

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 1

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)



Professor Alain Peyraube (Lyon, France, 2010)

Preface

It is a great pleasure and honour for us, the four editors, to assemble together in these two volumes a large number of new and original contributions to research in linguistics and related domains, many of which reflect the most recent developments in their area.

The idea of editing a volume in homage to our esteemed colleague, mentor and friend, Professor Alain Peyraube, initially stemmed from one of the four editors, Thekla Wiebusch, when she first mooted it to Redouane Djamouri and Hilary Chappell, as a worthy undertaking. The project only began to gain momentum, however, in 2006, after invitations had been sent out to colleagues all around the world, working in many different fields of linguistics and its kindred disciplines. As the number of contributors swelled to over 50, Cao Guangshun kindly accepted our invitation to collaborate, taking care of the large number of articles written in Chinese for this multilingual work. The two volumes of this *Festschrift* thus opportunely bring together a truly representative range of colleagues and friends who, as the case may be, may have worked together with Alain, may have been inspired by his research and writings, or who may even have felicitously had frequent occasion to discuss and debate ideas with him (if not all three).

The original aim was to present the *Festschrift* to Alain Peyraube on the occasion of his 65th birthday in November 2009. Given the enormity of the editorial task with 54 chapters in the book and their 54 authors to correspond with, a special ceremony was held instead during the 17th International Conference on Chinese Linguistics (IACL-17) held in Paris in July 2009, with the CRLAO as its host.

To Alain's surprise, we presented him with an elegant scroll containing the *Table of Contents* for what had become a project of increasing magnitude. Professor William S-Y. Wang, then of the City University of Hong Kong and our co-editor, Professor Cao Guangshun of the Chinese Academy of Social Sciences, both friends and colleagues of Alain of long-standing, gave speeches in his honour which they carried out with great panache and humour.

These two speeches, each in the form of a tribute, can be found below in the introductory section to the two volumes of this *Festschrift*. In fact, the *Festschrift* begins with a series of key photos which trace Alain's academic career in a quite lively and personal manner. These precede the transcriptions of the two speeches and reflect

the selection which was originally made for the slide presentation, specially created for the IACL-17 ceremony by Thekla Wiebusch, to the great delight of the audience — and we hope of our readers too.

The 54 chapters witness the diversity of scientific interests held by Alain, the outcome of a ceaselessly enquiring mind, characteristics that are by their nature highlighted in the biography and similarly strike the casual reader browsing through his list of publications at the end of the introduction.

Alain has always been up to date with the latest theories and has never ceased re-examining data in this light, whether it be from Archaic, Medieval or Modern Chinese, or even from Sinitic languages and Chinese dialects other than Standard Mandarin.

The chapters have been grouped together according to their broad subject matter in order to achieve a coherence for the diverse domains that have formed and restructured the discipline of linguistics in recent decades and which, in particularly the case of diachronic syntactic change and grammaticalization, the research of Alain Peyraube has been influential. These include cognition, Modern Chinese, Ancient and Classical Chinese, Medieval Chinese, Early Mandarin Chinese, diachrony, classification, language contact, typology and the Chinese dialects and script.

Our sincere hope is that this work will prove to be useful and relevant for very many years to come, and this to a wide readership not only in general and Chinese linguistics but across a broad spectrum of related domains as well.

We would also like to take this opportunity to thank Professor Ho Dah-an of the Institute of Linguistics at the Academia Sinica for his generosity in offering to publish this monumental work. Last, but certainly not least, we wish to express our gratitude to our colleagues who make up the indefatigable editorial team of *Language and Linguistics* for their excellent work, fastidiousness and professional approach during the extremely protracted period of gestation for the *Festschrift*: Professor Elizabeth Zeitoun, Ms Joyce Chun-yu Kuo, Ms Vicky Y.-K. Chen and their colleagues.

