花園莊東地甲骨 [*Les os et carapaces découverts à l'est du village de Huayuan sur les ruines de Yin*], 6 vols. Kunming: Yunnan People's Publishing House.

The Reduplication Mark in Archaic Chinese

Chrystelle Maréchal

CNRS-EHESS-CRLAO

Guo Moruo's remarks in the nineteen thirties on the emergence of the Chinese graphical sign of reduplication (二) used in oracle-bone inscriptions of the Shang period, have since then provoked reactions from other scholars. The aim of this paper is to sum up various points of view on the subject, and at the same time to provide a new interpretation of its initial usage.

Key words: Shang-Zhou periods, oracle-bones and bronze inscriptions, reduplication mark

Le développement des livres de rimes en dictionnaires: *Qièyùn* et *Kānmiù bǔquē Qièyùn*

Françoise Bottéro (蒲芳莎)

CNRS-EHESS-INALCO

C'est dans la perspective de l'histoire des dictionnaires chinois que j'aborderai ici l'étude du *Qièyùn* 切韻 et du *Kānmiù bǔquē Qièyùn* 刊謬補缺切韻. En tant qu'ouvrage fondateur de la tradition des livres de rimes, le *Qièyùn* occupe avant tout une place fondamentale dans l'histoire de la phonologie chinoise, mais il m'a paru intéressant d'étudier le rôle qu'il a pu jouer dans l'histoire des dictionnaires, étant donné la nouvelle méthode de classement qu'il offrait pour organiser les caractères. J'ai donc tout d'abord cherché à dégager les particularités de ce texte fondateur en partant des plus anciens fragments connus du *Qièyùn* original et en les comparant aux versions ultérieures. J'ai ensuite étudié le système des gloses propres aux *Qièyùn* original et celles des versions augmentées jusqu'au *Kānmiù bǔquē Qièyùn*, afin de mieux comprendre comment ont évolué les livres de rimes à cette époque et comment ils en sont venus à ressembler de plus en plus à de véritables dictionnaires.

Mots clé: rime, classement des caractères, tons, variantes graphiques

1. Le *Qièyùn*

1.1 Présentation générale du *Qièyùn* d'après la préface de Lù Fǎyán

La version originale du *Qièyùn* est aujourd'hui perdue, mais la préface de Lù Fǎyán 陸法言 a été reproduite dans plusieurs éditions ultérieures.¹ Cet ouvrage en cinq *juǎn* 卷, achevé en 601, est un «livre de rimes» élaboré par un groupe d'au moins huit lettrés et rédigé par le plus jeune, Lù Fǎyán (581?-618?). La plupart de ces lettrés,² y compris le père de Lù Fǎyán qui n'est pas nommé dans la préface,³ avait appartenu à

¹ *Guǎngyùn, Kānmiù bǔquē Qièyùn*, manuscrits de Dūnhuáng S. 2055, P. 2017, P. 2129, etc.

² À savoir Liú Zhēn 劉臻 (527-598), Yán Zhītuī 顏之推 (531-591), Lú Sīdào 盧思道 (531-582 ou 536-586), Wèi Yànyuān 魏彥淵, Lǐ Ruò 李若, Xiāo Gāi 蕭該 (ca. 535 - ca. 610), Xīn Déyuán 辛德源, Xuē Dàohéng 薛道衡 (540-609) (d'après la préface).

³ Lorsque l'empereur Wéndì 文帝 fut destitué en 600, il attribua son malheur au père de Lù Fǎyán, Lù Shuǎng 陸爽 (539-591), qui avait été son conseiller. Lù Shuǎng étant mort neuf ans plus tôt, l'empereur se vengea sur sa famille qu'il fit bannir de la fonction publique. Lù Fǎyán

l'Académie des lettres des Qí septentrionaux⁴ et étaient originaires du Nord. Tandis que Yán Zhītuī 顏之推 (531-591) et Xiāo Gāi 蕭該, à qui l'on doit, selon Lù Fǎyán, les principaux choix de prononciations, venaient du Sud. Dans la préface, Lù Fǎyán présente cinq livres de rimes antérieurs⁵ qui leur ont servi de base, et donne un aperçu des discussions qui ont eu lieu entre eux, vingt ans auparavant,⁶ dans sa demeure, sur les divergences entre prononciations du Nord et du Sud, anciennes et modernes, sur les différences de tons, de voisement d'une région à l'autre, etc. Ils décidèrent d'y remédier en établissant une norme des prononciations. C'est ainsi qu'ils ont réalisé un compromis entre prononciations littéraires du Nord et du Sud.⁷ De fait, le *Qièyùn* ne se voulait pas comme un dictionnaire, mais comme un ouvrage de référence pour la composition littéraire, classant les caractères selon les quatre tons, les 193 rimes, et les initiales, tout en indiquant leur prononciation par la méthode *fǎnqiè* 反切.

1.2 Les manuscrits de Dūnhuáng 敦煌

Pour expliquer la structure du *Qièyùn* perdu, on s'est tout d'abord appuyé sur le *Guǎngyùn* 廣韻 de Chén Péngnián 陳彭年 (1008), puis, à partir de sa découverte en 1947 au palais impérial de Pékin, sur le *Kānmiù bǔquē Qièyùn* 刊謬補缺切韻 de Wáng Rénxǔ 王仁珣 (706). Or, comme on va le voir, malgré une structure similaire, comprenant tons, rimes⁸ et initiales pour organiser les caractères, ces textes étaient sensiblement différents.

fut alors révoqué et il lui fut impossible de mentionner dans la préface le nom de son père comme ayant participé aux réunions à la base de la création du *Qièyùn*. Cf. Wang Lien-tseng (1957:55).

⁴ Le *Wénlín guǎn* 文林館 qui fonctionna de 573 à 577 sous le règne de Gāo Wēi 高緯.

⁵ Le *Yùnjí* 韻集 de Lǚ Jìng 呂靜, le *Yùnlüè* 韻略 de Xiàhóu Yǒng 夏侯詠, le *Yùnlüè* 韻略 de Yáng Xiūzhī 陽休之, le *Yīnpǔ* 音譜 de Lǐ Jìjié 李季節, le *Yùnlüè* 韻略 de Dù Táiqīng 杜台卿, tous aujourd'hui perdus. Comme on peut le voir à plusieurs reprises dans le tableau des rimes introduisant chaque chapitre du *Kānmiù bǔquē Qièyùn*, et notamment dans le premier chapitre sous les rimes n° 2 *dōng* 冬, 6 *zhī* 脂, 14 *jiē* 皆, 15 *huī* 灰, et 25 *shān* 刪, ces auteurs n'étaient pas toujours d'accord entre eux sur les distinctions à opérer entre les rimes. Lorsqu'il s'appuie sur un de ces auteurs pour distinguer ou regrouper des rimes, Lù le signale (voir le texte reproduit par Zhōu Zǔmó 1983:435 et 436).

⁶ Au début de l'ère Kāihuáng 開皇 (581-600) des Suí (589-618).

⁷ Fēng Yǎn 封演 des Táng note dans son *Fēng shì wénjiàn jì* 封氏聞見記, chap. *Shēngyùn* 聲韻: «À l'époque des Suí, Lù Fǎyán, Yán (Zhītuī), Wèi (Yànyuān) et les autres lettrés ont fixé les prononciations du Nord et du Sud et composé le *Qièyùn*. Avec ses 12 158 caractères, c'était un modèle pour la littérature. Comme il distinguait les rimes *xiān* 先 et *xiān* 仙; *shān* 刪 et *shān* ㄴ, tous (les lettrés) souffraient de sa rigoureuse exactitude 共苦其苛細».

⁸ Même si le nombre de rimes variait: 193 dans le *Qièyùn*, 195 dans le *Kānmiù bǔquē Qièyùn* (*idem* dans le *Tángyùn* 唐韻 de Sūn Miǎn 孫愐 (2^e version datée de 751)) et 206 dans le

Plusieurs manuscrits du *Qièyùn* ont été découverts à Dūnhuáng 敦煌, au début du XX^e siècle.⁹ Ces fragments provenant des collections Stein et Pelliot, nous permettent de nous faire une idée beaucoup plus précise de la structure originale du *Qièyùn* que ne le font le *Kānmìu bǔquē Qièyùn* ou le *Guǎngyùn*. Dans son ouvrage *Táng wūdài yùnnshū jí cún* 唐五代韻書集存 (1983), Zhōu Zǔmó 周祖謨 a étudié l'ensemble de ces textes retrouvés à Dūnhuáng ou provenant du palais impérial de Pékin et réuni six manuscrits qu'il considérait correspondre à des fragments du *Qièyùn* original. Il s'agit des six fragments P. 3798 (R° & V°), de P. 3695/3696 (R° & V°), de S. 6187 et de S. 2683/P. 4917. Mais du fait que P. 3695 et P. 3696 appartiennent au même manuscrit, tout comme S. 2683 et P. 4917, on ne possède en réalité que quatre manuscrits. Zhōu Zǔmó s'est appuyé sur le fait que ces versions regroupaient moins de caractères que les autres et qu'elles ne comprenaient ni «observations» *àn* 案 (按) concernant le texte de Lù Fǎyán, ni «caractères ajoutés» *jiā* 加. L'étude de ces manuscrits originaux m'a permis de repérer certaines particularités propres au *Qièyùn* de Lù Fǎyán, et de mieux comprendre la composition originale de ce texte aujourd'hui perdu.

1.3 Les caractéristiques propres au *Qièyùn* original

Les fragments représentatifs, selon Zhōu Zǔmó, le *Qièyùn* original possèdent un certain nombre de points communs qui les distinguent du *Guǎngyùn* et du *Kānmìu bǔquē Qièyùn*.

1° La première chose qu'il faut noter, c'est que le *Qièyùn* original ne comprenait pas de points ou cercles utilisés pour séparer les différents groupes d'homophones sous chaque rime.¹⁰ C'est le cas de P. 3695/P. 3696 et S. 6187,¹¹ mais aussi d'autres versions annotées postérieures telles S. 2071, S. 2055 ou encore P. 2017. Il me semble donc que les manuscrits S. 2683/P. 4917 et P. 3798, considérés par Zhōu Zǔmó comme des copies du *Qièyùn* original, sont en réalité des copies un peu plus tardives où ces points noirs ou rouges ont été ajoutés.¹²

Guǎngyùn 廣韻 de Chén Péngnián 陳彭年, le système phonologique restait le même, cf. Norman (1988:25) et Baxter (1992:38-39).

⁹ Plusieurs fragments ont également été retrouvés à Turfan, au début de XX^e siècle, mais je n'ai pas encore vérifié les textes originaux conservés à la Staatsbibliothek zu Berlin et à la Berlin-Brandenburgische Akademie.

¹⁰ Norman (1988:27) parle de *buttons* 'boutons' pour *niǔ* 紐, terme qui a été repris par Baxter (1992:822, note 58), alors que la tradition chinoise attribue à ce caractère le sens d'initiale.

¹¹ A l'exception de S. 6187, dans lequel on observe trois points rouges précédant numéros et rimes dans la marge, ces textes ne comprenaient pas non plus de couleurs.

¹² Dans P. 4917, la petite rime *zhàng* 丈 sous la rime *yǎng* 養 (chap. 3) regroupe le caractère *zhàng* 杖 qui est aussitôt suivi de 昶 *chǎng*, mais un point rouge 〇 entre eux permet de

2° Les rimes étaient exposées de manière bien visible, soit dans la marge supérieure (P. 3695/3696 et de S. 6187), soit dans une colonne à part ou au deux tiers laissée en blanc (P. 3798).

3° Les gloses précédaient toujours le *fānqiè*, suivi du nombre total de caractères homophones.

4° Il y avait très peu de gloses sémantiques. En effet, rimes et caractères courants n'étaient, en général, pas glosés. Ainsi, par exemple, dans P. 3798, la rime *dōng* 冬, les entrées *kōng* 空, *lóng* 籠, mais aussi *lái* 來, *wú* 吾, *āi* 哀 et *shī* 詩 de P. 3696 ne comprennent que des gloses phonétiques sans aucune explication sémantique.¹³ Cette particularité se retrouve dans S. 6187 (*xún* 尋徐林反, *chén* 沉除深反), P. 4917 (*yǎng* 養餘兩反), et S. 2683 (*shèn* 腎時忍反; *cǔn* 忖倉本反). Le *Qièyùn* n'était donc pas du tout conçu comme un dictionnaire, puisqu'il ne fournissait même pas de glose sémantique pour tous ces caractères.

5° Les caractères pouvaient s'enchaîner sans glose. Dans P. 3695, sous la rime *hāi* 哈 dans le 1^{er} chapitre, on trouve sous 'la petite rime' *cái* 裁 les quatre caractères homophones 財, 才, 材 et 赅, rangés sans glose, les uns à la suite des autres, de la façon suivante: 裁昨來反六 纔僅或作裁 財才材赅. Mais on trouve aussi beaucoup d'exemples de caractères qui se suivent sans glose et sans point séparant les groupes d'homophones, alors qu'ils n'ont pas la même initiale. Dans le même manuscrit P. 3695 sous la même rime *hāi* 哈, les deux caractères 哉 et 猜 se suivent alors que le premier se prononce *zāi* 哉 et le second *cāi* 猜.¹⁴ Sachant que de nombreux manuscrits comprennent des fautes susceptibles d'embrouiller les lecteurs peu avertis,¹⁵ on comprend alors l'importance de ces points qui permettaient de séparer clairement les groupes de caractères aux initiales différentes. Leur utilité semble avoir été ressentie assez vite pour que les auteurs des manuscrits S. 2693/P. 4917 et P. 3798, considérés par Zhōu Zǔmó comme des copies du *Qièyùn* original, éprouvent le besoin de les ajouter.

visualiser une frontière plus nette entre ces deux caractères aux prononciations différentes: 丈直兩反二 杖○和通丑兩反.

¹³ Ces caractères sont glosés comme suit: *dōng* 冬都宗反二; *kōng* 空苦紅反; *lóng* 籠慮紅反; *lái* 來落哀反六; *wú* 吾五胡反十; *āi* 哀烏開反三; *shī* 詩書之反.

¹⁴ *Zāi* 哉 est rangé sous 'la petite rime' *zāi* 灾祖才反三 et *cāi* 猜, correspondant à une 'petite rime', est glosée de la façon suivante: 疑倉才反二. Le cas des caractères qui se suivent sans glose se reproduit dans tous les *Qièyùn* originaux: P. 3798 (chap. 1, rime *zhōng* 鍾): *yōng* 庸 et 墉, puis 鏞 et 鄘; S. 6187 (chap. 2 rime *qīn* 侵): *jīn* 今 et *jīn* 齡; S. 2693/P. 4917 (ch. 3, rime *chǎn* 產): *jiǎn* 柬 et *jiǎn* 揀; etc.

¹⁵ Il y a, en effet, quelques exemples où une erreur de comptage peut facilement entraîner une erreur de prononciation. Ainsi, dans le cas de P. 3695, on peut lire sous la rime *hāi* 哈 que *gāi* 該 regroupe neuf homophones (該:古哀反九). Or dans le texte, il n'y en a que huit avec *gāi* 該: 荻垓菱郊殫剗咳.

6° Si dans le *Qièyùn* original, on trouve quelques exemples de variantes anciennes, «*gǔ zuò x*» 古作 *x*,¹⁶ ou de variantes contemporaines, «*huò zuò x*» 或作 *x*,¹⁷ en revanche, la formule «*sú zuò x*» 俗作 *x*, pour les graphies «vulgaires» ou «populaires» n'est semble-t-il jamais utilisée. On notera que dans P. 3695 (Zhōu 1983:55), les caractères sont écrits dans le style «courant» (*tōng* 通), et parfois même «vulgaire» ou «populaire» (*sú* 俗): *lái* 來, par exemple, s'écrit 來, tandis que *kāi* 開 s'écrit avec 井 au lieu de 开 à l'intérieur, etc. Cependant, comme il s'agit de copies, on ne peut pas être sûr qu'elles représentent le style original de Lù Fǎyán.

7° Il n'y a bien évidemment pas de commentaires ou d'annotations du type *àn* 案, propres aux versions augmentées du *Qièyùn*.¹⁸

Étant donné l'état fragmentaire des manuscrits du *Qièyùn*, il n'est guère possible de rétablir l'ordre original des petites rimes. Mais, il semble que l'auteur du *Kānmìù bǔquē Qièyùn* a reproduit assez fidèlement cet ordre que je laisse aux spécialistes le soin d'étudier.¹⁹

2. Les premières versions augmentées du *Qièyùn*

Parmi les manuscrits de Dūnhuáng, se trouve une dizaine de versions augmentées du *Qièyùn*. Elles correspondent aux fragments S. 2071, S. 2055, P. 3693/3694/3696(7)/S.6176, S. 5980, P. 3799, P. 2017, S. 6013, S. 6012, P. 4746 et S. 6156. Elles comprennent en général un nombre plus important de caractères et de gloses, avec éventuellement des annotations et des corrections.

1° La plupart des manuscrits ont des points ou cercles rouges bien visibles entre chaque groupe d'homophones. Mais il faut noter que certains textes, qui ont sûrement reproduit la présentation originale du *Qièyùn*, n'en ont pas: S. 2071, S. 2055, et P. 3799.

2° En général les rimes ne sont pas rehaussées dans la marge. A la différence du *Qièyùn* original, elles se suivent dans le texte et sont précédées d'une numérotation noire ou rouge ainsi que d'un espace blanc.²⁰ On remarque dans S. 2071 des points ou

¹⁶ Elles correspondent aux graphies sigillaires, comme dans l'exemple P. 3696, ch. 1, rime 6 *zhī* 脂: *méi* 眉古作 𪛗 (Zhōu 1983:49).

¹⁷ Elles représentent d'autres façons d'écrire le même 'mot' comme dans le cas de P. 3695, ch. 1, rime 16 *hāi* 哈: *cái* 纔或作 裁 (Zhōu 1983:55).

¹⁸ Dans P. 4917 (chap. 3, rime 35 *yǎng* 養), on remarque cependant que l'on a indiqué que le caractère *qiǎng* 搶 provenait du *Shìjì* (搶頭搶地出史記), ce qui pourrait suggérer une version un peu postérieure au *Qièyùn* original.

¹⁹ Voir Ozaki Yūjirō (1980:100-128) 切韻系韻書における韻の排列について.

²⁰ Dans P. 2017, qui représente la table des rimes aux quatre tons avec les 4 premières colonnes incomplètes du début du texte, la première rime devait être introduite par le caractère *píng* 平

sortes de croix dans la marge supérieure permettant de repérer les rimes noyées dans le texte.

3° Dans certains manuscrits les gloses précèdent le *fǎnqiè* (S. 2071, S. 5980, P. 3799), alors que dans d'autres, tantôt elles le précèdent, tantôt elles le suivent (S.2055, S. 6013). Ceci s'explique à mon sens par le fait que les auteurs d'annotations suivaient le texte original de Lù Fǎyán, et que lorsqu'ils rajoutaient une glose, ils devaient le faire à la suite du texte original. Ainsi, s'il n'y avait qu'une glose phonétique ou *fǎnqiè*, ils rajoutaient leur propre glose à la suite du *fǎnqiè*,²¹ donnant ainsi une impression de désordre.

4° Les auteurs des *Qièyùn* annotés ont augmenté le nombre de caractères. Ils les ont en général signalés par le terme *jiā* 加 «(caractères) ajoutés», ou *xīn jiā* 新加 «(caractères) nouvellement ajoutés», en précisant parfois qu'ils les avaient tirés du *Shuōwén* (出說文). Ils ont aussi ajouté des gloses. Certains auteurs ont extrait leurs gloses uniquement à partir du *Shuōwén* (P. 3693/3694/3696(7)/S. 6176), d'autres se sont également appuyés sur un plus grand nombre de textes, voire d'auteurs: Ēryǎ, Fāngyán, Shǐjì (P. 4746), Shàngshū (S. 6012), Yùpiān (P. 2011), Zhēng Xuán 鄭玄(云), etc.

5° Malgré un plus grand nombre de gloses, on remarque encore des suites de caractères sans glose (S. 2071, S. 2055, P. 3693/3694/3696(7)/S. 6176, S. 5980, P. 2017, P. 4746). Il faut noter que parmi ces manuscrits, S. 2071, S. 2055, P. 3799 n'ont toujours pas de points séparant les groupes d'homophones. Le manuscrit S. 2055, qui correspond à une version augmentée du *Qièyùn* par Zhāngsūn Nèyán 長孫訥言, dont la préface est datée de 677, comprend plusieurs exemples de caractères aux initiales différentes qui se suivent sans glose, tels *xī* 歛 et *qī* 鄴 (chapitre 1, rime 6 *zhī* 脂).²²

6° Les versions augmentées du *Qièyùn* comprennent souvent la formule «*sú zuò x*» 俗作 x, 'la graphie vulgaire s'écrit x', ou «*tōng zuò x*» 通作 x, 'la graphie courante s'écrit x'. Cette innovation pourrait bien remonter à Zhāngsūn Nèyán qui est l'auteur d'un *Qièyùn* augmenté, rédigé en 677 et dont nous possédons à la British Library un fragment de manuscrit sous la côte S. 2055. Dans sa préface, Zhāngsūn Nèyán précise que lorsqu'il était jeune, il a souvent consulté le *Modèle de caractères* (zìyàng 字樣) de Yán Shīgǔ 顏師古 (581-645). Ce texte, qui avait été rédigé sous ordre de l'empereur à

que l'on devine dans la marge supérieure.

²¹ Zhōu (1983:834-835) fait remarquer que dans tout le passage concernant la rime n° 8 *wēi* 微 du 1^{er} chapitre dans S. 2055, le *fǎnqiè* est donné en premier et le texte est identique à celui du *Kānmiù būquē Qièyùn*. Son explication est la suivante: le texte original de cette version du IX^e siècle étant abîmé, le copiste a dû le reproduire en s'appuyant sur le texte du *Kānmiù būquē Qièyùn* (706).

²² Le texte se présente de la manière suivante: 歛鄴 ... 取私反.

l'occasion de la réfection des Classiques sous l'ère Zhēngguān 貞觀 (627-649), distinguait les graphies tirées des Classiques des graphies courantes et vulgaires.²³ Dans le manuscrit fragmentaire S. 2055, on trouve, en effet, des exemples de ce type: 微:無非反. 妙. 通. 俗作敝六. 'Wēi 微: se prononce wú-fēi. Subtil. [Graphie] courante. La graphie vulgaire s'écrit: 敝. Six [caractères homophones].' (S. 2055, ch. 1, rime 8 wēi 微, Zhōu 1983:155). Zhāngsūn Nèyán a également pris soin d'indiquer, dans son texte, la structure graphique «correcte» des caractères, afin d'éviter des erreurs. Pour ce faire, il s'est naturellement appuyé sur le *Shuōwén*. C'est pourquoi, on voit souvent explicités dans ce manuscrit les constituants graphiques des caractères. Dès la première rime, *dōng* 東, Zhāngsūn Nèyán ajoute une glose tirée du *Shuōwén* à la suite du *fānqiè* de Lù et indique la structure graphique de *dōng* décrite par Xǔ Shèn: 東德紅反二. 桱(sic)說文春方也動也. 從日又云日在水²⁴中 «J'observe que d'après le *Shuōwén*, 東 [*dōng*] correspond au printemps, et représente le mouvement. [La graphie] a 'soleil' pour constituant (sémantique). D'après une autre interprétation, elle représente le 'soleil' dans 'l'eau.'» (S. 2055).²⁵ Dans les *Qièyùn* augmentés, on trouve aussi parfois la formule «zhèng zuò x» 正作 x 'la graphie orthodoxe s'écrit x'. Dans le manuscrit P. 4746, rédigé en style courant (*tōng* 通),²⁶ l'auteur a pris soin de signaler les graphies orthodoxes (*zhèng* 正) *guó* 國 et *huò* 或 pour ces deux caractères écrits en style courant dans le texte, c'est-à-dire avec 厶 à la place de 口.²⁷

7° Certaines versions augmentées du *Qièyùn* ont ajouté des observations. Celles-ci sont parfois précédées du terme *àn* 案 «j'observe» (écrit tantôt 桱, tantôt 按), comme on a pu le voir dans l'exemple de *dōng* 東 ci-dessus. Dans S. 2055, l'auteur écrit plusieurs fois *jīn wéi* 今爲 probablement pour indiquer un contraste avec une interprétation ancienne ou avec celle du *Shuōwén*: 僮古作童子今爲僕 «*tóng* 僮: désignait autrefois 'un enfant', aujourd'hui il représente 'un serviteur.'»; ou encore 終按說文絳絲也. 今爲終始字. 職隆反十 «*Zhōng* 終: d'après le *Shuōwén* c'est un 'fil de soie très tendu'. Aujourd'hui il représente le caractère qui a le sens de 'finir'. Se prononce *zhí-lóng*. Dix [caractères homophones].» (chap. 1, rime 1, *dōng* 東). Mais les auteurs ne citent pas toujours leurs sources, ou s'appuient parfois sur d'autres textes, voire d'autres auteurs: le *Zhèngmíng* 正名 et le *Ziyàng* 字樣 dans P. 3693, le *Fāngyán* 方言 ou Zhāng Yàn 張晏 (III^e

²³ Voir la préface du *Gānlù zìshū* 干祿字書 de Yán Yuánsūn.

²⁴ Comme on peut également le constater dans le manuscrit P. 2017 (Zhōu 1983:150/160, 226/228), il s'agit bien du caractère *shuǐ* 水 et non de *mù* 木.

²⁵ Toutes les autres versions du *Shuōwén* ont *mù* 木 à la place de *shuǐ* 水: *dōng* 東 動也. 从木. 官溥說, 从日在木中 (cf. Xú Xuàn 徐鉉 (1977: 6A 24b) et *Shuōwén jiězhì gūlín* (1988:2652 sq.))

²⁶ Voir le *Gānlù zìshū* 干祿字書 ch. 4, 1972, p.169 pour les définitions des différents styles «orthodoxe», «courant» et «vulgaire».

²⁷ Dans le *Kānmìu bǔquē* *Qièyùn*, *guó* et *huò* s'écrivent en style *zhèng* «orthodoxe»: 國, 或.

siècle)²⁸ dans P. 4746, etc. Dans certaines versions augmentées du *Qièyùn*, on a souligné les erreurs ou les insuffisances de Lù Fǎyán au moyen de la formule *Lù qiàn* 陸欠 «Lù est incomplet / Lù se trompe».²⁹ Dans le petit fragment P. 4746 de seulement 16 colonnes, on trouve au moins six signalements d'erreurs de Lù Fǎyán (cf. Zhōu 1983:232-234).

Les versions augmentées comprenant plus de caractères et de notes que celles du *Qièyùn* original, leurs auteurs ont dû chercher à en faciliter la consultation. Ce qui explique pourquoi tous ces textes ont soit des marques dans la marge supérieure pour repérer les rimes, guère visibles dans le texte (S. 2071), soit des points (rouges ou noirs) permettant de distinguer clairement les groupes d'homophones qui ont été allongés du fait des additions de notes et de caractères.

3. Le *Kānmiù bǔquē Qièyùn* 刊謬補缺切韻

On possède au moins trois manuscrits du *Kānmiù bǔquē Qièyùn* de Wáng Rénxǔ. Un manuscrit incomplet, provenant de Dūnhuáng et conservé à la B.N. sous la côte P. 2011, qui consiste en 21 feuillets abîmés et qui a été reproduit avec une transcription dans Zhōu (1983:246-358, 359-433). Une version complète du *Kānmiù bǔquē Qièyùn*, découverte en 1947 au Palais impérial de Pékin, également reproduite dans Zhōu (1983: 434-527)³⁰ (appelée *Sòng lián bá běn* 宋濂跋本). Et une autre version incomplète provenant également du Palais impérial, mais découverte en 1920 par Luó Zhènyù 羅振玉 (1866-1940) dans la bibliothèque de l'ex-empereur Pú Yí 溥儀 (1906-1967) (appelée *Nèifǔ běn* 內府本). Elle est reproduite dans Liú Fù 劉復 (1891-1934), 1963. Comme elle comprend beaucoup plus de caractères ajoutés, elle est considérée comme plus tardive (Wang Lien-tseng 1957:110; Liú Fù 1963:34). Enfin, le petit fragment P. 2129, conservé à la B.N., reproduit seulement la préface du *Kānmiù bǔquē Qièyùn* (Zhōu 1983:242-245).

²⁸ Zhāng Yàn est l'auteur du *Xī Hànsū yīnshì* 西漢書音釋 en 40 volumes. L'exemple dans lequel cet auteur est cité est le très court fragment P. 4746, chap. 5, rime 30, *dé* 德: *mò* 縹: 索案張晏云三股曰縹 (Zhōu 1983:232/234). Noter que dans le *Shuōwén* 13A 12a, ce caractère est aussi glosé 縹: 索也. 從糸黑聲. D'après le *Hànyǔ dà zìdiǎn* vol. 5, p. 3466, 縹: 繩索也. “玉篇, 糸部”: 縹索也...

²⁹ Ainsi, par exemple, dans S. 6012 (chap. 5, rime *zhí* 職): 蠱: 傷也. 見尙書陸欠 «*shì* 蠱, c'est blesser. Voir le *Shàngshū*. Lù est incomplet.»; ou encore (chap. 5, rime *duó* 鐸): 飢: 楚人食麥饘謂之飢見方言陸欠 «*Zuò* 飢: le porridge de blé que mangent les gens de Chǔ s'appelle *zuò*. Voir le *Fāngyán*. Lù est incomplet.». Le *Fāngyán* 方言 1. 31 (1993, p.9) propose la glose suivante: 饘飢食也. 陳楚之內相謁而食麥饘謂之饘, 楚曰飢. 凡陳楚之郊南楚之外相謁而飢, 或曰飢, 或曰飢; et le *Shuōwén* 5B 4b *zuò* 飢: 楚人相謁食麥曰飢从食乍聲 (Tāng Kějīng 湯可敬 1997:696).

³⁰ Selon Wang Lien-tseng (1957:51-150) et Lung Yu-Chun 龍宇純 (1968), cette version de 1947 comprend de nombreuses fautes.

Si Wáng Rénxǔ a repris l'ordre des rimes (et des petites rimes de Lù Fǎyán), il a ajouté deux rimes qui manquaient aux 2^{ème} et 3^{ème} tons: *yǎn* 𪛗 et *yán* 嚴, pour correspondre à la rime *yán* 嚴 du 1^{er} ton. Le *Kānmiù bǔquē Qièyùn* comprend donc un total de 195 rimes. A quelques exceptions près, Wáng Rénxǔ a reproduit les gloses phonétiques de Lù Fǎyán (Zhōu 1983, vol. 2: 888 sq). Mais comme on va le voir ci-dessous, il a en revanche opéré un certain nombre de modifications, et ajouté de nombreux caractères, qu'il a placés après chaque groupe d'homophones. D'après sa préface tous les ajouts ont été faits en rouge.³¹

1° Wáng Rénxǔ a définitivement adopté le système des points rouges avant les petites rimes pour séparer les groupes d'homophones.

2° Les rimes ne sont pas rehaussées dans la marge supérieure mais sont précédées d'une numérotation rouge. Dans la version du palais impérial reproduite dans Zhōu (1983:434-527), les rimes sont facilement repérables par leur numérotation placée dans la marge supérieure, alors qu'elles se trouvent elles-mêmes soit en début de colonne, soit dans le texte précédées d'un espace blanc pour les rendre visibles. Tandis que dans P. 2011 (Zhōu 1983:246-433), elles semblent toutes avoir été écrites en début de colonne, avec une numérotation rouge dans la marge.

3° Wáng inverse et donne systématiquement le *fānqiè* en premier, avant l'explication du sens du caractère et le nombre de caractères compris sous 'les petites rimes'.

4° Il fournit une glose pour chaque caractère.

5° Il n'y a par conséquent plus de caractères qui s'enchaînent sans glose.

6° Comme Zhāngsūn Nèyán, auteur du *Qièyùn* augmenté S. 2055 (en 677), Wáng indique les graphies correctes et précise si l'on a affaire à des graphies courantes, vulgaires, contemporaines, etc. C'est pourquoi, on trouve toutes sortes de formules parmi ses gloses: «*yì zuò x*» 亦作 *x*, «*yòu zuò x*» 又作 *x* 's'écrit également *x*', «*huò zuò x*» 或作 *x* 's'écrit parfois *x*', «*běn zuò x*» 本作 *x* '[la graphie] première s'écrit *x*', «*sú zuò x*» 俗作 *x* '[la graphie] vulgaire s'écrit *x*', «*jīn zuò x*» 今作 *x* 's'écrit aujourd'hui *x*', «*tōng zuò x*» 通作 *x* '[la graphie] courante s'écrit *x*', mais aussi comme dans les

³¹ Dans sa préface (manuscrit P. 2129), Wáng Rénxǔ donne les clés de son programme: 既字該樣式, 乃備應厄疑. 韻以部居, 分別清切. 舊本墨書, 新加朱書. 兼本闕訓, 亦用朱書. 其字有疑涉, 亦略注所從, 以決疑謬. 使各區析, 不相雜廁. «Les caractères doivent être dans la norme pour faire face aux difficultés et aux doutes. Les rimes sont classées et rigoureusement distinguées. Le texte original est écrit en noir, les ajouts à l'encre rouge. Les gloses qui étaient absentes dans le texte original sont également (ajoutées) à l'encre rouge. Pour ceux des caractères qui sont écrits de manière douteuse, [j'ai] aussi ajouté de brèves notes indiquant leurs constituants afin d'éliminer doutes et erreurs. Ainsi tout étant bien analysé, il n'y a pas de confusion possible».

Modèles de caractères: «*tōng sú zuò x*» 通俗作 x '[graphie] courante; [la graphie] vulgaire s'écrit x', «*shàng tōng zhèng zuò x*» 上通, 正作 x '(le caractère) ci-dessus correspond à la graphie courante, la graphie orthodoxe s'écrit x', etc. On trouve également la formule indiquant la structure graphique des caractères: «*cóng x*» 從 x 'a pour constituant x' comme par exemple pour *chí* 趁: «*趁...從多音支聲*» 'chí... a pour constituant *duō* qui se prononce *zhī* et sert de constituant phonétique» (autrement dit, livre de rimes oblige, l'auteur a ajouté en plus la prononciation du constituant phonétique) (P. 2011, ch. 1, rime *zhī* 支).

7° Wáng Rénxǔ formule ses propres critiques au texte de Lù et en signale les erreurs. Zhōu (1983, vol. 2: 875-876) en a relevé environ une dizaine. Dans le chap. 4, rime 56, *yán* 嚴, on trouvera, par exemple, la glose suivante: «*嚴: 魚俺反陸無此韻目失*» 'Yán se prononce *yú* + *ǎn*. Lù n'a pas cette rime, c'est une omission.' (Zhōu 1983:489), ou encore, dans le chap. 1, rime 33, *gē* 歌, sous le caractère *xuē* 韋: «*...陸無反語*» '... Lù n'a pas donné de *fāngqiè*.' (*ibid.*:459).

Avec le *Kānmìu būquē Qièyùn* la modification de la présentation du *Qièyùn* s'accroît et s'uniformise. En fournissant une glose sémantique pour chaque caractère, le *Kānmìu būquē Qièyùn* a résolu le problème d'ambiguïté concernant la prononciation des caractères qui se suivaient sans glose, se rapprochant ainsi du dictionnaire.

4. Les gloses dans les différentes versions du *Qièyùn* au *Kānmìu būquē Qièyùn*

4.1 Les gloses dans le *Qièyùn*

Comme on l'a noté, de nombreux mots courants parmi les rimes et petites rimes (*dōng* 冬, *kōng* 空, *lóng* 籠, *jiāng* 江, etc.), mais aussi les entrées (*lín* 林, *lái* 來, *wú* 吾, *āi* 哀, *shī* 詩, etc.) ne sont pas glosés dans les plus anciennes versions du *Qièyùn* (P. 3695/3696 R°&V°, S. 6187, P. 3798 et S. 2683/P. 4917). Lorsqu'elles existent, les gloses du *Qièyùn* sont extrêmement simples. On observe les types suivants de gloses:

- Des noms génériques du type: «X: 水名/木名/鳥名/竹名/魚名/草名/地名, etc.» ('X: nom de cours d'eau/nom d'arbre/nom d'oiseau/nom de bambou/nom de poisson/nom de plante herbacée/nom de lieu, etc.'):
 - 鷺: 鳥名 (P. 3798)
'*lóng* 鷺: nom d'oiseau'.
 - 邾: 地名; 在蜀. (P. 3695)
'*lái* 邾: nom de lieu, situé à Shǔ'
 - 瀉: 水名出新陽 (P. 3696).
'*wéi* 瀉: nom de rivière, prend sa source à Xīnyáng'.

La situation géographique des noms de lieux, de montagnes ou de rivières est généralement donnée.

- Des quasi-synonymes ou des paraphrases.³²

X: A

tú 途: 道 ‘chemin: voie’.

tú 圖: 畫 ‘tableau: dessin’.

X: AB

○ *lú* 壚: 黑田 (P. 3695),

‘*lú*: terre noire’.

○ *qīn* 駢: 馬行 (S. 6187) (*idem* dans S. 2071)

‘*qīn*: cheval au galop’.

○ *xín* [xún] 鄴: 古姓 (S. 6187)

‘*xín*: ancien nom de famille’.

X: ABCD

lái 騾: 馬高七尺 (P. 3696) (*idem* dans le *Kānmiù bǔquē Qièyùn* (Zhōu 1983:448).

‘*lái*: cheval haut de sept *chǐ*’.

- Quelques rares gloses descriptives ou analytiques:

○ *hú* 猢: 猢猻獸名似猿; 猢字土咸□³³ (P. 3695) (Zhōu 1983:51/61)

‘*hú*: comme dans *chánhú*, nom d’animal qui ressemble au singe. Le caractère 猢 [se prononce] *tǔ-xián*’.³⁴

○ *tú* 駢: 鳥名與鼠同穴. (P. 3695)³⁵ (*ibid.*)

‘nom d’oiseau qui habite dans les caves avec les souris’

○ *máng* 嗙: 語雜亂曰嗙. (rime *jiāng* 江, P. 3696) (Zhōu 1983:42/57)

‘On appelle *máng* des propos confus’

○ *jī* 姬: 王妻別稱

‘*jī*: autre appellation pour l’épouse du roi’. (rime *zhī* 之 P. 3695).³⁶

³² Dans ces exemples les gloses diffèrent un peu de celles du *Shuōwén*: «*tú* 圖: 畫計難也»; «*lú* 壚: 剛土也»; «*qīn* 駢: 馬行疾也»; «*xín* [xún] 鄴: 周邑也»; «*lái* 騾: 馬七尺爲騾»; quant à *tú* 途 il n’est pas inclus dans le *Shuōwén*.