The editors

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
<i>Ancient and Classical Chinese</i>	
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
<i>Medieval Chinese</i>	
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
<i>Early Mandarin Chinese</i>	
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: <i>Qièyùn</i> et <i>Kānmiù bǔquē Qièyùn</i> 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 𠩺 and 𠩻 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
--	------



Dressed up in soldier's clothes as a saltimbanque or 'travelling acrobatic performer' at a school fête, 2nd from right



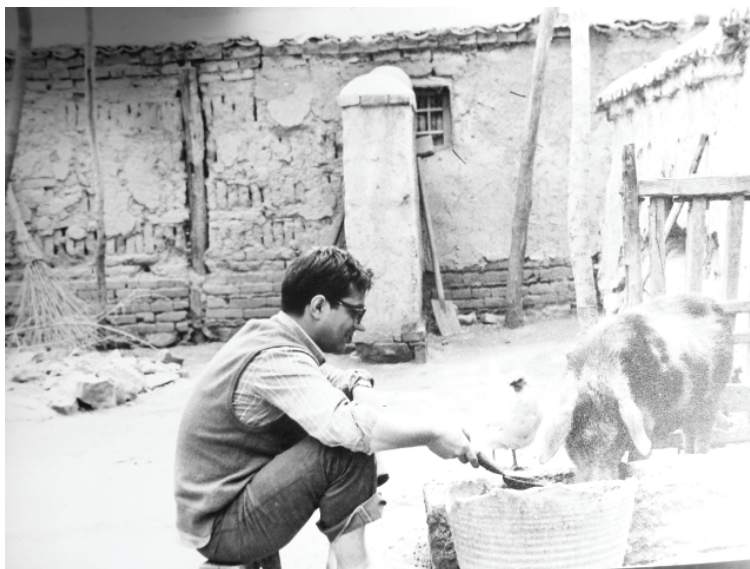
School year, 1950-1951, at Les Capucins Primary School, Bordeaux, 6 years of age, bottom row, 5th from right



School year, 1953-1954, at Les Capucins Primary School, Bordeaux,
9 years of age, 2nd row from bottom, 3rd from left



Working in internal combustion engine factory in 1974 內燃機總廠



Feeding the pigs at the Shang Lian Brigade of the Bei Xiao Ying
People's Commune in 1975 北小營公社上輦大隊



Open Door Schooling at the Four Seasons People's Commune in 1976
四季人民公社開門辦學



Alain Peyraube in his EHESS office at 54 boulevard Raspail, Paris



Public speech near the Trocadéro, Paris in mid-1980s



With the anthropologist, Marc Augé, and Ta-You Wu 吳大猷, the President of the Academia Sinica, 1983-1994



With Professors William S-Y. Wang 王士元 and Tsu-Lin Mei 梅祖麟



At Simatai, the Great Wall of China, during a workshop on Altaic and Chinese language contact, September 2009 司馬台長城景區



Conference on the Language of the Lao Qi Da and the Piao Tong Shi, Hangzhou, June 2010



Joint workshop in Paris of Taiwanese NSC and French ANR ‘DIAMIN’ teams
April 2010



Workshop in Tsing Hua on diachronic change in Early Southern Min
November 2010



Family photo Left to right; Eva (granddaughter), Marie (daughter), Zoë (granddaughter),
Alain and son-in-law Mickael Spitzberg

Speech in Honour of Professor Alain Peyraube

presented by Professor William S-Y. Wang

There are not many people with the surname BÈI 貝. Only two famous people come to mind. One is a Chinese architect, who has recently made a very visible contribution to the culture of France. The other is a French linguist, who continues to deepen our understanding of historical developments in Chinese grammar. The Chinese architect is 貝聿銘 Bèi Yù Míng (I. M. Pei), whose crystal pyramid for the Louvre stands not very far from where we are now. The French linguist is 貝羅貝 Bèi Luó Bèi who we are gathered here to honor today.

Alain's contribution to Chinese linguistics spans several decades, plumbing the depths in many central topics of grammatical evolution. These include locative constructions in 1980, double object constructions in 1986, passive constructions in 1989, and so on. Among these many studies, his monograph of dative constructions in 1988 won the prestigious Stanislas Julien Prize of the French Academy.

Alain's contributions extend beyond his own research, into many areas which stimulate the scholarship of others. He was a core member of a small group who created the International Association of Chinese Linguistics in 1992, which has done so much to shape our field. It is Alain who implemented the Young Scholar's Award, which has become such a defining feature of our organization. He contributes editorial guidance to many important publications of our field, including many years of wise counsel to the *Journal of Chinese Linguistics*, for which this editor is especially grateful.

Like his contributions, recognitions of his achievements are numerous. Among the high honors are a Hu Shih Visiting Professorship to Cornell University in 1995, membership of the European Academy of Arts and Sciences since 2006, corresponding member of Academia Sinica and Honorary Professor of Peking University since 2007.

However, as we all know, a true measure of a rich life includes more than just books and diplomas. Scholarship is also measured by the number of friends, colleagues, and students whose life you have touched, and who in turn come to honor you by joining in your Festschrift, and by gathering here today. Looking around this hall now, it is clear that Alain's is a rich life indeed!