³³ Caractère manquant qui devait correspondre à *fǎn* 反.

³⁴ Cette glose absente du *Shuōwén* pourrait provenir du dictionnaire *Yùpiān* 玉篇 de Gù Yěwáng 顧野王 (VI^e siècle). Elle a été reprise dans le *Kānmiù bǔquē Qièyùn* (cf. Zhōu 1983:253/263).

³⁵ L’origine de cette glose semble inconnue. Dans le *Shuōwén*, *tú* est mis pour *táotú* qui désigne naturellement un cheval: «駢: 駒駢也 et *táo* 駒: 駒駢, 北野之良馬» ‘excellent cheval des plaines du nord’, mais en aucun cas un nom d’oiseau, (10A 7a, TKJ 1997:1333)

³⁶ Comparer avec le *Shuōwén* (12B 1a, TKJ 1997: 1742): «姬: 黃帝居姬水, 以爲姓».

- Des impressifs du type: «X: A 兒/AB 兒/ABC 兒» (兒 est une variante de 貌), qui rappellent que l'on a affaire à un livre pour la composition littéraire et poétique:
 - *chōng* 懂: 行來兒. (P. 3798)³⁷
'descriptif d'indécision'
 - *zhōng* 徂: 征徂行兒
'comme dans *zhēngzhōng* descriptif d'une marche'.
 - *tōng* 炆: 火盛兒
'descriptif d'un feu vivace'.
- Enfin, dans certains cas, le caractère glosé est repris dans la glose pour en clarifier le sens: «X: XY».
 - *míng* 鳴: 鳴呼 (P. 3695).³⁸
'*míng*: comme dans [l'expression] chanter/crier'
 - *lóng* 隴: 喉隴' (P. 3798) (Zhōu 1983: 36/38).
'*lóng*: comme dans [l'expression] *hóulóng* (gorge)

Les gloses multiples sont extrêmement rares dans le *Qièyùn*. L'univocité du vocabulaire est la règle alors que c'est l'exception dans l'usage de la langue. On en trouve cependant un exemple dans la version augmentée P. 3799 (Zhōu:221/223) qui devait faire partie du *Qièyùn* original, puisque l'on y voit cité le nom de Xiāo Gāi 蕭該, l'un des auteurs du *Qièyùn*. Elle a été reprise dans le *Kānmiù bǔquē Qièyùn* (Zhōu 1983: 524) avec quelques petites différences représentées ici entre crochets:

xiào 芍: 蕭該云芍藥是藥草. 又香草可和食. 芍字張約反. 藥字良約反. 又芍陂在淮南. 七削反. 又蓮芍縣名[*KM*:0]在馮翊. 之若反. [*KM*:之言也] 又鳧茈草. 胡□ [*KM*: 又草鳧茈. 胡了反]

'*xiào* 芍: d'après Xiāo Gāi, 芍藥 est une plante médicinale, ou encore une plante odoriférante qui permet d'assaisonner les mets. Le caractère 芍 se lit *zhāng-yāo* [*zhāo*], le caractère 藥 *liáng-yāo*. Par ailleurs, 芍陂 se situe à Huáinán et se prononce *qī-xiāo* [*qiāo*], et 蓮芍 est un nom de district situé à Píngyì qui se prononce *zhī-ruò* [*zhuò*]. Il y a aussi la plante *fúcí*³⁹ qui se lit *hú*-□. [la prononciation est incomplète mais correspond sans aucun doute à celle du *KM*: 胡了反: = *xiào*].'

³⁷ Comparer avec le *Shuōwén* (10B 17b, TKJ 1997:1464): «懂: 意不定也».

³⁸ Comparer avec le *Shuōwén* (4A 25a, TKJ 1997:537): «鳥聲也».

³⁹ Cette dernière glose correspond à celle du *Shuōwén* pour *xiào* 芍: «芍: 鳧茈也» (TKJ 1997: 105).

Il est important de constater que, dans cet exemple, le *Qièyùn* donne également des prononciations locales ou dialectales plutôt que standards et littéraires.

La particule *yě* 也 s'avère quasi absente des gloses du *Qièyùn*.⁴⁰ Il est vrai que dans le *Qièyùn*, la signification des mots ou des caractères est secondaire et doit juste permettre de les distinguer les uns des autres, tandis que les prononciations (*fānqiè*), elles, sont primordiales.

4.2 Les gloses dans les versions augmentées du *Qièyùn* jusqu'au *Kānmiù būquē Qièyùn*

Les auteurs de versions augmentées ont en général reproduit les gloses ainsi que les absences de gloses de Lù Fǎyán. Mais, on a vu qu'ils avaient parfois introduit de nouvelles gloses, lorsque que Lù n'en avait pas données, et ajouté toutes sortes de notes. Si l'on compare *zhī* 支 dans S. 2071, S. 2055 et dans le *Kānmiù būquē Qièyùn* (abrégé en KM), outre le nombre toujours plus important de caractères (9, 10, 15), on trouve les différentes gloses suivantes:

- Chap. 1, rime n° 5 *zhī* 支
 S.2071: «章移反. 九» (Zhōu:74)
 '(Se prononce) *zhāng-yí*; 9 [caractères homophones]'.
 S.2055: «章移反. 十. 按說文. 去竹之枝也. 從又持半竹» (Zhōu:152)
 '(Se prononce) *zhāng-yí*; 10 [caractères homophones]; d'après le *Shuōwén*: branche que l'on a séparé de la tige...'.
 KM: «章移反. 計. 十五» (Zhōu:438)
 '*Zhī*: (se prononce) *zhāng-yí*, mesure; 15 [caractères homophones]'.
- Pour *fēi* 非 et *yī* 衣, chap. 1, rime n° 8 *wēi* 微, on trouve les gloses suivantes:
 S.2071:⁴¹ «*fēi* 非: 不» «*yī* 衣: 0».

⁴⁰ Il y a bien sûr quelques exceptions: «郃: 地名在始平, 或作𡵚也» '*Tāi* 郃: Nom de lieu, situé à Shipíng, s'écrit également *tāi* 𡵚' (P. 3695); «郃: 炎帝之後, 姜姓所封, 周棄外家國... 右扶風𡵚縣是也. 《詩》曰: 有郃家室» (TKJ 1997:865); «愔: 靖也» '*yīn* 愔 c'est être apaisé' (S. 6187); «黔: 黑而一曰黃黔首眾也» '*qián* 黔: Noir et [jaune] d'après (jaune) une autre source *qiánshǒu*, 黔首 désigne le peuple.' (S. 6187). Cette glose a été reprise dans le *Kānmiù būquē Qièyùn* (Zhōu 1983:468) sans faute comme ici où 'jaune' devrait précéder '*yī yuē* 一曰', et sans '*yě* 也': «黔: 黑而黃一曰黔首 眾» (SW 10A 26a); «黔: 黎也. 从黑今聲. 秦謂民爲黔首, 謂黑色也. 周謂之黎民. 《易》曰: 爲黔喙.» (TKJ 1997:1398).

⁴¹ Le manuscrit S.2071 devait être plus proche du *Qièyùn* original: il n'avait pas de points et il n'a pas vraiment ajouté de gloses. Comparer dans le chap. 3, rime 26 *xiǎn* 獮 entre:

S. 2071: «演: 廣以淺反. 三»; «衍: 達»; et
 P. 3693 R°: «演: 廣以淺反. 三. 按說文作此演 長流»; «衍: 達 按說文. 水朝宗於海. 故從水行».

- S.2055: «*fēi* 非: 不是» «*yī* 衣: 服裝» ‘vêtement; jupe’.
 KM: «*fēi* 非: 不是» «*yī* 衣: 服» ‘vêtement’.

Dans le texte original S. 6187 (chap. 1, rime n° 46 *qīn* 侵, Zhōu 1983:63), le caractère *zhēn* 針 n’est pas glosé. En revanche, dans S. 2071 (Zhōu 1983:91), on lit: «針: 案文作鍼» ‘*zhēn*: d’après le (*Shuōwén*) *wén* s’écrit *zhēn* 鍼’. Quant au *Kānmiù bǔquē Qièyùn* (Zhōu 1983:467), il reprend la variante graphique donnée dans S. 2071, mais sans citer le *Shuōwén* et surtout il fournit une glose: «針: 縫衣具. 或作鍼» ‘*zhēn*: instrument pour coudre, s’écrit également *zhēn* 鍼’.

Wáng Rénxǔ a systématiquement ajouté des gloses là où il n’y en avait pas:

- S. 6187: «尋: 徐林切. 七» (Rime n° 46 *qīn* 侵 (Zhōu 1983:63))
 ‘*Xún*: (se prononce) *xú-lín*, 7 [caractères homophones]’.
 KM: «尋: 徐林切. 八尺曰. 十二» (Zhōu 1983:467)
 ‘*Xún*: (se prononce) *xú-lín*. Une mesure de huit *chǐ* se dit [*xún*]. 12 [caractères homophones]’.
 S. 6187: «林: 力尋反. 八» (*ibid*)
 ‘*Lín*: (se prononce) *lì-xún*, 8 [caractères homophones]’.
 KM: «林: 力尋反. 攢(sic)木. 八»
 ‘*Lín*: (se prononce) *lì-xún*, série d’arbres, 8 [caractères homophones]’.

De manière générale, Wáng Rénxǔ ne s’appuie pas sur le *Shuōwén* pour expliquer le sens des caractères:

- P. 3696: «*ér* 而: 如之反. 十二» (*idem* dans S. 2071; Zhōu 1983:50/60).
 S. 2055: «*ér* 而: 如之反. 十二. 頰毛也. 象毛之形. 周禮曰: 作其鱗之而» (Zhōu 1983:155/165).
 KM: «*ér* 而: 如之反. 詞. 十三» (Zhōu 1983:441).

Alors qu’il n’y a qu’une glose phonétique pour *ér* 而 dans le *Qièyùn* original P. 3696 et dans la version augmentée S. 2071, l’auteur de la version augmentée S. 2055 cite le *Shuōwén* selon lequel *ér* représente les poils des joues (les favoris). Ce faisant, il fournit une explication graphique de *ér* (ou plutôt de *ér* 𠂔 ‘favoris’) qui ne correspond absolument pas au sens courant de *ér* 而. Wáng Rénxǔ remplace cette interprétation par la simple glose *cí* 詞 et lui reconnaît ainsi sa fonction grammaticale. De même dans:

- S. 2055: «終: 按說文. 綵絲也. 今爲終始字. 職隆反十»
 ‘*Zhōng*: d’après le *Shuōwén*, c’est un ‘fil de soie très tendu’.

Aujourd'hui il représente le caractère qui a le sens de 'finir'. (Se prononce) *zhí-lóng*. 10 [caractères homophones]'. (chap. 1, rime 1 *dōng* 東; Zhōu 1983:150/161).

KM: «終: 職隆反. 畢. 十一» (Zhōu 1983:436).⁴²
'*zhōng*: (se prononce) *zhí-lóng*, terminer'.

L'auteur de S. 2055 reprend la glose graphique du *Shuōwén* mais éprouve en même temps le besoin d'expliquer à sa manière le sens courant de *zhōng* 'finir'. Alors que Wáng Rénxǔ fournit une glose à la fois simple, originale et efficace: *bì* 'terminer'. S'il s'appuie parfois sur le *Shuōwén* pour clarifier la graphie des caractères, Wáng se détache des auteurs antérieurs qui fournissaient parfois de longues explications tirées du *Shuōwén*. Là où Zhāngsūn Nèyán explique longuement: «*dōng* 東... 按 (sic) 說文. 春方也. 動也. 從日. 又云日在水中». (S. 2055) (voir la traduction plus haut), Wáng se contente de dire: «東: ...木方». 'correspond au point cardinal bois'. À quelques exceptions près, les gloses de Wáng Rénxǔ sont extrêmement simples, et suffisent à identifier les caractères. Contrairement à certaines versions augmentées,⁴³ il n'y a guère de gloses multiples dans le *Kānmiù bǔquē Qièyùn*.

De même si plusieurs textes augmentés tels P. 3693/3694/3696/S. 6176, S. 2055 (voir l'exemple *dōng* 東 donné ci-dessus) fournissent de nombreuses gloses en *yě* 也, car leurs auteurs se sont inspirés de gloses tirées d'autres textes et, en particulier, du *Shuōwén*, ce n'est pas tant le cas du *Kānmiù bǔquē Qièyùn*.

5. Conclusion

Le *Qièyùn* de Lù Fǎyán était donc différent du *Kānmiù bǔquē Qièyùn* et beaucoup plus rudimentaire sur le plan des gloses sémantiques. Il est probable qu'à l'origine il ne comprenait pas de points ou cercles entre les groupes d'homophones, comme le suggèrent les fragments P. 3695/3696 R° & V°, S. 6187, ainsi que S. 2071, S. 2055, P. 3799. L'absence de points marquant les groupes d'homophones dans les plus anciens manuscrits et dans certaines des versions augmentées du *Qièyùn* me semble indiquer que Lù Fǎyán n'avait pas imaginé une véritable méthode pour retrouver les caractères

⁴² *Zhōng* n'était sûrement pas glosé dans le *Qièyùn* original. Nous n'avons malheureusement aucun moyen de le vérifier étant donné qu'il ne subsiste aucun fragment original de ce passage, mais le fait que dans la version augmentée S. 2055, le texte commence par 'd'après le *Shuōwén*' et dans le *Kānmiù bǔquē Qièyùn* on ait une autre glose constituant en quelque sorte des preuves de l'absence de glose pour *zhōng* dans le *Qièyùn*.

⁴³ Comme, par exemple, dans P. 3799: «若 順也. 善也. 辭也 ...» '*ruò*: se conformer, excellent, prendre congé...' (chap. 入聲, rime 27 *yào* 藥; Zhōu:221/223).

dans son texte. Son but était avant tout de fournir, en les organisant, les prononciations ‘standards’ des caractères afin d’aider les lettrés dans leur travail de composition littéraire.

La séquence des tons, rimes et initiales offrait cependant une nouvelle méthode de classement qui facilitait la consultation des caractères. En effet, comparé au système de classement par clés trop subjectif de Xǔ Shèn 許慎 (100),⁴⁴ il était plus simple de retrouver un caractère de prononciation connue dans le *Qièyùn*.⁴⁵

Le fait que l’on ait découvert tant de différentes versions du *Qièyùn* non seulement à Dūnhuáng, mais aussi à Turfān, suggère qu’il devait être très utilisé, et probablement à des fins personnelles.⁴⁶ Il faut noter, par ailleurs, qu’un grand nombre de copies du *Qièyùn* retrouvées à Dūnhuáng étaient écrites sur des feuillets «en tourbillons», *yèzi běn* 葉子本. Ce type de support, qui se rapproche du livre, facilitait énormément la consultation car l’on pouvait feuilleter les pages qui étaient collées uniquement sur leur bord droit à une longue feuille de rouleau et écrites recto-verso.

L’ajout de points entre groupes d’homophones a dû se faire assez vite, puisque P. 3798 et S. 2683/P. 4917, considérés par Zhōu Zǔmó comme des versions proches du *Qièyùn* original, les comprennent déjà. Les points constituent en fait une nette amélioration de la présentation du texte. En délimitant clairement les différents groupes d’homophones, ils permettaient d’éviter d’éventuelles confusions entre prononciations de caractères qui appartenaient à des groupes d’homophones différents tout en se suivant sans glose. En même temps, ils facilitaient la recherche: on repérait plus vite les petites rimes avec leurs *fānqiè* sous lesquelles était rangé le caractère recherché. C’est pourquoi, cette aide à la consultation me paraît plutôt avoir été développée par les utilisateurs du *Qièyùn*.

Profitant à la fois d’une nouvelle méthode de classement des caractères bien plus efficace et d’une présentation matérielle mieux adaptée à la consultation, les auteurs successifs des versions augmentées ont développé un nouveau style de texte à cheval entre le simple livre de rimes et le dictionnaire. Ils y ont inclus le plus possible de caractères et d’informations ou gloses tirées du *Shuōwén*, livre de référence par excellence pour les textes classiques à cette époque. Les références au *Shuōwén* apparaissent à partir de la version annotée de Zhāngsūn Nèyán (677) (P. 2017), mais les

⁴⁴ A moins d’en connaître le texte par cœur, il était parfois très difficile de retrouver un caractère dans le *Shuōwén*.

⁴⁵ Le *Qièyùn* original devait comprendre entre 11 000 et 12 000 caractères, soit un nombre inférieur aux 12 158 décrits par Fēng Yǎn 封演 dans son *Fēng shì wén jiàn jì* 封氏聞見記 (deuxième moitié du VIII^e siècle), voir ci-dessus note 7.

⁴⁶ C’est ce que donnent à penser certaines copies bâclées ou écrites à la va-vite telles que S. 2071 et S. 2055, par exemple.

gloses ne sont pas encore systématiques et il n'y a pas toujours de points séparant les groupes d'homophones.⁴⁷ C'est pourquoi, on peut dire que Wáng Rénxǔ a grandement contribué à introduire les livres de rimes dans la voie des dictionnaires, en utilisant les points rouges, et surtout en fournissant une glose sémantique pour chaque caractère dans son *Kānmiù bǔquē Qièyùn*. Dans le *Qièyùn* lorsque les caractères étaient glosés, ils l'étaient en dehors de tout contexte traditionnel (classique ou graphique), selon leur sens courant. Dans la perspective de Lù Fǎyán, il ne s'agissait ni de suivre les commentateurs des Classiques, qui expliquaient les mots en contexte, ni de suivre Xǔ Shèn qui les expliquait en fonction de leur graphie. Il lui suffisait de les distinguer les uns des autres par de simples gloses lorsque cela s'avérait nécessaire. Wáng Rénxǔ a repris cette méthode qu'il a appliquée de manière systématique, tout en tenant compte de certaines notes tirées du *Shuōwén* ajoutées par ses prédécesseurs.

Avec ses gloses simples et décontextualisées, le *Kānmiù bǔquē Qièyùn* se rapprochait plus du modèle du dictionnaire, tel qu'on l'entend aujourd'hui. Cependant, il était loin de donner les principales acceptions des caractères pour remplir parfaitement cette tâche. Le *Tángyùn* 唐韻 de Sūn Miǎn 孫愐 (720 et 751) et le *Guǎngyùn* 廣韻 de Chén Péngnián (1008)⁴⁸ s'en rapprocheront plus. Mais le *Tángyùn* ne fournissait pas toujours le sens des mots courants et restait trop lié au contexte des Classiques,⁴⁹ tandis que le *Guǎngyùn* 廣韻 avait choisi d'inclure beaucoup de caractères rarissimes qui ne correspondaient pas à des mots courants⁵⁰.

Si le nouveau système de classement par tons, rimes et initiales ne pouvait pas définitivement remplacer le système des clés de Xǔ Shèn, notamment à cause des caractères de prononciation inconnue que l'on ne peut chercher que par leur graphie, il s'avérait malgré tout beaucoup plus efficace pour retrouver les caractères de prononciation connue, surtout lorsqu'il était combiné au support sur feuillets en

⁴⁷ On se souvient que non seulement les versions augmentées S. 2071, S. 2055, P. 2017 n'ont pas de points, mais qu'elles ont aussi des groupes de caractères sans glose, tout comme P. 3693/3694/3696(7)/S.6176, S. 5980, P. 2017 et P. 4746.

⁴⁸ Chén Péngnián (陳彭年), qui a produit le *Guǎngyùn*, est également l'auteur d'une version révisée et augmentée du *Yùpiān* 玉篇 (premier dictionnaire chinois à fournir les différentes acceptions des caractères dans les Classiques) intitulé *Dà guǎng yì huì Yùpiān* 大廣益會玉篇.

⁴⁹ Ainsi, par exemple, ni *yǐn* 飲 'boire', ni *liù* 六 'six' ne sont glosés dans le *Tángyùn*, comme on peut le voir dans le texte de Zhōu 1983:681 et 687; en revanche Sūn Miǎn s'appuie sur un très grand nombre de textes classiques et anciens pour confirmer ses gloses.

⁵⁰ Voir W. Boltz (2007). C'est d'ailleurs la critique qui a été faite par Fēng Yǎn (封演) (VIII^e siècle) aux auteurs de livres de rimes qui ont suivi le *Qièyùn*: «Par la suite, Sūn Miǎn (孫愐) et les autres ont extrait les caractères rares des *Livres de caractères* (字書) et les ont ajouté au *Qièyùn* sans comprendre que c'était inutile pour la composition littéraire, tel que l'avait voulu Lù.» cf. supra note 7.

tourbillons. Les livres de rimes se sont ainsi engagés dans la voie des dictionnaires et ont ouvert de nouvelles perspectives pour le recensement des mots de la langue en Chine, au risque cependant d'entraîner des confusions entre dictionnaires des prononciations et dictionnaires des significations, censés fournir les différentes acceptions des mots.

Références

- Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin & New York: Mouton de Gruyter.
- Boltz, William. 2007. Characters and words in Middle Chinese: lexical data of the Middle Chinese rime dictionaries. *Cahiers de Linguistique – Asie Orientale* 36.1: 37-60.
- Chen, Pengnian (陳彭年) (Song dynasty), and Hsin-hsiung Chen (陳新雄). (eds.) 2001. *Xin Jiao Songben Guangyun* 新校宋本廣韻 [Nouvelle révision du Guanyun des Song]. Taipei: Hungyeh Publishing Co., Ltd.
- Ding, Fubao (丁福保) (1874-1952). 1988[1928, 1932, 1959]. *Shuowen Jiezi Gulin* 說文解字詁林 [La forêt des explications sur le Shuowen jiezi], 20 vols. Beijing: Zhonghua Book Company.
- Feng, Yan (封演) (Tang dynasty). 1933. *Feng Shi Wenjian Ji* 封氏聞見記 [Notes de Maître Feng sur ce qu'il a vu et entendu]. Beijing: Yenching University Library.
- Jiang, Liangfu (姜亮夫). 1955. *Ying Ya Dunhuang Yun Ji* 瀛涯敦煌韻輯 [Collection des rimes de Dunhuang]. Shanghai: Shanghai Press.
- Liu, Fu (劉復) et al. 1963. *Shi Yun Huibian* 十韻彙編 [Compilation des dix ouvrages de rimes]. Taipei: Student Book.
- Lung, Yu-Chun (龍宇純). 1968. *Tang Xie Quanben Wang Renxu Kanmiu Buque Qieyun Jiao Jian* 唐寫全本王仁昫刊謬補缺切韻校箋. Hong Kong: The Chinese University of Hong Kong.
- Norman, Jerry. 1988. *Chinese*. Cambridge & New York: Cambridge University Press.
- Norman, Jerry, and W. South Coblin. 1995. A new approach to Chinese historical linguistics. *Journal of the American Oriental Society* 115.4:576-584.
- Ozaki, Yûjiro (尾崎雄二郎). 1980. *Chugokugo On'inshi no Kenkyû* 中国語音韻史の研究 [Etude sur l'histoire de la phonologie de la langue chinoise]. Tokyo: Sobunsha.
- Sugimoto, Tsutomu (杉本つとむ). 1972. *Kanji Nyûmon: Kanroku Jisho to Sono Kôsatsu* 漢字入門：干祿字書とその考察 [Initiation aux caractères chinois: le Ganlu zishu]. Tokyo: Waseda University Press.
- Takata, Tokio. 2004. The Chinese language in Turfan with a special focus on the Qieyun fragments. *Turfan Revisited: The First Century of Research into the Arts*

- and Cultures of the Silk Road*, ed. by D. Durkin-Meisterernst et al., 333-340. Berlin: Dietrich Reimer Verlag.
- Wang, Lien-tseng. 1957. Un dictionnaire phonologique des T'ang: Le «Ts'ie yun corrigé et complété» de Wang Jen-hiu. *T'oung Pao* 45:52-150.
- Xu, Xuan (徐鉉). 1977[1963]. *Shuo Wen Jie Zi* 說文解字 [*Expliquer les graphies et analyser les caractères*]. Beijing: Zhonghua Book Company.
- Yan, Yuansun (顏元孫) (?-714). *Ganlu Zishu* 干祿字書 [*Le livre des caractères pour un poste officiel*]. *Houzhi Buzu Zhai Congshu* 後知不足齋叢書 n° 2, comp. by Yanjue Bao 鮑延爵 (1830).
- Yu, Nae-wing (余迺永). 1993. *Xin Jiao Huzhu Songben Guangyun* 新校互註宋本廣韻 [*Nouvelle révision avec renvois du Guanyun des Song*]. Hong Kong: The Chinese University of Hong Kong Press.
- Zhou, Zumo (周祖謨). 1983. *Tang Wudai Yunshu Jicun* 唐五代韻書集存 [*Collection des livres de rimes conservés de l'époque des Tang et des Cinq dynasties*], 2 vols. Beijing: Zhonghua Book Company.

The Development of Rhyme Books into Dictionaries

Françoise Bottéro

CNRS-EHESS-INALCO

The *Qieyun* 切韻 (601) occupies an important place in the history of Chinese phonology, since it is the first known systematic work to give pronunciations for Chinese characters. As it offers a new and original way of classifying characters, it also plays an important part in the history of Chinese lexicography. In this paper, I shall give a detailed description of the particularities of the oldest *Qieyun* fragments found in Dunhuang and compare these fragments with later augmented versions down to *Kanmiu buque Qieyun* (706). I shall show how the addition of semantic glosses, which were scarce in Lu Fayan's (581?-618?) *Qieyun* (most rhymes and current expressions went unglossed), as well as of dots, used to separate groups of homophonous characters under each rhyme, came to transform rhyme books into "real dictionaries", providing more and more information on the written word.

Key words: rhymes, classification of characters, tones, graphic variants

L'étymologie graphique et l'héritage des concepts fondamentaux de la culture chinoise

Shun-chiu Yau (游順釗)

CNRS-EHESS-INALCO

Cet article réfute l'étymologie, idéologiquement orientée, du caractère *he* 和 'harmonie' diffusée par les médias chinois, selon laquelle il serait composé des graphies pour 'céréale' 禾 et 'population' 口. Il expose l'état de la question sur l'origine graphique de ce caractère et montre que, selon tous les travaux qui lui ont été consacrés, sa forme originale est en réalité celle d'un instrument de musique et que l'idée d'harmonie sociale qui en dérive procède de celle de polyphonie musicale.

Mots clé: étymologie, clé culturelle

À la cérémonie d'ouverture des Jeux Olympiques à Pékin en août 2008, le caractère *he* 和 'harmonie' a été présenté à des millions de spectateurs et téléspectateurs peu familiers de l'écriture chinoise et est devenu en une nuit un de ces très rares idéogrammes connus dans le monde entier, comme le *zhong* 中 de 'Chine'. Ce caractère, qui est le composant principal dans la formation de mots chargés de signification pour la culture chinoise, tels que 'paix' (*he-ping* 和平) et 'équilibre' (*zhong-he* 中和), est aussi porteur de valeurs sociales et morales universelles, particulièrement en ces temps de crises et de tensions internationales.

Les raisons expliquant un tel choix pour la cérémonie d'ouverture des Jeux Olympiques sont multiples. Outre qu'il sert la promotion de relations pacifiques entre les peuples et les nations, le caractère *he* 和 constitue un *leitmotiv* de la presse chinoise depuis 2007, particulièrement depuis que le président Hu Jintao a fait de la nécessité de «promouvoir l'harmonie sociale» le thème principal de son rapport présenté en octobre de cette même année devant le 17^e congrès du Parti.

Quelques mois avant le congrès, Hu Jintao se rendit au chevet d'un membre du Parti hospitalisé du nom de Fang Yonggang 方永剛, connu pour être un «théoricien innovateur modèle». À cette occasion, celui-ci se lança dans une exégèse du caractère *he* 和, disant en s'adressant au président: «Je pense que votre théorie sur l'harmonie surpasse toutes celles de vos prédécesseurs. Je l'exposerai de la façon suivante: le caractère *he* 和 'harmonie' est composé de deux parties, *he* 禾 (épi) et *kou* 口 (bouche).

He 禾 signifie ‘céréale’, il représente la nourriture. *Kou* 口 représente la population. En d’autres termes, l’harmonie signifie permettre à toute la population de manger à sa faim. C’est le ‘droit de vivre’ que nous avons proposé.» Hu Jintao approuva de la tête à plusieurs reprises, ce qui signifie que, de son point de vue, *he* 和 ‘harmonie’ signifie être correctement nourri.

La visite de Hu Jintao à l’hôpital fit l’objet d’un reportage de l’agence Chine nouvelle qui parut à la une du *Quotidien du peuple* de Pékin et du quotidien *Guangming ribao* de Shanghai le 3 avril 2007. Les propos de Fang Yonggang furent ensuite repris par toute la presse chinoise (j’ai moi-même eu connaissance de cet épisode par un hebdomadaire de Hong Kong). En raison de sa grande diffusion en Chine et auprès des Chinois d’outre-mer, cette étymologie idiosyncratique et sauvage pose un problème à ceux qui travaillent sur ou s’intéressent à l’étymologie graphique chinoise.

Rappelons d’abord les faits historiques. L’ouvrage de référence de Yu Xingwu (1996:1426) résume de la façon suivante l’évolution de la forme du caractère 和 *he*: «La forme ancienne du caractère est composée de 侖 et de 禾, 侖 étant la forme simplifiée de 𪛗, qui elle-même est la forme originale de 𪛖. Le *Shuowen* 《說文》 interprète 𪛖 comme 調 ‘équilibrer, moduler’ et dit qu’il se lit comme 和 qu’il interprète comme 相應 ‘répondre’. En réalité, 和 est la forme simplifiée de 𪛖, les deux formes étant à l’origine interchangeables, et 𪛖 est dérivé de 𪛗».¹

Un large consensus sur l’étymologie graphique de ce caractère s’est établi sur la base d’un article que lui a consacré Guo Moruo ([1931, 1952], 1982) et dont Yu Xingwu fait de longues citations dans son *Recueil de commentaires* cité ci-dessus. Guo Moruo écrit :

«Dans les inscriptions oraculaires, on trouve le caractère *he* 𪛖 ... Luo Zhenyu 羅振玉 a raison de dire qu’il est ‘composé de 𪛗’. En effet, il est composé de 𪛗 [plus la partie inférieure du caractère], qui est une forme iconique de la flûte de Pan.»

«Zhuangzi 莊子 dans son *Qiwulun* 《齊物論》 dit: ‘*Ren lai* 人籟, la flûte humaine [flûte à trois trous] est faite de tubes de bambou assemblés.’ Elle correspond à la forme de la graphie 𪛗, qui est la clé du caractère *he* 𪛖.»

«Si nous comprenons le caractère 𪛗, nous comprenons le caractère 𪛖. À l’origine 𪛖 a dû désigner un instrument de musique. De l’harmonie des sons

¹ Le texte original de cette citation est: “字從‘侖’，從‘禾’，而‘侖’實為‘𪛗’之省，此即《說文》‘𪛖’之初文。《說文》訓‘𪛖’為‘調’，謂‘讀與和同’。《說文》訓‘和’為‘相應’。實則‘和’即‘𪛖’之省，初本同文。而‘𪛖’則為‘𪛗’之孳乳字。”

dans la musique est dérivé le sens de 's'harmoniser', 's'accorder', tandis que de la résonance musicale est dérivé le sens de 'répondre'. De même, le caractère *yue* 樂 'musique' désignait au départ la cithare chinoise, *qin* 琴, mais par extension, il en est venu à signifier aussi bien la musique que l'harmonie sociale. Par la suite, le sens dérivé devint courant, tandis que le sens original devenait obsolète. On n'a plus reconnu que le sens de 'musique' et d' 'harmonie sociale' dans le caractère de *yue* 樂 et oublié qu'il portait aussi l'image d'un instrument de musique. De même, on ne connaît plus que le sens de 's'harmoniser', 's'accorder' dans le caractère *he* 和 et on a oublié quelle forme d'objet il recèle. Or *he* 和 représente bel et bien un instrument de musique. Sa forme graphique originale est 龠, sa forme 和 est tardive. Ainsi à partir de la graphie 龠 nous pouvons connaître 龠 et, réciproquement, à partir de 龠 nous connaissons 龠». ²

D'après cette étude de Guo Moruo, et de l'avis de bien d'autres spécialistes, il est clair qu'à l'époque des inscriptions oraculaires, au XIII^e siècle avant notre ère, le caractère qui s'écrit aujourd'hui 和 s'écrivait avec le radical 龠 qui représente une flûte à plusieurs tubes. Par extension, il en vint à signifier l'harmonie dans les sons et la musique. Aussi les gens de la dynastie des Zhou orientaux en vinrent-ils très tôt à déduire de l'harmonie des sons la philosophie suivante: «Quand les sons se ressemblent tous, il n'y a rien à entendre. Quand les couleurs sont toutes les mêmes, il n'y a pas d'art». ³ La graphie *he* 禾 à gauche dans 和 est un composant phonétique qui ne doit pas être lu pour sa valeur sémantique. On peut citer d'autres cas semblables, comme ceux de *dan* 啖 'manger, nourrir' et *ti* 啼 'pleurer ; chanter (en parlant des oiseaux)' dont les composants respectifs *yan* 炎 'feu, chaud' et *di* 帝 'empereur' à droite n'ont rien à voir avec la signification du caractère dans son ensemble. Personne ne dira que *dan* 啖 signifie 'feu dans la bouche' et *ti* 啼 'empereur dans la bouche'. Ce qu'a dit Fang Yonggang à Hu Jintao de la forme et de la signification du caractère *he* 和 'harmonie' nous fait penser à la façon dont Su Dongpo 蘇東坡 se moqua jadis de l'analyse graphique du caractère *bo* 波 'vague' que proposait Wang Anshi 王安石 («les vagues

² Le texte original de cette citation est: “卜辭有龠字，羅釋龠，謂「從龠者」是矣…諦視之，實乃從彡象形，象形者象編管之形也……莊子齊物論云「人籟則比竹是矣」，籟爲比竹，與龠之字 形正相一致……知龠則知龠。龠之本義必當爲樂器，由樂聲之諧和始能引出調義，由樂聲之共鳴始能引出相應義，亦猶樂字之本爲琴，乃引伸而爲音樂之樂與和樂之樂也。引伸之義行而本義轉廢，後人只知有音樂和樂之樂，而不知有琴絃之象，亦僅知有調和應和之和而不知龠之爲何物矣。然龠固樂器名也。當以龠爲正字，和乃後起字。蓋由龠可以知龠，由龠亦可以返知龠也。”

³ Citation du *Guoyu* 《國語》 [Discours des royaumes], chapitre *Zhengyu* 《鄭語》 [Discours de l'État de Zheng], ouvrage historique de l'époque *Chunqiu* (722-481 avant notre ère).

sont la peau de l'eau» 波爲水之皮) en présentant lui-même une analyse ridicule du caractère *hua* 滑 'lisse' («la douceur est l'os de l'eau» 滑爲水之骨).⁴ Ce genre d'étymologie déformée et vulgaire du caractère 'harmonie' pousse à négliger son contenu philosophique mis en avant par des générations de savants. Ce caractère signifie l'harmonie des contrastes, une harmonie qui ne peut être obtenue par la seule satisfaction de besoins matériels.

Utiliser des caractères hors de leur contexte et travestir leur étymologie à des fins intéressées ne date pas d'hier. Au XVIII^e siècle déjà, les Jésuites interprétaient certains caractères comme des preuves que les Chinois de l'Antiquité avaient connu l'enseignement de la Bible. Dans une *Lettre de Pékin sur le génie de la langue chinoise et la nature de leur écriture symbolique comparée à celle des anciens Egyptiens*, attribuée au Père Cibot et publiée en 1773, on lit par exemple que «le caractère de *barque*, *vaisseau* 船 est composé de la figure de *vaisseau* 舟, de celle de *bouche* 口 et du chiffre *huit* 八, ce qui peut faire allusion au nombre de personnes qui étaient dans l'Arche [de Noé] ... Le caractère *lan* 婪, *convoiter* est composé de deux *mu* 木, *arbre*, au milieu desquels est le caractère de *nü* 女, *femme*; ... cela s'accorde bien avec le péché d'Eve». ⁵ Si l'auteur cité prend la précaution de préciser «qu'on ne doit prendre ce (qu'il dit) que comme des conjectures», d'autres après lui se sont montrés plus affirmatifs. J'ai moi-même entendu l'étymologie biblique de *chuan* 船 de la bouche de prédicateurs de Hong Kong.

Loin de ces interprétations *pro domo*, l'étymologie populaire, par laquelle «le sujet parlant se fondant sur certaines ressemblances formelles rattache consciemment ou inconsciemment une forme donnée à une autre forme avec laquelle elle n'avait aucune parenté génétique»,⁶ se rencontre dans tous les types de langage oraux, écrits et gestuels.

À l'époque où je faisais mes recherches sur la création du langage auprès des sourds chinois au début des années 80, j'avais observé que le geste pour WC s'effectuait de la manière suivante: écarter les trois derniers doigts d'une main en extension pointant obliquement vers le haut et index et pouce formant un arc. Lorsque je demandai l'origine de ce geste à mes informateurs, ils me répondirent, très sûrs d'eux, que son étymologie était transparente. Selon eux, les trois doigts en extension formaient le pan d'un toit, tandis que le cercle formé par le pouce et l'index représentait un trou. Le geste était pour eux l'image réaliste d'un lieu d'aisance traditionnel comme on en

⁴ Le caractère 滑 'lisse' est composé de la clé de l'eau à gauche et de l'élément 'os' à droite. Le caractère 波 'vague' est composé de la clé de l'eau à gauche et de l'élément 'peau' à droite.

⁵ Cité par Étiemble. 1966. *Les Jésuites en Chine, La querelle des rites (1552-1773)*, Paris: Julliard, 194-196. Les caractères ont été ajoutés par moi pour une meilleure compréhension du texte.

⁶ Jean Dubois et al. 1973. *Dictionnaire de linguistique*. Paris: Larousse.

rencontrait alors partout en Chine. Malheureusement cette interprétation n'avait rien à voir avec la réalité. Documents à l'appui, je découvris en effet que ce geste chinois est une traduction manuelle du mot anglais WC, imbriquant les deux lettres manuelles W et C.⁷

Un autre exemple d'étymologie gestuelle fantaisiste dans la langue des signes chinoise est celui du geste pour Shanghai qui s'effectue avec les deux auriculaires accrochés l'un à l'autre. Selon les sourds de Shanghai que j'ai interrogés, ce geste représentait une ancre, et, comme Shanghai est un port, il signifiait le port de Shanghai, et par extension la ville elle-même. L'explication la plus probable de ce geste est cependant qu'il représente l'initiale manuelle S, non pas telle qu'on la connaît aujourd'hui en Chine,⁸ mais telle que la représente la langue des signes anglaise, les écoles de sourds de Shanghai ayant été jusqu'au milieu du XX^e siècle tenues par des missionnaires anglais.⁹

Même si elles sont erronées, les étymologies populaires de ce genre ne font de mal à personne. Il n'en va pas de même lorsque l'explication de l'origine d'un caractère sert des fins politiques. L'exégèse de *he* 和 'harmonie' par Fang Yonggang conforte en effet la ligne officielle de la propagande chinoise qui est de faire correspondre aussi étroitement que possible la substance idéologique de l'harmonie avec l'idée de gains matériels et, ce faisant, de privilégier la réforme économique au détriment du progrès politique et social. C'est la raison pour laquelle il importe d'en exposer la fausseté avant qu'elle ne se répande dans le public non sinophone, toujours friand de petites anecdotes amusantes sur l'écriture chinoise.

⁷ Yau (1992:175-176).

⁸ Yau (1978:116).

⁹ Yau (1992:196-197).

Références

- Guo, Moruo (郭沫若). 1982. *Shi he, yan* 釋和言. *Guo Moruo QuANJI* 郭沫若全集 • 考古編 [*Euvres complètes de Guo Moruo*, Vol. I: *Archéologie*]. Beijing: Science Press.
- Jiang, Meiming (姜昧茗). 2002. *Shilun Xian Qin zhixue zhong he gainian de fazhan* 試論先秦哲學中「和」概念的發展 [À propos du développement de la notion de *he* 'harmonie' dans la philosophie pré-Qin]. *Lishi Wenxian yu Wenhua Yanjiu* 歷史文獻與文化研究 [*Recherches sur les documents historiques et la culture*], 218-239. Wuhan: Hubei Lexicographical Publishing House.
- Yau, Shun-chiu. 1978. *Les signes chinois, langage gestuel standard des sourds chinois*. Paris & Hong Kong: Langages Croisés.
- Yau, Shun-chiu. 1992. *Création gestuelle et débuts du langage*. Paris & Hong Kong: Langages Croisés.
- Yu, Xingwu (于省吾). (ed.) 1996. *Jiagu Wenzi Gulin* 甲骨文字詁林 [*La forêt des commentaires des graphies sur os et carapaces*]. Beijing: Zhonghua Book Company.

Graphic Etymology and Heritage of Fundamental Concepts of the Chinese Culture

Shun-chiu Yau

CNRS-EHESS-INALCO

This article disproves the ideologically motivated etymology of the character *he* 和 “harmony” as publicized in the Chinese press, according to which the character is regarded as consisting of two parts, *he* 禾 “grains on stalk”, foodstuffs, and *kou* 口 “mouth”, the populace. In other words, harmony is construed to mean letting the whole population have enough to eat. This is an attempt to distort its graphical etymology so as to fit a specific ideological pattern, thus degrading the notion of social harmony to the mere satisfaction of a biological need. The fact is that in Oracular Bones, the present character 和 had a radical 龠 representing a wind instrument with arranged tubes; the meaning was extended to harmony in sound and music. Evidently the people in the Western Zhou Dynasty had, that early on, derived from harmony in sound and music the following philosophy: “When sounds are all the same, then there is no music to hear. When colours are all the same, then there is no design.”

Key words: etymology, cultural key

Graphic Designs: A New Approach to Palaeographic Analysis*

Ken-ichi Takashima (高嶋謙一)

University of British Columbia

The term “graphic design” refers to a kind of mental plan the original Chinese character makers must have had, and the scribes who succeeded them are presumed to have accepted, modified, or even rejected to create afresh. The initial job was to devise ways to express words or morphemes graphically. By analyzing graphic component(s) and a configuration, it is often possible to understand how the plan may have been laid out. From such a vantage point as described, this paper takes up the graph 𠂔 which has not yet been correctly interpreted. The paper proposes that it be transcribed as *shou* 首 ‘head’ used to stand for the word *dao* 導 ‘to lead’, while another scribe created the graph 𠂔, 𠂔, and 𠂔. Each of these forms subsequently came to be used in different contexts. This is examined in the perspectives of grammatology, morphology, and syntax, but the main theme of the paper is to discuss how by paying particular attention to the design or plan of a graph we can obtain a better way to reading texts in the oracle-bone script.

Key words: graphic design, palaeography, oracle-bone inscriptions

1. Introduction

This paper examines a few palaeographs that are associated, either directly or indirectly, with the words meaning “head (首),” “road (道),” “to lead (導),” and “dagger axe (戈)” in Shang oracle-bone inscriptions (abbrev. hereinafter as OBI). How they are related will be made clear in due course, but the major thesis of the paper deals with the question of what I refer to as “graphic designs” on the part of the character makers whose legacies were succeeded by the Shang, Zhou, and later scribes.

* An earlier version of this paper was presented at the Early China Seminar, Columbia University, New York, March 8, 2008. On that occasion I put a playful title to the paper calling it “Reconstructing the Lost ‘Head’ 𠂔 (首): A Palaeographical Odyssey,” but more seriously the present title more accurately describes the main thesis of the paper. I am grateful to Professors John Major, David Branner, Li Feng, Chen Kuang-yu, Chen Chao-jung, and Ms. Zou Changhua (a writer and poet with interest in *guwenzi*) for their comments, questions, and criticism.

Methodologically, the first problem to be considered is how to accurately transcribe the palaeographs into characters with known sound and meaning. A good many of them, when analyzed into stroke-by-stroke components (筆畫構成) or graphemes (字素), can be transcribed into the *kaishu* 楷書 forms. This may be referred to as “direct transcription” (abbrev. DT) (直接隸定) which is deemed the basic work in historical palaeography. In earlier stages of such transcription work, DT¹ is often referred to as *guwen liding* 古文隸定, a processual work whose history goes back to the Former Han when the so-called *kedou* 科斗 ‘tadpole’ script was transcribed into the then current *lishu* 隸書 ‘clerical script’.² Once the palaeographs are transcribed into the clerical-script forms with their regular correspondences with the *kaishu* forms, we can examine them in functional terms individually such as phonetic and/or signfic in both narrow and broad sense.³

¹ The so-called *kuanshi* 寬式 ‘laxed way (of transcription)’ should not, in principle, be included under the notion of DT because it is a compromise between the precise and imprecise, as well as between the known and unknown (and hence an object of inquiry). For example, if one says that the graph 𠂔 is *yǎn* 衍 ‘extend; superfluous’ according to the *kuanshi* transcription, that would be going against DT at least on two counts: 𠂔 is different from 𠂔 or 水; also the element 亅 is not in the original. Thus, DT should strictly be 𠂔.

² References to this sort of processual work appear in the spurious preface to the *Shangshu* 尚書 attributed to Kong Anguo 孔安國 (fl. 80 B.C.), but the preface is known to have been faked sometime in the Wei-Jin 魏晉 period (3rd-4th c.). Nevertheless, the relevant passage which reflects an understanding of such a processual work, contemporary with the said period reads as follows: ... 科斗書廢已久, 時人無能知者; 以所聞伏生之書, 考論文義, 定其可知者, 爲隸古定, 更以竹簡寫之 ‘the “tadpole” script already fell out of use long ago, and there was none of the contemporary who knew how to read it. On the basis of what I have heard in the writing of Fu Sheng [i.e., the 29 *pian* of the *Jinwen Shangshu* 今文尚書], I have examined and discussed the meaning of the [ancient] characters, and have determined those that can be known, fixing them as transcriptions of the ancient characters in the clerical script and, furthermore, copied them onto the bamboo slips’. Descriptions similar to this are also found in Lu Deming’s 陸德明 (556-627) preface (序錄) to his *Jingdian shiwen* 經典釋文 (*juan* 1), and in the *Shangshu yinyi* 尚書音義 section (*juan* 3) where he comments on the expression *li gu* 隸古 as follows: 隸古謂用隸書寫古文 ‘By *li gu* is meant to use the clerical script to write the *guwen*’. As for the expression *liguding* 隸古定 in the spurious preface quoted above, Qiu Xigui (1988:78) notes that some scholars think the 隸古 and 定 should not be *liandu* 連讀 ‘read together’, but here Qiu reads them together. It would certainly be hard to make a coherent sense if we follow such an alternate reading. The current use of the term *liding* 隸定, rather than 隸古定 or even **kaiding* 楷定, has such a historical background as described above.

³ For example, we can examine how strict or lenient a certain phonetic is to represent another word for which it was originally intended (e.g., *jie*/**kradx*/**krij* 皆 ‘all’ for the word *kai*/**khra*d(?)/**krij*(?) 楷 ‘regular, neat; model’) and how broad a certain semantic category can encompass. The semantic category ranges from subsuming only a few characters to the extreme case under which superfluous embellishments might have been intended. This by itself is a

But how do we go about doing this sort of analytical work when no transmitted authorities exist? The question needs to be considered in the light of relevant palaeographical sources and larger inscriptional context, specifically paying attention to the syntactic environments in which such transcribed graphs — frequently historically discontinuous ones (abbrev. HDC) — appear. This should be considered in conjunction with a morphological analysis of the restored words in HDC. Such an approach often proves to be potent in decipherment but, if not, at least susceptible of obtaining some answer to the question.

Thus, the *liding* work is the first step we must take in order to read any text in ancient scripts. It is a disciplinary practice efficacious to adopt as has been advocated by Noel Barnard (1996:xx; cf. also 1973:32) because “[it] keeps us in direct contact with the archaic character structure” so that we may be able to “avoid the unwitting use of erroneous or unacceptable transcriptions of problematic graphs.” This is easier said than done, but if pursued rigorously and scrupulously it does have its value as the present paper attempts to show. Particularly, I should like to develop Barnard’s “archaic character structure” in terms of what I call “graphic design” (形設/造字意圖) or “graphic intent” (構形之旨/構形意圖) as might have been in the minds of those who created characters. As a way of illustration, the paper discusses various graphs that are relevant to the analysis of the OBI graph 𠂔, suggesting its transcription, in its final analysis, into the *kaishu* form as: 首 *shou*/**sthjəg*^w?/**hlju*?,⁴ the word for “head” which is used to stand for the word *dao*/**dəg*^w*s*/**lus* 導 ‘to lead, guide’.

2. The *liding* 隸定 and its ramified problems

The actual *liding* work is sometimes rather complex. There is the problem of sorting out and establishing such things as follows:

rather complex issue requiring independent investigation. Here the “superfluous embellishment” may be comparable to the Chinese terminology *shibi* 飾筆. This is taken to function neither phonetically or semantically, but in terms of a writing system one has to be careful in making such a judgment. Liu Zhao (2006:26), for instance, thinks that the grapheme 𠂔 is a *shibi*, but I disagree. This grapheme and its simplified form 一 can be construed as a graphic element used to convey a “binding” sign in several graphs such as, for instance, in 𠂔 (= *di* 帝; cf. *di* 締 ‘to bind’), a topic I will be discussing later.

⁴ Where appropriate I have provided the reconstructed OC forms — the first one is Li Fang-Kuei’s (1971) and the second one is Baxter’s (1992), both of which have been modified with little consequences.

- (1) variant graphs (異形 or 異體)
- (2) structural variants (異構)
- (3) alternate forms (或體) or transformation (變形)
- (4) old-seal forms (古籀 or 大篆) as distinct from small-seal forms (篆文 or 小篆)
- (5) *guwen* 古文 ‘ancient forms’ as distinct from current forms (今文), where “current” refers to the clerical script (隸書) of Han time
- (6) “odd” characters (奇字), where “odd” is in reference to the *guwen* 古文 in that they are *guwen* but different from it (古文而異者)
- (7) the so-called vulgar forms (俗體) as distinct from orthodox forms (正體)
- (8) abbreviated forms (省體, 省形)
- (9) embellishing or ornamental forms (飾文) sometimes including the examples with only ornamental strokes (飾筆) detected

In addition to the above,⁵ we have to elucidate such processes as follows:

- (10) graphic “merger” (形合) and “split” (形裂) in the history of Chinese script
- (11) graphic “misinterpretation” (形誤) or “reanalysis” (形之再析) in the history of Chinese script⁶

These should be examined in the successive stages of the evolution of the Chinese script from Shang OBI, through Zhou (including *Chunqiu* and *Zhanguo*) bronze inscriptions and bamboo-tablet and silk writings of the *Zhanguo* period, down to the Han clerical script. Since it is impossible to address all these complexities in this paper, I have narrowed down the *liding* work and the related issues to a smaller scale, limiting our main corpus to the Shang OBI (ca. 13th-11th c. B.C.).

3. Analysis of the graph 𠄎

Sōran (#3695, p.453) gives about ten different interpretations (*jishaku* 字釋) for the bone graph 𠄎. At first sight, one of them might be regarded as the *liding*, namely, 冒 (Chang Ping-ch’üan 1967:701). However, this is a hybrid of *liding* and “non-transcription.” That is, the bottom element, *mu* 目, is no doubt a case of the *liding*, but the upper element

⁵ Discussed in more detail in Xu Zaiguo (2002) and Huang Xiquan (1990), the former more extensively than the latter. Also, Liu Zhao (2006) is quite informative.

⁶ Theoretically speaking, (11) is not limited to something entirely diachronic. It could be a synchronic process as well; that is, in terms of “graphic design” of a character there could have been multiple character makers synchronically or geographically with different traditions. It is reasonable to assume the possibility of mutual influences among themselves.

cannot be said to be a transcription at all as it is the same as the bone graph. Therefore, this does not qualify to be a *liding*. Let us also consider the bone graph 𠄎 transcribed as 𠄎 (Yen Yi-p'ing 1951:230-231). As before there is no problem for the lower element being *mu* 目, but the “claw” element 爪 goes back to 𠄎 which is quite different from 𠄎. I find no evidence for the forms 𠄎 and 𠄎 being variant graphs (異形 or 異體), structural variants (異構), or alternate forms (或體)/transformation (變形), odd forms (奇字), or a vulgar/orthodox pair (俗體/正體). So the graph 𠄎 is not acceptable as a correct *liding* either. Our task, then, is to do a historical palaeography of the bone element 𠄎 by asking how it should be transcribed in the *kaishu* form when combined with the 𠄎 (= *mu* 目 ‘eye’) element.

Seemingly the most natural interpretation is to consider the element 𠄎 being the same as that found in such graphs as 𠄎 (雨), 𠄎 (𩇔), 𠄎 (雪), 𠄎 (𩇔), etc. Ding Shan (1960:36-37, quoted also in Yu Xingwu 1996:1.579) is the first scholar to transcribe the bone graph 𠄎 as 𩇔. Ding surmises that this HDC 𩇔 stands for the word *meng* 濛 by analyzing 𩇔 as: 从雨从目, 象雨點迷目形 ‘(the graph 𩇔) consists of “rain” and “eye”, the form of rain drops obscuring eyesight’. In support of this interpretation, he quotes a passage from the *Shijing*, Dongshan 東山 (Mao 156) “零雨其濛” (Drizzling rain obscures vision). This interpretation hinges on the *huiyi* 會意 analysis of the bone graph, further invoking a completely different graph, 濛, from which the reading is obtained. We cannot dismiss this reading out of hand, because some independent, albeit not completely unassailable, evidence might be available. This is to regard the graph 𠄎 not as a *huiyi* but as a *xingsheng* 形聲 character with the phonophoric *mu* 目 (𠄎) and the signific *yu* 雨 (𠄎 is assumed for the time being to be an abbreviated form [省體] of 𠄎). That is, *mu* in OC is something like **mjək^w/*mjuk*, whereas *meng* is **munj/*monj*. Here, the nasal initial and the rime are not entirely disparate,⁷ though the codas are problematical. So Ding Shan’s interpretation has shortcomings. But there could have been a graph that combined the element 濛 or 蒙⁸ and 戈 ‘dagger axe’. In other words, within the graphic configuration of 𠄎, the element 𠄎 may have been phonetic and the 𠄎 (戈) element signific. However, in terms of “graphic design” presumably intended by the maker of this graph, such a combination would be hard to comprehend because the significance of the 𠄎 (戈) ‘dagger’ is left unexplained. The only way to discover what the significance might have been is to conduct an independent study of the graphic function of the dagger grapheme. This is presented in Appendix. The table provided in Appendix shows that none of the 45

⁷ Depending on how one interprets the rime of **munj/*monj*, it is not impossible to reconstruct **-anj^w*. If so, the **-ə-* and **-a-* vowels may not constitute a strict distinction in writing, though morphologically they seem to make a significant difference noticeably in minimal pairs (Pulleyblank 1973:118ff).

⁸ 濛 and 蒙 are mere variant with the same sound (*meng*) and meaning (obscure).

characters found in the *SW* that have the 戈 (𠂇) grapheme to which Xu Shen assigns the phonetic function. It is thus reasonable to assume that the 𠂇 (𠂇) grapheme used in the graph 𠂇 was semantically motivated and that the 𠂇 constituent may have served as phonophoric — read according to Ding Shan *meng*/**muj*/**moŋ*. The *liding* of such a graph could be reconstructed like: (*𠂇>) *𠂇. As will be explained shortly, however, there are problems with this analysis and, hence, this reconstruction.

As opposed to the “rain” interpretation of the grapheme 𠂇, there is yet another possible interpretation offered by Li Hsiao-ting (1965:14.4481). He suggests the *liding* of the OBI graph 𠂇 is *guo* 𠂇 (𠂇) ‘cut off head’. Although this character is not listed under its own heading in the *SW*, it is found under *guo* 𠂇 (𠂇) as a variant (或體): 𠂇 軍戰斷耳也。...从耳或聲。𠂇 𠂇或从首。 (In war, they cut the [left] ear [of an enemy as proof for slaying him]. ... [the graph] consists of ear, and the 或 element is phonetic. As for 𠂇, namely, 𠂇, it sometimes consists of 首 ‘head’ [rather than ear].) Li Hsiao-ting explains the graph 𠂇 as follows: 契文𠂇字作𠂇, 从此象纓繫首之形 ‘the graph 𠂇 in the bone inscription is written as 𠂇 which consists of the head tied with a cord’. I find this explanation very attractive particularly from the vantage point of the “graphic design” we are presently considering. That is, the semantic significance of *ge* 𠂇 (𠂇) is transparently revealed in that the dagger axe was used as a cutting instrument for severing the head. By contrast, Ding Shan’s interpretation seen in the above paragraph — namely, the graph 𠂇 could have been read like 濛/蒙 *meng*/**muj*/**moŋ*, and the presumed *liding* would be reconstructed like *𠂇 — seems to resist any cogent explanation for the semantic significance of the dagger grapheme. That is, the “graphic design” of *𠂇 is very hard to interpret because no vital or organic relationship between 濛 and 戈 or 𠂇 seems available. Furthermore, there are the phonological difficulties encountered in the **mjak*^w/**mjuk* 𠂇 associated with **muj*/**moŋ* 濛/蒙.

To give my own interpretation, the *liding* of the OBI graph 𠂇 is not 𠂇, but a structural variant (異構 — §2, [2]) of *shou* 首/𠂇 (<*zhuanwen* 𠂇). Some scholars take the grapheme 𠂇 (𠂇) in the graph 𠂇 as identical with the 𠂇 grapheme in such a character as *liu* 𠂇 (流) ‘flow’.⁹ The grapheme in question is indeed shared by 𠂇 and 𠂇 which are also phonologically comparable: 流, being a *youbu* 幽部 word, can be reconstructed like OC **rjæg*^w/**C-rjuw* and, since *shou* 首 (𠂇) is reconstructed like

⁹ One would think that such characters as *shu* 梳 ‘comb’ and *shu* 疏 ‘coarse grain’ also share the phonophoric 𠂇 (𠂇), but they belong to the *yubu* 魚部 with the *-*ag*/*-*a* rime. I have therefore excluded them. I do not know what might have happened between the *youbu* 幽部 and *yubu* 魚部 or what to make of the relationship. One possibility is the “merger” sensitive to the time period in which the same grapheme 𠂇 (𠂇) was used to write the words *shu* 梳 ‘comb’ and *shu* 疏 ‘coarse grain’.

*sthjəg^w?/*hlju?, it is plausible that the grapheme 𠂔, whose *kaishu* form being 川,¹⁰ served as the phonetic. However, from the standpoint of the graphic design once again, the grapheme 𠂔 (川) must have represented amniotic fluid (羊水) since it occurs underneath the upside down or born-head-first infant 子 (𠂔) coming out of the womb in the case of 流, to wit: 𠂔. But in the case of 𠂔 (首), the same grapheme 𠂔 must have stood iconically or indicatively for “hair.” The absence of the grapheme 𠂔 or any of its variants such as 𠂔, 𠂔 in the bone form 𠂔 suggests that 𠂔 had nothing to do with “water” represented by the same grapheme or its variants. This is very important because the possibility we entertained that 𠂔 is related to “rain” is ruled out, even though 𠂔 is a constituent element of the graph for rain (𠂔). One can also observe the elucidation with the 𠂔 grapheme to make the grapheme 𠂔 (𠂔) unambiguously clear that it represents “head.” In other words, what was originally the grapheme 𠂔 representing “flowing water” in such OBI graphs as 𠂔 (州), 𠂔 (徇 or 衍), 𠂔 (災), etc. and the grapheme 𠂔 representing “hair” in the *zhuanwen* 𠂔 (> 首) **merged** (cf. §2, [10]). Here, however, the latter grapheme (𠂔) was further reanalyzed as having the phonetic function. If such a phonetic reanalysis is not an entirely appropriate explanation for the addition of the grapheme 𠂔 to 𠂔, the element 𠂔 (川) with its one reading of *rjəg^w/*C-rjuw was specified or “boasted” (i.e., given prominence). The grapheme 𠂔 in the OBI form 𠂔 — no *zhuanwen* but can easily be isolated as 𠂔 > 𠂔 — also became reanalyzed as “amniotic fluid with an upside-down or born-head-first infant”¹¹ with the “boasted” specification of the phonetic. The analysis just presented is supported by the fact that in OBI the head is also written like 𠂔 and, very interestingly, 𠂔, as well. This form has the hair grapheme on top of the head.¹² If so, such a reanalysis (cf. §2, [11]) or the phonetic “boasting,” as it were, may have been due to reasons that are probably diachronic than geographical.¹³ In fact, it is various issues in historical palaeography that we have so far been concerned and have accounted for how the graphic configuration 𠂔 could have yielded the *kaishu* form *shou* 首 ‘head’. Furthermore, I would suggest that the three vertical lines in 𠂔 are the

¹⁰ This is at least one possibility of the *liding* of 𠂔; another possible development is 𠂔, the details of which need further study. Here, however, there are at least two instances of graphic split (cf. §2, [10]).

¹¹ In terms of “graphic design,” the grapheme 𠂔 and the grapheme 𠂔 in the OBI script seem to have been quite distinct, the former representing liquid spattering/sprinkling (as in amniotic fluid) and the latter flowing (e.g., 𠂔 (州), 𠂔 (徇 or 衍), 𠂔 (災), etc. as mentioned above).

¹² Many specialists take this graph with the additional detail of a mouth as a variant of 𠂔. But their usages are different: the 𠂔 stands for the word *tou* 頭 ‘head’, while the 𠂔 is used as a proper noun (cf. the beginning of §3.1 below; also *Sōrui* 100.3).

¹³ This is not a foregone conclusion. In fact, if archaeological findings could yield more inscrip-tional data from different locations, they might show that the reanalysis in question was moti-vated by the infusion of different scribal traditions.

same grapheme as 𠂔 occurring on top of the bone graph for “head” in the OBI graph 𠂔. These vertical lines must have represented “hair.” On the other hand, the horizontal line on top of the grapheme 𠂔 (yielding 𠂔) may be interpreted as a “binding” sign. Such a sign is also found in the following graphs:

- 𠂔 (a variant of 𠂔; 祭 ‘to offer burnt sacrifice’ [sc. bound brushwood])
- 𠂔 (朋 ‘a coupled jade string’)
- 𠂔 (a variant of 𠂔;¹⁴ 帝 ‘Di, the Binder/Unifier’)
- 𠂔 (𠂔, a drawing of manacle with the [binding:] tying sign —; “to manacle” is represented by *zhi* 執 — OBI form 𠂔)
- 𠂔 (車 ‘chariot’)
- 𠂔 (𠂔 ‘bamboo tablets’; not really a horizontal line in the OBI graph, but there are two that are connected to each other, i.e., “a coupled string”)
- 𠂔 (a variant of 𠂔?, i.e., 眾 ‘masses’)
- 𠂔, 𠂔 (并 ‘combine’)

Li Hsiao-ting’s interpretation seen in the previous paragraph, i.e., 纓繫首之形 ‘drawing of the head tied with a cord’, is similar to what I have just presented. However, it is surprising that in spite of this good analysis Li (1965:8.2813) interprets the graph 𠂔 itself as *shi* 視 ‘see, examine’. This is because he made a wrong identification of the graph 𠂔 with a *guwen* form 𠂔. The right hand element of this is clearly 示 and cannot be traceable to 𠂔. Moreover, our interpretation differs from Li’s in other details, i.e., committing to interpret the vertical lines on top of the head as hair, and the horizontal bar on its top signifies the hair being bound, and this for ease in carrying the severed head. On the other hand, if the graph 𠂔 is indeed *shou* 首 ‘head’, how then do we explain the difference between this form and two other OBI graphs for *shou*, namely, 𠂔 and 𠂔? It is because, I would suggest, the “graphic design” was very different. The graph 𠂔 is a severed human head suspended from the above, while the latter graphs are not. This also explains the fact that the graph 𠂔 occurs as an element underneath the dagger-axe graph, 𠂔, while the other two head graphs never occur in that structural environment.

As for the reading of the graph 𠂔, there are two possibilities: one is that it represents the word *guo*/**k*^w*rək*/**k*^w*rik* 馘/馘 ‘cut off head’ and another one is *jia*/**krit*/**krit* 戛 ‘pike’ (an instrument used to sever the head at the neck). The former has some support from the *SW*,¹⁵ while the latter has support from the *DT*, though the structural environment

¹⁴ The bone graph is more frequently written like this rather than 𠂔; however, the horizontal line seems comparable to yet other variant graphs such as 𠂔 and 𠂔 in which the middle part indicates “binding.”

¹⁵ The nature of support is not entirely unequivocal because 或 is not a DT of 𠂔 (戈): the extra

in which the graph 𠂔 occurs (i.e., underneath the dagger-axe graph, 𠂔) is reversed in 𠂔 (𠂔). It is quite possible that this particular placement of 戈 under 𠂔 was the result of such a structural configuration as placing it underneath would produce the same graph as 𠂔. The words expressed by *jia*/**krit*/ **krit* 𠂔 and *guo*/**k*^w*rək*/**k*^w*rik* 𠂔 were different, and it would be no surprise that the “graphic intent” of the scribe were also different. As a supporter of the “synchronic evidential approach,”¹⁶ I tend to opt for the latter *jia* 𠂔 reading. Also, the predominant use of *jia* as a noun including a proper noun, as opposed to that of *guo* as a verb in transmitted texts seems to favor this reading. In order to interface the palaeographical analysis given above with linguistic analysis, we need now to turn to the context in which these graphs occur.

3.1 Syntactic context of the graphs 𠂔 and 𠂔 / 𠂔

We begin with some representative cases of the latter graphs that seem to represent the head (𠂔 / 𠂔) of some animal rather than that of human,¹⁷ but the eye is appreciably distinctive as it is in 𠂔/首 and 𠂔/首.

- (1) 旬𠂔 (有) 𠂔 (崇), 王疾首 (𠂔) 中日 (時) 雪. (Heji 13613)
In the decameron there was an (ancestral) curse; His Majesty (ailed:) had ailment in his head. Around noon time it snowed.
- (2) 疾首 (𠂔) 不佳丁. (Heji 13617)
... (is ailing:) is having ailment in the head; it is not (due to) (ancestor) Ding.
- (3) 貞子 疾首 (𠂔) 延 (Heji 13614)
Tested [the following proposition to seek for sapience from the numen of the

element 𠂔 is contained in the character 或, quite possibly having the phonophoric function, indicating the sound *guo*/**k*^w*rək*/ **k*^w*rək*, rather than *jia*/**krit*/**krit*.

¹⁶ The reader is referred to Takashima (2005) for a fuller exposition of this approach. To mention what it is very briefly here, it is used in contrast with Wang Guowei’s *erchong zhengjufa* 二重證據法 ‘two-layered evidential approach’ (Wang 1925) tacitly followed by a great majority of specialists, but with serious problems because information culled from transmitted OC texts are often anachronistic and distorted as I have tried to show in the paper just referred to. The synchronic evidential approach comprises two general aspects, of which the first is the use of excavated materials to augment and correct, if necessary, the material handed down by tradition — with the proviso that the excavated material be given precedence. The second is to explore and discover such phenomena as might have been hitherto unknown in the received traditions.

¹⁷ The graphic unit occurs in a few graphs, but one form, 𠂔, is of interest as this is commonly identified with *kui* 夔 ‘a mythological animal; ape-like creature’ or *nao* 獠 ‘a monkey’.

turtle]: Zi ... (is ailing:) is having ailment in the head and might continue (to suffer from it)

- (4) 甲辰卜出貞王疾首 (𠄎) 亡延. (Heji 24956, 24957)
Divining on the *jiachen* day, Chu tested [the following proposition to seek for sapience from the numen of the turtle]: His Majesty (is ailing:) is having ailment in the head, (but) will (certainly) not continue (to suffer from the ailment).
- (5) 丁亥卜般貞誓享首 (𠄎) 𠄎于雇. (Heji 13619)
Divining on the *dinghai* day, Que tested [the following proposition to seek for sapience from the numen of the turtle]: X will make an offering of the head (at or from?) Y and Gu.

The evidence is quite clear in that the word for “head” is consistently written as given in the parentheses. As for example (5), the translation is not certain, but we do know that the Shang steam cooked human heads in the *yan* 甗 bronze vessel.¹⁸ By contrast, the following examples show that the graph 𠄎 (including a few variants) with the three vertical lines representing hair is used as a proper noun:

- (6) 甲戌卜般貞翌乙亥王𡗗 (除)¹⁹ 首 (𠄎) 亡禍. (Heji 6032o)
Divining on the *jiaxu* day, Que tested [the following proposition to seek for sapience from the numen of the turtle]: On the next *yihai* day His Majesty will (be able to) expel/remove the Shou; there shall be no misfortunes.
- (7) 翌庚辰王往𡗗 (除) 首 (𠄎). (Heji 6033r)
On the next *gengchen* day His Majesty should go (in order to:) so that he will (be able to) expel/remove the Shou.
- (8) 貞庚申我伐易日. 庚申明霧. 王來𡗗 (除) 首 (𠄎) 雨小. (Heji 6037o)

¹⁸ The team of the archaeologists of Anyang Work Station of the Institute of the Chinese Academy of Social Sciences (2004:361).

¹⁹ Many specialists follow Yu Xingwu’s (1943:3.23a) interpretation that *tu/*dag/*la* 𡗗 (途) is a loan for the homophonous 屠 ‘to slaughter’. It seems to work well for this example and the rest of examples quoted hereunder. However, it is extremely awkward to apply this interpretation to such an example (Heji 67) as: 貞王勿往𡗗眾人 ‘Tested [the following proposition to seek for sapience from the numen of the turtle]: His Majesty should not go to slaughter (?) the royal masses and men’. It would make a far better sense to take the graph 𡗗 as standing for a word that can be taken in a neutral meaning such as *chu/*diag/*lrja* 除 ‘to do away with’ that encompasses both “withdraw” and “expel,” an interpretation deemed fit in different context. Applied to the above example, we would get: “... His Majesty should not go to withdraw the royal masses and men.”

Tested [the following proposition to seek for sapience from the numen of the turtle]: On the *gengshen* day when we launch an attack [against the Shou people], it will turn sunny. In the morning of the *gengshen* day it was misty. When His Majesty returned after expelling the Shou, (the rain was light:) it was raining lightly.

- (9) 王臣固曰 𠂔 (除) 首 (𠂔) 若. (Heji 11506r)

A servitor of His Majesty, having prognosticated, declared: ... (will be able to) expel the Shou people, and it will meet with (Di's) approval.

Although the graphs 𠂔 and 𠂔 are commonly considered to have represented the same word, *shou* 首 'head' (cf. fn. 12), we cannot confirm such was really the case. Xu Zhongshu (1988:994) says “《說文》以存髮之𠂔與不存髮之𠂔分別二部。二字實爲一字。” (*Shuowen* divides the graph 𠂔 with hair and the graph 𠂔 without hair into two separate headings, but the two graphs in fact **represent** just one word.) What Xu says might be correct, but as seen in the above examples, the two Shang predecessors (𠂔 and 𠂔) are clearly distinguished.²⁰ Unfortunately, whether this difference in usage reflects different words with different pronunciations cannot be determined. Xu Shen himself does regard them as being equivalent, to wit: 𠂔頭也, 象形 ‘𠂔 is head; it is a pictograph’ and 𠂔同, 古文百也, 𠂔象髮 ‘𠂔 is the same as 𠂔 which is the *guwen* form; 𠂔 (sc. 𠂔) depicts hair’. Apparently based on this *Shuowen* explanation, Zhu Ao 朱翱, a Song phonetic commentator to the *xiao Xu ben* 小徐本, has given the *fanqie* spelling 書九切 for both 首 and 𠂔. We seem to be at the mercy of this reading ever since.

Paying attention now to the context in which the graph 𠂔 occurs, we will find a satisfying solution to the problem of exactly what word the graph may have expressed. Let us consider the following examples:

²⁰ There are yet other graphs inscribed like 𠂔 (Heji 916o) and 𠂔 (Heji 22092). The latter is an abbreviation (省形) of 𠂔 (Huadong 304) according to Huang Tianshu (2005:59) who transcribes it as 首 'head'. Since both these forms occur after the verb *ji* 疾 'have sickness in ...', I believe Huang is correct. However, it is uncertain that the forms 𠂔 and 𠂔 should be transcribed as 首. Although further work is needed, I would tentatively transcribe it as *tou* 頭 'head'. The hypothesis I have in mind is that the bottom portion of these represents a transformation (變形) of *dou* 豆 (𠂔, 𠂔). As for the graph 𠂔 (to be distinguished from 𠂔, the bottom portion being phonophoric), it may be considered as a variant of 𠂔 based again on the fact that they occur after the verb 𠂔 which we have interpreted as standing for the word *chu* 除 'to do away with, expel; withdraw'.

- (10) 貞王往省牛. (Heji 11176)
Tested [the following proposition to seek for sapience from the numen of the turtle]: His Majesty should go to inspect cattle.
- (11) 貞勿往省牛. (ibid.)
Tested [the following proposition to seek for sapience from the numen of the turtle]: [His Majesty] should not go to inspect cattle.
- (12) 貞𠄎 (首 = 導) 牛百. (ibid.)
Tested [the following proposition to seek for sapience from the numen of the turtle]: [His Majesty] should (lead:) herd cattle numbering one hundred.
- (13) 寅卜𠄎貞𠄎 (首 = 導) 三千人伐. (Heji 7345)
Divining on the ... *yin* day, RUK²¹ tested [the following proposition to seek for sapience from the numen of the turtle]: (We) should lead three thousand men to launch an attack [against an enemy].
- (14) 王其𠄎 (首 = 導) 戌𠄎. (Heji 28026)
His Majesty should lead the guard and vanguard.
- (15) 戊午卜爭貞𠄎 (首 = 導) 羊于𠄎. (Heji 11199)
Divining on the *wuwu* day [55], Zheng tested [the following proposition to seek for sapience from the numen of the turtle]: (We) should (lead:) herd sheep from/to²² RUK.

Examples from (10) to (12) occur on a piece of bone, and the contents are related to each other, (10) and (11), in particular, forming a positive and negative *duizhen* 對貞. Xu Zhongshu (1988:369) says that the above inscriptions are on a *tongban* 同版 ‘same piece’ relationship and speculates that the 𠄎 stands for the word *xing* 省 as in *xingshi* 省視 ‘to pay a visit, inspect’. This is hard to accept as there already is the graph for *xing* written like 𠄎. There is also the graph for *shi* written like 𠄎 (e.g., Heji 36932). Yu Xingwu 于省吾 (1996:1.578-581) also cites the view of a couple of other specialists who suggest the same interpretation as Xu Zhongshu’s, but the person in charge of the note (按語) of this entry in Yu (1996:581) rightly rejects it, offering, instead, that the 𠄎 refers to the method of disposing sacrificial victims (用牲之法). This is equally unacceptable. Failure to

²¹ RUK stands for “Reading Unknown.”

²² I am inclined to interpret the directional feature of the particle *yu* 于 here as meaning “from,” rather than “to.” That is, the Shang divined if they should herd sheep homeward, to their capital, from the place called 𠄎 (Reading Unknown, RUK). However, this interpretation needs to be verified with other inscriptions and in the context of how the herding and inspection of livestock and other related activities may have contributed to the Shang economy.

analyze this graph in a palaeographically scrupulous way underlies these unacceptable views. I believe that the identity of this graph with *shou* 首 ‘head’, as we have done, is the key to dispel the layer of the dense fog, so to speak, shrouding example (12). Understood in this way, (13), (14) and (15) all make good sense.

3.2 Morphological examination

The possibility of the 𠂔 being equal to 首 (導) would increase if we can establish some morphological relationship between 首 *shou*/*sthjæg^wʔ/*hljuʔ, 道 *dao*/*dæg^wʔ/*luʔ on the one hand and 導 *dao*/*dæg^ws/*lus ‘to lead, guide’ on the other. Let us consider:

首 <i>shou</i> /*sthjæg ^w ʔ/*hljuʔ ‘head’	
	} hard to establish relationship, but see below
道 <i>dao</i> /*dæg ^w ʔ/*luʔ ‘road, way’	
	} easier to establish relationship; see below
導 <i>dao</i> /*dæg ^w s/*lus ‘to lead, guide’	

As indicated above, the first pair must somehow be related morphologically but difficult to account for how the derivation might have worked. To start from the second pair of 道 ‘road’ and 導 ‘to lead, guide’ first, it is not so hard to see the derivation because one can interpret the final *-s as a causative suffix meaning “to cause ... to be on the road” yielding “to lead, guide.”²³ Schuessler (2007:207) has also pointed out the possibility that “... the latter [導] is a LOC [Later Old Chinese] general tone C derivation from ‘way’.”