曹廣順教授在貝羅貝先生六十五歲 生日聚會上的賀詞

非常感謝主辦者給我這次機會。我實在覺得無法全面評價貝羅貝先生的學術成就。

那麼，什麼是貝羅貝呢？除了一臉燦爛的笑容之外，我們最後得出的結論就是他“非常法國”。法國是什麼呢？用三項特徵來概括：巧克力、紅酒和美食。我認為這三項法國特徵都集中在貝羅貝先生身上。到什麼程度？這本來是不應向外人道的。我僅在這裡舉兩個小例子。

十幾年以前有一次我和貝先生在黃正德教授家裡，黃先生有事要出去，讓我們自便，自便的內容包括客廳裡茶几上的半盒巧克力。等黃先生回來的時候，貝先生已經一個人把那半盒巧克力自便完了！他還向黃先生道歉，說忘了應該留一點。

說到法國紅酒，大家一定要記住，絕不能在貝羅貝先生面前說它不好。你要說不好，他會跟你拼命！

這就是我認識的貝羅貝。以上是一些玩笑話。接下來我從三個方面來說說貝羅貝先生。

首先，貝先生是傑出的科研組織者和科研領導人。大家都知道，貝羅貝先生是從事漢語史研究的。而大家不知道的是，他在中國對自然科學基金講歐洲的基金管理，他在比利時給歐洲人講中國的科技政策。更好的是，將來他能在中國講中國的科技政策，在歐洲講歐洲的科學基金管理。我這麼講是因為貝羅貝先生在法國和歐洲有舉足輕重的地位，特別是在語言學研究領域。

第二，貝羅貝先生是一位非常傑出的研究者。在漢語史研究的眾多問題上，貝羅貝先生都有傑出的成就。我有次跟他開玩笑說，多年來我寫文章，他要麼是我論據主要的支持者，要麼就是我的主要的對手。貝羅貝先生所有的成績，使他在研究漢語史和在研究中國語言學的學者中，都具有很高的聲望。

在這兒，請允許我說一句題外的話。鑒於貝羅貝先生在語言學研究中的傑出貢獻，我們中國社會科學院語言研究所聘請貝羅貝先生作客座研究員。我們衷心的希望貝羅貝先生今年下半年有機會可以到我們研究所接受聘請，我們更希望在今後的研究工作中通過他能有更多的收穫。

第三，貝羅貝先生最重要的一個特徵是他真是一個好人。在座的不論是他的朋友還是他的同事、他的學生，大概都得到過他的幫助。中國人說，好人一生平安。貝羅貝先生 65 歲了，我們衷心的希望貝羅貝先生保持一個健康的身體，有旺盛的精力，能夠在語言學研究上作出更大的貢獻。

謝謝！

A Short Biography of ALAIN PEYRAUBE

Born in the wake of the liberation of Bordeaux and the Gironde during World War II — on precisely the 22nd November 1944 — our colleague, Alain Peyraube, has led nothing less than a remarkable and fortunate life — viewed from every aspect.

Alain is someone who incarnates multiple lives in the one person, in view of the astounding diversity and breadth of his career. Alain is not only a pioneer of research into the diachronic linguistics of Chinese, being the author of three books and over one hundred scholarly articles, but for most of his career to date has also been involved in promoting modern Chinese culture and literature in the west, including a series of translated works for the Parisian publisher, Flammarion, and the writing of more than 200 articles and compte-rendus in these domains. His engagement in post-Cultural Revolution Chinese politics and, in its wake, the intellectual and artistic revival in China, is seen in his regular contributions to the French newspaper *Le Monde* in the 1980s up until 1992, not to mention his involvement with Chinese cinema, presenting two films at successive Cannes film festivals.¹ Neither should the publication of four collaborative works, which include two book translations on dramatic periods of late 20th century Chinese political history, such as the 1978 Peking Spring and Tian'anmen Square in 1989, be overlooked in any comprehensive coverage of Alain Peyraube's intellectual life.