The first pair of *shou*/*sthjæg^wʔ/*hljuʔ 首 and *dao*/*dæg^wʔ/*luʔ 道 may allow a few possible interpretations, but at first sight no cogent explanation seems available, nor is there any specialist that I am aware who offered a specific explanation. I shall just mention here what I think is a promising interpretation. First, Baxter & Sagart (1998:54) have observed “... alternations of *m-, *l-, and *n- with *hm-, *hl-, and *hn- also seem to be morphological in nature; these may involve a different prefix (perhaps simply *h-).” Indeed, both phonologically and graphically 首 and 道 seem related, but *prima facie* the problem is that their meanings are not encouraging. However, Downer (1959:280) has noticed that the *qusheng* derivation of the *shangsheng* word *shou* 首, reconstructed as *shou*/*sthjæg^ws/*hljus, means “to point the head towards” This could be the key, and it is quite instructive to find that Schuessler (2007:48, 207) observed the derivation process

²³ This is not the only example, and we find such other non-*qu* and *qu* pairs as follows:

bin/*pjɪn 賓 ‘guest’ and *bin*/*pjɪns 儼 ‘to treat ... as a guest’
tian/*din 田 ‘cultivated field’ and *tian*/*dins 𦵏 ‘to cultivate land’
yu/*ɲjag/*ɲ(r)ja 魚 ‘fish’ and *yu*/*ɲjag/*ɲ(r)jas 𩺰 ‘to fish’

of the *shangsheng* word 道 ‘road, way, method’ could be “endoactive noun ‘the thing which is doing the conducting’ derived from the following word [referring to *dao* 道/導].” The semantic characterization “the thing which is doing the conducting” corresponds to 道 ‘road, way’, but derived from 導 in the sense of “to lead” for which Schuessler himself has glossed “to go along, bring along, conduct.” Thus, it would seem that the semantic link between 首 and 道 in the first word pair, with which we began our inquiry in this section, is that the “head” (首) is a noun, and so is the “road, way” (道), but the latter specifically in the literal meaning of “heading” or, following Schuessler, “the thing which is doing the conducting.” Since the road is where the act of “conducting takes place, it is a place to “to point the head towards” There is also a *qusheng* reading glossed as “to speak, talk,” but it is quite plausible that the underlying meaning here is “to speak with instruction,” showing a person to head towards some place by way of the road. This is tantamount to “direct the hearer into listening the speaker with the offer of advice,” or simply “to counsel.” If this is correct, then the semantic relationship between the words 首 ‘head’ and 道 ‘road, way’ seems not so inexplicable or irregular.

To end with only a few examples from the transmitted OC texts that may still retain such an underlying meaning of the word 道, we can give as follows:

- (16) 非先王之法言不敢道; 非先王之德行不敢行. (*Xiaojing* 4)
 Unless it is not the exemplary words of the former kings, I will not dare (speak to you with instructive bent:) give you instruction; unless it is not the virtuous deeds of the former kings, I will not dare put them into practice.
- (17) 牆有茨不可掃也, 中冓之言不可道也, 所可道也言之醜也. (*Shijing*, Mao 46, Qiang youci)
 There are thorns on the wall, but they cannot be removed; words spoken in the inner chamber cannot be (spoken with instructive bent:) counseled; although it could be, it would be shameful to do so.
- (18) 有斐君子, 終不可諠兮, 如切如磋者, 道學也. (*Liji*, Daxue 2.2)
 Here is our elegant and accomplished prince who cannot ever be forgotten; as for (what the *Shi* says about) as we cut and as we file, it is to (speak with instructive bent:) tell us the work of learning.

4. Conclusion

The palaeographical analysis of the OBI graph 𠂔 presented in §3 has led me to suggest that its *liding* is 首 *shou*/*sthjægʷ/*hljuʔ which as far as I know is a new interpretation. Furthermore, I have used the concept of “graphic design” ascribable ultimately to

the creator of the original graph. I have applied it to the analysis of the graph 𦣻 with the suggestion that its DT is better taken as *jia/*krit/*krit* 戛 ‘pike’ (used to sever the head at the neck), rather than *guo/*kʷrək/*kʷrik* 馘/馘 ‘cut off head’. This particular interpretation is something I am not entirely confident of, but it is one for which the “synchronic evidential approach” (cf. fn.16) seems to give a slight edge over the other. If the OBI graph 𦣻 can be identified with the word *shou* 首 ‘head’ as this paper contends, what is the difference between it and the graphs 𦣻 and 𦣻 also identified with the same word? The examples examined in this paper, from (1) to (5), suggest that the graph 𦣻 is used as a normal, “biological” head, whereas the graph 𦣻 is used as standing for the word *dao* 導 ‘to lead’ in the examples from (10) to (15). In so far as the graphic design is concerned, the 𦣻 is a iconic representation of the head of some animal, and the 𦣻 is a severed human head with its hair bound by a cord, the head being represented by the eye in the same way as in the small seal form of 𦣻.²⁴ The OBI graph 𦣻 with the hair on top, corresponding to the small seal form 𦣻, is used as a proper noun, though we are at the mercy of Zhu Ao 朱翱, a Song phonetic commentator to the *xiao Xu ben* 小徐本, who has given the *fanqie* spelling 書九切 for both 首 (𦣻) and 𦣻. Zhu Ao, however, is based on Xu Shen’s own definition that 首 is the same as 𦣻.

²⁴ On March 8, 2008, at the Early China Seminar meeting at Columbia University, to which I presented an earlier version of this paper, Zou Changhua 鄒長華 pointed out a structural transformation is observable from the OBI form of the eye, 𦣻, to the small-seal form of it, 𦣻. That is, what appears as a natural depiction of the eye with its horizontal orientation transformed to the form with its vertical orientation. She cited such a bronze form of the head with the “eye” component as 𦣻 (周早井侯簋) changed to 𦣻 (周中大簋), and by the time of the latter the eye element became vertically orientated. Thus, she asked if the vertically orientated eye might be traced back to a different source (來源). This is an interesting question, one which I previously considered in Takashima (2001). This 2001 paper has provided many graphs that had gone through such transformation in orientation, suggesting the development of a “cosmographical” template that defined the vertical shape. Eventually a great majority of Chinese characters, a.k.a. *fangkuaizi* 方塊字 ‘rectangular shaped characters’ resulted.

Appendix: The Graphic Significance of the Dagger 𠂔 (戈) Element

A good way to determine the function of a particular graphic element is to look at all the characters that contain the said element in the *Shuowen* (abbrev. *SW*) and find out how they are analyzed by Xu Shen. In this way we would be in a position to know the general tendency of the function of the graphic element being used as either phonophoric or signific. Given below is a table showing this, but a basic assumption here is that the reading of 戈 is *ge/*k^war/*k^waj*.

Character	戈 as phonophoric	戈 as signific	Note
wu 戊 (戊)	—	(+)	象人脅
cheng 成 (成)	—	+	从戊丁聲; 𠂔 古文成, 从午
yue 戉 (戉)	—	+	从戊 丩 聲, but since the bone graph of 戊 is 𠂔, it may well be different from the 戈 (𠂔) element.
zhao 肇 (肇)	—	(+)	段注: 許原書無篆體; 李舟《切韻》云从戈乍聲
ji 𢇛 (𢇛)	—	+	从戈軌
jia 戛 (戛)	—	+	从戈从百
rong 戎 (戎)	—	+	从戈从甲
kui 戣 (戣)	—	+	从戈癸聲
gan 戠 (戠)	—	+	从戈旱聲
shu 戔 (戔)	—	+	从人持戈
zhan 戰 (戰)	—	+	从戈單聲
die 戔 (戔)	—	+	从戈呈聲
guo 或 (或)	—	+	从口从戈
yu 域 (域)	—	(+)	或, 或从土
jie 戣 (戣)	—	+	从戈雀聲
kan 戎 (戎)	—	+	从戈今聲
qiang 戕 (戕)	—	+	从戈月聲
lu 戔 (戔)	—	+	从戈𠂔聲
kan 戣 (戣)	—	+	从戈甚聲
yan 戣 (戣)	—	+	从戈寅聲
zai 戔 (戔)	—	+	从戈才聲
jian 戣 (戣)	—	+	从戈晉聲

<i>jian</i>	戔 (𠂔)	—	+	从从持戈
<i>wu</i>	武 (𠂔)	—	+	止戈爲武
<i>ji</i>	戠 (𠂔)	—	+	从戈昌聲
<i>zhi</i>	戢 (𠂔)	—	+	从戈从音
<i>can</i>	戠 (𠂔)	—	+	从二戈
<i>qi</i>	戚 (戚)	—	(+)	从戌未聲
<i>wo</i>	我 (我)	—	+	从戈手
<i>yi</i>	義 (義)	—	+	从我从羊
<i>guo</i>	職 (職)	—	(+)	从耳或聲
<i>guo</i>	𦣻 (𦣻)	—	(+)	職, 或从首
<i>xu</i>	戌 (戌)	—	(+)	从戌含一
<i>jie</i>	戒 (戒)	—	+	从升持戈
<i>hua</i>	𠂔 (𠂔)	—	+	从𠂔从戈
<i>wei</i>	威 (威)	—	(+)	从女从戌
<i>xian</i>	咸 (咸)	—	(+)	从口从戌
<i>yong</i>	𦣻 (𦣻)	—	+	勇或从戈用
<i>mie</i>	威 (威)	—	(+)	或从戌
<i>huo</i>	𦣻 (𦣻)	—	(+)	从川或聲
<i>zhi</i>	戢 (戢)	—	(+)	从大戠聲
<i>zang</i>	臧 (臧)	—	(+)	从臣戠聲
<i>bo</i>	𦣻 (𦣻)	—	(+)	从二或
<i>dai</i>	戴 (戴)	—	(+)	从異戠聲
<i>xi</i>	戲 (戲)	—	+	从戈盧聲

It is striking that none of the above 45 characters with the grapheme 戈 that are regarded as having the phonetic function. There are some characters for which Xu Shen's explanation of the graphic significance of 戈 is not entirely unequivocal, and they are shown in brackets as “(+)” in the above table. However, most of them consist of 戈 and another grapheme that Xu Shen takes them as a unit, e.g., 戌, 或 (for 職 and presumably for 𦣻 as well), 戌, 戠 (for 戢), and 戠 (for 臧).

References

- Barnard, Noel. 1973. *The Ch'u Silk Manuscript: Translation and Commentary*. Canberra: The Australian National University.
- Barnard, Noel. 1996. *The Shan-fu Liang Ch'i Kuei and Associated Inscribed Vessels* (in association with Cheung Kwong-yue). Taipei: SMC Publishing Inc.
- Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin & New York: Mouton de Gruyter.
- Baxter, William H., and Laurent Sagart. 1998. Word formation in Old Chinese. *New Approaches to Chinese Word Formations: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, ed. by Jerome L. Packard, 35-76. Berlin & New York: Mouton de Gruyter.
- Chang, Ping-ch'üan (張秉權). 1967. Jiaguwen zhong suo jian ren di tongming kao 甲骨文中所見人地同名考 [Examination of the same appellation used for the personal and place name seen in oracle-bone inscriptions]. *Qingzhu Li Ji Xiansheng Qishi Sui Lunwenji* 慶祝李濟先生七十歲論文集 [Symposium in Honor of Dr. Li Chi on his Seventieth Birthday], Vol. 2, 687-776. Taipei: Qinghua Xuebaoshe.
- Ding, Shan (丁山). 1960. *Shang Zhou Shiliao Kaozheng* 商周史料考證 [Verification of the Historical Sources of the Shang and Zhou]. Shanghai: Longmen Lianhe Shuju.
- Downer, G. B. 1959. Derivation by Tone-Change in Classical Chinese. *Bulletin of the School of Oriental and African Studies* 22.2:258-290.
- Huang, Tianshu (黃天樹). 2005. Huayuanzhuang dongdi jiagu zhong suo jian de ruogan xin ziliao 花園莊東地甲骨中所見的若干新資料 [Several new materials discovered on the tortoise shells unearthed east of the Huayuan Village]. *Journal of Shaanxi Normal University (Philosophy and Social Sciences Edition)* 2005.2:57-60.
- Huang, Xiquan (黃錫全). 1990. *Hanjian Zhushi* 汗簡注釋 [Commentaries on Han Tablet Writings]. Wuhan: Wuhan University Press.
- Li, Fang-Kuei (李方桂). 1971. Shangguyin yanjiu 上古音研究 [Studies on Archaic Chinese phonology]. *Tsing Hua Journal of Chinese Studies*, New Series 9.1-2:1-61. (Reprinted as a monograph in Beijing: The Commercial Press, 1980. This includes a separate lecture ms. and an article, the former is entitled "Zhongguo shangguyin shengmu wenti" 中國上古音聲母問題 [Notes on Archaic Chinese initials], a lecture delivered at the Chinese University of Hong Kong on 25 August, 1970; and the latter is entitled "Jige shanggu shengmu wenti" 幾個上古聲母問題 [On some Archaic Chinese initials], from *Memorial Volume Dedicated to President Chiang Kai-shek*, 1143-1150. Taipei: Academia Sinica, 1976).
- Li, Hsiao-ting (李孝定). 1965. *Jiagu Wenzi Jishi* 甲骨文字集釋 [The Collected Interpretation on Oracle Bone Inscriptions], 15 vols. Taipei: Institute of History and Philology, Academia Sinica.

- Li, Yanong (李亞農). 1951. Yinqi zashi 殷契雜釋 [Miscellaneous Interpretations on the Yin Inscriptions]. *Kaogu Xuebao* 考古學報 [*Acta Archaeological Sinica*] 5:225-245.
- Liu, Zhao (劉釗). 2006. *Guwenzi Gouxingxue* 古文字構形學 [*Grammatology of Palaeographs*]. Fuzhou: Fujian People's Publishing House.
- Pulleyblank, Edwin G. 1973. Some new hypothesis concerning word families in Chinese. *Journal of Chinese Linguistics* 1.1:111-125.
- Qiu, Xigui (裘錫圭). 1988. *Wenzixue Gaiyao* 文字學概要 [*An Introduction to the Study of Writings*]. Beijing: The Commercial Press.
- Schuessler, Axel. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- Shirakawa, Shizuka (白川靜). 1955-62. *Kōkotsu Kimbungaku Ronsō* 甲骨金文學論叢 [*Collections of Studies on the Oracle-Bone and Bronze Inscriptions*], 10 vols. Mimeographed; privately published. Vol. 1, March 1955; Vol. 2, May 1955; Vol. 3, August 1955; Vol. 4, December 1956; Vol. 5, September 1957; Vol. 6, December 1957; Vol. 7, May 1958; Vol. 8, August 1958; Vol. 9, December 1958; Vol. 10, June 1962. (Some of these were later published under the title *Kōkotsu Kimbungaku Ronshū* 甲骨金文學論集. Kyoto: Hōyū Shoten, 1973. But, of course, the pagination is different.)
- Sōran Kōkotsu Moji Jishaku Sōran* 甲骨文字字釋綜覽 [*A Comprehensive Overview of the Interpretations of the Words in Oracle-Bone Inscriptions*]. Compiled by Michio Matsumaru (松丸道雄) & Ken-ichi Takashima (高嶋謙一). Tokyo: Tokyo University Press, 1994.
- Sōrui Inkyo Bokuji Sōrui* 殷墟卜辭綜類 [*A Comprehensive Overview of the Oracle-Bone Inscriptions Classified*]. Compiled by Kunio Shima (島邦男). Tokyo: Kyūko Shoin, 1971.
- Takashima, Ken-ichi. 2001. A cosmography of Shang oracle-bone graphs. *Actes du Colloque International Commémorant le Centenaire de la Découverte des Inscriptions sur os et carapaces*, ed. by Shun-chiu Yau, 37-62. Paris: Centre Recherche Linguistiques sur l'Asie Orientale, EHESS.
- Takashima, Ken-ichi. 2005. How to read Shang oracle-bone inscriptions: a critique of the current method. *Bulletin of the Museum of Far Eastern Antiquities (BMFEA)* (2004) 76:22-43. (Because *BMFEA* is behind the schedule, it is back dated as 2004. However, I use "2005" to refer to it as it is substantially the same as what I initially presented at a seminar held on February 10, 2005 at the Swedish Collegium of Advanced Study in the Social Sciences, Uppsala.)
- The team of the archaeologists stationed at Anyang Work Station of the Institute of the Chinese Academy of Social Sciences (中國社會科學院考古研究所安陽工作站). 2004. Anyang Yinxu Liujiazhuang bei 1046 hao mu 安陽殷墟劉家莊北 1046 號墓 [Tomb No. 1046 Liujiazhuang North in the Yin Ruins in Anyang]. *Kaoguxue Jikan*

- 考古學集刊 [*Papers on Chinese Archaeology*] 15:359-390.
- Wang, Guowei (王國維). 1925[1968]. *Gushi Xinzheng* 古史新證 [*New Evidence for Ancient History*]. *Wang Guantang Xiansheng Quanjì* 王觀堂先生全集 [*Collected Works of Wang Guantang*], Vol. 6. Taipei: Wenhua Publishing Company.
- Xu, Zaiguo (徐在國). 2002. *Liding Guwen Shuzheng* 隸定古文疏證 [*Marginal Notes on the Verification of the Transcriptions of Palaeographs*]. Hefei: Anhui University Press.
- Xu, Zhongshu (徐中舒). (ed.) 1988. *Jiaguwen Zidian* 甲骨文字典 [*Dictionary of the Oracle-Bone Graphs*]. Chengdu: Sichuan Lexicographical Publishing House.
- Yen, Yi-p'ing (嚴一萍). 1951. Zhongguo yixue zhi qiyuan kaolue 中國醫學之起源考略 [A Cursory Examination of the Origins of Chinese Medical Science]. *Dalu Zazhi* 大陸雜誌 [*The Continent Magazine*] 2.8:20-22; 2.9:14-17.
- Ye, Yusen (葉玉森). 1934. *Yinxu Shuqi Qianbian Jishi* 殷虛書契前編集釋 [*Collected Interpretations of the Inscriptions in the Inscriptions from the Yin Ruins*, Vol. I: “*Qianbian*”]. Shanghai: Datong Shuju. (Reprinted in Taipei: Yee Wen Publishing Company, 1966.)
- Yu, Xingwu (于省吾). 1940-43. *Shuangjianchi Yinqi Pianzhi* 雙劍詒殷契駢枝 [*Desultory Notes on the Yin Inscriptions from My Studio Shuangjianchi ‘Double-Sword Wedge’*], 3 vols. Reprinted in Taipei: Yee Wen Publishing Company.
- Yu, Xingwu (于省吾). 1980. *Shi dun* 釋盾 [*Interpreting the Word for “Shield”*]. *Guwenzi Yanjiu* 古文字研究 3:1-6. Beijing: Zhonghua Book Company.
- Yu, Xingwu (于省吾). (ed.) 1996. *Jiagu Wenzi Gulin* 甲骨文字詁林 [*A Forest of Exegeses of the Oracle-Bone Graphs*], 4 vols. Xiaosui Yao (姚孝遂) contributing notes (按語). Beijing: Zhonghua Book Company.

Palaeography Meets Linguistics: Analysis of 𠂇 and 𠂆 in the Oracle-Bone Inscriptions^{*}

Anne O. Yue (余靄芹)

University of Washington

The graphic origin of 𠂇 is controversial. This article brings forth a proposal based on the functions of 𠂇 and 𠂆 as presented in the 41,956 inscriptions of the *Jiaguwen Heji* 甲骨文合集. It has been generally assumed that 𠂇 appears in Period I and is replaced by 𠂆 afterward. This is, however, oversimplified. Actually both 𠂇 and 𠂆 appear in Period I but are largely complementary in function: 𠂇 normally functions as a verb (as a sacrificial verb close to 2,000 times, as existential verb about 1,500 times and as stative verb around 170 times) while 𠂆 functions as a verb far less often (as sacrificial verb about 200 times and as existential verb just 60 times but not as stative verb). On the other hand, 𠂆 functions mainly as a noun (about 300 times) while 𠂇 does not serve in this capacity. Both function as a connector, but 𠂇 appears far more frequently, over 100 times, than 𠂆 with merely 5 times. 𠂇 even becomes grammaticalized as an affirmative marker with close to 100 occurrences while 𠂆 is not observed to have gone through this process yet.

After Period I the occurrence of 𠂇 dramatically drops—in Period II it appears only about 50 times as a sacrificial verb, which is just 1/40 of its tokens in Period I, fewer than 30 times as existential verb, which is 1/50 of Period I, and merely twice as a connector. After Period II it no longer appears and we can assume that it is already replaced by 𠂆, which continues to serve as a noun with nearly 500 tokens, as connector about 420 times, as existential verb nearly 460 times and as sacrificial verb some 3,800 times. In addition, it now appears as a stative verb over 130 times and becomes grammaticalized as an affirmative marker over 30 times.

It is clear that 𠂇 and 𠂆 are largely complementary in usage during Period I, after which 𠂆 gradually replaces 𠂇. Why is it that 𠂆 can replace 𠂇? There is no disagreement that the graphic origin of 𠂆 is the shape of the right hand. Our conclusion is that the graphic origin of 𠂇 also derives from the shape of a hand plus a reference line below it, namely, 𠂇 + 一 where 𠂇 represents the shape of a hand and 一 the symbol for verbal usage. Semantically related pairs of oracle-bone graphs consisting of a noun, and a verb that is formed through use of a referential

^{*} The first draft of this paper, with the title ‘𠂇 and 𠂆 in Periods I and II of the Oracle-Bone Inscriptions’, was presented at the 6th International Conference on Classical Chinese Grammar held in Xi'an, Shaanxi, China, August 14-16, 2007. Special thanks are due to Ken-ichi Takashima 高嶋謙一, who answered many of my queries while I was writing the draft of this paper. Needless to say, I am solely responsible for what is presented here.

base line added to the noun to signify verbal function include 𠂔 and 𠂕, 𠂖 and 𠂗, 𠂘 and 𠂙, 𠂚 and 𠂛, 𠂜 and 𠂝 or 𠂞, and 𠂟 and 𠂠. It is no wonder that 𠂡 and 𠂢 constitute such a pair.

Key words: oracle-bone graphs, graphic origin, graphic structure, graphic function, 𠂣, 𠂤

1. Introduction

The source and origin of the graph 𠂥, written in different styles as 𠂦 (as in *Jiaguwen Heji* 19828, 19956, 19957 反, 19958, 205760, etc.), 𠂧 (as in *Jiaguwen Heji* 19903, 20012, 20348, 22079 乙, etc.), 𠂨 or 𠂩, in the oracle-bone inscriptions (henceforth OBI) have been controversial. Commenting on the style 𠂩, Guo Moruo 郭沫若 considers it a variant of 𠂪 but does not comment on its graphic origin:

𠂥即常見之𠂥字，乃又字之異，字形亦無可說。‘𠂥 is the frequently seen graph 𠂥, a variant of 𠂪, nothing can be said of its graphic form.’ (Guo 1983 別一:九下)

Hu Houxuan 胡厚宣 conjectures that 𠂥 may be a loan graph for 侑:

𠂥爲武丁時常見之祭名。以其又每用爲‘有來自東’[後下 37.2]‘五十有六’[前 4.81]‘禽有犬’[前 6.24.5]之‘有’，推之或當爲侑之借字。‘𠂥 is the frequently seen name of a sacrifice. Since it is also used for 有 in such expressions as “有來自東”，“五十有六”，“禽有犬”，it should be a loan graph for 侑.’ (Hu 1945 廈甲:一上)

Wu Qichang 吳其昌, probably taking the explanation of Sun Yirang 孫詒讓 of 𠂥 as 之, quotes the *Shuowen* 說文 description to relate the two:

𠂥，出也。象艸過𠂦枝莖漸益大有所之也。一者地也。‘𠂥, to emerge, symbolizing leaves going through a plant gradually becoming larger as if they were going somewhere.’ (Wu 1959:4-11)

Wu Kuang 吳匡 (1989:892) analyzes the graph into two parts: 𠂦 + 𠂧, the former designating the early form of 坎 ‘a pit’ while the latter signifying the ground, with the entire graph having the meaning of 坎 ‘a pit’. Jin Xiangheng 金祥恆 (1967:2854-2857), despite the fact that 𠂪 or 𠂥 appears in the 1st rime category while 侯 appears in the 4th in Duan Yucai 段玉裁’s *Shuowen Liushu Yinjun Biao* 《說文六書音均表》, reads 𠂥

as 侯 just because 又 and 侯 share the same rime /ou/ in Zhou Deqing's 周德清 *Zhongyuan Yinyun* 《中原音韻》十六去聲尤侯韻. Huang Xiquan 黃錫全 (1981:196-204) painstakingly correlates the four different shapes of 𠩺 that appear in Period I with the shapes of a bovine head, while at the same time pointing out that variant 1, 𠩺, occurs in inscriptions of the Shi 自 diviner group and variant 2, 𠩺, in those of the Bin 賓 diviner group. Tang Yuming 唐鈺明 and Ken-ichi Takashima 高嶋謙一 follow suit. While Tang analyzes the graph into two parts, the upper part representing the bovine head and the bottom line an offering plate or container (Tang 1992:69) Takashima argues that “the graph 𠩺 is a stylized form of the top portion of 𠩺...” (Takashima 2005:18). Both Huang and Takashima present phonetic relation to justify their claim, 𠩺 and 𠩺 being from the archaic rime category 之, although the initials differ and there is no *xiesheng* 諧聲 connection between the two.

Despite the indeterminacy of the graphic source of 𠩺, its usage is well understood, as summarized by Chang Ping-Chuan 張秉權:

𠩺是有字，在這裡作為有無之有解，但在卜辭中又有作為祭名的侑，和作為再又之又講的。因為這個字祇見於早期及文武丁時的卜辭中，有著時代性的，所以我把它定為𠩺... 𠩺是又字，亦即右字，古文右和又為一字... 在早期卜辭中，又僅作為祐講。在晚期卜辭中，則兼有「又」「祐」「有」等意義。

‘𠩺 is the word 有, understood as the 有 of 有無, but in the OBI it is also used for the sacrifice 侑 and the 又 of 再又. Since this graph is only seen in the OBI of the early period and of the time of Wen Wu Ding, it is period bound. Therefore I settle it as 𠩺... 𠩺 is the word for 又 or 右, which are the same word in ancient script... In the early OBI, 又 only stands for 祐 whereas during the later periods it comprises the meaning of “又”, “祐” and “有”.’ (Chang 1959:31-33)

We shall now examine the functions of both 𠩺 and 𠩺 in the OBI, focusing on Period I when the difference between them is most conspicuous and on Period II for observing the transition of change, before speculating on the graphic origin of 𠩺.

2. Period I

During Period I the graphs 𠩺 and 𠩺 seem to be in complementary distribution in accordance with their functions, although the occurrence of 𠩺 (in at least 3,781

inscriptions) far outnumbers that of 𠄎 (in at least 565 inscriptions).¹ This is the reason why scholars in the field generally claim that 𠄎 occurs in Period I but is replaced by 𠄎 after that. While this is true, it is not the whole story. At the same time, 𠄎 occurs mainly as a verb or as an affirmative marker derived from its verbal function but not as a noun, while in comparison 𠄎 occurs as a verb far less frequently (in about 260 inscriptions) but it occurs as a noun (in about 300 inscriptions) while there are no cases where it occurs as an affirmative marker during this period.

The familiar expression 受𠄎 𠄎 ‘receive abundant assistance’ always has the graph 𠄎 precede the graph 𠄎 and never the other way around. There is also the expression 受 𠄎 but never 受𠄎, although there are rare cases of 𠄎𠄎.

In the following sections, we shall give exhaustive lists of the occurrence of 𠄎 and 𠄎 in the *Jiaguwen Heji* 甲骨文合集 (henceforth *Heji*) in their different functions outlined above.²

3. Occurrence of 𠄎

3.1 𠄎 as verb

3.1.1 𠄎 as sacrificial verb

𠄎 occurs most frequently as a sacrificial verb (henceforth Vsac) with the general meaning of ‘to make a sacrificial offering’. The total number of inscriptions with its occurrence is close to 2,000, with 97+% appearing in Period I and about 2.6% in Period II, including the three variant graphs described by Huang Xiquan. None occurs after Period II.

These inscriptions are from a great variety of diviner groups, especially the following (where the name of diviner group is followed by inscription number from the *Heji*):³

Que 𠄎: 190, 418, 446, 559, 562, 656, 679, 682 𠄎, 767, 849, 924, 970, 1027, 1075, 1151, 1254, 1272, 1336, 1402, 1534, 1536, 1542, 1543, 1545, 1677, 1679, 1692, 1696, 1703, 1773, 1780, 1785, 1868, 1906, 1907, 2362, 2363, 2378, 2518, 02874, 2975, 3453, 3467, 3534, 4021, 6431,

¹ We use the loose term ‘inscription’ as the unit for frequency count, which does not show the actual number of occurrences since some inscriptions consist of a set of examples while others have a single one.

² The lists of occurrence can only be taken as a very rough approximation since the calculation was done manually and since there remain many problems with respect to the interpretation of the OBI.

³ We cite examples only of those with the name of the diviner group given explicitly. Therefore the total number of examples is rough and conservative.


- 7323, 7604, 7829, 9504, 11484, 11497, 11917, 12921, 14512, 14732, 14807, 15021, 15045, 15051, 15616, 16204, 17511, 17540, 17547, 39535, 39776, 39857, 40454, 40638
- Bin 旁: 6, 94, 326, 339, 955, 978, 1140, 1144, 1191, 1415, 1524, 1528, 1533, 1551, 1571, 1597, 1661, 1854, 1901, 1908, 1909, 1910, 1918, 1924, 1928, 1935, 1940, 2019, 2056, 2484, 2545, 2556, 2725, 2830, 2924, 2978, 3077, 3461, 3465, 3597, 4052, 5732, 8995 𠂔, 10125, 10408, 10942, 13658, 14522, 14573, 14832, 14927, 14928, 14974, 15016, 15077, 15087, 15943, 17649, 19348, 39514, 39593 反, 39599, 39605, 40452
- King: 982, 1027, 1522, 1565, 1566, 1843, 1845, 1848, 1849, 1929, 1999, 2290, 2403, 2406, 2448, 2553, 3312, 4141, 5000, 14755 反, 16201, 19771, 19777, 19806, 19812, 19832, 19862, 19866, 19867, 19871, 19913, 19914, 19915, 19949, 19961, 19962, 20015, 20034, 20037, 20056, 20463, 20980, 39683
- Zheng 爭: 380, 408, 409, 444, 680, 682, 728, 952, 954, 1046, 1052, 1143, 1147, 1157, 1158, 1552, 1554, 1653, 1654, 1934, 2136, 2292, 2454, 2496, 3731, 4051, 4759, 5127, 5622, 6943, 6947, 9530, 10306, 14161 反, 14521, 14854, 15022, 40467
- Huan 亘: 766, 1115, 1534, 2108, 2133, 3183, 3768, 4277, 5647, 7275, 7954, 14513, 14931, 14965, 15003, 40429
- Fu 扶: 19798, 19817, 19882, 19883, 19904, 19908, 19921, 19928, 19931, 20098, 20100, 20101, 21120
- Dun 𠂔: 339, 977, 1275, 1807, 2003, 3824, 14409, 39599, 39715, 40462

and occasionally:⁴

- Shi 𠂔: 19880, 21284, 21376, 21290, 21284, 21376, 21290
- Shao 𠂔: 19826, 19884, 19890, 19920, 20022, 20032
- 𠂔: 2483, 6097, 9965, 10091, 13726
- Nei 內: 209, 1747, 3461, 14514(?)
- 𠂔: 33
- Er 耳: 967
- Fu 𠂔: 3905
- Xie 叶: 20153
- Wei 韋: 39649
- [Chu 出: 365]⁵

⁴ Graphs with unknown reading are given without *pinyin* romanization.

For example:

- (1)  乙巳卜殼貞出于祖乙一牛。用。 06945(15)
 ‘Divining on yi-si day Que tested, make offering to Ancestor Yi of one bovine. Use.’

Besides, there are countless inscriptions with no name of diviner group designated, which we shall call ‘unnamed’.⁶ Apart from these unnamed ones, Que and Bin top the list of diviner groups during Period I.

For Period II there are the following diviner groups in descending order of frequency:⁷


⁵ Square brackets indicate that the diviner group was not clearly given but was assumed by the *Jiaguwen Heji Shiwen* 甲骨文合集釋文.

⁶ These divinations are the most frequent in number in almost 1,000 inscriptions, for example: 25, 102, 175, 324, 377, 385, 428, 430, 501, 698, 942, 1141, 1146, 1162, 1205, 1253, 1277, 1369, 1372, 1374, 1375, 1379, 1403, 1421, 1422, 1424, 1432, 1472, 1475, 1479, 1481, 1508, 1509, 1510, 1513, 1520, 1526, 1527, 1531, 1537, 1680, 1682, 1683, 1684, 1685, 1686, 1687, 1707, 1713, 1714, 1715, 1784, 1786, 1787, 1789, 1791, 1800, 1828, 1829, 1830, 1831, 1834, 1835, 1837, 1840, 1841, 1844, 1850, 1851, 1912, 1915, 1925, 1942, 1998, 2005, 2011, 2033, 2046, 2051, 2601, 2098, 2104, 2107, 2110, 2111, 2112, 2114, 2130, 2132, 2135, 2164, 2169, 2173, 2174, 2179, 2180, 2181, 2182, 2185, 2192, 2303, 2340, 2348, 2349, 2350, 2352, 2374, 2375, 2376, 2377, 2381, 2385, 2387, 2390, 2391, 2401r, 2405, 2407, 2408, 2409, 2410, 2411, 2412, 2414, 2451, 2455, 2457, 2458, 2511, 2519, 2548, 2550, 2552, 2554, 2555, 2577, 2582, 2583, 2585, 2588, 2827, 2829, 2870, 2872, 2873, 2877, 2878, 2880, 2881, 2905, 2907, 2946, 2972, 2973, 2974, 3016, 3112, 3113, 3255, 3397, 3414, 3462, 3463, 3466, 3472, 3507, 3509, 3515, 4047, 4048, 4065, 4324, 4593, 4917, 5029, 5544, 5652, 6077, 6109, 6144, 6167, 6209, 6336, 7026, 7259, 7301, 7311, 7862, 7865, 7920, 8724, 8746, 8949, 9101, 9102, 9220, 9615, 9827, 9856, 9861, 9862, 9863, 9881, 10079, 10114, 10130, 10139, 10940, 11500, 12442, 12627, 12642, 12658, 12665, 13277, 13520, 13562, 13706r, 13865, 14130, 14157, 14344, 14346, 14361, 14375, 14410, 14412, 14413, 14414, 14415, 14417, 14418, 14419, 14509, 14510, 14516, 14517, 14519, 14520, 14658, 14659, 14710, 14711, 14712, 14714r, 14716, 14724, 14726, 14728, 14729, 14731, 14766r, 14768, 14775, 14784, 14814, 14819, 14846, 14868, 15001, 15002, 15013, 15065 乙, 15068, 15056, 15080, 15850, 16019, 17302, 17615r, 17997, 18948r, 19315, 19828, 19848, 19861, 19903, 19956, 19958, 19963, 19972, 19978, 20012, 20576, 21264, 21540, 21297, 22044, 22065, 22068, 22072, 22075, 22078, 22079 乙, 22088, 22098, 22219, ...

⁷ Parentheses with enclosed question mark ? indicate possible interpretation of an inscription with limited context. For example, #23620 reads: 癸酉卜出貞出母...在六[月]. We consider 出 a Vsac and neither a Vex nor an affirmative marker (see §3.3 below) within the given context.

Chu 出: 22546, 22580, 22739, 22740, 22742, 22823, 22824, 22825, 23064, 23620(?), 23717, 24230, 24413, 24945, 25038, 25039, 25040, 25041, 25043, 25937, 26094(?), 40962, 41111, 41112, 41113
 Da 大: 22548, 22559, 22741, 23719, 24433, 24610, 24941, 25025, 25029, 25042, 26765, 40926, 41005, 41158
 ?? 𠂔: 23805, 25942
 Xiong 兄: 25035, 41023
 He 何: 27150, 27153
 Zhong 中: 23059
 Xing 行: 25068
 Unnamed: 22543, 22565, 22595, 22960, 24951, 25044, 25954, 26826, 40912

For example:

- (2)  乙亥卜中貞曰其出于丁車三宰。九月。 23059(2)
 ‘Divining on *yi-hai* day Zhong tested and stated: in making offering to Ding it is of three penned sheep. Ninth month’.

3.1.2 出 as existential verb


There are also many instances (in close to 1,500 inscriptions) of 出 occurring as the verb of existence or possession (henceforth Vex), 98% of which appear in Period I, barely 2% in Period II and none afterward. We include within this category where 出 is followed by both an NP and a VP.

For Period I there are inscriptions which include the following diviner groups in descending order of frequency:

Que 𠂔: 94, 376, 672, 776, 1107, 2438, 3032r, 3458, 4102, 4194, 4499 乙, 5382, 5775, 6087, 6441, 6536, 6577, 6653, 6947, 7085, 7137, 7239, 7312, 7685, 8492, 10109, 10125, 10315, 13615, 13616, 13682, 13767, 13778, 13886, 13888, 13926, 13928, 13930, 13931, 13936, 14149, 14575, 17145, 17235, 17296, 17364, 17409, 17419, 39493, 40386, 40470, 40638
 Bin 旁: 6, 849, 1163, 2924, 2954, 3222, 3679, 4059, 4083, 4735, 5354, 6092, 6566, 6672, 7089, 7111, 7214, 9621, 10768, 10935, 12831, 13696r,

	13843, 13868, 13880, 13925, 13932, 13996, 15485, 15486, 16339, 16951, 16952, 16994, 17410, 17633, 39675, 40228, 40605
Zheng 爭:	137, 223, 672, 779, 780, 891, 1869, 3096, 3271, 3738, 4248, 4569, 4618, 4978, 5370, 5637, 6016, 6076, 9575, 10047, 10067, 10136, 11485, 11499, 13420, 13465, 13651, 13674, 13788, 13876, 16471, 16989, 17078, 17301, 17452, 19441, 39530
Huan 亘:	94, 2659, 3762, 4271, 4272, 5545, 7121, 7571, 7629, 13644, 13771, 39781
Dun 咎:	136, 3830, 4264, 4888, 7112, 7191, 8894r, 10137, 13750, 13878, 16950, 17410, 19184
Nei 內:	536, 2940, 3813, 4519, 5828, 6572, 10964, 13505, 13759r, 13883
King:	3413, 4144, 9650, 13715, 20181, 20232, 40817
Yong 永:	6065, 9233, 39768
𠄎:	3865, 6454
Wei 韋:	3852, 3847, 3861
Si 兑:	6, 40375
Fu 扶:	20196
Gong 晉:	19322
Ye 叶:	20535
Shi 𠄎:	21071
Unnamed:	48, 526, 698, 915, 1121, 1130 乙, 1136, 1137, 2123, 2134, 2164, 2251, 2333, 2652, 2667, 2831, 2837, 2936, 2993, 3123, 3295, 4037, 4086, 4125, 4126, 4127, 4145, 4226, 4247, 4276, 4281, 4331, 4398, 4498, 4500, 4568, 4570, 4571, 4617, 4641, 4654, 4656, 4796, 4902, 4979, 4980, 5402, 5455, 5456, 5494, 5530, 5582, 5750, 5839, 6066, 6067, 6075, 6086, 7076, 7092, 8307, 8339, ...

For example:

- (3)  出來自南氏龜。 7076(30)
 'There are turtles brought coming from the south.'⁸

As observed above, Que, Bin and Zheng are the most popular diviner groups during Period I.

⁸ Another interpretation is to have 出來 as a unit with 出 as an affirmative marker (see §3.3 below for the term 'affirmative marker') of 來: 'there came turtles brought from the south.'

For Period II, there are far fewer examples (barely 30 inscriptions) and far fewer diviner groups:

Chu 出: 23531, 23532, 23533, 23709, 23803, 24146, 24650, 26096, 41091, 41228
 Zhong 中: 23650, 26097, 26098
 ?? 𠄎: 24132, 24866
 Unnamed: 40998, ...


For Period I there are numerous inscriptions with unnamed diviners but few for Period II.

3.1.3 出 as stative verb

出 occurs as a stative verb (hereafter Vst) with the meaning of ‘to be abundant’ only during Period I in at least 22 inscriptions (787, 795, 9503, 9504, 9507, 9517, 9541, 9552, 9572, 9759, 9818, 9819, 9909, 9930, 9950r, 10031, 13505...) in the expression 出年 ‘bumper crop’ with the popular diviner groups Zheng, Que, Bin and the King as well as unnamed ones.

Zheng 爭: 787, 9504, 13505
 Bin 𠄎: 9818
 King: 10031
 Que 𠄎: 795
 Unnamed: 9552, 9759, 9909, 9930, ...

For example:

- (4)  癸亥卜爭貞我黍受出年。一月。 00787(7)
 ‘Divining on *gui-hai* day, Zheng tested our millet receives bumper crop. First month.’

Another instance with 出 as Vst in the given sense is 出𩇑 ‘abundant catch’ in a few inscriptions with unnamed diviners:

Unnamed: 6384, 6387, 6487, 39879, ...

For example:

- (5) 貞弗其受出𩇑. 39879(1)
‘Tested: will not receive abundant catch.’

The expression 受出又 ‘to receive abundant assistance’ will be discussed in §4.1.2.

3.2 出 as affirmative marker

There are cases where 出 is grammaticalized into an ‘affirmative marker’ bleached of lexical meaning, although this grammatical meaning can easily be perceived as derived from the lexical meaning ‘to exist, to have’ that implies affirmation of existence. These can be identified by 出 occurring after a verb and before another noun, or after a verb and before a verb phrase. We shall describe the former as V + [出+N]_{NP} and the latter V + [出+VP]_{NP}, with both [出+N] and [出+VP] functioning as single units of an NP. Both types occur exclusively in Period I.

3.2.1 [出+N]_{NP}


There are at least 60 inscriptions in Period I with [出+N]_{NP} appearing with various named and unnamed diviner groups:

Que 𩇑: 924, 5805, 10344, 14198, 17055
Huan 亘: 667, 7571
Zheng 爭: 9575, 17171
Bin 𠂔: 10048, 17168
Dun 𠂔: 46
Wei 韋: 8985
Chu 出: 26010
King: 39683
Unnamed: 47, 115, 1253, 2163, 3167, 4003, 4421, 6956, 7072, 7869, 9576, 10151, 13646r, 14659, 15818, 39686, ...

Such [虫+N]_{NP} structures include as many as 23 expressions:

虫自 [1253, 5805, 5806, 6813, 8985, 8987, 10048, 17055, 17168, 17169, 17171, 17172]	虫祖 [914r, 924, 4421, 6956, 13646, 39683, 39686]
虫邑 [635r, 4003, 7072, 8987]	虫妣 [3167, 3171, 4421, 10407]
虫家 [13584 正甲, 13584 正乙, 13588]	虫众 [46, 47]
虫兄 [2163, 14198]	虫田 [9575, 9576]
虫芻 [115]	虫妻 [667]
虫母 [924]	虫任 [3521]
虫丁 [6571]	虫壳 [7571]
虫泉 [8368]	虫水 [10151]
虫土 [10344]	虫貝 [11425]
虫屯 [13515]	虫老 [13758]
虫示 [14659]	虫酒 [15818]
虫宗 [26010]	

For example:

- (6)  貞乎子室出于虫祖宰。 00924(9)
'Tested: call Zi-an to make offering to the Ancestor of penned sheep.'

3.2.2 [虫+VP]_{NP}

When 虫 is followed by a VP, it may be the case that the VP is functioning as an NP, as observed in the expression 虫不若 where 虫 serves as the verb and 不若 serves as its object, for example:

- (7) 王不佳虫不若。 376(4)
'It is not [that] the King has no approval [from the Supreme Being].'
(8) 辛未卜旁貞王虫不正。 5354(1)
'Divining on *xin-wei* day, Bin tested: the King has no properness.'


There are also cases where 虫+VP together constitutes a single unit of NP, namely, [虫+VP]_{NP}, where 虫 occurs after another verb:

- (9) 壬寅卜余収出往田于不乎从帛弌。 891r(2)
 ‘Divining on the *ren-yin* day, I mobilize (< raise) those who had gone to hunt at Pi and call them to follow Cai(?) and Kou(?).’⁹

The most common [出+VP]_{NP} is 出取 occurring after the verb 氏 (‘to bring’) in at least 10 inscriptions, mostly with unnamed diviners; for example:

Unnamed: 8720, 9070, 9072, 9073, 9105r, 9109, 9126, 9127, ...

For example:


- (10)  弗其氏出取。 8720(2)
 ‘Will not bring what is obtained.’

There are just a few cases with the popular diviner groups Que and Zheng:

Que 𠄎: 9050, 09075

Zheng 爭: 9069

For example:

- (11)  貞𠄎眾𠄎氏出取。 09050(1)
 ‘Tested: Ni(?) and Ke(?) will bring what is obtained.’


⁹ Translation according to Takashima.

3.3 出 occurring with units of numbers

出 very often occurs with numbers but only during Period I. This function as a connective sees at least over 100 inscriptions with various diviner groups, for example:

Que 𠩺:	897, 901, 903, 1863, 6830, 9681r, 10349, 11697 正乙, 13998, 14002, 15257, 17055, 17579, 17610
Bin 旁:	896, 11365, 17581, 17611, 17612
Huan 亘:	94, 641, 15734 𠩺, 17663
Zheng 爭:	137, 6834, 13331, 39720
King:	21081, 39707, 40535
Dun 𠩺:	6650, 17580
Yong 永:	17628
Unnamed:	325, 499, 898, 1445, 1699, 3263, 3499, 3524, 6621, 7101, 7327, 7434, 8452, 8752, 10200, 10372, 10830, 10831, 11060, 11065, 11298, 11300, 11366, 11641, 11642, 11643, 11648, 13753, 13985, 14017, 14092, 15834, 17152, 39591, 40198, ...

For example:

- (12)  乙亥卜𠩺貞來乙亥酹下乙十伐出五卯十宰。二旬出一日乙亥不酹雨。五月。 903(3)
 ‘Divining on *yi-hai* day, Que tested: [on the] coming *yi-hai* day, *you-cut* to Xia-Yi ten decapituri and five, cut open 10 penned sheep. Two ten-day weeks and one day [on] *yi-hai* [day] did not *you-cut*, rain. Fifth month.’¹⁰

There are one or two exceptions where we find such inscriptions in Period II, for example:

¹⁰ According to Takashima, *you-cut* refers to the manner of cutting flesh or meat in preparation for sacrifice.


- (13) 乙巳出於母辛宰出一牛。十月。 40912(1)
 ‘[On] *yi-si* day make offering to Mother Xin of penned sheep and one bovine. Tenth month.’

4. Occurrence of 𠂔

There is prolific occurrence of 𠂔 in the OBI, mostly as a Vsac. We shall first start with its characteristic nominal function.

4.1 𠂔 as noun

There are at least 800 inscriptions with 𠂔 functioning as a noun with the meaning of ‘assistance’ except for one instance where it indicates the position of ‘the right side’ in the following inscription:

- (14)  亞立其于 𠂔 利。 28008(5)
 ‘If Ya stands to the right [it] is beneficial.’

其于左利。 28008(6)
 ‘If to the left [it] is beneficial.’

𠂔 as a noun occurs at the highest frequency during Period III, with almost 400 inscriptions or 48%, followed by Period I with at least 300 inscriptions or about 39%. Period IV sees some 60 inscriptions or 7%, Period V has some 50 inscriptions or 7% while Period II has a single inscription. In the overwhelming cases 𠂔 appears in the expression 受 𠂔 ‘to receive assistance’ and next in the fixed phrase ‘to receive abundant assistance’, which appears almost exclusively as 受出 𠂔 during Period I but as 受 𠂔 𠂔 afterwards. The expression 受方 𠂔 ‘to receive assistance from the *fang*’ is also popular.¹¹ Examples are listed in the following sections.


¹¹ In the context of assistance, *fang* refers probably to the ‘Four *fang*’ or ‘Four Directions’, which are Powers or Spirits to be reckoned with.

4.1.1 受𠬞

Inscriptions with 受𠬞 ‘to receive assistance’ number at least 90 in Period I and have the following diviner groups besides the unnamed ones:¹²

Que 𠬞: 6204, 6316, 6317, 6318, 6320, 6321, 6498, 6886, 8008, 8426, 9472, 17283, 39926
 King: 5323, 20191, 20510, 20524, 20527, 20530, 20608, 20613, 20628
 Bin 𠬞: 561, 1364, 07627, 07628
 Zheng 爭: 6664, 7440
 Dun 𠬞: 6223
 Unnamed: 1171, 1463, 4199, 4285, 4692, 6420, 6432, 6474, 6500, 6506, 6541, 6719, 7013, 7539, 8644, 8787, 12579, 14190, 14191, 14257, 14259, 14260, 14266, 14378, 14671, 16294, 16296, 16302, 16305, 16306, 19656, 20513, 20516, 20525, 20612, 20616, 21073, 21542, 39896, 40398, 40402, 40446, 40594, ...

For example:

- (15)  甲辰卜爭貞我伐馬方帝受我𠬞。一月。¹³
 6664(8)
 ‘Divining on *jia-chen* day, Zheng tested: [if] we attack Ma-*fang* the Supreme Being will bestow us assistance. First month’

We did not find any example for Period II, and for Period III there are some 300 inscriptions overwhelmingly with unnamed diviners as expected. Only a handful instances can be identified with the diviner group 𠬞:

𠬞: 27215, 27391, 30885
 Unnamed: 26910, 26912, 29013, 26915, 26916, 26917, 26920, 26922, 26955, 26956, 26962, 26980, 26992, 27018, 27020, 27023, 27037, 27040, 27055, 27059, 27110, 27111, 27121, 27130, 27133, 27180, 27184, 27190, 27195, 27196, 27202, 27209, 27217, 27226, 27227, 27251,

¹² In the following inscriptions, #7627 and #7628 have 受萑𠬞, #20608 has 受𠬞 𠬞, #20628 has 受西土𠬞, #8426 has 弗受朕事𠬞, #6664, #7440 and #14671 have 帝受我𠬞, #1463 has 受祖𠬞, #6719 has 受侯𠬞, #14257, #40402 and 40446 have 受我𠬞, and #16302 受余𠬞.



¹³ In the context of warfare, *fang* refers to hostile tribes from neighboring regions.

27262, 27266, 27279, 27280, 27281, 27284, 27288, 27324, 27347,
27354, 27366, 27370, 27381, 27383, 27384, 27416, 27575, 27651,
28154, 28274, 29004, 29487, 29575, 29587, 29603, 29614, 29617,
29619, 29649, 30678, 30683, 30684, 30729, 30743, 30747, 30763,
30811, 30917, 30935, 30987, 31188, 31190

For example:

(16)



癸亥卜狄貞重  至王受  。 30085(1)


‘Divining on *gui-hai* day, Da(?) tested: it is RUK arriving
that the King receives assistance.’¹⁴

There are over 50 inscriptions in Period IV and about 10 in Period V, all with unnamed diviners of course:

Period IV: 31973, 31974, 31976, 31977, 31979, 31981, 31987, 32609, 32613,
32616, 32642, 32717, 32844, 32893, 32967, 32974, 32975, 32976,
32977, 33018, 33019, 33020, 33022, 33029, 33042, 33043, 33069,
33071,¹⁵ 33074, 33116, 33117, 33347, 34055, 34684, 34685, 34686,
41495, 41505, 41536, 41573, 41579, ...

Period V: 35395, 35719, 36023, 36123, 36517, 36535, 36765, 37157, 37371, ...

4.1.2 受出 versus 受

There are close to 150 inscriptions of 受出  in Period I with the following diviner groups:

Zheng 爭: 547, 548, 4597, 6087, 6214, 6226, 6227, 6229, 6230, 6234, 6291, 6337,
6338, 6412, 6479, 6487, 6489, 6490, 6496, 07556

¹⁴ RUK stands for ‘reading unknown’.

¹⁵ #33071 has 受侯又.

Que 𠩺: 32, 540, 6172, 6409, 6482, 6483, 6484, 6485, 6486, 6513, 6518, 6519, 6534, 6550, 10198, 16280, 39864, 39887
 Bin 旁: 4596, 6169, 6232, 6259, 6292, 6297, 6413, 6417, 6471, 6520, 6524, 6525
 Huan 亘: 6205, 11942
 King: 3415
 Dun 𠩺: 6650
 Unnamed: 63, 2164, 4001r, 5146, 6162, 6168, 6178, 6213, 6224, 6233, 6235, 6236, 6238, 6241, 6275, 6276, 6277, 6278, 6282, 6355, 6422, 6432, 6433, 6499, 6501, 6541, 6553, 6554, 7297, 7311, 7315, 7447, 7525, 7558, 7623, 8474, 8495, 8502, 8515, 8518, 8519, 8520, 8647, 9814, 9972, 9973, 11619, 16273, 16274, 16275, 16276, 16278, 16279, 16281, 16282, 16283, 16284, 16285, 16286, 16287, 16288, 16289, 16291, 16308, 16309, 19241, 39869, 39870, ...

For example:

(17)




乙卯卜𠩺貞王从望乘伐下危受𠩺𠩺。 32(1)
 ‘Divining on *yi-mao* day Que tested [if] the King follows Wang Cheng to attack Xia-wei [he] will receive abundant aid.’

whereas there are in all close to 130 inscriptions with 受𠩺𠩺 — only one instance in Period II (#40982), 85 in Period III and 41 in Period V but none in Period I or Period IV. Named diviner groups of Period III include just the following two while the rest are unnamed:

Da 大: 27283
 狄: 27459, 28011, 28140, 29645, 30748, 31108
 Unnamed: 26991, 27151, 27228, 27287, 27396, 27881, 28089, 28093, 30974, 31110, ...

For example:

- (18)  癸酉卜狄貞卯三牢...受 λ = 31108
 ‘Divining on *gui-you* day, Da(?) tested cut open three penned bovine... receive abundant assistance.’

All divinations in Period V are unnamed.

Unnamed: 35356, 36123, 36124, 36125, 36126, 36168, 36169, 36315, 36344, 36350, 36482, 36507, 36511, 36515, 36522, 38289, 38731, ...

The above account, captured in Table 1, clearly shows the complementary distribution of 𠩺 versus λ in the function of serving as a Vst, with the former limited to Period I and the latter to subsequent periods.

Table 1: 𠩺 versus λ as Vst


Period	受 λ 年 = 22 [diviner groups]	受 λ 𠩺 = 1	受 λ λ = 148 [diviner groups]	受 λ λ = 127 [diviner groups]
I	22 [Zheng = 3; Bin, King = 2; Que = 1]	1	148 [Zheng = 20, Que = 18, Bin = 12, Huan = 2, King = 2, Dun = 1]	0
II			0	1
III			0	85 [狄 = 6, 大 = 1]
IV			0	0
V			0	41

4.1.3 方𠂔

Expressions with 方𠂔, including 土方𠂔 ‘aid against Tu-*fang*’, 舌方𠂔 ‘aid against Gong-*fang*’ and 𠂔方𠂔 ‘aid against Zhou?-*fang*’, appear only in Period I in some 30 plastrons with the following diviner groups:¹⁶

?? 𠂔:	64
Zheng 爭:	8480
Que 𠂔:	8501, 8511, 8512
Bin 𠂔:	8506
King:	20613
Unnamed:	440, 8478, 8479, 8481, 8482, 8483, 8502, 8503, 8505, 8517, 8616, 8617, 8649, 8650, 40032, 40033, 40555, 40990, ...

For example:

- (19)  貞我弗其受舌方𠂔。 8505
‘Tested, we will not receive assistance against Gong-*fang*.’

4.2 𠂔 as verb

While 𠂔 functioning as a verb occurs exclusively during Period I, 𠂔 used as a verb, on the other hand, occurs throughout the 5 periods although in low frequency during Period I.

4.2.1 𠂔 as Vsac

There are over 3,600 inscriptions with 𠂔 occurring as a Vsac. Apart from Period V that has only 20 or so instances or 0.58%, the lowest frequency is observed during Periods I and II, with just 5% (close to 200) each; while the highest, 73% (over 2,600 inscriptions) is observed during Period III and 15% (over 500 inscriptions) during Period IV.

As expected, Periods IV and V have only unnamed diviners.

¹⁶ Literally ‘方𠂔’ should refer to ‘aid/assistance from the *fang*’, however, since the *fangs* refer to enemy tribes it is more logical to interpret the said expression as ‘aid/assistance against the *fang*’ according to Takashima.

For Period I we have the following diviner groups but with few inscriptions for each:

Zi 子:	21665, 21671, 21676, 21679, 21384(?)
Yu 余 ¹⁷ :	21586, 21595, 21664, 21686
Fu 𠄎:	21626, 21667, 21668, 21677
Wo 我:	21586, 21663, 21673
Zheng 爭:	4670, 14207r
Chuan 𠄎:	19967, 21086
Fu 扶:	20112, 40840
Chuan 𠄎:	21635, 21684
Bin 旁:	4671
Que 𠄎:	8947
Huan 亘:	15932
Shi 𠄎:	19890
Gu 𠄎:	20324
King:	20966
Zhi 𠄎:	21287
Ya 亞:	22312
Unnamed:	436, 964, 19815, 19818, 19819, 19821, 19825, 19831, 19837, 19839, 19863, 19876, 19938, 19940, 20398, 20737, 21538, 22067, 22092, 22098, 22164, 22165, 22175, 22201, 22206, 22249, 22274, 22294, 22322, 22384, 39574, ...

It can be observed from the above list that as many as 16 diviner groups are involved; however, more than half of these are lesser known ones while the more popular ones prefer to use the form 𠄎 as we have seen. We can clearly see the preference in Table 2:

Table 2: Vsac with 𠄎 versus 𠄎 in Period I within the same diviner groups

Vsac	Total number	Diviner group					
		Que	Bin	King	Zheng	Huan	Shi
𠄎	1940	70	64	43	38	16	7
𠄎	200-	1	1	1	2	1	1


Moreover, the overwhelming majority of the Period I cases are with unnamed diviners (no fewer than 160 inscriptions).

¹⁷ The inscriptions with Yu read 余卜 instead of 余貞.

For Period II the distribution of diviner groups — some more examples than others — is similar to the scenario of Period I with 𠩺 as verb. Here we have the following diviner groups while unnamed ones take up the most examples:

Lü 旅:	22558, 22573, 22611, 22791, 22886, 22889, 23029, 23172, 23176, 23403, 23405, 23640, 24969, 24976, 25049, 25057, 25058, 25059, 25061, 25062, 25063, 40917, 40986
Ji 即:	22554, 22610, 22887, 23173, 23227, 22945, 23048, 23071, 23171, 23173, 23227, 23418, 23501, 23520, 23605, 25055, 25060
Da 大:	22616, 22759, 22967, 23085, 23109, 23177, 23346, 23480, 24649, 25047, 25048, 40926, 40984
King:	22567, 22888, 22963, 22966, 22968, 25064, 22969, 22971, 23006, 23114, 24666
Xing 行:	22605, 22760, 23155, 23264, 23347, 23348, 23349, 23411, 23648, 25056, 25074
?? 𠩺:	22577, 23481, 23591, 25902
Xi 喜:	22753, 22826, 23061
?? 𠩺:	22925
Yin 尹:	23345
Chu 出:	25036
Unnamed:	22617, 22884, 22961, 23404, 23408, 23413, 23417, 25037, 25065, 25075, 26908, 40915, ...

For example:

- (20)  甲寅卜旅貞翌乙卯其又于祖乙宰。 22886(2)
 ‘Divining on *jia-yin* day Lü tested next *yi-mao* day make offering to Ancestor Yi of penned sheep.’

Period II is most interesting in representing the transition of the replacement of 𠩺 with 𠩺 as Vsac and the diviner group Da 大 is the best example. For diviner groups that span both Period I and Period II, we see the beginning of the use of 𠩺 as Vsac by popular diviner groups in Period I, as illustrated in Table 2, whereas in Period II we observe greater acceptance of 𠩺 as Vsac among the lesser diviner groups with Da 大 as

the mid-point and 𠄎 as already over the hill and Xing on the verge of complete turnover (see Table 3 below).

Table 3: Vsac with 𠄎 versus 𠄎 in Period II within the same diviner groups

Vsac	Total number	Diviner group			
		Chu 出	Da 大	𠄎	Xing 行
𠄎	50+	24	14	2	1
𠄎	200-	1	13	4	11

Period III tops the list of examples with no fewer than 2,600 inscriptions by unnamed diviners and just a handful by named diviner groups:

He 何: 27042, 27088, 27263, 27321, 27456
 Peng 彭: 27264, 27265, 27543, 27875, 28238
 ?? 貴: 27385, 27607, 27721, 27932
 ?? 狄: 27054, 27244, 30757(?)
 Gu 壹: 27542, 27877
 Da 大: 27376
 Jiao 教: 27734
 Unnamed: 26910, 26919, 26922, 26924, 26936, 27040, 27089, 27090, 27093, 27095, 27100, 27149, 27251, 27326, 27383, 27402, 27444, 27501, 27571, 27622, 28106, 29618, 30681, 31678, 41311, ...

For example:

(21)



貞惠牢又丁。 27326

‘Tested: it is penned bovine to be offered to Ding.’

For Period IV there are sizeable examples (500+ inscriptions) but none with named diviner groups:

Unnamed: 32012, 32057, 32113, 32171, 32302, 32447, 32508, 32512, 32665, 33006, 33180, 33291, 33694, 33880, 34116, 34163, 34168, 34183, 34337, 41660, ...


Period V has very few examples (20 or so) and all with unnamed diviners:

Unnamed: 35356, 36123, 36420, ...

4.2.2 𠄎 as Vex

There are not as many instances of 𠄎 occurring as Vex when compared with those of it as Vsac — not quite one-sixth as many (530+ versus 3600+ inscriptions). Over half of them appear during Period III — 300 inscriptions or 56%. There are 90+ inscriptions in Period IV or 18%, 60+ each in Period II and Period I or close to 13% and merely 7 in Period V or 1+%.

Whereas Period IV and Period V examples are exclusively with unnamed diviners, such as:


- (22)  癸酉貞日夕又食佳若。 33694
'Tested on *gui-you* day the sun in the evening has eclipse; it is approval.'

Period III examples are also overwhelmingly from unnamed diviners as expected and the content mostly concern calamity (又戕) or rain (又雨, 又大雨), among others.

Unnamed:

- 又戕: 26888, 26894, 26898, 26905, 27820, 27878, 27879, 27966, 27975, 27979, 27980, 28012, 28036, 28058, 28065, 28067, 28089, 28094, 28536, ...
- 又雨: 27021, 27254, 27804, 28021, 28108, 28180, 28226, 28228, 28244, 28252, 28257, 28258, 28259, 28265, 28267, 28275, 28967, 29214, 29548, 29656, 29723, 29914, 29984, 29985, 29986, 29987, 29988, 29989, 29990, 29991, 29992, ...
- 又大雨: 26961, 27499, 28021, 28085, 28220, 28252, 28293, 28977, 28422, 28628, 29789, 29995, 29996, 30007, 30008, 30009, 30010, 30011, 30012, 30013, 30014, 30015, 30016, 30017, 30018, 30019, 30020, 330022, 30023, 30024, ...

For example:

- (23)  貞車乙卯酉又大雨。 30826
‘Tested: it is on *yi-mao* day to *you-cut*; there is heavy rain.’

A small number of inscriptions describe the existence of animals such as boars or *mi*-deer:

Unnamed: 28309, 28310, 28317, 28380, 29425, ...

We found only five inscriptions with three named diviner groups:

?? 狢: 27459, 28011, 29092

?? 貫: 27725

Gu 亓: 27302

In contrast to the great number of inscriptions rendered by a great variety of diviner groups found with 𠩺 occurring as Vex, there are just a few inscriptions by several named diviner groups with 𠩺 occurring as Vex during Period I and the rest are by unnamed diviners:

Fu 扶: 20805

Yu 余: 21586

Fu 𠩺: 21610

Chuan 徂: 21635


Zi 子: 21583

Unnamed: with 又𠩺 (‘disaster’): 19996, 20406, 21302, 21303, 21306, 22318, 22398, 22399, 22400, 22401, ...; with 又𠩺 (‘mishap’): 22246, 22247, ...; others: 21052, 40035, ...

What is more interesting is that although the frequency of occurrence during Period II is about the same, there are more inscriptions with named diviner groups:

King: 22765, 22782, 24659, 24664, 24769, 26036, 26187, 26188, 26190
 Da 大: 23581, 23585, 24215, 24439, 24868, 40926
 ?? 𠄎: 22592, 26186, 26189, 26092, 22592, 26092
 Lü 旅: 22539, 23690, 24962
 Ji 即: 24756, 22542
 Xing 行: 23671, 23671
 Chu 出: 23705
 Yin 尹: 23027
 Unnamed: 22541, 24123, 224135, 24150, 26099, 22541, 24150, 24398, ...

For example:

- (23)  丙午卜即貞又氏羌翌丁未其用。 22542(2)
 ‘Divining on *bing-wu* day, Ji tested, there is/are brought Qiang, next *ding-wei* day will use.’

For Periods I and II, the encroachment of 𠄎 upon 𠄎 is also evident among diviner groups that use both 𠄎 and 𠄎 as Vex as given in Table 4.

Table 4: Vex with 𠄎 versus 𠄎 within the same diviner groups


Vex	Period I		Period II	
	Total number	Diviner groups	Total number	Diviner groups
𠄎	1470	King = 7; Fu = 1	30	𠄎 = 2; Chu = 10
𠄎	60+	Fu = 1	60+	𠄎 = 6; Chu = 1; King = 9

The diviner groups Chu and 𠄎 are active mainly during Period II: whereas Chu prefers the use of 𠄎 to 𠄎 𠄎 prefers the opposite as already seen in Table 3 with respect to Vsac and here with Vex. The King of course refers to different kings according to the different Periods. Since the name of the king as a diviner is never given we cannot claim that King so-and-so uses 𠄎 during Period I and later begins to use 𠄎 also. All we can presume is that His Majesty the King, whoever he is, uses 𠄎 almost exclusively for both Vex and Vsac (see Table 3 above) during Period II.

In Period I we see 3 or 4 inscriptions of 𠄎 𠄎 where 𠄎 functions as Vex:

Unnamed: 4073, 8336, 8337, (8338)¹⁸

For example:

- (24)  貞其出𠬪。十二月。 4073(2)
‘Tested: will have assistance. Twelfth month.’


But there are just 2 instances of 𠬪 𠬪 where the first 𠬪 functions as Vex and the following 𠬪 its object:

Shao 𠬪: 20248 𠬪𠬪[卜]𠬪𠬪𠬪人𠬪 𠬪。
𠬪 20249 𠬪寅卜𠬪𠬪人𠬪 𠬪。

In subsequent periods, there are 4 examples of 𠬪 𠬪 in Period III and 2 in Period V but no example of 出𠬪 after Period I.¹⁹ For example:

Unnamed: 29274, 30656, 31189, 31846.

For example:

- (25)  貞𠬪冊至𠬪。大[吉]。 30656
‘Tested: RUK the Record came (will) have assistance. Great (auspiciousness).’

This means that 出 is not seen as a verb occurring before 𠬪 after Period I.


¹⁸ The *Shiwen* considers #8338 to contain the graph 出 before 𠬪, namely, the inscription reads ... RUK ... [出] 又在泚.

¹⁹ The *Shiwen* lists #22619 and #27303(2) containing ‘...又又’ but the ‘又又’ part is unclear in both. The 2 inscriptions, #36731 and #39472 contain ‘...又又’ but we are not sure if it is not preceded by ‘受’.

4.2.3 𠂔 as Vst

𠂔 does not occur as Vst with the meaning of ‘to be abundant’ in Period I while it does occur as Vst in Period III just about 7 times in inscriptions with the expression 𠂔年 by unnamed diviners: 28198, 28199, 28200, 28211, 28216(1), 28216(2), 28224.

For example:

- (26)  ...𠂔耜喪藿...其受又年。 28200
 ‘...we should not till the land of Sang and Guan...will receive good harvest.’²⁰

Examples of 𠂔 as Vst in the expression 受𠂔 𠂔 have already been discussed in §4.1.2.

4.3 𠂔 as affirmative marker


𠂔 does not occur as affirmative marker during Period I or Period II but occurs during Period III in about 32 inscriptions, Period IV in one inscription and Period V in 2 inscriptions, exclusively in the pattern [𠂔+N]_{NP} and none of [𠂔+VP]_{NP}. Such examples include 𠂔車 (27628), 𠂔燭 (27974), 𠂔𠂔²¹ (27995), 𠂔石 (28180), 𠂔𠂔 (28317, 28320), 𠂔𠂔 (28317, 28356, 28357, 28364, 28365, 28368), 𠂔史 (30524), 𠂔𠂔²² (30536, 30539, 30540), 𠂔司 (35362), 𠂔邑 (36429), 𠂔虎 (28300, 28301), 𠂔豕 (28305, 28366), 𠂔鹿 (28327, 28332, 28333, 28335, 28336, 28337, 28339, 28345, 33367), and 𠂔兕 (28392, 28393, 28394, 28395), all occurring in inscriptions with unnamed diviners.

For example:

²⁰ Translation according to Takashima.

²¹ The graph 𠂔, which we could not locate, should have 1 dot inside the top U shape and 1 dot and not 2 dots to each side.

²² The graph 𠂔, which we could not locate, should have 3 small vertical strokes over the ‘示’.

- (27)  貞𠂔𠂔𠂔。 28368(2)
‘Tested: it is the *mi*-deer to catch/be caught.’

4.4 𠂔 occurring with units of numbers

During Period I there are only 5 inscriptions with 𠂔 occurring with numbers:²³

Bin 𠂔: 13443 𠂔, 17525

Xi 喜: 40894

Unnamed: 20548, 21465

The highest frequency of such usage is seen in Period V with 336 inscriptions, followed by Period III with 43, Period IV with 37 and Period II with just 7 inscriptions. Such examples include dates, duration of time, and amount, with very few named diviner groups:

Xing 行: 22550, 22551, 24275

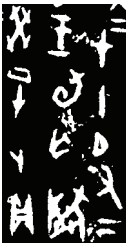
Lü 旅: 24424

Peng 彭: 26907, 27857, 32663

He 何: 27042, 27138

Unnamed: 24593, 26911, 26914, 26915, 32057, 32063, 33371, 35411, 35574, 35818, 37411, 37471, 37861, 37970, 41307, 41456, ...

For example:

- (28)  癸巳卜貞王旬亡𠂔在十月又二。 37970(2)
‘Divining on *gui-si* day tested: the King (during) 10-day-week has no misfortune. On the tenth month and the second day.’

²³ The *Shiwen* lists #17526 as containing the graph 𠂔, but the inscription is too blurred to decipher with certainty.

A summary given in Table 5 confirms our observation all along of the gradual encroachment of 𠂔 upon 𠂔 with respect to individual diviner groups.

Table 5: Connective with 𠂔 versus 𠂔 within the same diviner groups

Connective	Period I		Period II	
	Total number	Diviner groups	Total number	Diviner groups
𠂔	100+	Bin = 5	2	Chu = 1
𠂔	5	Bin = 2	7	

The diviner group Chu, with preference for 𠂔 although occasionally uses 𠂔 for Vsac or Vex, does not use it for a connective.

5. Conclusion

Let us compare the distribution of diviner groups with respect to 𠂔 serving as Vex to that where 𠂔 serves as Vsac. In both cases, inscriptions by named diviner groups are relatively few, but there is one significant difference between Period I and Period II. During Period I inscriptions by named diviner groups are uniformly few, with mostly one (as Vex) or at most 4 (as Vsac) for each diviner group, whereas during Period II the number of inscriptions per diviner group is uneven, some diviner groups having as few as a single instance while others as many as 9 (as Vex) or 23 (as Vsac).

This suggests that the use of 𠂔 in lieu of 𠂔 probably starts with unnamed diviner groups and a number of not so popular diviner groups such as 余, 𠂔, 子, 我, 𠂔, 𠂔, 亞, 𠂔, 𠂔, 扶, 𠂔 and 喜 during Period I, but with popular diviner groups such as 𠂔, 𠂔, 𠂔, 𠂔 and the King showing the beginning of the acceptance of 𠂔 as Vsac or Vex in one or two examples. Then the use of 𠂔 in all capacities dominates the scene from Period II onward. Among the popular diviner groups of Period II, only Chu 𠂔 favors the use of 𠂔 while the King falls for 𠂔.

Table 6 captures the scenario of the use of 𠂔 versus 𠂔 in their various functions during Period I and Period II, displaying the complementary aspect of their functions during Period I as well as the gradual encroachment of 𠂔 upon 𠂔.

Table 6: Complementary functions of 𠩺 versus 𠩻 during Period I and Period II
 (Bold type of diviner group names indicates cross reference;
 # refers to inscriptions with 𠩺 or 𠩻)

		𠩺	𠩻
Period I		as Vsac 祭祀動詞	
	# of e.g.	Total # of 𠩺 = 1940+	Total # of 𠩻 = 200-
		Diviner groups	Diviner groups
	70	𠩺	
	64	𠩻	
	43	王	
	38	爭	
	16	亘	
	13	扶	
	10	𠩺	
	7	自	
	6	勻	
	5	𠩺	子
	4	內	余, 𠩺
	3		我
	2	永	爭, 扶, 𠩺, 𠩻
	1	耳, 叶, 簠, 韋, 𠩺, [出]	𠩻, 𠩺, 亘, 自, 𠩺, 王, 𠩺, 亞
		as Vex 存在動詞	
		Total # of 𠩺 = 1470	Total # of 𠩻 = 60+
	52	𠩺	
	39	𠩻	
	37	爭	
	13	𠩺	
	12	亘	
	10	內	
	7	王	
	3	永, 𠩺, 韋	
	2	𠩺	
	1	扶, 𠩺, 叶, 自	扶, 余, 𠩺, 𠩻, 子
		= Vst (受𠩺年, 受𠩺𠩻, 受𠩻𠩻)	
		Total # of 𠩺 = 171	Total # of 𠩻 = 0
	21	𠩺, 爭	
	13	𠩻	


	2	王, 亘	
	1	𠂔	
	as Affirmative Marker (亘祖 = [亘+N] _{NP} , 亘取 = [亘+VP] _{NP})		
		Total # of 亘 = 100	Total # of 𠂔 = 0
	7	𠂔	
	3	爭	
	2	𠂔, 亘	
	1	王, 𠂔, 韋, 出	
	as Connective 連詞		
		Total # of 亘 = 100+	Total # of 𠂔 = 5
	14	𠂔	
	5	𠂔	
	4	亘, 爭	
	3	王	
	2	𠂔	𠂔
	1	永	喜
	as Noun		
		Total # of 亘 = 0	Total # of 𠂔 = 300+
	34		𠂔
	23		爭
	17		𠂔
	11		王
	2		𠂔, 亘
	1		𠂔
Period II	as Vsac 祭祀動詞		
	# of e.g.	Total # of 亘 = 50+	Total # of 𠂔 = 200
		Diviner groups	Diviner groups
	25	出	
	23		旅
	17		即
	14	大	
	13		大
	11		王, 行
	6		尹
	4		𠂔
	3		喜
	2	𠂔, 兄, 何	

	1	中, 旅, 行	出, 𠂔
	as Vex 存在動詞		
		Total # of 𠂔 = 30-	Total # of 𠂔 = 60+
	10	出	
	9		王
	6		𠂔, 大
	3	中	旅
	2	𠂔	即, 行
	1		出, 尹
	as Vst (受 𠂔 𠂔)		
	1	Total # of 𠂔 = 0	Total # of 𠂔 = 1
	as Connective 連詞		
		Total # of 𠂔 = 2	Total # of 𠂔 = 7
	3		行
	1	出	旅


The close relationship between 𠂔 and 𠂔 leads us to the conclusion that they are graphically and semantically of the same origin. 𠂔 as symbol and graph for the ‘right hand’ is added a reference bar below it to form its verbal usage of ‘to have, to exist’ which is derived from the inference of ‘in the hand’ giving rise to the concept of possession. The verbal usage is extended to include the act of offering as a form of sacrificial rite. ‘To offer’ in its primitive form involves the use of one’s hands. The connection between the nominal and the verbal usage is quite natural. The reference bar may be taken as the frame of the body to which the hand is attached or it may simply be construed as a sign for the feature [+abstract, +verbal].

Moreover, the use of a reference bar below a graph with concrete meaning to generate a graph with abstract meaning has several other examples:

Sun Yirang (1986), referring to the bronze graph corresponding to 𠂔, mentions that 金文有足跡形 ‘in the bronze inscription [it] has the shape of a footprint’ and that 綜考金文甲文, 疑古文 𠂔 爲足止, 本象足跡而有三指... ‘summarizing my investigation of bronze and oracle-bone scripts, [I] suspect that the ancient graph 𠂔 is a toe, originally resembling a footprint and has three references...’ while Luo Zhenyu 羅振玉 (1915 中 63 葉下) explains the graph 𠂔 as follows: 說文解字: ...“一者地也” 按卜辭从止从一, 人所之也 ‘*Shuo Wen Jie Zi*: ... “一 is the ground.” According to the oracle-bone inscription [it] derives from 止 and 一, [it is] where one goes’. Yu Xingwu 于省吾 (1963:122) summarizes 𠂔 as ‘resembling toes walking on the ground’. It is easy to see that 𠂔 symbolizes the big toe while 𠂔 signifies ‘to go’ with the reference bar representing the ground. For example:

- (29)  己卯卜爭貞王乍邑帝若我从之唐。 14200 正(1)
 ‘Diving on *ji-mao* day, Zheng tested: the King builds a settlement, the Supreme Being approves, I follow to go to Tang.’