Up until 2012, Alain concurrently held the triple posts of *Directeur de Recherches de classe exceptionnelle* at the Centre National de la Recherche Scientifique (CNRS), *Directeur d'Etudes* at the Ecole des Hautes Etudes en Sciences Sociales (EHESS) and *Directeur de l'Institut des Etudes Avancées* in Lyon.² The enormous range of his interests is also seen in his involvement at both the European and international levels where he sits on many executive decision-making committees that play a major role in shaping future directions and strategies for research in Europe. These include notably

¹ These were the films 'Raise the Red Lantern' (*Dà hóng dēnglóng gāo-gāo guà* «大紅燈籠高高掛») directed by Zhang Yimou 張藝謀 (1991) based on the novel *Qīqiè Chéngqún* 妻妾成群 by the author Su Tong 蘇童; and 'Farewell my Concubine' (*Bàwáng biéjī* «霸王別姬») directed by Chen Kaige 陳凱歌 (1993) which was based on the novel of the same name by Lilian Lee — Li Bihua 李碧華. The latter film won the Palme d'Or at Cannes in 1993, the first Chinese film to win this award.

² These French titles roughly translate as 'Senior Research Fellow, Outstanding Class', at the French National Center for Scientific Research; 'Chair Professor' at the Graduate School of Humanities and Social Sciences in Paris, and Director of the Institute for Advanced Studies in Lyon. Alain Peyraube still holds the last two posts.

the Scientific Committee of the prestigious European Research Council, of which he was a founding member in 2007; the Standing Committee of the European Science Foundation and Fellow of the Academia Europaea in 2006. Internationally, the strong ties that he has established over more than 40 years with research centres in linguistics in China and Taiwan have led to his appointment as a Corresponding Member of the Academia Sinica of Taiwan in 2002, as Honorary Professor of Peking University in 2007 and as Honorary Research Fellow at the Chinese Academy of Social Sciences in 2009.

His career is all the more remarkable: Alain's father was a pointsman,³ working for the French National Railway Lines, the SNCF in Bordeaux,⁴ situated in the southwest of France, while his mother worked as a charlady to support Alain throughout his high school years. Yet even this seemingly less privileged background is quite deceptive of the unusually rich political and cultural milieu in which the family lived. His father, Fernand Peyraube, was a militant trade unionist who, during World War II, had also served as member of the French Resistance, bringing fellow fighters over the snowy paths of the Pyrénées mountains to safety in Spain and Spanish Republicans to relative freedom in France.

The family life of the Peyraubes was split between the city of Bordeaux and the region of Gascony to the south, more specifically the area corresponding the ancient viscounty of Béarn. Indeed, the older generations on both sides of Alain's family were staunchly Gascon, while the Gascon language, which belongs to the Occitan branch of Romance (Indo-European), was the language of familiarity particularly in evidence during visits to the grandparents' homes in Saint Médard and Orthez during holidays. After overnighing with one set of grandparents in the town of Orthez, the family would stay in the small village of Saint Médard, where after many years of sharecropping and later renting land, Alain's maternal grandfather had been able to buy a farm, still owned by the family today.

His mother, Angèle Peyraube, was a fine cook, winning many prizes for the regional delicacies she produced, including one of best-known — *foie gras*. As part of Alain's cultural upbringing, we can hardly fail to overlook that the southwest of France is an area that has enjoyed a long tradition of gourmet cuisine, and is equally well-known for its classical red wines and high quality produce. Suffice it to mention here the *confits*, *magrets*, and *terrines* of Gascony and Périgord, the *grands crus* wines of the Bordeaux region (St-Emilion, Médoc, Graves to name but a few) and the brandy of Armagnac, for all of which he is undoubtedly a connoisseur in the highest degree.

³ *Aiguilleur* in French comes under the name of 'pointsman' in British English or 'switchman' in American English.

⁴ SNCF is the acronym for the *Société nationale des chemins de fer français*.

On their return to Bordeaux, it was the national language, French, which dominated as the language mainly used with the outside world, being the formal language of education and the workplace. This is no doubt what Alain would describe as classic diglossia to his doctoral classes in linguistics today. Interestingly, however, Alain spoke more French with his father, a reflection of the fact, perhaps, of his father's involvement in, and enthusiasm for, the external world of union politics. — It was this combination of his Alain's father's political leanings, the family background and his experience of country life on their farm which was to instantly strike a chord of recognition as a 'true-blue' member of the proletariat when Alain was later to visit China for the first time in 1973, as the climactic end to the Cultural Revolution drew near. Another coincidence is hard to overlook: his father's occupation of pointsman, *bāndàogōng* 扳道工 in Chinese, also happened to be the same as that of the hero, Li Yuhe 李玉和, in one of Jiang Qing's (江青) eight model revolutionary operas approved for performance in China during this time, namely, *Hóng Dēng Jì* 紅燈記, The Legend of the Red Lantern.