The graph 大 is generally taken to have the meaning ‘to be big’. Tian Wuzhao 田吳炤 further explains that 大 象人正立形, 大 象人側立形 ‘大 resembles the shape of a man standing facing the front, 大 resembles the shape of a man standing sideways’ (《說文二徐箋異》);²⁴ while the graph 立 is formed by adding a horizontal line at the bottom. Xu Kai 徐鍇 comments that 大, 人也; 一, 地也, 會意 ‘大 is man, 一 is the ground, [forming] an associative compound’ and the graph 立 stands for the word 立 ‘to stand’ with the extended meaning of 位 ‘to be positioned’ or 蒞 ‘to be present, to arrive’. Thus 立 signifies a verbal meaning while 大 is a noun in its root meaning with the reference bar underneath it to represent the ground. For example:

- (30)  庚戌卜設貞王立黍受年。 9525(1)
 ‘Divining on *geng-xu* day, Que tested: [if] the King is present at the planting of the millet, [we] will receive bumper crop.’


While 屮 and 屮 are variant graphs with the meaning of ‘grass shoot’, the graph 屮, according to the *Shuo Wen Jie Zi*, signifies: 生, 進也, 象艸木生出土上 ‘生 is to get ahead, symbolizing grass and plants growing from the earth’. Yu Xingwu (1996:1325) elaborates: 甲骨文从中从一, 一者地也, 正象艸出於土之形 ‘The bone graph [屮] derives from 屮 and 一 [where] 一 is the ground, resembling the shape of grass emerging from the earth’. ‘Grass shoot’ and ‘to germinate’ form another pair with a reference bar

²⁴ Explanation taken from Yu (1996:209, 231).

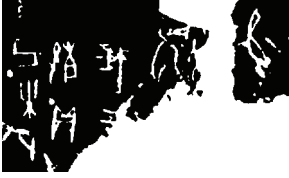

signifying the ground and the verbal form. For example:

- (31)  丁酉卜癸貞帚好受生。 13925 正(1)
'Divining on *ding-you* day, Bin tested; Lady Hao has given child birth.'

Wang Guowei 王國維 describes 奠 as resembling the shape of a wine jug: 酉, 象尊形, while Luo Zhenyu (1915 中 73 葉下) states that 奠 象尊有薦, 乃奠字也 '奠 resembles a jug with a coaster, [it] is the word for 奠' and according to the *Shuo Wen Jie Zi*: 奠, 置祭也 '奠 is to make a sacrifice'. In the OBI, 奠, besides being a place name and a personal name, also means 'to perform a sacrifice.' Presumably libation is involved, so that borrowing the shape of a wine jug with a reference bar indicating verbal usage seems most appropriate. For example:

- (32)  貞奠于丘索。 780(5)
'Divined: perform a sacrifice at/to 丘索。'

The graph 𠂔 signifies 'an army, division of an army'. Luo Zhenyu summarizes it well: 即古文師字, 金文與此同 '[it] is the ancient word for 師, same in the bronze script'. 𠂔 or 𠂕, according to Yao Xiaosui 姚孝遂, is the verbal counterpart of 𠂔: 甲骨文師旅之師作 𠂔, 為名詞, 師旅之止舍則於 𠂔 下增添橫畫作 𠂕 或 𠂖, 乃動詞. '師, "an army", is 𠂔 in the bone script, [it] is a noun, for "stationing an army" add a horizontal stroke below 𠂔 to create 𠂕 or 𠂖, and [it] is a verb'. The verb 'to station (an army)', read as 次, can easily be seen as derived from the noun with a verbalizing reference bar. For example:

- (34)  己未卜設貞我于雉  。 7352(5)
 ‘Divining on *ji-wei* day, Que tested: we station at Zhi.’

The five pairs of graphs cited above give ample support to our construal of 𠂔 and 𠂔 having the same root grapheme but different grammatical functions of nominal versus verbal. To be more precise, the verbal form is derived from the root grapheme designating a noun through the use of a reference bar. The function of this reference bar is to bestow the feature [+verb], simultaneously with certain semantic feature such as a referential frame.

This assumption of the paleographic origin of 𠂔 and 𠂔 gives natural explanation for their close relationship in meaning, in function as well as in phonetic form. If 𠂔 is taken to be or to contain the symbol of a bovine head, how can one explain this close relationship between 𠂔 and 𠂔? Why would a graph designating a bovine head be replaced by a graph designating the hand? As for the difference in style between 𠂔 and 𠂔, since 𠂔 is also used as a sacrificial verb, it is imaginable that its graphic shape is stylized to convey a sense of formality, compared to the free and simple shape of 𠂔. However, exactly because of its simple, free running style popular with unnamed diviners, it finally takes over the formal and more ornate 𠂔. The existence of stylistic variants for 𠂔 suggests that perhaps at the time of Period I there are already graphic variants in use, depending on the regions.

References

- Chang, Ping-Chuan (張秉權). 1959. 小屯. 第二本. 殷虛文字丙編考釋. 上輯二 [Xiaotun. Di Er Ben. Yin Xu Wenzì Bing Bian Kaoshi. Shang Ji Er]. Taipei: Institute of History and Philology, Academia Sinica.
- Guo, Moruo (郭沫若). 1933. 卜辭通纂 [Buci Tongzuan]. Tokyo: Bunkyo Shoten. [Re-edition: 1983. Beijing: Science Press; Beijing: People's Publishing House.]
- He, Jinsong (何金松). 1990. 釋𠩺 [Shi 𠩺]. 中南民族學院學報 [Zhongnan Minzu Xueyuan Xuebao] 1993.3:121-125.
- Hu, Houxuan (胡厚宣). 1944-45. 甲骨學商史論叢 [Jiaguxue Shang Shi Luncong], Series 1 (Vols.1-4) (1944); Series 2 (Vols.1-2) (1945). Chengdu: Cheeloo University.
- Huang, Xiquan (黃錫全). 1981. 甲骨文‘𠩺’字試探 [Jiaguwen ‘𠩺’ zi shitan]. 古文字研究 [Guwenzi Yanjiu] 6:195-206. Beijing: Zhonghua Book Company.
- Jin, Xiangheng (金祥恆). 1962. 釋又𠩺 [Shi 又𠩺]. 中國文字 [Zhongguo Wenzì] 7:773-784. Also in 金祥恆先生全集 [Jin Xiangheng Xiansheng Quanjì], Vol. 3, 1033-1052. Taipei: Yee Wen Publishing Company, 1990.
- Jin, Xiangheng (金祥恆). 1967. 甲骨文𠩺字音義考 [Jiaguwen 𠩺 zi yin yi kao]. 中國文字 [Zhongguo Wenzì] 47.6:2854-2857.
- Keightley, David N. 2000. *The Ancestral Landscape*. Berkeley: Center for Chinese Studies, University of California.
- Liu, Jize (劉紀澤). 1923. 釋𠩺市旦百 [Shi 𠩺市旦百]. 國學叢刊 (東南) [Guoxue Congkan] 1.2. Nanjing: Southeast University.
- Luo, Zhenyu (羅振玉). 1915. 殷虛書契考釋 [Yin Xu Shuqi Kaoshi].
- Nivison, David. 1977. The pronominal use of the verb *yu* (giug: 𠩺, 𠩺, 有) in Early Archaic Chinese. *Early China* 3:1-17.
- Nivison, David. 1978-79. Reply to Takashima. *Early China* 4:30-36.
- Sun, Yirang (孫詒讓). 1986. 名原 [Ming Yuan]. Ji'nan: Qilu Press. (First edition: 1905, woodblock print.)
- Takashima, Ken-ichi (高嶋謙一). 1978-79. Existence vs. reference: decipherment of the word *yu* 𠩺, 又, 有 in Pre-classical Chinese. *Early China* 4:19-29.
- Takashima, Ken-ichi (高嶋謙一). 1980. The Early Archaic Chinese word *yu* in the Shang oracle-bone inscriptions: word-family, etymology, grammar, semantics and sacrifice. *Cahiers de Linguistique – Asie Orientale* 8:81-112.
- Takashima, Ken-ichi (高嶋謙一). 1988. An emphatic verb phrase in the oracle-bone inscriptions. *Bulletin of the Institute of History and Philology Academia Sinica* 59.3:653-694.
- Takashima, Ken-ichi (高嶋謙一). 2005. Placement of inscriptions on oracle-bone plastrons as a guide to decipherment. *Asiatische Studien* 59.1:11-29.

- Tang, Yuming (唐鈺明). 1992. 出、又考辨 [出又 kaobian]. 古文字研究 [Guwenzi Yanjiu] 19:401-407. Also in 著名中年語言學家自選集：唐鈺明 [Zhuming Zhongnian Yuyanxuejia Zixuanji: Tang Yuming], 68-76. Hefei: Anhui Education Press, 2002.
- Tsai, Che-mao (蔡哲茂). 1974. 說出𠂔又𠂔 [Shuo 出𠂔又𠂔]. 中國文字 [Zhongguo Wenzhi] 51:1-11.
- Wang, Yuxin (王育新). 1985. 釋有 [Shi 有]. *Journal of Keshan Teachers College* 1985.3.
- Wu, Kuang (吳匡). 1989. 釋出 [Shi 出]. *Proceedings of 2nd International Conference on Sinology: Section on Linguistics and Paleography*, 891-894. Taipei: Academia Sinica.
- Wu, Qichang (吳其昌). 1934-37. 殷虛書契解詁 [Yinxu shuqi jiegou]. 武漢大學文哲季刊 [Guoli Wuhan Daxue Wenzhe Jikan] 3.2, 3.3, 3.4, 4.2, 4.4, 5.1, 5.4, 6.3. [Re-edition: 1959. Taipei: Yee Wen Publishing Company.]
- Yen, Yi-p'ing (嚴一萍). 1973. 說出 [Shuo 出]. 中國文字 [Zhongguo Wenzhi] 47:87-97.
- Yen, Yi-p'ing (嚴一萍). 1976. 甲骨古文字研究 [Jiagu Guwenzi Yanjiu], Vol. 1. Taipei: Yee Wen Publishing Company.
- Yen, Yi-p'ing (嚴一萍). 1978. 甲骨學 [Jiaguxue], Vol. 2. Taipei: Yee Wen Publishing Company.
- Yen, Yi-p'ing (嚴一萍). 1990. 凡將齋所藏殷虛文字考釋 [Fanjiangzhai Suo Cang Yinxu Wenzhi Kaoshi]. Taipei: Yee Wen Publishing Company.
- Yu, Xingwu (于省吾). 1963. 詩經中止字的辨識 [Shijing zhong zhi zi de bianshi]. 中華文史論叢 [Zhonghua Wenshi Luncong] 3:121-132.
- Yu, Xingwu (于省吾). (ed.) 1996. 甲骨文字詁林 [Jiagu Wenzhi Gulin]. Beijing: Zhonghua Book Company.
- Yu, Zhen (玉振), and Rong Xiao (曉蓉). 1997. 甲骨文“之”、“又”新探 [Jiaguwen “zhi”, “you” xin tan]. 喀什師範學院學報 [Hashe Shifan Xueyuan Xuebao] 1997.1: 71-72.

Radicals as Ontologies: Concept Derivation and Knowledge Representation of Four-Hoofed Mammals as Semantic Symbols

Chu-Ren Huang (黃居仁)¹, Ya-Jun Yang (楊雅君),
and Sheng-Yi Chen (陳聖怡)

*The Hong Kong Polytechnic University*¹

Research on linguistic ontology can shed light on how human beings conceptualize as well as provide powerful tools for knowledge engineering. We show in this paper that the linguistic ontology of Chinese is conventionalized by the radical/semantic symbol of its writing system and a study of this conventionalized system leads to many interesting theoretical discoveries. The *Shuowen Jiezi* (Xu Shen 121), the oldest preserved dictionary of Chinese, is organized according to the radical forms as semantic symbols. Our hypothesis is that for characters sharing the same radical, their meanings must be related to the basic concept of the semantic symbol in a systemic way. In this current research, we focus on the small semantic field of four-hoofed mammals: 羊 ‘caprid’, 鹿 ‘cervid’, 牛 ‘bovine’ and 馬 ‘equine’. We describe the relations between the derived characters and each basic concept to construct a conventionalized ontology, headed by basic concepts expressed by the semantic symbols. Our analysis of the semantic symbol ontologies for these four-hoofed mammals shows that they have similar conceptual structures, which are strongly motivated by their functions in human society. In particular, we show that the conceptual dependencies between the basic concept of a radical and the meanings of the derived characters can be explained by an enriched version of the generative lexicon.

Key words: ontology, radical, generative lexicon

1. Introduction

The system of radicals (*yǐfú*, often directly translated as ‘semantic symbols’) in the Chinese writing system offers a unique opportunity for a systematic and comprehensive comparison between formal and linguistic ontologies. Previous studies adopt either a WordNet-based representation (Wong & Pala 2002, and Hsieh 2006) or a SUMO (The Suggested Upper Merged Ontology)-based mapping (Chou 2005). In one study, Chou & Huang (2010) suggest that the family of Chinese characters sharing the same radical

can be linked to a basic concept by qualia relations, as formalized in Pustejovsky's (1995) generative lexicon theory based on the original definitions of Aristotle. The four original qualia aspects which form the basis of our analysis are formal, constitutive, agentive, and telic. This approach has great implications of accounts for radicals as a linguistically conventionalized ontology. In this paper, we take this approach further and try to account for each group of words sharing the same radical as representing a domain ontology headed by one basic concept. In particular, we examine in detail four animal radicals: 羊 (*yáng*, CAPRID), 鹿 (*lù*, CERVID), 牛 (*niú*, BOVINE) and 馬 (*mǎ*, EQUINE). Among these four animals, 羊, 牛, and 馬 are domesticated and serve specific functions in human society. They are highly related to the daily lives of human beings. One interesting research issue is to see if the derived concepts of these four animal radicals reflect the differences in interaction between the various animals and human.

In this paper, it is crucial to differentiate between the distinct denotations of three related terms: radical, semantic symbol, and basic concept. By radical, we refer to the form-meaning pair that serves as the essential component of the Chinese writing system. The form of a radical is the semantic symbol, while the meaning is the basic concept. In Saussure's terms, the basic concept is the signified, the semantic symbol is the signifier, while the radical is the sign composed of both the signifier and the signified. It is important to note that the difference between radicals and semantic symbols is not well conventionalized and they are often used interchangeably. However, for our study, such a difference is crucial.

Our theoretical foundation is Pustejovsky's (1995) qualia structure, while we refer to the *Shuowen Jiezi* 說文解字 (Xu Shen, 121 CE) as empirical evidence for our analyses whenever possible. In the *Shuowen Jiezi*, Chinese characters are classified into 540 radicals according to the editor Xu Shen's ideas. In this study, we assume that each radical represents a basic concept and that all derived characters are conceptually dependent on that basic concept. Our study aims at accounting for the exact nature of these conceptual dependencies. Combined with previous work, we suggest that conceptual extensions from the basic concept encoded by a radical can be classified into seven main types: formal, constitutive, telic, participant, participating, descriptive (static/active) and agentive.

2. From Hantology to radical-driven ontologies

Our research relies crucially on the accessibility of HANTOLOGY (Chou 2005, and Chou & Huang 2010), which allows us to search for characters according to radicals or ontological concepts. HANTOLOGY is a system expressing the relation of HANZI and meaning clusters (<http://hantology.sinica.edu.tw>). Our radical-driven ontology system

extended the basic structure of HANTOLOGY, which maps the meanings of 540 radicals in *Shuowen Jiezi* with IEEE SUMO. We use the results from analyzing derivative concepts to construct an ontology for each radical. Our current working interface allows easy querying of the existing database as well as recording of new entries.

2.1 Matching radical with ontological concepts

The radical-concept map in Hantology can be retrieved in two ways:

- (i) Search the SUMO concepts classification. Choose a SUMO concept, and then this concept and its lower SUMO concepts will show up on the interface.
- (ii) Search by semantic symbols, the form of radicals. Key in a radical form, and users can get the data of that radical directly.

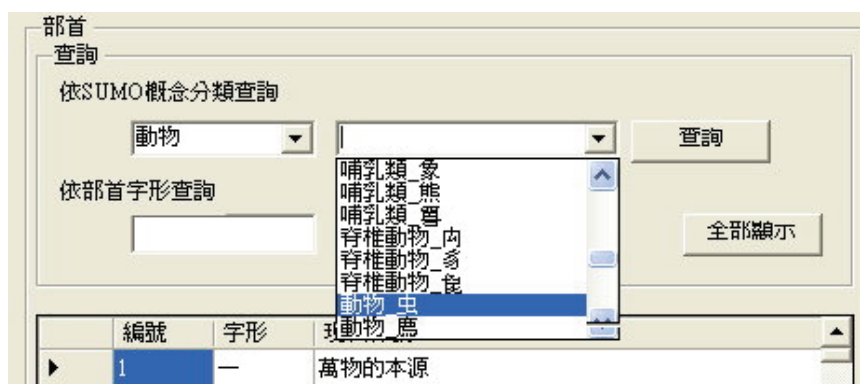


Figure 1: Radical search in the semantic ontology system

2.2 Basic concept

We consult the definitions of each radical in the *Shuowen Jiezi* and analyze the meanings of the characters derived from the same radical to posit the basic concept for each radical. In addition to being compatible with the meaning defined in the *Shuowen Jiezi*, we also expect that the meanings of all derived characters must be derivable with a small set of semantic relations from this basic concept. For example, the basic concept of 羊 is ‘mammal with hoofs’. This interpretation of the radical system allows for the possibility that the concept represented by the semantic symbol differs from the concept it stands for as an independent character. This possibility will be well supported and illustrated by the examples discussed later in this paper.

2.3 Description of concept derivation

Based on the definition in the *Shuowen Jiezi*, we attempt to discover generalizations of the relationship between the meanings of derived characters and the basic concept of a radical. We use Pustejovsky's qualia structure as our framework and add new relations only when the analysis based on the definitions in the *Shuowen Jiezi*, cannot be satisfactorily characterized. The result is that the concept derivation of HANZI radicals falls under seven categories, expanded from the original four qualia aspects, i.e. formal, constitutive, agentive, and telic:

- (i) Formal: related by *kind of* relation and can be further divided into five sub-categories: 'sense', 'material', 'characteristic', 'proper name' and 'atypical'. The 'sense' category can be further divided into five sub-categories: 'sight', 'hearing', 'touch', 'smell', and 'taste'.
- (ii) Constitutive: related by *part of* relation and can be further divided into three sub-categories: 'part', 'member', and 'group'.
- (iii) Telic: meaning of a character related to the basic concept in terms of function or usage.
- (iv) Participant: a character is classified into this category when its meaning as defined in the *Shuowen Jiezi* refers to the basic concept as a participant involved in a specific event.
- (v) Event participating: according to different events, concepts are divided into six smaller categories: 'action', 'state', 'purpose', 'function', 'tool', and 'other'. This category differs from the category of 'participants' in that the concepts refer to a (sub-) event and not to a participant. The concepts of this category describe the properties of the events referred to, but do not refer directly to proper sub-events or participants.
- (vi) Descriptive: the concepts of this category are related by broad descriptions which do not refer to a specific event and can be further divided into two categories: 'active' and 'static' according to the target of description.
- (vii) Agentive: the meaning of the character is related to the basic concept by virtue of how it comes into being (e.g. being born or being produced).

2.4 Concept derivation with multiple concepts

We discovered that a character meaning can be derived from more than one basic concept, in a way different from the traditional category of *huiyi* 會意 'meaning association', where meaning is established through the inference of a relation. In other words, a character may contain more than one semantic symbol. Under 'telic' and

‘participating’, we add a column for ‘related semantic symbol’ to show and link the related derivation concepts. For example, the character 羴 is explained as ‘西戎羊種人也’ (‘people of Xirong who herd CAPRIDS’ (goats/sheep)) and involves two basic concepts that are represented in the character: CAPRID and HUMAN, since the Chinese character form of 羴 contains both 羊 (CAPRID) and 人 (HUMAN). Our ontology system links 羴 with its related semantic symbol 人 to offer cross-referencing in order to build a more realistic ontology of the conceptual convention.

說文常用字

編號	字形	部首	現代釋義	說文釋義	說文釋義2	說文2	說文釋義3	說文3
3192	群	羊	獸畜相...	羴也				
1411	美	羊	味道可口	甘也				
▶ 1047	羴	羊	古代西...	西戎牧...				
6515	羴	羊						
6515	羴	羊						

字：羴

基本概念1 | 基本概念2 | 說文釋義

基本概念1

基本概念：羊(有蹄哺乳動物) ☐ 與部首相同

衍生概念：

☐ 物質 ☐ 感官 ☐ 視覺 ☐ 聽覺 ☐ 味覺

☐ 觸覺 ☐ 嗅覺

☐ 材質 ☐ 特性

☐ 專名 ☐ 非典型

☐ 組成 ☐ 功用 ☐ 參與者 ☒ 事件

相關意符： ☐ 連結

動作 ☐ 狀態 ☐ 目的

功能 ☐ 工具 ☐ 其他

☐ 描述狀態 ☐ 產生

新增下層概念 清除

儲存

Figure 2: The classification of HANZI semantic symbols

3. Domain ontologies of four-hoofed animals as conventionalized by radicals

Among the 540 radicals of the *Shuowen Jiezi*, there are 12 radicals representing hoofed mammals, 5 with sizable domains, still productive in Modern Chinese, and 7 with very small or ‘single element’ domains, not used productively any more.

Table 1: List of radicals related to hoofed mammals in the SWJZ

Radical			SW-Rad. No.	Derived Hanzi in SWJZ ¹	Zoological kinds represented in the SW ²	Corresponding radical No. in <i>Kangxi Zidian</i> (1715)
馬	<i>mǎ</i>	equine/horse	370	Det. 115+8	7	187
牛	<i>niú</i>	bovine/cattle	19	Det. 45+1	1	93
犛	<i>lí, máo</i>	yak	20	Rad. 3+1	1	-
兕	<i>sì</i>	rhinoceros	367	P 1+1	1	-
羊	<i>yáng</i>	sheep/goat	114	Det. 26+2	19	123
羴	<i>huán</i>	mountain sheep/goat	376	P(?) 1	1	-
豕	<i>shǐ</i>	porcine/pig	362	Det. 22+1	12	152
豕/豸	<i>yì</i>	wild boar	363	Det. 5+5	1	153
彘	<i>jì</i>	(head of pig)	364	Rad. 5		58
豚	<i>tún</i>	suckling pig	365	2+1	S	-
鹿	<i>lù</i>	cervid	372	Det. 26+6	14	198
麩	<i>zhì</i>	unicorn	371	Rad. 4+2		-

The five important basic concepts represented by the radicals above are EQUINE (*mǎ* 馬), BOVINE (*niú* 牛), CAPRID, i.e. sheep or goat (*yáng* 羊), PORCINE (and boar) (*shǐ* 豕) and CERVID (*lù* 鹿). The minor domains represented are unicorn (*zhì* 麩); rhinoceros (*sì* 兕), yak (*lí, máo* 犛), mountain CAPRID (*huán* 羴), wild boar (*yì* 豕/豸), and young pig (*tún* 豚).

¹ We would like to thank an anonymous reviewer for providing the complete table and for agreeing to let us include this table in our paper. Any remaining errors are our own.

In this table, radicals marked with ‘P’ are mere ‘pictograms’ (or matrograms in terms of Boltz 1994). The first number under ‘derived’ HANZI represents the total number of entries (i.e.; independent, derived characters) derived from a radical in the SWJZ, the second number refers to written variants only. Radicals marked with ‘Det.’ (determinative) have a true radical function in the perspective of this article: all or many characters in their domain contain the radical as a semantic element. Radicals marked with ‘Rad.’ are represented as a graphic form in the characters in their domain, but they do not play any semantic role, and some are phonetic elements.

² As according to Guo, Needham & Cheng (1999:119-120).

In our study, we deal with four of the five productive radicals representing hoofed animals. As pig/boar (*shǐ* 豕) plays a central role in the domestic life of Chinese, we intend to deal with it separately. Our study involves the classification of the relation between character meanings and the basic concepts of characters as well as a subsequent generalization of the domain ontology after all conceptual relations have been mapped.

3.1 CAPRID domain ontology conventionalized by the radical 羊

Table 2: Distribution in all the relation categories under the radical CAPRID

Formal	16
Participating	1
Constitutive	1
Agentive	4
Formal + Constitutive	1
Formal + Telic	2
Total Number of Characters	25

The basic concept of 羊 derives a range of character meanings in only five relations, namely ‘formal’, ‘participating’, ‘constitutive’, ‘agentive’, and ‘telic’. Among these five classes, the most prevalent conceptual derivation is ‘formal’, which in turn contains three dominant sub-classes: ‘sense’, ‘characteristic’ and ‘proper name’. Two-thirds of the characters with the radical 羊 are derived from this conceptual relation.

Figure 3 is the concept deriving illustration of the radical 羊, with the top concept of CAPRID.

We observe the following generalizations:

(i) The concept cluster belonging to ‘formal’ mainly describes the age, color, and sex of CAPRIDS. For example, ‘𦍋, 羊未足歲也’ (*zhào* (lamb): ‘CAPRID which is less than one year old’) expresses a concept involving the age of the CAPRID. ‘𦍋, 黃腹羊也’ (*fàn* (Mongolian gazelle): ‘CAPRID which has a yellow/brown belly’) involves both the constitutive part of a CAPRID’s belly and its visual attribute. Mainly for domesticated animals, gender is also an important concept, e.g., ‘𦍋, 牡羊也’ (*fēn* (ram): ‘male CAPRID’) and ‘𦍋, 牝羊也’ (*zāng* (ewe): ‘female CAPRID’).

(ii) A smaller cluster denotes events that involve CAPRIDS and is classified as ‘participating’, e.g., ‘𦍋, 西戎羊種人也’ (*Qīāng*: ‘people of Xirong who herd CAPRIDS’). The concept of a particular type of human being is defined by referring to their relation with CAPRIDS. We categorize this as participant-goal. Note that it could be argued that the basic concept should be HUMAN because 𦍋 combines two semantic

symbols 羊 and 人. However, as mentioned earlier, we made a commitment to describe the concept classification according to the *Shuowen Jiezi* unless it can be proven to be incorrect.

(iii) There is a small ‘agentive’ cluster, describing the event and manner of ‘birth’. For example, ‘𦍋，五月生羔也’ (*zhù* (5-month(s)-lamb): ‘lamb born in May’). In addition, there are characters related to castration, such as ‘羴，騷羊也’ (*yí* (stag): ‘castrated CAPRID’). We classify this as ‘agentive’ since it describes an event at the birth of the CAPRID.

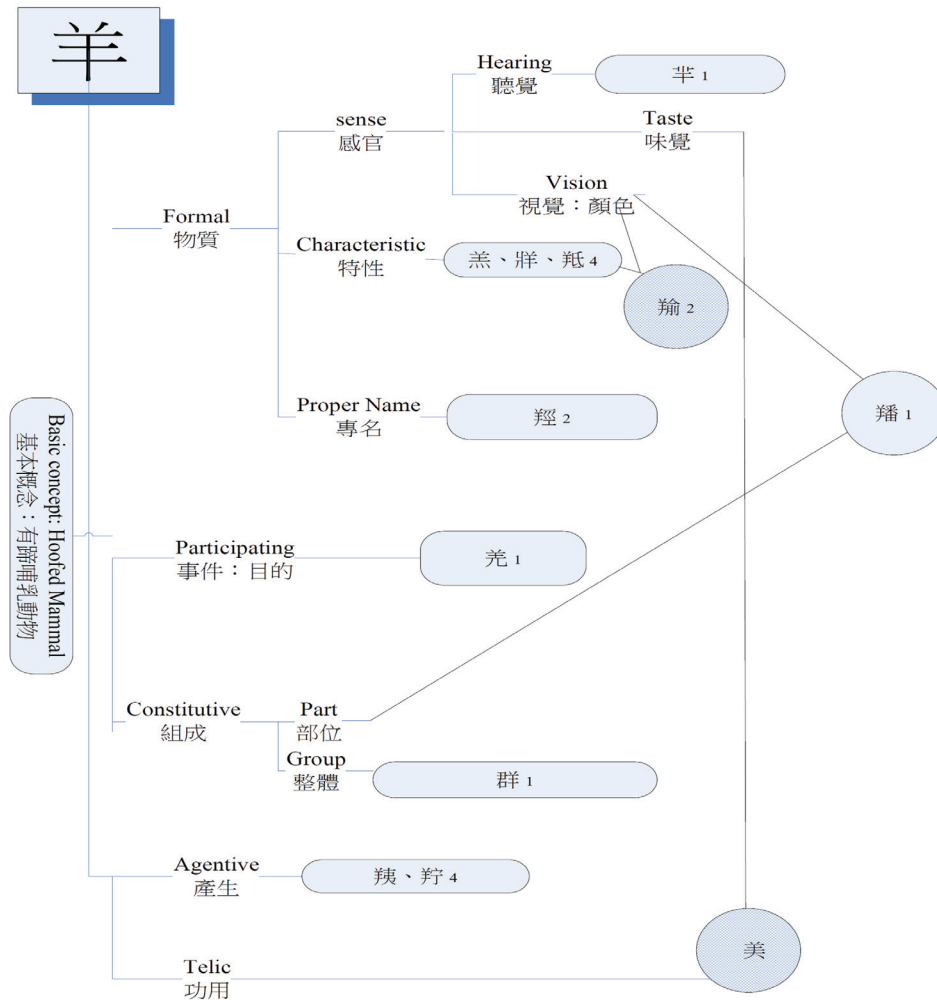


Figure 3: The CAPRID Domain Ontology Conventionalized by the Radical 羊

Some generalizations are given below.

- (i) Among the dominant conceptual relations of ‘formal’, most belong to the ‘proper name’ and ‘characteristic’ sub-classes, e.g., ‘麋，鹿屬’ (*mí* (moose): is a kind of CERVID’ (proper name) and ‘麀：牝鹿也’ (*yōu* (hind) ‘female CERVID’), which describes the sex of the animal and belongs to the ‘characteristic’ sub-class. There are also concepts combining both the ‘proper name’ and ‘characteristic’ sub-classes. For example, ‘麋，麋牡者’ (*chén* (male moose) ‘male moose’) describes both the sub-kind (moose) and the gender of the animal.
- (ii) There is also an ‘atypical’ category in the concepts derived from the radical ‘鹿’, e.g., ‘麇，山羊而大者細角’ (*yán* (Siberian Ibex): ‘big (mountain) goat with thin horns’). A goat is not a kind of CERVID, so this could be a miss-classification, either by convention or by Xu Shen.
- (iii) There is a single example of a concept derived from ‘participating’: ‘麗，旅行也’ (*lì* (beautiful, elegant): ‘traveling’). However, etymology and conceptual convention of this character cannot be clearly defined.

3.3 BOVINE and EQUINE domain ontologies conventionalized by the radicals 牛 and 馬

The BOVINE and EQUINE domain ontologies are much richer than the CAPRID and CERVID domain ontologies, which shows that these two animals are more central to ancient Chinese society and play salient roles in direct human experience. The typical concepts related to BOVINES is being used to till the land and the concept related to EQUINES is being used in transportation. Because of the close and rich first hand information, there are also descriptive events referring to experiences involving these two animals. For instance, the concepts of ‘to scare’ and ‘being scared’ are represented by 驚 (*jīng*) and derived from the drastic and vivid event involving the startling of a horse.

3.3.1 BOVINE domain ontology conventionalized by the radical 牛

Table 4: Distribution in all the relation categories under the radical BOVINE

Formal	22
Participating	10
Constitutive	4
Telic	2
Agentive	1
Formal + Constitutive	3
Formal + Participating	1
Descriptive + Participating	1
Total Number of Characters	44

Six categories of concepts are derived from the basic concept of 牛. They are ‘formal’, ‘descriptive’, ‘participating’, ‘constitutive’, ‘telic’, and ‘agentive’. About half of the characters with the radical 牛 belong to the category ‘formal’. Characters in the categories of ‘participating’, ‘telic’, and ‘agentive’ describe the events BOVINES are associated with in Chinese society and human experiences. We observe the following generalizations:

- (i) The concept cluster belonging to ‘formal’ describes the color, age, gender, and the sound of BOVINES. For example, ‘犛，牛黃白色’ (*piǎo*: ‘BOVINE which is yellow/brown and white’) describes the color of a BOVINE. ‘犛，四歲牛’ (*sì*: ‘BOVINE at the age of four’) describes the age of a BOVINE. The examples of the concept of gender are ‘牡，畜父’ (*mǒu* (male (BOVINE)): ‘father of livestock’) and ‘牝，畜母’ (*pìn* (female BOVINE): ‘mother of livestock’). Besides visual characteristics, auditive features are also observed by human beings and reflected in characters: ‘牟，牛鳴也’ (*móu*: ‘the voice of a BOVINE’). BOVINES are domesticated and people live in close contact with them, so it is not uncommon to hear the sound of BOVINES.

Figure 5: BOVINE domain ontology conventionalized by the radical 牛

- (ii) The concept cluster in the ‘participating’ class reveals the role BOVINES play in Chinese society: ‘犁，耕也’ (*lǐ*: ‘till the land’) indicates that the function the BOVINE serves in human society is to till the land; ‘犒，以芻莖養圈牛也’ (*chú*: ‘feed BOVINE with forage’) and ‘牯，牛馬牢也’ (*gǔ*: ‘BOVINE and EQUINE pen’) indicate that BOVINES are domestic animals.
- (iii) Another important role BOVINES play in Chinese human society is being sacrificed for religious purposes, e.g., ‘牲，牛完全也’ (*shēng* (sacrificial animal): ‘the whole, intact BOVINE [used for offering]’). This function can be found in the ‘telic’ category.
- (iv) There is one example in the ‘agentive’ class: ‘騊，騊牛也（牛被割掉睪丸）’ (*jīè* (bullock/ox): ‘castrated BOVINE’). In order to manage the BOVINES more conveniently and easily, farmers castrate them.

3.3.2 EQUINE domain ontology conventionalized by the radical 馬

Table 5: Distribution in all the relation categories under the radical EQUINE

Formal	39
Descriptive	20
Participating	21
Constitutive	3
Telic	2
Agentive	3
Formal + Constitutive	20
Formal + Participating	5
Constitutive + Participating	1
Total Number of Characters	114

The domain ontology of EQUINE is quite similar to the domain ontology of BOVINE. The concept cluster conventionalized by the radical 馬 falls into six classes, ‘formal’, ‘descriptive’, ‘participating’, ‘constructive’, ‘telic’, and ‘agentive’. The basic difference between BOVINES and EQUINES is their function. BOVINES are used for farming and religious offerings while EQUINES are used for transportation.

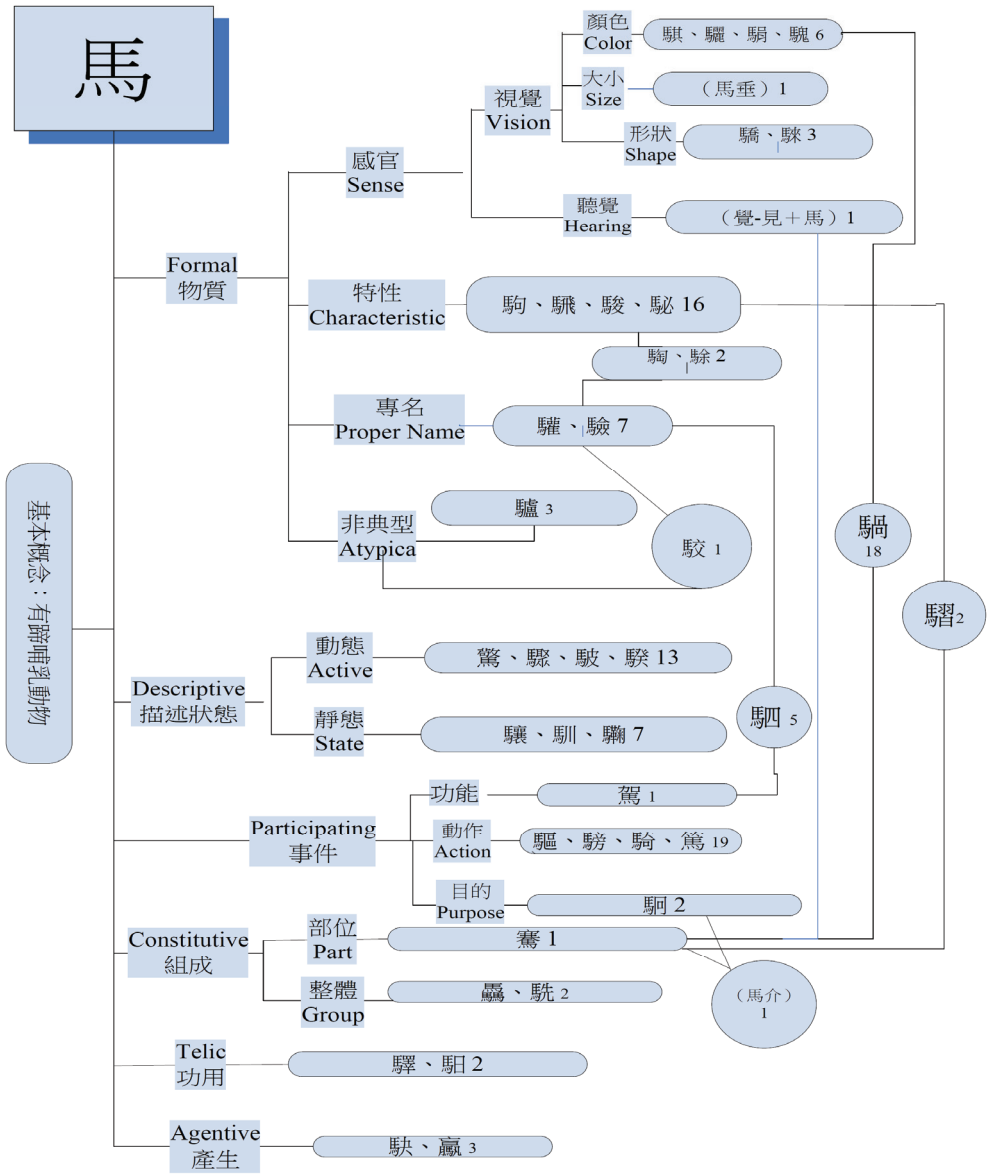


Figure 6: EQUINE domain ontology conventionalized by the radical 馬

Five generalizations are given below.