Alain belongs to the few per cent of his generation who broke through all the social barriers, as the title of this book highlights. Most of his classmates left school at age 14, having completed their *certificat d'études* (Diploma of Studies) and the minimum schooling of 7 years, from there to proceed into different kinds of apprenticeships.

Alain attended a primary school in the Capucins area of Bordeaux which largely catered to the children of Spanish migrants, Republicans fleeing from Franco's regime, and this is where he learned to speak Spanish fluently. As one of the few native French-speaking pupils and with his natural talents, Alain could hardly help shining at the top of his class. Despite his parents' misgivings, Alain's teacher encouraged him at age 11 to sit for the entrance exams for the academic stream of high schools. He won one of the small number of 120 scholarships for the entire Bordeaux region which enabled him to attend one of the municipal *collèges*, the scholarship providing him with all his textbooks for free.

His parents, still uncertain that he was suited for a white collar career, advised him to sit for the competitive entrance exams for government positions that were announced during the period of his high school studies, including the PTT post office (*Poste, Téléphone et Télécommunications*) and the SNCF railways. Although he successfully passed nearly all these entrance exams, much to his father's disappointment — his father had after all set his heart on a teaching career as the highest possible goal his son could ever possibly attain as a way of breaking with the family tradition, Alain did not qualify for just *one* entrance exam — that for gaining a scholarship to *Ecole Normale* or Teachers' College.

After obtaining the French *baccalauréat* or matriculation diploma, Alain immediately enrolled for university in Bordeaux and, as an independent young man,

took on a full-time position as a live-in house warden, working at a series of secondary-level boarding schools in the Bordeaux region. This was his first ‘double life’ in the sense of leading two almost separate existences at once, and was to set the pattern for the years to come. Although he was rarely able to attend lectures at the University of Bordeaux, given his 40-hour working week, he undertook, and successfully completed, two concurrent Bachelor of Arts degrees (*licences*) with majors in two brand new subjects introduced into the curriculum at that time — Linguistics and Chinese, graduating in both in 1968. In the year that followed, 1969, he obtained his *Maîtrise*, or Master’s degree, in Chinese with a minor thesis on “Talks at the Yanan forum on literature and art, May 1942” by Mao Tse-Tung (毛澤東).

In 1970, he departed for Paris and took up a position as *collaborateur technique* (technical research assistant) in the Centre de linguistique chinoise under the directorship of Professor Alexis Rygaloff. These were the early golden years of the centre which was to evolve into today’s Centre de recherches linguistiques sur l’Asie orientale (CRLAO), a centre that has continuously served as Alain’s ‘research headquarters’ and which owes a tremendous debt to him for his unswerving support of all its endeavours. In the 1970s, the centre was attached to the Sorbonne with seminar courses offered by Rygaloff, a specialist in the grammar and phonology of Chinese. The discussions were lively, the course being attended regularly by Claude Hagège and André Haudricourt. During this period, Alain also worked as research assistant on the Franco-German “Sòng project” for his former professor of Classical Chinese in Bordeaux, Yves Hervouet, who had responsibility for editing and compiling the detailed résumés that were to make up a bibliography of the Song period.

The French government was one of the earliest of the western nations to re-establish diplomatic ties with China in 1964 under the presidency of Charles de Gaulle. Subsequent to this, the first group of French exchange students were sent to China in 1965 but for the most part had their stay aborted due to the onset of the Cultural Revolution. As a consequence, the second group of students only set foot in China some 9 years later in 1973, including Alain in their ranks who stayed until 1975, studying at both the Beijing Institute of Languages (北京語言學院 *Běijīng Yǔyán Xuéyuàn*) and at Peking University (北京大學 *Běijīng Dàxué*).

He was thus able to witness at first-hand this turbulent political period in China’s recent history, since Peking University was the base of the Gang of Four, who were at this time engaged in continuously launching political campaigns, as announced in articles by Liang Xiao 梁效 in the People’s Daily (人民日報 *Rénmín Rìbào*), the Red Flag (*Hóng Qí* 紅旗) and in other newspapers. With other western (and Chinese) students, he took part in work experience sessions known as *kāi mén bàn xué* 開門辦學 ‘Open Door Schooling’, living in the countryside and working in factories which were

intended to give university students, as members of the intellectual class, hands-on experience and appreciation of the hardships of proletarian and peasant life.