- (i) About two-fifths of the characters with the radical 馬 can be classified by the

relation ‘formal’. The concept clusters belonging to ‘formal’ mainly describe the color, age, and characteristics of a horse, e.g., ‘驪，馬深黑色’ (*lí* (black): ‘dark black horse’), and ‘騮，黃馬黑喙’ (*guā*: ‘brown horse with black mouth’). There are many examples crossing the ‘formal’ and ‘constitutive-part’ categories. ‘駒，馬二歲’ (*jū* (yearling): ‘two-year-old horse’) is an example describing the age of horses. Some characters describe the qualities of horses, such as strong, good, wild, and fat. ‘駉，馬彊也’ (*zhī*: ‘horse which is strong’) and ‘驍，良馬也’ (*xiāo*: ‘good horse’) are two of the examples.

- (ii) Characters in the ‘descriptive-active’ category mainly describe the manner of an EQUINE when it runs or walks, e.g., ‘驟，馬疾步也’ (*zòu* (fast, sudden): ‘a horse which walks fast’) and ‘驚，馬亂馳也’ (*wù* (gallop, rush about): ‘a horse runs without following certain directions’). Characters in the ‘descriptive-state’ class describe the posture or manner of an EQUINE, e.g., ‘駐，馬立也’ (*zhù* (stop, be stationed at): ‘a horse which stands’) and ‘馴，馬順也’ (*xún* (tame): ‘a horse which is tame’).
- (iii) Because of the close relationship between EQUINES and HUMANS, there are many events involving a horse in the ‘participating’ class. ‘驅，驅馳’ (*qū* (to whip, to run a horse): ‘to gallop a horse’) describes an event where the horse is ridden by a person who causes the horse to gallop. ‘駟，馬載重難也’ (*zhēn*: ‘a horse carrying a heavy load and having difficulty walking’) indicates that the horse is involved in the carrying event.
- (iv) From characters in the ‘telic’ class, we can see that the main function the EQUINE served in human society of Ancient China was transportation, e.g., ‘驛，置騎也’ (*yì* (posting house): ‘a place where riders are given their mounts’).

4. Conclusion

We found that the semantic primitives differentiating the basic concepts of these hoofed mammals involve descriptions of the appearance of these animals:

- (a) 羊 CAPRID represents the shape of (an animal with) four legs and a tail.
- (b) 牛 BOVINE represents the shape of (an animal with) horns that triangulate with a tail.
- (c) 馬 EQUINE represents the shape of (an animal with) a head with a mane, tail, and four legs.
- (d) 鹿 CERVID represents the shape of (an animal with) horns and four legs.

Based on the source of their meanings, it is natural that the concepts derived from the four characters are predominantly classified under the ‘formal’ aspect, especially in the ‘vision’ and ‘characteristic’ sub-categories. Note that there are also many examples which straddle two categories, ‘formal’ and ‘constitutive part’, reflecting the fact that the basic concept definitions involve constitute parts.

Among the four radicals, the CERVID radical derives the least number of diverse concepts. This is a direct reflection of the human experience of the speaking community when Chinese characters were conventionalized. At that time, BOVINE, CAPRID (both goat and sheep) and EQUINE have already been domesticated and figured prominently in people’s daily lives. Hence there were many more objects and events which were related to these animals or which can be described based on people’s shared experience with these animals. We can say that the concept system of the characters also reflects the social and cultural structure of the speakers at the time of its conventionalization.

Indeed, the distribution of the concepts derived from each radical differs. For example, the BOVINE and EQUINE radicals both have a distribution that covers ‘formal’, ‘descriptive’, ‘participating’, ‘constitutive’, ‘telic’, and ‘agentive’. However, CAPRID is different, perhaps owing to the fact that the CAPRID were domesticated mostly for food, while BOVINES and EQUINES served functions in farming and transportation and were thus involved in more human activities.

It is also important to note that religion and rituals played a central role in human experience at that time. This can be observed from the definitions of the derived concepts. We observe these functions from the ‘telic’ category and find that people used BOVINES when making religious offerings, e.g., ‘牲，牛完全也’ (*shēng*: ‘whole, perfect BOVINE’ (used for offerings)). Horses, however, did not seem to be offered as sacrifices but rather belong to a developed system of transportation, e.g., in ‘驛，置騎也’ (*yì* (posting house): ‘a place where riders are given their mounts’).

In conclusion, our study of the four types of hoofed mammals supports our original thesis that the basic concepts as conventionalized by radicals represent domain ontologies of a cluster of concepts derived and marked by that radical. It is important to note that these conventionalized ontologies reflect human experience and knowledge at the time of conventionalization. Hence we see that the domain ontologies of domesticated and non-domesticated animals differ from each other, while further distinctions can be made between animals used for food and those for labor. This direct experience and knowledge are reflected in the qualia used to derive these concepts.

References

- Boltz, William G. 1994. *The Origin and Early Development of the Chinese Writing System*. New Haven: American Oriental Society.
- Chen, Sheng-Yi, Chu-Ren Huang, and Ya-Jun Yang. 2008. The knowledge system of radicals: A study on the representation and cultural implications of yu4 (jade) and shi2 (rock). Paper presented at the 4th International Conference on Literature and Information Technology (ICLIT 2008), November 11-13, 2008. Hong Kong: City University of Hong Kong.
- Chou, Ya-Min. 2005. *Hantology: A Chinese Character-based Knowledge Framework and its Applications*. Taipei: National Taiwan University dissertation.
- Chou, Ya-Min, and Chu-Ren Huang. 2010. Hantology: Conceptual system discovery based on orthographic convention. *Ontology and the Lexicon: A Natural Language Processing Perspective*, ed. by Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari & Laurent Prévot, 122-143. Cambridge & New York: Cambridge University Press.
- Guo, Fu, Joseph Needham (Li Yuese), and Qingtai Cheng. 1999. *Zhongguo Gudai Dongwuxue Shi [History of Ancient Chinese Zoology]*. Beijing: Science Press.
- Hsieh, Shu-Kai. 2006. *Hanzi, Concept and Computation: A Preliminary Survey of Chinese Characters as a Knowledge Resource in NLP*. Tübingen: Universität Tübingen dissertation.
- Huang, Chu-Ren, Sheng-Yi Chen, and Ya-Jun Yang. 2008. Yifu zhishi xitong yanjiu: 'wuguan lei' yifu de gainian yansheng yu zhishi biao Zheng. *Proceedings of the 9th Workshop on Chinese Lexical Semantics (CLSW 2008)*, 95-109. Singapore: Singapore National University.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge: MIT Press.
- Wong, Shun Ha Sylvia, and Karel Pala. 2002. Chinese characters and top ontology in EuroWordNet. *Proceedings of the First Global WordNet Conference*, ed. by U. N. Singh, 224-233. Mysore: Mysore University.
- Xu, Shen. 121. *Shuowen Jiezi*.
- Xu, Zhongshu et al. 1992. *Hanyu da Zidian [The Big Chinese Dictionary]*. Taipei: Jian-Hong Press.

The Last Text of the Last Pazih Speaker

Paul Jen-kuei Li (李壬癸)

Academia Sinica

This is a new Pazih text collected from our very last Pazih speaker after *Pazih Dictionary* and *Pazih Texts and Songs* were published in 2001 and 2002 respectively. It contains a few new lexical forms and expressions not recorded before. It also reflects a historical fact that the Thao people had occupied the Puli Basin before the western plain tribes migrated to Puli in the early 19th century. As based on the informant's memory, the aboriginal people (Thao) still came to Auran, a Pazih settlement, to collect land rent from the Pazih people around 1940 even though there was no title deed for the land. This text may be of interest to linguists, anthropologists and historians.

Key words: Pazih, text, land rent

1. Introduction

Pazih is one of the fourteen still extant Formosan languages,¹ but it is one of the few Formosan languages not very well studied yet. The only publications on the language to date are six articles, “The Pazez-Kahabu language” by Ferrell (1970), “An investigation of the Pazez language in Auran, Puli” (in Chinese) by Lin (1988), “The Pazez language” (in Japanese) by Tsuchida (1993), “Notes on Pazez phonology and morphology” by Blust (1999), “Some aspects of Pazez syntax” by Li (2000), and “Nominalization in Pazih” by Li (2002), and three monographs, *The Pazez Language* (in Chinese) by Lin (2000), *Pazih Dictionary* by Li & Tsuchida (2001)² and *Pazih Texts and Songs* by Li & Tsuchida (2002).³ The very last speaker of Pazih, Phoann Kim-giok (潘金玉) passed away in October 2010. Any new language data or study on the language should be a welcome contribution.

The Pazih texts Tsuchida and I published in 2002 contains 8 texts collected by myself, 4 texts by Tsuchida, 3 texts by Asai, and 1 text by Ino. Both Ino and Asai

¹ There are still a couple of speakers of Kaxabu, another dialect of Pazih.

² It is reviewed by Blust (2003) and Zeitoun (2002).

³ Both Lu (2003) and Lee (2007) adopted an optimality theory approach to reduplication in Pazih as part of their unpublished MA thesis and PhD dissertation respectively, as based on second-hand sources.

collected their texts in the first half of the 20th century, while Tsuchida and I collected ours in the second half of the 20th century. The language was no longer actively spoken when Tsuchida and I started to work with our informants. Our last speaker of Pazih had not used it for about half a century. Fortunately she had a good memory of it. Blust (1999) also collected his Pazih data from her.

I was fortunate to be able to collect a new Pazih text from her in September 2003. There are a few new lexical forms, such as *dadaran* ‘affair’ and *sukarum* ‘enter’, and some new expressions not recorded before. The text is an account of aboriginal people coming to collect the so-called “rent” from the residents of Puli during the Chinese New Year when she was a teenager in 1930s. I (Li 2011:113-116) have also collected a Thao text, which tells about the same tradition. Both texts seem to indicate that the aboriginal peoples, including Bunun, Seediq, and Thao, had larger territories in that area before the peoples from the western plains tribes moved to Puli Basin beginning in 1823 (Chang 1951:88, Appendix II).

2. Pazih text

In this text, I shall follow the same format as in the earlier publication, *Pazih Texts and Songs*, giving interlinear word by word and free translations in both Chinese and English. Information for the informant, date and place of recording is given before the text.

Pazih Text. mupudah dini ki daaxu

山地人來這裡 (收租) The Aborigines Came Here (to Collect Rent)

Told by: Pan Jin-yu (Phoann Kim-giok in Taiwanese) (潘金玉), female, age 89.

Date: 9 September, 2003.

Place: Auran, Puli

Recorded by: Paul Li and Amy Pei-jung Lee

1. uhuza	xasəbuza	isit	yu	turu	a	kawas,	isia
formerly	six	ten	and	three	LIG ⁴	year	at.that.time
從前	六	十	和	三		年	那時

⁴ Abbreviations used in this paper are as follows: AF: Agent-focus; ASP: aspect; CAUS: causative; COLL: collective; CONJ: conjunctive; EXC: exclusive; FUT: future; GEN: genitive; INV: invisible; LF: Locative-focus; LIG: ligature; LOC: locative; NEU: neutral; NOM: nominative; PF: Patient-focus; PL: plural; PRF: perfective; PRG: progressive; RED: reduplication; SG: singular; STA: stative; TOP: topic; VBZ: verbalizer.

nahada di binayu puzah a daaxu.
 have LOC mountain come LIG aborigine
 有 山 來 山地人
 六十三年前，那時有山地人（邵族）來這裡（收租）。
 Sixty-three years ago, there were aboriginal people who came from the mountains.

2. daaxu niam a p<in>alaləŋ-an auran a daxə.
 aborigine our LIG PRF-live-LF Auran LIG land
 山地人 我們的 住處 烏牛欄 地
 我們所住的烏牛欄過去是山地人的土地。
 The land of Auran was the residence of the aborigines.

3. pialay ka binayu a saw ma-laləŋ di auran dini. taxaw
 begin mountain LIG people AF-live LOC Auran here until
 開始 山 人 住 烏牛欄 此地 直到
 niam a apu-ʔapu-an pudah dini ma-laləŋ.
 our LIG RED-grandparent-LF come here AF-live
 我們的 祖先 來 此地 住
 從前山地人（邵）住在烏牛欄這個地方，直到我們的祖先才來這裡住。
 In the past the mountain people (Thao) had lived in Auran here until our ancestors
 came here to live.

4. liaka ma-raxiw di binayu lia ki daaxu.
 then AF-escape LOC mountain ASP NOM aborigine
 然後 逃 山 山地人
 後來山地人逃到山上去了。
 Then the aboriginal people escaped to the mountains.

5. darahən a yasia.
 aborigine LIG 3PL.INV
 山番 他們
 他們是山地人。
 They are the mountain (aboriginal) people.

6. taxaw razəm a dali, mu-puzah m-ara niam a dini a
 till New.Year LIG day AF-come AF-take our LIG here LIG
 到 過年 日子 來 拿 我們的 此地

rumut, tupalis yu inusat.

meat cake and wine

肉 糕 和 酒

到了過年的日期，他們來我們這裡收取我們的肉、年糕和酒。

On the New Year's Day, they came here to collect our meat, rice cake, and wine.

7. puzah taxaw dini m-ara.

come reach here AF-take

來 到 此地 拿

他們來這裡收取那些（東西）。

They came here to collect them.

8. yaku ka isit supad a kawas rəəzaw.

1SG.NEU TOP ten four LIG year only

我 十 四 年 只

（那時）我才十四歲。

I was fourteen years old (then).

9. yaku isia haapət maa-hadas.

1SG.NEU at.that.time like COLL-play.around

我 那時 愛 一起遊玩

我那時很愛跟大家一起玩。

At that time I was fond of playing around with other kids.

10. razəm ka nahada batan m-ukusa maa-hadas di daran.

New.Year TOP have friend AF-go COLL-play.around LOC road

過年 有 朋友 去 遊玩 路

過年時有朋友到道路上去一起玩。

During the new year, there were friends going to play around on the roads.

11. mi-kita yasia daraxən puzah dini m-iliw namisiw a

AF-see 3PL aborigine come here AF-carry their LIG

看 山地人 到 此地 挑 他們的

a-iliuʔ-ən balax a isia.

FUT-carry-PF stuff LIG that

挑 東西 那

（我）看到他們山地人到這兒來挑他們所要挑的那些東西。

(I) saw them, the aborigines, came here to carry all the stuff with poles on their shoulders.

12. mi-kita yamin raki-rakih-an.
 AF-see 1PL.EXC RED-child-LOC
 看 我們 小孩
 他們看到我們這些年輕人。
 Then (they) saw us children.
13. liaka mi-kita adan a saw mamaləŋ, mu-lubulup isia pudah
 then AF-see one LIG people man AF-follow him come
 然後 看 一 人 男 跟隨 他 來
 masu-karun di niam a xumak.
 go-inside LOC our LIG house
 裡面 我們的 家
 然後（我）看到一個男人，（我）跟著他進入我們的房子。
 (I) saw one man and followed him entering our house.
14. naki a ina ka mi-kita isia a saw ka, “riak xəxən.
 my LIG mom TOP AF-see that LIG people TOP good very
 我的 母親 看 那 人 好 很
 alu-ʔalu, alu su-karum di karum”
 RED-come come get-in LOC inside
 來 來 進裡面 裡面
 我母親看到那個人就說，「很好。來來，請進來。」
 My mother saw that man and (said), ‘Very good, come on, come on in.’
15. liaka mi-kiliw yaku m-ukusa m-ara adan a tupalis tuxubus
 then AF-call 1SG.NEU AF-go AF-take one LIG cake sweet
 然後 叫 我 去 拿 一 糕 甜
 a tupalis yu adan a saypu a tupalis.
 LIG cake and one LIG dried.radish LIG cake
 糕 和 一 蘿蔔 糕
 然後（她）叫我去拿一塊甜年糕和一塊蘿蔔糕。
 Then (she) called me to take one piece of sweet cake and one piece of dried radish cake.
16. mu-baxa yasia dusa a tupalis.
 AF-give them two LIG cake
 給 他們 兩 糕
 （我）給他們兩塊年糕。
 (I) gave them two pieces of rice cake.

17. naki a ina ka, m-ukusa m-ara rumput, pataru a rumput
 my LIG mom TOP AF-go AF-take meat chicken LIG meat
 我的 母親 去 拿 肉 雞 肉
 yu baruzak a rumput.
 and pig LIG meat
 和 豬 肉
 我母親去拿肉，雞肉和豬肉。
 My mother went to take meat: chicken meat and pork.
18. yu naki a aba ka, m-asu dusa a inusat. liaka suax-ən
 and my LIG dad TOP AF-bring two LIG wine then pour-PF
 並且 我的 父親 帶 兩 酒 然後 倒
 di nasia b<in>uriak a patakan a isia.
 LOC their make<PRF> LIG bamboo LIG that
 他們的 做 竹 那
 並且我父親帶兩瓶酒，然後倒進他們竹製的管子中。
 And my father brought two bottles of wine and then poured it into their containers
 made of bamboos.
19. liaka ma-hata-hatan lia ki darahən a isia.
 then STA-RED-happy ASP NOM aborigines LIG that
 然後 高興 山地人 那
 然後那些山地人都很高興。
 Then those aborigines were very happy.
20. liaka ma-rukat lia.
 then AF-go.out ASP
 然後 出去 了
 然後他們就出去了。
 Then they left.
21. adan a kawas ka, mu-puzah inan.
 one LIG year TOP AF-come again
 一 年 來 又
 一年（後）他們又來了。
 (They) came again a year later.

22. razəm dali ka, mu-puzah inaŋ.
 New.Year day TOP AF-come again.
 過年 日 來 又
 過年那天他們又來了。
 They came again during the New Year's Day.
23. haiki dusa a kawas mi-kita aku.
 seem two LIG year AF-see 1SG.NOM
 似乎 二 年 看 我
 我似乎看到了兩年。
 I seemed to have seen it for two years.
24. taxu ma-səəzəm a dali, liaka haapət ma-baza yamisiw
 till STA-long.time LIG day then like STA-know their (INV)
 到 久 日 然後 想要 知道 他們的
 maluhusu a dadaran lia ki namisiw a saw.
 in.that.way LIG affair ASP NOM their LIG people
 那樣 事情 他們的 人
 經過很久之後，他們那些人很想知道他們以前的那種事。
 After a long time, their people would like to know such a kind of affair.
25. liaka mi-kiliw yaku m-ukusa di uhuni p<in>a-laləŋ-an a
 then AF-call 1SG.NEU AF-go LOC now PRF-live-LF LIG
 然後 叫 我 去 現在 住過的
 dida namisiw daxə a didua.
 there their land LIG there
 那裡 他們的 地 那裡
 然後他們請我到他們現在住的那個地方去。
 Then (they) invited me to go to their present residence over there.
26. liaka mu-kawas pa-dudu yaku.
 then AF-talk VBZ-ask 1SG.NEU
 然後 說話 問 我
 然後他們問我。
 Then (they) talked and asked me.
27. “ma-baza niam a apuʔapu-an m-ukusa di imu-nan
 STA-know our LIG ancestors AF-go LOC 2PL-LOC
 知 我們的 祖先 去 你們處

m-ara isia ka, asay a dadaran,” la ka naki.⁵
 AF-take that CONJ what LIG affair ? ? my
 拿 那 什麼 事情 我的
 (他們說,) 「(我們)知道我們的祖先到你們住的地方收那些東西,可是
 我們並不知道是什麼事情。」他們這樣問我。
 (They said, “(We) know that our ancestors went to your place to collect (rent), but
 we don’t know what kind of affair it was,” they said to me.

28. “asay a dadaran ka, ma-mərək aku.
 what LIG affair TOP STA-not.know 1SG.NOM
 什麼 事情 不知 我
 「什麼事情,我並不知道。
 “I don’t know what kind of affair it was.

29. mi-kita rəəzaw mata-dusa a razəm.
 AF-see only times-two LIG New.Year
 看 只 二次 過年
 我只見過兩次那樣的過年。
 I saw such a new year only twice.

30. mi-kita rəəzaw aku. asay a dadaran ka, ma-mərək
 AF-see only 1SG.NOM what LIG affair CONJ STA-not.know
 看 只 我 什麼 事情 不知
 aku.” sən naki.
 1SG.NOM say my
 我 說 我的
 我只是看到而已。我並不知道什麼事情。」我那樣說。
 I simply saw it. I don’t know what kind of affair it was,” I said.

31. liaka m-itun lia ki nasia a adaŋ a mamaləŋ a saw.
 then AF-rise ASP NOM 3PL.GEN LIG one LIG man LIG people
 然後 起來 他們的 一 男 人
 後來他們當中有個男人站起來。
 Then one of the men stood up.

⁵ The structure of *la ka naki* is not clear.

32. m<a>axa-ʔasay a saw.
do-PRG-what LIG people
在做什麼 人
(我不知道)他是在做什麼的。
(I don't know) what he was.
33. m<a>axa-ma-taru a m<a>axa-ʔasay a mamaləŋ a saw
do-PRG-STA-big LIG do<PRG>what LIG man LIG people
做大 做什麼 男 人
m-itun kizəx ma-kuas.
AF-rise stand AF-talk
起來 站 講
有一個可能是重要的男人站起來說話。
Someone who might be in an important position stood up and talked.
34. “uħuza ka, niam a saw ma-laləŋ di nimu a
Formerly TOP our LIG people AF-live LOC your LIG
從前 我們的 人 住 你們的
auran a rətən na dua.
Auran LIG village GEN there
烏牛欄 村莊 那裡
「從前我們的人就住在你們烏牛欄庄那裡。」
“In the past, our people lived in your village of Auran.
35. nimu a mia-dua ka, kuaŋ a kari-karit sa paxuruma
your LIG go-there TOP not LIG RED-fields plant
你們的 那邊 沒 田園 種
piaxun, dukun, tawtaw, yu yaadisaysay.
millet taro peanut and various.things
小米 芋 花生 和 各種東西
你們那裡沒有田園可以種小米、芋頭、花生以及其他各種(植物)。
At your place, there were no fields for planting millet, taros, peanuts, and various other things.
36. niam a saw ka, mə-kən ya-yasia rəzaw, ini mə-kən sumay.
our LIG people TOP AF-eat RED-those only not AF-eat rice
我們的 人 吃 那些 只 不 吃 飯
我們的人只吃這些東西，不吃飯。
Our people only ate those things, but not rice.

37. ma-mərək paxuruma mulasi, liahaka puzah di binayu liamin.
 STA-unable plant rice therefore come LOC mountain we
 不會 種 稻 因此 來 山 我們
 我們不會種稻子，因而到山上來。
 (We) didn't know how to plant rice, so we came to the mountains.
38. liaka pa-palaləŋ-ən nimu auran a saw.
 then CAUS-live-PF 2PL.GEN Auran LIG people
 後來 使住 你們的 烏牛欄 人
 後來我們才讓你們的人住在烏牛欄。
 Later (we) made you live in Auran.
39. isia a dadaran uzay nimu a haziah, yamin talima haziah.
 that LIG affair not 2PL.GEN LIG bad 1PL.EXC.NEU self bad
 那 事情 不 你們的 壞 我們 自己 壞
 那件事不是你們不好，而是我們自己不好。
 As for that, it is not you who were bad, but that it is ourselves who were bad.
40. ma-mərək ba-bizu, ma-mərək mi-kita ba-bizu
 STA-unable RED-words STA-unable AF-see RED-book
 不會 字 不會 看 書
 (我們)不識字，不會看書。
 We did not recognize words or read books.
41. asay a dali ka, ma-mərək asay a ilas. asay a
 what LIG day TOP STA-not.know what LIG month what LIG
 什麼 日 不知 什麼 月 什麼
 dali ka, riak a paxuruma mulasi.
 day TOP good LIG plant rice
 日 好 種 稻米
 (我們)不知何時何日何月，什麼時候要種稻子。
 We did not know what days or months were good for planting rice.
42. haka ma-baza aku.”
 really STA-know 1SG.NOM
 真的 知 我
 (以上)是我確實知道的。」
 I know that's true.”

43. niam a puri a ma-taru a saw isia liaka, “hă,
 our LIG Puli LIG STA-big LIG people that later Oh
 我們的 埔里 大 人 那 後來
 maluhusu laila.”
 in.that.way say
 原來如此 說
 埔里的長官（鎮長）後來說，「啊，原來如此！」
 The mayor of Puli later said, “Oh! Is that so!”
44. darahən a yasia a saw liaka maluhusu rəəzaw.
 aborigines LIG they LIG people later in.that.way only
 山地人 他們 人 後來 原來如此 只
 後來他們山地人說，「確實如此。」
 The aborigines later said, “It’s only like that.”
45. ini mi-kita yamin haziah a saw, mi-kita yamin riak
 not AF-see 1PL.EXC poor LIG people AF-see 1PL.EXC good
 沒 看 我們 不良 人 看 我們 好
 （你們）沒有瞧不起我們，（你們）對我們不錯。
 (You) didn’t look down upon us; (you) treated us well.
46. razəm m-ukusa ka, mu-baxa yamin riak a ka-kan-ən
 New.Year AF-go TOP AF-give 1PL.EXC good LIG RED-eat-PF
 過年 去 給 我們 好 食物
 a yaadisaysay: rumut, inusat, tupalis.
 LIG various.things meat wine cake
 各種 肉 酒 糕
 過年時（我們）到（你們的地方）去，（你們）給我們各種好的食物：肉、
 酒、年糕。
 At new year (when we) went (to your place), you gave us various good things to
 eat: meat, wine and rice cake.
47. adaŋ a kawas, adaŋ a kawas, mu-baxa yamin.
 one LIG year one LIG year AF-give 1PL.EXC
 一 年 一 年 給 我們
 每年（你們都）給我們。
 (You) gave (them) to us every year.

48. riak xəxən mi-kita yamin binayu a saw.”
good very AF-see 1PL.EXC mountain LIG people
好 很 看 我們 山 人
(你們)待我們山地人很好。」
(You) were very nice to us mountain people.”

49. taxaw dini ki naki a k<in>awas.
till here NOM 1SG.GEN LIG speech<PRF>
到 此地 我的 講
我的話到此為止。
My speech ends here.

References

- Blust, Robert. 1999. Notes on Pazeh phonology and morphology. *Oceanic Linguistics* 38.2:321-364.
- Blust, Robert. 2003. Review of Paul Jen-kuei Li & Shigeru Tsuchida (2001), *Pazih Dictionary*. *Oceanic Linguistics* 42.2:531-540.
- Chang, Yao-chi (張耀錡). 1951. *Pingpuzu Sheming Duizhaobiao* 平埔族社名對照表 [A Comparative Name-list of Pei-Po-Fan's Villages through the Historical Ages]. Taipei: Historical Research Commission of Taiwan Province.
- Ferrell, Raleigh. 1970. The Pazeh-Kahabu language. *Bulletin of the Department of Archaeology and Anthropology, National Taiwan University* 31-32:73-96.
- Lee, Amy Pei-jung (李佩容). 2007. *A Typological Study on Reduplication in Formosan Languages*. Colchester: University of Essex dissertation.
- Li, Paul Jen-kuei (李壬癸). 2000. Some aspects of Pazeh syntax. *Grammatical Analysis: Morphology, Syntax, and Semantics: Studies in Honor of Stanley Starosta*, ed. by Videia P. De Guzman & Byron Bender, 89-108. Honolulu: University of Hawai'i Press.
- Li, Paul Jen-kuei (李壬癸). 2002. Nominalization in Pazih. *Language and Linguistics* 3.2:227-239.
- Li, Paul Jen-kuei (李壬癸). 2011. *Thao Texts and Songs*. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei (李壬癸), and Shigeru Tsuchida (土田滋). 2001. *Pazih Dictionary*. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei (李壬癸), and Shigeru Tsuchida (土田滋). 2002. *Pazih Texts and Songs*. Taipei: Institute of Linguistics, Academia Sinica.
- Lin, Ying-chin (林英津). 1988. Bazehaiyu: Puli Ailan diaocha baogao 巴則海語: 埔里愛蘭調查報告 [An investigation of the Pazeh language in Puli]. *Taiwan Fengwu* 台灣風物 [The Taiwan Folkways] 39.1:176-200.
- Lin, Ying-chin (林英津). 2000. *Bazehaiyu* 巴則海語 [Pazih Language]. Taipei: Yuanliou.
- Lu, Michael Shun-chieh (呂順結). 2003. *An Optimality Theory Approach to Reduplication in Formosan Languages*. Taipei: National Chengchi University MA thesis.
- Tsuchida, Shigeru (土田滋). 1993. Pazeh go パゼツヘ語 [Pazeh language]. *Gengogaku Daijiten* 言語学大辞典 [The Sanseido Encyclopaedia of Linguistics] 5:302-310. Tokyo: Sanseido.
- Zeitoun, Elizabeth. 2002. Review of Paul Jen-kuei Li & Shigeru Tsuchida (2001), *Pazih Dictionary*. *Language and Linguistics* 3.2:481-488.

Some Issues in the Study of Chinese Poetic Prosody

Hongming Zhang (張洪明)^{1,2} and Chenqing Song (宋晨清)³

*University of Wisconsin-Madison*¹

*Nankai University*²

*State University of New York, Binghamton*³

This paper begins by defining poetic prosody and discussing various building elements of Chinese poetic prosody, such as rhythmic grouping, tonal assignment, and rhyming patterns. It then distinguishes between poetic prosody itself and the performance of metered texts in order to provide a better understanding of tonal prosody and the nature of rhyming.

Key words: Chinese poetry, prosody, rhythmic pattern, tonal assignment

1. Introduction

Poetic prosody, *shilü* 詩律 in Chinese, has been a point of discussion in the fields of both literature and linguistics. Both literary scholars' and linguists' interest in this topic can be explained by the fact that prosody as seen in human poetry is indeed both literary and linguistic: certain sound properties in human language are utilized in creative literary architecting and result in distinctive patterns. While literary scholars often focus on the esthetic beauty of the patterns or poets' personal creations and choices,¹ linguists direct their attention more towards researching the linguistic materials that build these patterns. Questions such as what linguistic elements are chosen to form the patterns, to what degree they are subject to creative literary manipulation, and what is universal or language specific in the formation of the patterns used in different poetic traditions around the world all interest us enormously because linguists believe that the answers to these questions bear significant importance on our understanding of language itself. We have approached the subject covered in this paper from the standpoint of linguists and our scope of study is in Chinese poetry.

In this paper, we will not take on specific questions about meter, such as the formation of *lǜshī* 律詩, the rhythms of *Shijing* 詩經, or the patterns in Shen Yue's 沈約

¹ This is not to say that poetic prosody studies in literature do not concern the regulation of patterns. Rather such studies do not necessarily refer to linguistic principles and properties.

poems, which of course are fascinating and deserving of scholarly endeavor. The reason for this is that we feel, in the study of poetic prosody, there currently remains much confusion and inconsistent usage of basic concepts, hindering the advancement of our research at present and in the future. Therefore, this paper has been written in an effort to clarify confusion, define concepts, and lay common ground for future discussions of interesting topics in this area.

We will start this paper by defining poetic prosody, and will then discuss various building elements of Chinese poetic prosody, including rhythmic grouping, tonal assignment, and rhyming patterns. Afterwards we will distinguish between poetic prosody itself and the performance of metered texts, such as chanting, reciting, and singing. Finally, we will take a detailed look at tonal prosody and the nature of rhyming.

2. Linguistic elements in poetic prosody

In most literature on this topic, there are several English terms that are often used very interchangeably: poetic prosody, meter, and rhythm. In Chinese, we also find words that have very closely related meanings, such as *shilü* 詩律, *yunlü* 韻律, *jielü* 節律, *yinlü* 音律, and *shenglü* 聲律. What further confuses readers is that the English words are often translated into the different Chinese terms by different scholars, some assume that these words have the same denotation while others distinguish them in one way or another. We start our discussion here by doing some philological analyses of these English and Chinese terms. First, let's start with the word "prosody". Several definitions of this word are given in (1):

- (1) a. Merriam-Webster
 - i. the study of versification; especially the systematic study of metrical structure;
 - ii. a particular system, theory, or style of versification;
 - iii. the rhythmic and intonational aspect of language.
- b. The American Heritage Dictionary of the English Language
 - i. the study of the metrical structure of verse;
 - ii. a particular system of versification.
- c. *The English Prosody* by Asa Humphrey (1847)
 - i. prosody is the fourth and last part of grammar, and treats the construction of verse, and comprises all connected with poetical composition, in distinction from that of prose (p.1);
 - ii. prosody treats of versification, and teaches the rules for poetical composition, and all pertaining thereto; viz. the elementary and

component parts of verse; the different orders or kinds of verse, with their different forms and metrics; also, reading and scanning of verse, description of poems (p.2).

These definitions of prosody share at least one point: poetic prosody is the study of some linguistic forms of verse. To some these forms include the metrical structure only, to others, the forms include other formal properties of verse as well.

The term meter refers to something more specific as compared to prosody. Meter can refer to the systematically arranged and measured rhythm in verse. It can be the rhythm that continuously repeats a single basic pattern, e.g. iambic meter; it can also be a complex rhythm that is composed of several basic beats, e.g. a pentameter that contains five strong-weak beats. It is obvious that meter is only one of the elements studied in prosody.

Rhythm compared to the first two terms is not as technical. In daily language, rhythm is used often on various occasions, such as the rhythm of a song, the rhythm of city life, and the rhythm of a novel. Among these usages, the concept of rhythm always involves the fluctuation or variation of a certain element and the contrast of the different levels of this element. In languages with stress, words' rhythm is created by the strong and weak beats of the stress. In poetry, rhythm often is the acoustic effect of the meter. Fabb & Halle (2008) argue that rhythm and meter are different. Their observation is given in the following example:

(2) Fabb & Halle (2008:8)

Pléasure néver is at hóme	
(* * (* * (* * (*	0 ⇒
(* * (* * (*	1 ⇒
(* * (*	2 ⇒
*	3
At a tóuch swéet pléasure méltech	
(* * (* * (* * (*	0 ⇒
(* * (* * (*	1 ⇒
(* * (*	2 ⇒
*	3

Fabb and Halle point out that the stressed syllables do not always fall in the strong positions of the trochaic tetrameter, nor do the unstressed syllables always appear in the weak positions. The overall effect of the beat is a combination of both the word stress and the regular weak-strong alternation of the meter.

After briefly introducing the English terms, we are back to the Chinese terms, which share one morpheme *lǜ* 律.

(3) 律

《說文解字·彳部》：“律，均布也。從彳，聿聲。”

[In *Chi* radical of *Shuowen jiezi*: “*lǜ* 律 means to completely cover. It follows the form of 彳, and is pronounced as 聿.”]

段玉裁注：“律者，所以範天下之不一而歸於一，故曰均布也。”

[Duan Yucai’s annotation: “*lǜ* 律 is by which we set a model for what is not unified in the world to be unified. Therefore, *Shuowen jiezi* says it means completely cover.”]

《爾雅·釋詁》：“律，常也，……法也。”

[In *Gu* explanation of *Erya*: “*lǜ* 律 means constancy...means law.”]

《廣韻·術韻》：“律，法也。”

[In *Shu* rhyme of *Guangyun*: “*lǜ* 律 means law.”]

The most prominent meaning of *lǜ* 律 is “regulation”. Therefore *shilǜ* 詩律 is often defined as “the rules of versification”. However, in this broad definition, the rules are not necessarily the prosodic rules. Often Chinese literary critics include the requirement of *duizhang* 對仗 [parallelism] in the scope of *shilǜ* 詩律. Besides the rules for *duizhang* 對仗, there are other subsets of rules in *shilǜ* 詩律. These rules regulate other linguistic elements, which are the building blocks of the prosody of a poem. Some of these elements are used in Chinese as well as in other literary traditions, some are found only in Chinese. The first type of element is rhythm, which is closest to the Chinese term *jie* 節.

(4) 節

《爾雅·釋樂》：“和樂謂之節。”

[In *Yue* explanation of *Erya*: “Harmonizing music is called *jie* 節.”]

邢昺疏：“節，樂器名，謂相也。”

[Xing Bing’s annotation: “*jie* 節 is the name of a musical instrument. It refers to *xiang* 相.”]

Jie 節 was used as a musical term before it was used in poetry or linguistics. The *Erya* 爾雅 entry refers to it as the musical instrument that controls the rhythm (beat and tempo) of music. The word also referred to the action of rhythm control in a musical performance. The rules of rhythm, or meter, is *jielǜ* 節律.

In the scope of prosody, there is also a set of rules that regulate rhyming. These rules stipulate which vowel or coda consonant a rhyming word should have when a line

is supposed to end with a rhyming word and when to change to a different rhyme. We would like to use the Chinese word *yun* 韻 for rhyme or rhyming and, for the rules of rhyming, naturally *yunlü* 韻律.

(5) 韻

宋本《玉篇》引《聲類》：“音和曰韻也。”

[Song edition *Yupain* quotes *Shenglei*: “Sounds that are harmonious are called *yun* 韻.”]

《文心雕龍·聲律》：“異音相從謂之和，同聲相應謂之韻。”“雙聲隔字而每舛，疊韻雜句而必睽。”

[In *Shenglü* chapter of *Wenxin diaolong*: “Different sounds follow each other. This is called harmony. Identical sounds echo each other. This is called *yun* 韻.”]

However, in the modern Chinese language, *yunlü* 韻律 also has the meaning of prosody in general.

(6) 韻律

《漢語大詞典》：“聲韻和節律。指詩詞中的平仄格式和押韻規則。”

[*Hanyu da cidian*: “*Shengyun* 聲韻 and *jielü* 節律 refer to the *ping-ze* tonal patterns and the rules for rhyming in poetry.”]

We suggest that scholars studying poetic prosody ought to be cautious when they use the term *yunlü* 韻律, because of the ambiguity that it has here. In fact, we are reluctant to use *yunlü* 韻律 as prosody in our studies. It is very hard to find a better word than *yunlü* 韻律 for the rules of rhyming, and therefore the term should be given to this concept as a priority.

In the study of Chinese poetic prosody, we should not forget the rules that regulate the assignment of tonal categories. This unique prosodic characteristics represented by the Chinese regulated verse *lǜshi* 律詩 are the most studied elements in Chinese prosody. We would like to term the rules that govern the arrangement of tones in Chinese poetry *shenglü* 聲律. Notice we avoid the term *ping-ze* 平仄, because dividing the four ancient tones into the two *ping* and *ze* categories is only one way of manipulating tones in verse. Theoretically and practically in versification there are other tonal manipulations and *ping-ze* rules are only one set of rules among all *shenglü* 聲律 rules.

In most Chinese scholars' work on Chinese poetic prosody, the number of characters or syllables is a separate topic, distinct from rhyming, tonal assignment, and parallelism. This is a practice inherited from the traditional way of categorization using

yan 言 (number of characters). It is not hard to observe the fact that in a particular verse type the lines mostly consist of a certain number of syllables. For example, *Shijing* 詩經 is characterized by its four-syllable lines, and the regulated verses have five-syllable and seven-syllable lines. Should there be a set of rules that regulate the number of syllables in a line among the rules of prosody? To answer this question, we have to ask: is the number of syllables in a line essential in creating prosody? If the answer is yes, then there should be a set of distinct rules. On the other hand, if the answer to this question is no—the position that we are taking here—then there is no need for any separate rules for *yan* 言. We argue that the concept of *yan* 言 is not a primary element of poetic prosody because it can be determined by the rules of rhythm. The examples in (7) illustrate that *yan* 言, or the number of syllables in a line, is an outcome of a particular rhythm. Describing a particular line of verse in terms of its *yan* 言 is neither precise nor sufficient and it would be more accurate to describe it using rhythm.