Significantly for his academic career, Alain was yet again fortunate in his specific request and strong wish for Professor Zhu Dexi (朱德熙) to be his doctoral supervisor during his studies at Běidà (Peking University). Professor Zhu, who was actually on secondment at the Wenwu Publishing House (文物出版社) graciously accepted and offered to return to Běidà to conduct bi-weekly meetings with Alain, his first foreign doctoral student after the end of the Cultural Revolution. Few doctoral students have ever been in such a unique and enviable position of having one of the doyens in the field of modern Chinese syntactic studies act as both linguistic consultant and mentor!

Tightly clasping his university diploma from Běidà from its *gōng-nóng-bīng* period (工農兵 a worker-peasant-soldier diploma), Alain returned to Paris at the end of 1975 and successfully defended his doctoral thesis on *Les constructions locatives en chinois moderne* (Locative constructions in Modern Chinese) at the Université de Paris VIII in 1976 under the supervision of Professor Alexis Rygaloff. From 1976 onwards, Alain took up an appointment with the National Center of Scientific Research (CNRS) as an *attaché de recherche* (Assistant Research Fellow), and was soon promoted to a *chargé de recherche* (Research Fellow) in 1981 and then a *directeur de recherche* (Senior Research Fellow) in 1985.

Nineteen seventy-seven was the year of Alain's first visit to Taiwan where he worked at Academia Sinica (中央研究院) for 4 months, signing important research agreements between the Institute of History and Philology (歷史語言研究所) and the CRLAO. For some, the more memorable event may have been the live appearance of a foreigner on Chinese TV in the personage of Bèi Luóbèi 貝羅貝, that is, none less than our Dr. Peyraube, performing a *xiàngshēng* (相聲) or Chinese comic dialogue, a quite extraordinary event for the time.

Embarked upon the second doctoral degree that was required in France during this period (a *doctorat d'état*), and finding that the formalist grammar paradigm had reached a momentary stalemate during the debate on generative versus interpretive semantics, Alain, with customary aplomb, decided to change tack and do research on Chinese historical linguistics. This was the second turning point in his career: the inspiration was a seminal article by Professor Mei Tsu-Lin (梅祖麟) on the evolution of the particle *le* 了, which saw him travelling to the United States in 1981-1982, despite his almost non-existent English competence that might have somewhat deterred other less venturesome personalities.

Alain spent several months at Cornell University, discussing linguistics with Professor Mei nearly every day and working on the vast holdings on Chinese Linguistics held at the University Library in the Charles W. Wason East Asian Collection. It was

Professor Mei Tsu-Lin who was thus instrumental in introducing him to Chinese diachronic linguistics and, it follows, to the important classical texts of Chinese. Mei Tsu-Lin also put Alain into contact with linguists at the Chinese Academy of Social Sciences in Beijing where historical linguistics had become the mainstream direction in the post-Cultural Revolution era with its renewal of research in China.

In 1984, Alain successfully defended his second doctoral degree at the Université de Paris VII with a masterpiece on the evolution of dative constructions from the Shang oracle bone texts of the 14th century BC to 18th century Modern Chinese. This was entitled in French: *Syntaxe diachronique du chinois — Evolution des constructions datives du 14^{ème} siècle avant J.-C. au 18^{ème} siècle*. On its publication as a monograph in 1988 by the Collège de France, he was awarded the prestigious Stanislas Julien prize by the French *Académie des Inscriptions et Belles-Lettres*.

During the first visit of any Chinese premier and leader of the Communist Party to the West — the state visit of Hua Guofeng 華國鋒 in October 1979, Alain acted as the official interpreter for the French President, Valéry Giscard d'Estaing, as he did on numerous other occasions for various government ministers. This was an historic occasion, as Premier Hua chose to set foot in Europe first of all in France, thereby honouring the long and special relationship between these two countries.

From the 1980s onwards, Alain continued to regularly visit the Institute of Linguistics at CASS in Beijing to discuss common research interests with Professors Liu Jian (劉堅), Jiang Lansheng (江藍生) and Cao Guangshun (曹廣順) whose positive influence on his own intellectual development in matters of historical linguistics has been immeasurable. These trips were interspersed with research stays at the Institute of History and Philology at Academia Sinica in Taiwan where he held many theoretical discussions with Professors Ting Pang-Hsin (丁邦新) and Mei Kuang (梅廣), a second important institute where he has built up a strong network of close work relationships. The Project on Linguistic Analysis, directed by Professor William S-Y. Wang (王士元) at UC Berkeley became another crucial venue for him for linguistic debate and exposure to the latest research in linguistics and its closely related domains, during Alain's visits to the USA.

This early foundation period of his research career is characterized by the use of a functional framework in historical and diachronic linguistics with a strong cross-linguistic perspective. Going far beyond the study of dative constructions, Alain considered many issues related to the mechanisms of syntactic and semantic change in Chinese from the period of the first recorded inscriptions (14th century BC) up until the modern period (18th century), as his publications list will attest. He has continued this research axis, delving into major considerations of contemporary diachronic linguistics of analogy, reanalysis (grammaticalization) and external borrowing due to language contact, for

which the research impact of his sharp analyses will be lasting and go beyond many geographical and theoretical borders.

From this period onwards, his commitments in the areas of research, teaching, editing, research administration and evaluation have risen exponentially, as too his publications and international lectures, conference talks and seminars. Suffice it to say that after this first period from 1983 to 2000 when Alain served as Director, after even earlier serving as Interim and then Co-director, of the Centre de recherches linguistiques sur l'Asie orientale (CRLAO), the late 1990s saw Alain branching out from matters purely Chinese to work on the national research stage in France as the Deputy Director of the Division of Humanities and Social Sciences of the CNRS from 1997 to 2001, then as Director of the Division of Humanities and Social Sciences at the French Ministry of Higher Education and Research from 2001 to 2007.

Apart from his research interests in Chinese diachronic syntax and semantics, Alain began to take an even broader outlook in the 1990s to work within a cognitive approach to analyze linguistic problems, and then significantly, on the origin and evolution of languages. He launched the Eurocores project of the European Science Foundation entitled 'The origin of man, language and languages (OMLL)' (or in French "Origine de l'Homme, du Langage et des Langues") which represented a major leap forward in the internationalization of French research and an objective that he had steadfastly followed from the outset of his career. Directed by Professor Jean-Marie Hombert at the Université de Lyon II, OMLL was a truly interdisciplinary project which encouraged close interaction between the CNRS and researchers from outside of France in the disciplines of linguistics, biological anthropology, paleo-anthropology, molecular and population genetics and neurosciences. In other words, it was a large scale project that brought together top researchers in the domains of the Humanities, Social Sciences and Life Sciences to tackle the intriguing and intertwined questions of the origin of humans and language. In the wake of these research pursuits, more recently Alain further branched into the study of the typology of East Asian languages and its bearing on the question of the degree of linguistic diversity — and thus the degree of divergence from predicted typological features — for Sinitic languages.

His influence on his 'home' research laboratory of the CRLAO is hard to overlook in the manner in which Alain has helped to build it up in terms of researchers, facilities and the drive for research funding. In his indefatigable manner, he continues to keep a sharp eye on the CRLAO, long after relinquishing the director's post to take up these even higher duties in the hierarchy of French research administration, and also at the level of the European Union, where he continues to serve on many executive committees.

Alain's intrinsic diplomatic skills and *savoir faire*, his refusal to resign in the face of seemingly insurmountable difficulties or even to cave in at the slightest adversity are

List of Publications of ALAIN PEYRAUBE

Monographs and edited works in linguistics

- 現代漢語語法講義 *Xiandai Hanyu Yufa Jiangyi* [Course in Modern Chinese Grammar]. Bordeaux: Université de Bordeaux III, 1978, 201 pp. (in Chinese).
- *Les constructions locatives en chinois moderne*. Paris: Editions Langages croisés, 1980, 337 pp.
- *Syntaxe diachronique du chinois — évolution des constructions datives du 14^{ème} siècle avant J.-C. au 18^{ème} siècle*. Paris: Collège de France (Collection “Institut des Hautes Etudes Chinoises”), 1988, 351 pp. (Stanislas Julien prize in 1989 awarded by the Académie des Inscriptions et Belles-Lettres).
- *In Honor of Mei Tsu-Lin: Studies on Chinese Historical Syntax and Morphology*. Paris: EHESS, 1999, 362 pp. (co-edited with Chaofen Sun).
- *Encyclopédie des Sciences du Langage: Dictionnaire des langues*. Paris: Presses Universitaires de France, 2011, 1705 pp. (co-edited with E. Bonvini, Joëlle Busutil).

Monographs and edited works in other domains

- *Le printemps de Pékin, Oppositions démocratiques en Chine. Novembre 1978 - mars 1980*. Paris: Gallimard (Collection “Archives”), 1980, 249 pp. (co-authorship using the nom de plume of ‘Victor Sidane’). (Spanish edition translated by T. Bilbao, *La primavera de Pekin*, Mexico: Folio Ediciones, 1982, 211 pp.).
- *Le tremblement de terre de Pékin*. Paris: Gallimard (Collection “Au vif du sujet”), 1991, 588 pp. (in collaboration with J.-P. Béja & M. Bonnin).