- (7) a. 郭李紛爭爲非， (Kong Rong 孔融, *Liu yan shi* 《六言詩》)
遷都長安思歸。
[Guo and Li keep struggling with each other and commit wrong deeds.
The capital city was moved to Chang'an and people are longing to return.]
- b. 靈連蜷兮既留， (Qu Yuan 屈原, *The Nine Songs*, “Yun zhong jun”
爛昭昭兮未央。 《九歌·雲中君》)
[The god has halted, swaying, above us, shinning with a persistent
radiance.]²

The two couplets in (7) each consists of six syllables per line, and therefore can both be called *liu yan* 六言. In fact in many works discussing the development of the so called *liu yan* 六言 poems, the lines in *Chuci* 楚辭 such as the couplet in (7b) are often said to be the origin of the *liu yan* 六言 style, represented by the lines like (7a). However, upon closer examination of the rhythm of the two couplets, it is not hard to detect a difference. In (7a), there are three groups of syllables in each line, with two syllables in each group, while in (7b) the first three syllables obviously form one group. This analysis is supported by the fact that in the second and third position of the second line of (7b), there is reduplication word *zhaozhao* 昭昭. Rhythmically it is unlikely that the reduplication would straddle two groups. Fabb & Halle (2008) point out that the meter primarily controls the number of rhythmic groups and therefore indirectly controls the number of syllables. The rhythm of the two couplets in (7) can be roughly (and non-technically) represented in (8):

² The translation is from David Hawkes' 1959. *Ch'u Tz'u: The Songs of the South*, Oxford University Press.

- (8) a. (XX) (XX) (XX), (XX) (XX) (XX)
 b. ((XXX) X) XX, ((XXX) X) XX)³

The rhythmic grouping in (8) is sufficient to provide the information of total syllable count per line. The meter in (8a) contains three complete binary groups. A simple calculation will thus show the total line length to be six syllables. Therefore, once we have described the setting of the primary prosodic element, rhythm, we can derive the number of syllables. Meanwhile a description of the rhythm carries much more information about the line than the syllable counting method.

On the other hand, the *yan* 言 counting method could hinder analysis of poetic prosody because it might blind us from detecting similarities between poetic forms. Lines of different length in terms of syllable count might have similar or exactly the same meter. This is shown by the example in (9).

- (9) = (2) Fabb & Halle (2008:8)
- | | |
|-----------------------------------|-----|
| Pléasure néver is at home | |
| (* * (* * (* * (* | 0 ⇒ |
| (* * (* * (* | 1 ⇒ |
| (* * (* | 2 ⇒ |
| * | 3 |
| At a touch swéet pléasure méltech | |
| (* * (* * (* * (* | 0 ⇒ |
| (* * (* * (* | 1 ⇒ |
| (* * (* | 2 ⇒ |
| * | 3 |

These two lines have exactly the same meter because they are from the same poem. The difference in the number of syllables is explained by the setting in the meter that allows incomplete groups at the right edge of the lines. We suggest that in Chinese poetry, poems with lines containing different numbers of syllables would be better understood if we were to start seeking similarities in their rhythmic structures.

So far, we have recognized three linguistic elements that build prosody: rhythm (meter) *jie* 節, rhyme *yun* 韻, and tone *shengdiao* 聲調. It is not true that these three elements are all present in any poem's prosody. It is also incorrect to exclude other linguistic elements playing a role in the prosody. Moreover, these three elements can interact in building a poem's prosody. For example, tonal alignment in the Chinese

³ The rhythmic parsing is adapted from the suggestion of an anonymous reviewer.

regulated verse interacts with both rhythm and rhyme. Tonal requirements are put on certain rhythmic positions, e.g. the second, fourth or the sixth syllables in a line. In the regulated verse *ping* tone words also rhyme only with *ping* tone words.

Poetic prosody manifests itself as patterns existing in individual poems. The patterns are built by the elements including rhythm, rhyme, and tone according to certain rules. Although we call them rules, they are not rules in the sense of pure human stipulation or manipulation. Because the elements are linguistic, the rules that built the patterns have to follow the linguistic principles of the elements. The reason why the esthetic beauty of a poem's prosody is appreciated by not only the poet but by all those readers who are speakers of the language, is that the prosody has been constructed with unconscious linguistic knowledge shared by the poet and the readers. On the other hand, poets do have room to creatively play with the elements, as long as the manipulation is within the range that linguistics allows. That is why a poet can choose to rhyme every line, every other line or every three lines but the poet cannot violate the linguistic principles that decide which words rhyme with each other.

3. Poetic prosody and performance factors

Poems are written and performed in various forms. Many poems in *Shijing* 詩經 are said to have been lyrics from folk songs before being collected, edited, possibly transformed and made into an anthology. In the Spring and Autumn Period and the Warring States Period, the poems were also performed at court with musical accompaniment (see Gu 1926, Qu 1962, and Wang 1974, among others.). Moreover, in the "Great Preface" to *Shijing*, the relationship between poetry and music is explained as "言之不足, 故嗟歎之, 嗟歎之不足, 故歌詠之." (If words are not sufficient, then sigh them forth; if sighing is not sufficient, then sing them.) Thus, poetry and music have been inseparable in history. Besides singing, poems in Chinese history have other forms of performance, including chanting and reciting (Zhu 1935). However, these forms of performance each have their own prosody and, therefore, individual rules of prosody. The rules governing chanting prosody may be called *songlü* 誦律, and the rules for reciting *dulü* 讀律. It goes without saying that music has its own rhythm, melody and other building elements that function according to their own rules. One common error in the study of Chinese poetic prosody is to confuse the prosody of the poem with the prosody of its performance. In this section, we will try to prove that these different prosodies are distinct although often interacting systems.

3.1 Rhythm in poetic prosody itself and rhythm in performance

It is often the case, in the discussion of poetic rhythm, that the question of the music that once accompanied it or is assumed to have accompanied the poems is brought up. For example, in the study of the four-syllable *Shijing* 詩經 prosody, many scholars claimed that the rhythm should be slow and elegant because the music was so (Ge 1956, Yuan 1999, Guo & Guo 2003, Beijing Normal University Department of Chinese 2008, etc.). And some people further hypothesize that the music must have had the (XX) (XX) rhythm as well (Guan 2007, He 2009, Zhou 2009, etc.). We believe, however, that there is no such compulsive correlation between the two rhythms. Our viewpoint is supported by evidence from modern songs. The song in (10) below is only one example that contains many mismatches of the rhythmic grouping between the lyrics themselves and the music.

- | | |
|---|--|
| (10) 噢嘿媽媽 | [Oh, hey mom |
| 那天你再次爲我悄悄流下淚 | You shed tears for me quietly that day |
| 你可知道它已(化 作)(傷痛) | You should know they already turned |
| | into wounds |
| 滴滴落在我心扉 | Every drop fell in my heart |
| 永遠都不會忘記 | I will never forget |
| 你看我時那難捨的眼神 | Your eyes that couldn't bare my leaving |
| 我不會 | I will not |
| | |
| 我愛他 | I love him |
| 他是我心中的(那 只)(蝶) | He is the butterfly in my heart |
| 飛呀 | Flies |
| 飛到我心裡面(化 成)(繭) | Flies into my heart and becomes a cocoon |
| 不知道 | I don't know |
| 還要多久才能(叫 醒)(我) | How long I will wait until it awakens me up] |
| (Wang Rong 王蓉, <i>Father and Mother</i> 《爸爸媽媽》) | |

The parentheses in the text above mark the rhythmic grouping of the lyrics themselves. This grouping exists on its own and can be detected by reading the words only. The double straight lines in the text indicate a major break in the rhythm of the music that accompanies the lyrics in the song. The rhythmic misalignment is proven by the misalignment of the parenthesis and the double straight lines.

In reciting and chanting performances of poems, rhythms of the performance also have their own distinctive characteristics that are not necessarily a part of the rhythm of

the poems. For example, in Chinese poem chanting tradition, the *san yan* 三言⁴ poems have often been recited and chanted at a faster tempo with a ballad or folksong-type overtone. This tradition started very early as *san yan* folksongs date back to the Han 漢 and Six Dynasties.

- (11) a. 舉秀才，不知書。舉孝廉，父別居。
[Someone is promoted as a *Xiucai* (excellence in knowledge and talent), but he does not know the Book of Documents. Someone is promoted as *Xiaolian* (outstanding filial piety and integrity), but his father lives separately from him.]
(Ge Hong 葛洪, *Baopuzi* 抱朴子)
- b. 我府君，道教舉。恩如春，威如虎。
[My lord promotes the principle and teaching. His kindness is like the spring and his might is like a tiger.]
剛不吐，弱不茹。愛如母，訓如父。
[He does not fear the powers, nor does he take advantage of the weak. He loves (the people) like a mother and disciplines (them) like a father.]
(*Yuefu shiji* 樂府詩集, “Jingzhao yao” 京兆謠)

Even in modern chanting practices, such poems are often chanted in this rhythm.⁵ Despite the strong tradition of chanting and reciting, researchers should be discouraged to conclude that the fast tempo rhythm is inherent in the rhythm of the poem itself. Rather, the chanting and reciting practice is only one possible cultural choice of the interpretation of the poem. The poems can logically be chanted with a slow and elegant rhythm, which in the Chinese tradition is used to perform the *Shijing* 詩經 poems or five-syllable regulated verses.

3.2 Tone in poetic prosody itself and melody in performance

In the scholarly works on the origin of the Chinese regulated verse (Jiang 1982, Zhu 1984, Chu 1990, Zhang 2003 and 2006, and Goh 2004), a music-origin argument has been made for the introduction of tonal manipulation in poetry: in the post Han era, music and poetry started to become separate art forms, then tonal manipulation was introduced into the composition as a compensation for the loss of acoustic beauty that

⁴ Strictly speaking, a lot of such poems are *san yan* 三言 and *qi yan* 七言 mixtures.

⁵ We have recordings of different peoples' chanting of such poems, e.g. Li Bai's famous *Jiang jin jiu* 將進酒.

used to accompany the poems. Whether that argument is strong or not goes beyond the scope of this paper. But it is clear that tones and musical melody have been associated in the tradition of poetic composition and literary criticism. Y. R. Chao (1956) introduced the practices of tonal chanting, singsong and the tonal composition of music, in each of which ancient tonal pitch values serve as a basis for the pitch melody of the performance. For example, in the tonal composition tradition the rule is for the even tones to go with long notes or to extend over descending notes of small range, while oblique tones go with the shorter and higher notes as well as quickly changing or widely skipping notes. However this tradition has led some researchers to believe tonal manipulation in Chinese poetry and music melody are just two sides of the same coin.

The position of this paper is that the prosody of poetry and the prosody of its performance are distinct systems. Tonal manipulation and music melody are distinct as well. First, tonal manipulation in Chinese regulated verse is based on the bifurcation of four tones. In other words, the four tones are divided into two categories: even tones and oblique tones. This bifurcation might have a phonetic and or phonological basis, however the reason why the four ancient tones were divided into two groups and not three groups deserves some thought. If we were to take a survey of the world's metrical verses, we would easily find that despite the very different languages, the specific prosodic rules, and the literary traditions of these verses, there is always a bifurcation of syllables. Fabb & Halle (2008) also observe that most metrical systems not only group the syllables but also impose conditions on the groups. "For the purpose of stating the conditions, the syllables that compose the lines are always partitioned into two classes." In the Chinese regulated verse, the partition is done according to the tonal classes, namely even versus the other three tone groups. In languages that have word stress, then stressed versus unstressed may be the condition, knowing that the bifurcation is not unique to Chinese but a rather universal principle in poetic prosody. On the other hand, the pitch variation in music melody does not follow the same principle of bifurcation or binary contrast. The music notes are not divided into two groups, nor does the organization of the notes in forming a melody make crucial reference to the binary grouping of the notes.

Another piece of evidence against the tonal composition is that when lyrics written in Chinese are sung in a certain musical melody, the lexical tones are lost and the musical melody is the only pitch that is perceived. The acoustic realization of tones is pitch; the acoustic realization of music melody is also pitch. When a singer is singing a song with Chinese lyrics, the expression of the two pitches must be adjusted because the singer can only sing one frequency. The result of the adjustment logically can have many possibilities. However, in reality the lexical tones will not continue to be perceivably distinct. In order to test this, we created two sets of lyrics for the same melody in Chinese. These two sets of lyrics have the exact same segmental material, meaning the same

consonants and vowels. The syllables in the two sets of lyrics differ only in tone. We randomly asked Chinese native speakers to sing the two lyrics, then randomly played the recordings to other native Chinese speakers. Example (12) shows these two lyrics sung to the melody of the French song *Frère Jacques*:

(12) a. lyrics 1:

1=F 4/4

1 2 3 1 | 1 2 3 1 | 3 4 5 - | 3 4 5 - |

打 倒 列 強 打 倒 列 強 救 中 國 救 中 國

[Down with the invading powers. Down with the invading powers. Save China. Save China.]

5 6 5 4 3 1 | 5 6 5 4 3 1 | 1 5 1 - | 1 5 1 - |

殺 盡 一 批 軍 閥 殺 盡 一 批 軍 閥 震 世 界 震 世 界

[Kill all the warlords. Kill all the warlords. Shock the world. Shock the world.]

b. lyrics 2:

1=F 4/4

1 2 3 1 | 1 2 3 1 | 3 4 5 - | 3 4 5 - |

大 刀 獵 槍 大 刀 獵 槍 酒 中 過 酒 中 過

[Big knives and hunting guns. Big knives and hunting guns. Spend time with wine. Spend time with wine.]

5 6 5 4 3 1 | 5 6 5 4 3 1 | 1 5 1 - | 1 5 1 - |

紗 巾 一 披 俊 髮 紗 巾 一 披 俊 髮 真 士 傑 真 士 傑

[Cover the beautiful hair with a veil. Cover the beautiful hair with a veil. Real heroes. Real heroes.]

The Chinese speakers who heard the recordings of the songs could not distinguish lyrics 1 from lyrics 2. This result shows that the distinctive lexical tones in Chinese lose their distinctiveness when musical melody supersedes them.

Tonal composition might be a real tradition in Chinese poetry and music, but it is very wrong to assume tonal rules and music melody rules to be the same in any theory about poetic prosody. Besides singing, Y. R. Chao also introduced other forms of performance, including chanting and reciting. These forms of performance were taught by masters to students orally from generation to generation. Meanwhile, in these performances, improvisation was allowed and was very frequent. As a result, today many people chant and recite the same poems very differently. We should take the rules of these forms of performance as separate prosodic systems from the prosody of the poems themselves. Tones are linguistic elements that are manipulated within the range that all human poems fall in, while pitch change in chanting is not subject to this limitation.

4. Rules for tonal assignment in Chinese poetic prosody

One feature that has been closely associated with Chinese metrical verse is tonal prosody, or a set of rules that assigns tones to different prosodic positions in a poem. The rules for the regulated verse, or *lǜshī* 律詩, are the most well studied among all tonal prosodic rules. Because of the classic status of the regulated verse, its *ping-ze* tonal rules and the related terms used in the tonal prosody (e.g. *ba bing* 八病 “the eight defects”) have been assumed to have existed not long after the discovery of the four tones and the first movement for metrical verse in Chinese literary history (Mair & Mei 1991, Zhang & Luo 1996, Yuan 1999, Fu 1999, Lin 1999, etc.). In this section, we will argue against this position.

As argued in §3.2, metrical systems bifurcate syllables and this bifurcation is used in the statement of metrical conditions. For example, in a stress language, stressed and unstressed syllables are two natural classes, and metrical rules in a particular prosody will first form metrical grouping in a line, and then require stressed and unstressed syllables to appear in certain metrical positions respectively. The bifurcation is largely universal while the particular nature of the binary distinction is language specific. The Chinese *ping* 平 versus *ze* 仄 binary tonal grouping is one realization of the bifurcation requirement in metrical systems. However, historically the development of this specific bifurcation required the time as well as ingenuity and creativity of poets. The ancient four tones did not readily form two groups as did the stressed and unstressed syllables. However, the bifurcation of *ping* versus *ze* tones is so influential that many scholars oversimplified the process through which the bifurcation or the tonal prosody in the standard regulated verse was finally developed and formulated. It was mistakenly assumed that Shen Yue, who advocated the four tones and the tonal manipulation in poetry, was also the founder of the *ping-ze* rules. For example, Wang Li (1979) attributed the idea of *ping-ze* to Shen Yue. However, studies of Shen’s own poetic composition show that he did not follow the prosodic regulations of the classical regulated verse, including the simple *ping-ze* rules. It was sometimes argued that Shen was not capable of complying with the prosodic theories that he advocated very earnestly. Shen’s own writing unmistakably promotes tonal manipulation in metrical poems:

- (13) “欲使宮羽相變，低昂互節，若前有浮聲，則後須切響，一簡之內，音韻盡殊；兩句之中，輕重悉異。”
[One would want to alternate the *gong* and *shang* notes, making the low and high pitches tempering with each other. If in the previous lines, there is a floating sound, then in the later lines there must be a cut-off echoing sound. In the same stanza, sounds and rhymes should be completely different. In a

couplet, the choices of heavy and light should be distinct.]
(Shen Yue 沈約, *Song shu* 宋書, “Xie Lingyun zhuan lun” 謝靈運傳論)

However, the interpretation of the sentences in (13) is often extended to including the *ping-ze* rules. A more conservative and literary reading of the text only supports the manipulation of sounds, including tones, and a pursuit of acoustic “contrast”, which is represented by the words like *shu* 殊 and *yi* 異, both meaning “difference”. As for the content of such contrasts, the text itself does not provide concrete evidence. It was not found in Shen’s other writing that he clearly argued for the contrast between the *ping* tone on the one side and the other three tones on the other side. In sum, Shen intuitively felt the principle of contrast in poetic prosody, but did not design the particular *ping-ze* bifurcation that we see in the regulated verses. Therefore, the belief that Shen’s own poetic composition failed to embody his poetic prosody theory should never have existed in the first place, if this theory was different from the known *ping-ze* rules. Rather, a careful examination of Shen’s poems gives strong support to his primitive ideal of “contrast” that lacks *ping-ze* bifurcation. Example (14) is a poem that illustrates the principle of contrast and not the principle of *ping-ze*.

- (14) 漢池水如帶 [Alone the Han moat, water [stretches] like a girdle⁶
去平上平去
巫山雲似蓋 Above Shamanka Mountain, clouds are like a canopy
平平平上去
瀨汨背吳潮 Up rolling rapids, you will leave behind the tides of Wu
入入去平平
潺湲橫楚瀨 And, swirling smoothly, cross the shoals of Ch’u
平平平上去
一望沮漳水 Once having gazed upon the waters of Chu and Chang
入去平平上
寧思江海會 You will then miss this gathering besides the Chiang and sea?
平平平上去
以我寸草心 With all my heart, that measures but an inch
上上去上平
從君千裡外 I will follow you beyond a thousand *li*
平平平上去
(Shen Yue 沈約, *Farewell Banquet for Hsieh Tiao, on the Eve of Departure*
餞謝文學離夜)

⁶ Translation by Richard Mather. 2003. *The Age of Eternal Brilliance: Three Lyric Poets of the Yung-ming Era*, Brill Academic Publisher.

In (14) we have marked the tonal category of each syllable. If we apply the standard tonal rules in the classic regulated verse, then the poem apparently has defects. For example in the first line, the second and the fourth syllables are both *ping* tone and thus violate the *ping-ze* alternation rule. If we forgo the *ping-ze* rules, and examine the tones, we find that the poem has a very well designed, intricate tonal pattern. First of all, the second, fourth, sixth and eighth lines have exactly the same tonal arrangement, namely *ping-ping-ping-shang-qu*. If we conduct a statistical analysis, we can easily obtain a significant test result, which underpins the conclusion that such an arrangement is extremely unlikely to be the result of chance. In other words, Shen must have intentionally manipulated the tones in these four lines. Secondly, in each line, the second and fifth syllables have different tones. This is not a binary contrast but it is undeniable this is a contrast, or *yi* 異.

The poem in (14) is one example from Shen Yue's compositions, and a thorough study of all his poems would yield more patterns and generalizations based on the patterns. Before doing that, it is already clear that Shen's tonal prosody was different from the regulated verse's, which only emerged after Shen's time and after the experimentation of various poets. Returning to the universal principle of bifurcation in metrical poems, we can at least give one reason why Shen's tonal prosody did not become the standard practice and was to be changed and replaced by the standard rules for regulated verse that rely on the contrast between the *ping* and *ze* tones: if binary contrast, or the bifurcation approach to syllables, is universal or at least preferred, then after this period of experimenting with various possible manipulations, we find that the trend moved towards the approach of bifurcation, which has a sufficient phonetic and phonological basis.⁷

5. The nature of rhyming

In §1 of this paper, when we clarified the linguistic elements that are utilized in poetic prosody, we touched upon rhyming. In this section, we will delve further into this

⁷ Although we take the position that the binary contrast is universal in metrical poetry, we also admit the existence of various types of non-metrical poetry, some of which are based on line length, some on syntactic parallelism, and some on other formal features. One anonymous reviewer brought up the question about possible metrical contrast before the development of the tonal prosody system, for example in the poems in *Shijing* 詩經 or the *Book of Change* 易經. Although the question deserves more research, in this paper we would like to refrain from making strong assertions about the metrical nature of these earlier poems. In other words, it could be the case that the earlier poems are non-metrical and thus do not utilize any metrical binary contrast. It is also possible that they are metrical and that we just have not found strong evidence supporting this yet.

issue and explore the nature of rhyming. By nature we refer to the linguistic nature of the phenomenon called rhyming. To be specific, we will discuss the linguistic conditions surrounding what can be defined as rhyme.

Most Chinese scholars argue that if two syllables rhyme, then they should have a similar main vowel and coda. This position is held by Wang Li (1980), who argues that we should not equate the *yunzhe* 韻轍 “rhyming group” with *yunmu* 韻母 “rhyme”. To Wang, one rhyming group may contain several rhymes, and words that have different rhymes yet belong to the same rhyming group can rhyme with each other. This view-point holds that the similarity between rhymes is a sufficient condition for the rhyming phenomenon. Wang’s position has been echoed by many scholars. Zhengzhang (2003) claims that, because words could rhyme without having identical vowels, we can reconstruct different vowels for the same OC rhyming group. Similarly, Pan (2000) assumes that phonetic similarity could be sufficient for rhyming, and therefore in his reconstruction **-ak* could rhyme with **-a*. These assumptions about rhyming come from certain observations in modern languages. For instance, Zhengzhang (2003) points out that in modern Chinese Mandarin *yin* 因 [in] belongs to the *renchen zhe* 人辰轍 rhyming group and therefore rhymes with words like *ren* 人 [zən] and *chen* 辰 [tʂən]. This observation is certainly valid. Based on such observations, Zhengzhang argues that rhyming does not require identity in vowel quality. He also insists that in order to follow the “one vowel one rhyme group” principle, we would have to stipulate *yin* 因 as [iən] in Mandarin, which is definitely not how the word is pronounced today.

However, at the theoretical level, the similarity condition for rhyming is not very desirable, because the term “similarity” lacks scientific definition and therefore is often subject to a very subjective interpretation. Besides the theoretical consideration, the “similarity condition” also faces challenges from rhyming facts. Below, we investigate some rhyming facts. Example (15) shows some rhymed texts:

(15) a.	<u>Pinyin form</u>	<u>SR form</u> ⁸	<u>UR form</u>
紅軍不怕遠征難	nan	[nan]	/nan/
萬水千山只等閒	xian	[ɕien]	/ɕian/
五嶺逶迤騰細浪			
烏蒙磅礴走泥丸	wan	[uan]	/uan/
金沙水拍雲崖暖	nuan	[nuan]	/nuan/
大渡橋橫鐵索寒	han	[χan]	/χan/
更喜岷山千里雪			
三軍過後盡開顏	yan	[iɛn]	/ian/

⁸ See the following page for SR form and UR form.

[The Red Army does not fear the difficulty of a long march
 Taking the thousands of waters and mountains as nothing
 The Five Ridges wind their way, like tiny surging waves
 The Wumeng Mountain is majestic, (but to us) is a rolling mud ball
 The Jinsha River whips the misty cliff, which is warm
 A bridge lies across the Dadu River, and its iron chain is cold
 The boundless snow peaks of the Min Mountain make us even happier
 Every soldier in the army smiled with joy after we passed the Min
 Mountain]

(*Long March* 七律 • 長征, 毛澤東 Mao Zedong)

b.	<u>Pinyin form</u>	<u>SR form</u>	<u>UR form</u>
我的兒汾河灣前去打雁	yan	[iɛn]	/ian/
天到了這般時，不見回還	huan	[χuan]	/χuan/
將身兒坐置在窑門以外			
等我兒他到來同把飯餐	can	[ts ^h an]	/ts ^h an/
[My child went to hunt the geese in front of Fenhewan It is so late now, but I haven't seen him back I am sitting outside of the house gate Waiting for my child to come home and eat dinner together] (Traditional Peking Opera 傳統京劇, <i>Fenhewan</i> 汾河灣)			

The first example in (15) is a traditional regulated style poem composed by Mao Zedong using the standard traditional rhyming schema. The second sample is an extraction from a traditional Peking opera, and the rhyming schema follows the traditional *yanqian* group or 言前轍. *SR* stands for surface form, and *UR* stands for underlying form. In (15) we listed the surface forms as well as the underlying forms of the rhyming syllables. The surface forms are “similar” while the underlying forms are identical. Past arguments for the similarity condition for rhyming, as in Zhengzhang (2003), are made on the surface values of the rhyming syllables. These syllables are similar but the similarity is very hard to define. In other words, it needs to be explained why [a] and [ɛ] and [i] and [ə] are similar enough to rhyme. If similarity is only defined in an ad hoc way, then scholars supporting the similarity condition need to provide an explanation for the consistent rhyming intuition of native speakers. On the other hand, the underlying forms of the rhyming words given in (15) seem to point to a more promising solution. The main vowels and codas of these rhyming words are identical at the underlying level. Can we therefore argue that the nature of rhyming is “identicalness at the underlying level”?

Before we rush to accept the hypothesis that the nature of rhyming is “identicalness at the underlying level”, more rhyming cases indicate the truth isn’t as simple as we expected. Let’s look at the English poem given in (16) below:

- (16) (Rhyming pattern [Rhyme Royal]: ababbcc) **SR** **UR**
 Out of the air a voice without a face
 Proved by statistics that some cause was *just* [dʒʌst] /dʒʌst/
 In tones as dry and level as the place
 No one was cheered and nothing was *discussed* [diskʌst] /diskʌs-**d**/
 Column by column in cloud of *dust*
 They marched away enduring a belief
 Whose logic brought them, somewhere else, to grief
 (*The Shield of Achilles*, Wystan Hugh Auden, 1907-1973)

In this English poem, *just* and *discussed* rhyme with each other. Their surface forms have identical vowels and coda, and their underlying forms don’t. Therefore this poem is a counter example to the hypothesis that the nature of rhyming is “identicalness at the underlying level”. Should this discrepancy be attributed to the language specific difference between Chinese and English? Our answer is no, because examples of the same nature such as (16) can also be found in Chinese. These are rhyming examples in the *erhua* 兒化 categories. In (17), we show two *erhua* rhyming groups, namely the *xiao yanqianr zhe* 小言前兒轍, and the *xiao renchenr zhe* 小人辰兒轍.

- (17) a. *xiao yanqianr* 小言前兒 includes
 /a/ /ua/ /ia/ /ai/ /uai/ /an/ /uan/ /yan/ /ian/ + /er/
 b. *xiao renchenr* 小人辰兒 includes
 /o/ /e/ /i/ /u/ /y/ /ie/ /ye/ /ei/ /uei/ /en/ /in/ /uen/ /yen/ + /er/

Below are several examples to support the two rhyming groups given in (17).

- (18) a. 小言前兒 **UR** **SR**
 白蓋兒 /-ai/ + /er/ [-ar]
 白干兒 /-an/ + /er/ [-ar]
 大褂兒 /-ua/ + /er/ [-uar]
 撒歡兒 /-uan/ + /er/ [-uar]
 b. 小人辰兒 **UR** **SR**
 雞兒 /-i/ + /er/ [-ie:r]
 今兒 /-in/ + /er/ [-ie:r]

hoping to clear the road for future research. Our goal is to strengthen the research of poetic prosody by isolating and analyzing the various linguistic elements, which are the building materials of poetic prosody. Eventually, studies on various Chinese metrical poems will test the correctness of our claims here, and these studies and their results are exactly what we wish this paper to stimulate.

References

- Allen, W. Sidney. 1973. *Accent and Rhythm*. Cambridge: Cambridge University Press.
- Beijing Normal University, Department of Chinese. (eds.) 2008. *Zhongguo Gudai Wenxueshi [A History of Ancient Chinese Literature]*. Beijing: Beijing Normal University Press.
- Chao, Yuen Ren 1956[2006]. Tone, intonation, singsong, chanting, recitative, tonal composition, and atonal composition in Chinese. *Linguistic Essays by Yuenren Chao*, 596-611. Beijing: The Commercial Press.
- Chen, Matthew Y. 1979. Metrical structure: evidence from Chinese poetry. *Linguistic Inquiry* 10.3:371-420.
- Chen, Matthew Y. 1984. Unfolding latent principles of literary taste: poetry as a window onto language. *Tsing Hua Journal of Chinese Studies*, New Series 16.1:203-240.
- Chu, Binjie. 1990. *Zhongguo Gudai Wenti Gailun [Genres of Chinese Literature Throughout History]*. Beijing: Peking University Press.
- Fabb, Nigel, and Morris Halle. 2008. *Meter in Poetry: A New Theory*. Cambridge & New York: Cambridge University Press.
- Fu, Xuancong. 1999. *Zhongguo Shixue Dacidian [Dictionary of Chinese Poetry]*. Hangzhou: Zhejiang Education Press.
- Ge, Xianning. 1956. *Zhongguo Shishi [History of Chinese Poetry]*. Taipei: Chinese Cultural Publishing Committee.
- Gibbon, Dafydd, and Helmut Richter. 1984. *Intonation, Accent and Rhythm*. Berlin & New York: Walter de Gruyter.
- Goh, Meow Hui 2004. *Wang Rong's Poetics in the Light of the Invention of Tonal Prosody*. Madison: University of Wisconsin-Madison dissertation.
- Gu, Jiegang. 1926. *Gu Shi Bian [Discriminations on Ancient History]*, Vol. 3. Beijing: Pu She.
- Guan, Enhao. 2007. Lun bianzhong he yuewu dui Shijing siyan shiti de yingxiang [On the influence of bianzhong and music to the four-syllable style of *Shijing*]. *Movie Review* 2007.7:96-97.
- Guo, Yingde, and Changbao Guo. 2003. *Zhongguo Gudai Wenxueshi [A History of Ancient Chinese Literature]*. Chengdu: Sichuan People's Publishing House.

- Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge & New York: Cambridge University Press.
- Halle, Morris, and Samuel Jay Keyser. 1966. Chaucer and the study of prosody. *College English* 28.3:187-219.
- Halle, Morris, and Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge: MIT Press.
- Hayes, Bruce. 1983. A grid-based theory of English meter. *Linguistic Inquiry* 14.3:357-393.
- Hayes, Bruce. 1995. *Metrical Stress Theory*. Chicago: University of Chicago Press.
- He, Dan. 2009. Shijing siyanti xingshuai tanlun [The rise and decline of the four-character verse of *Shijing*]. *Journal of Zhejiang University (Humanities and Social Sciences)* 2009.3:131-139.
- Hogg, Richard M., and C. B. McCully. 1987. *Metrical Phonology: A Coursebook*. Cambridge & New York: Cambridge University Press.
- Humphrey, Asa. 1847. *The English Prosody: With Rules Deduced from the Genius of our Language, and the Examples of the Poets*. Boston: Crocker & Brewster.
- Jiang, Shuge. 1982. *Shixue Guanglun [A Broad Discussion of Poetry]*. Beijing: China Social Sciences Press.
- Kiparsky, Paul. 1977. The rhythmic structure of English verse. *Linguistic Inquiry* 8.2:189-247.
- Kiparsky, Paul. 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2:85-138.
- Kiparsky, Paul, and Gilbert Youmans. (eds.) 1989. *Rhythm and Meter*. San Diego: Academic Press.
- Lin, Jiali. 1999. *Shenyue Yanjiu [Studies on Shen Yue]*. Hangzhou: Hangzhou University Press.
- Mair, Victor H., and Tsu-Lin Mei. 1991. The Sanskrit origins of recent style prosody. *Harvard Journal of Asiatic Studies* 51.2:375-470.
- Merriam-Webster. 1983. *Webster's Ninth New Collegiate Dictionary*. Springfield: Merriam-Webster.
- Pan, Wuyun. 2000. *Hanyu Lishi Yinyunxue [Chinese Historical Phonology]*. Shanghai: Shanghai Education Press.
- Patel, Aniruddh D. 2008. *Music, Language, and the Brain*. Oxford & New York: Oxford University Press.
- Qu, Wanli. 1962. Lun *Guofeng* fei minjian geyao de benlai mianmu [*Guofeng* poems are different from the original folk songs]. *Bulletin of the Institute of History and Philology Academia Sinica* 34:477-491.
- Selden, Raman. 1995. *The Cambridge History of Literary Criticism*, Vol. 8. Cambridge & New York: Cambridge University Press.

- Thaut, Michael H. 2008. *Rhythm, Music, and the Brain*. New York: Routledge.
- The American Heritage Dictionary of the English Language*. 2001. Boston: Houghton Mifflin.
- Wang, Ching-hsien. 1974. *The Bell and the Drum: Shih Ching as Formulaic Poetry in an Oral Tradition*. Berkeley: University of California Press.
- Wang, Li. 1979. *Hanyu Shilüxue [Chinese Poetics]*. Shanghai: Shanghai Education Press.
- Wang, Li. 1980. *Hanyu Yinyun [Chinese Phonology]*. Beijing: Zhonghua Book Company.
- Yip, Moira. 2002. *Tone*. Cambridge & New York: Cambridge University Press.
- Yuan, Xingpei. (ed.) 1999. *Zhongguo Wenxueshi [A History of Chinese Literature]*. Beijing: Higher Education Press.
- Zhang, Hongming. 1987a. Hanyu jintishi shenglü moshi de wuzhi jichu [The property of tonal prosody in recent style poetry]. *Zhongguo Shehui Kexue [Social Sciences in China]* 1987.4:185-196.
- Zhang, Hongming. 1987b. Yuyan de duibi yu shilü de bijiao [Comparison within Languages and Poetics]. *Fudan Xuebao [Fudan Journal] (Social Sciences Edition)* 1987.4:51-56.
- Zhang, Hongming. 2003. Tonal prosody of poetics in the Six Dynasties. Paper presented at the 2nd International Symposium of Historical Chinese and Classical Chinese. New York: Columbia University.
- Zhang, Hongming. 2006. The consistency of Shen Yue's poetic metrical theory and poem composition: a linguistic perspective. Paper presented at the Annual Meeting of the Association for Asian Studies. San Francisco, USA.
- Zhang, Hongming. 2008. Yongming style poetics: a linguistics perspective. Paper presented at the 16th International Conference on Chinese Linguistics. Beijing: Peking University.
- Zhang, Hongming. 2009. On linguistic properties in Chinese poetic prosody. Paper presented at the 17th International Conference on Chinese Linguistics. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale.
- Zhang, Hongming, and Wenjing Li. 2011. Yu Xin wuyanshi shenglü kaocha [Study of metrical patterns in Yü Hsin's pentasyllabic verse]. *Wenxue yu Wenhua [Literature and Culture Studies]* 4:66-76. Tianjin: Nankai University Press.
- Zhang, Peiheng, and Yuming Luo. 1996. *Zhongguo Wenxueshi [History of Chinese Literature]*. Shanghai: Fudan University Press.
- Zhengzhang, Shangfang. 2003. *Shanggu Yinxi [Phonological System of Archaic Chinese]*. Shanghai: Shanghai Education Press.
- Zhou, Weifeng. 2009. Biqi yinlu yu *Shijing* siyanshi tishi de shengcheng [The development of four-syllable poetry in *Shijing*]. *Zhongguo Yunwen Xuekan [Journal of Chinese Verse Studies]* 2009.1:1-4.

- Zhu, Guangqian. 1984. *Shi Lun* [*On Poetry*]. Beijing: SDX Joint Publishing Company.
- Zhu, Qianzhi. 1935. *Zhongguo Yinyue Wenxue Shi* [*History of Musical Literature*]. Shanghai: The Commercial Press.

LANGUAGE AND LINGUISTICS

LANGUAGE AND LINGUISTICS is a department of the INSTITUTE OF LINGUISTICS, ACADEMIA SINICA. It furthers the Institute's objective of excellence in research, scholarship, and education by publishing pioneering articles and monographs in Linguistics and other related fields. All publications have been reviewed rigorously according to academic standards.

Copyright © 2013 by INSTITUTE OF LINGUISTICS, ACADEMIA SINICA
128, Sec. 2, Academia Road, Nankang, Taipei 11529, Taiwan, R.O.C.
<http://www.ling.sinica.edu.tw>
All rights reserved.

LANGUAGE AND LINGUISTICS MONOGRAPH SERIES 50

Edited by Guangshun Cao, Hilary Chappell, Redouane Djamouri
and Thekla Wiebusch

Breaking Down the Barriers: Interdisciplinary Studies in Chinese Linguistics and Beyond (2 Volumes)

《語言暨語言學》

《語言暨語言學》隸屬於中央研究院語言學研究所，以出版語言學及相關領域之最新研究成果為宗旨。所有出版品均經過嚴格學術審查。

《語言暨語言學》專刊系列之五十

綜古述今 鉤深取極 (共二冊)

曹廣順 曹茜蕾 羅 端 魏婷蘭 編輯

出版：中央研究院語言學研究所
發行：中央研究院語言學研究所
115 台北市 南港區 研究院路 2 段 128 號
<http://www.ling.sinica.edu.tw>
版權所有 翻印必究

印刷：大光華印務部
初版：中華民國一〇二年九月 [September 2013]
定價：新台幣 2500 元 / US\$130 (全套二冊，不分售)

ISBN: 978-986-03-7678-4 (平裝)
GPN: 1010201631