Translations from Chinese

- *Procès politiques à Pékin: Wei Jingsheng, Fu Yuehua*. Paris: Maspero, 1981, 243 pp. (co-authorship using the nom de plume of ‘Victor Sidane’).
- Liu Qing, *J'accuse devant le tribunal de la société*. Paris: R. Laffont, 1982, 261 pp. (in collaboration with the “Collectif pour l'étude du mouvement démocratique en Chine”).
- Lao She (老舍), *Gens de Pékin*. Paris: Gallimard, NRF, 1982, 291 pp. (in collaboration with P. Bady, F. Moreux, T. Li & M. Valette-Hémery). Reprinted under the title of *Histoire de ma vie*. Paris: Gallimard, Folio, 2002, 115 pages. Translation of several short novels.

- *Les opéras des bords de l'eau. Théâtre Yuan, 13^{ème} – 14^{ème} siècles*. Paris: Association pour l'Analyse de Folklore, 1983, 189 pp. (in collaboration with M. Coyaud & A. K. Leung). Translation of *Shuǐhǔ Xìjù* 水滸戲劇.
- *La remontée vers le jour*. Aix: Alinéa, 1988, 270 pp. (Translation of a collection of short stories: “Le glacier” by Zhang Chengzhi (張承志), pp. 113-132, in collaboration with Baiyun). Translation of *Dàbǎn* 大坂.
- Li Ang (李昂), *La femme du boucher*. Paris: Flammarion, 1992, 199 pp. (in collaboration with H.-F. Vizcarra). New edition with the title of *Le meurtre du mari*, Paris: Denoël, 2004. Translation of *Shāfū* 殺夫.

Articles and book chapters in linguistics (partial list)

• 1970-1980

- La réforme de l'écriture en Chine. *Tel Quel* 64:91-103, Hiver 1975.
- Adverbiaux et compléments de lieu en chinois. *Cahiers de linguistique – Asie orientale* 1:43-60, 1977.
- La linguistique en Chine après la Révolution culturelle. *Journal Asiatique* CCLXV.1-2:185-211, 1977.
- L'enseignement du chinois aux étrangers à l'Institut des langues de Pékin. *Cahiers de Linguistique, d'Orientalisme, et de Slavistique* 10:157-173, 1977. English translation by John DeFrancis: The teaching of Chinese to foreign students at the Peking Foreign Languages Institute, *Journal of the Chinese Language Teachers Association* 13.1:26-36, 1978.
- Les syntagmes prépositionnels de lieu adverbiaux en Mandarin-problèmes d'analyse syntaxique. *Cahiers de linguistique – Asie orientale* 3:25-33, 1978.
- Bilingual education in the People's Republic of China. *Studies in Bilingual Education*, ed. by R. Lord & B. K. Tsou, 61-66. Hong Kong: Heinemann Educational Books, 1979.
- La langue commune et le Pékinois. *Langues formelles, langues quotidiennes: quelques langues d'Asie, Journée d'études*, ed. by A. Cartier, 51-61. Paris: U.E.R. de linguistique générale et appliquée, Université R. Descartes, 1980.
- Les “approximatifs” chinois: *chabuduo, jihu, chayidianr*. *Cahiers de linguistique – Asie orientale* 6:49-62, 1979.

- **1980-1990**

- The dative construction in Cantonese. *Computational Analyses on Asian and African Languages* 16:29-65, 1981.
- Les termes d'adresse dans les opéras des Yuan aux 13^{ème} et 14^{ème} siècles. *Cahiers de linguistique – Asie orientale* 11.2:3-36, Décembre 1982.
- Réforme et modernisation de la langue chinoise. *Language Reform: History and Future*, ed. by I. Fodor & C. Hagège, 191-209. Hamburg: H. Buske, 1983. (in collaboration with C. Hagège & G. Métailié).
- Les structures en *ba* en chinois vernaculaire médiéval et moderne. *Cahiers de linguistique – Asie orientale* 14.2:193-213, 1985.
- 漢語拼音正詞法評析 Hanyu pinyin zhengcifa pingxi [Critical analysis of the 'pinyin' spelling system]. 漢語拼音正詞法論文選 *Hanyu Pinyin Zhengcifa Lunwenxuan* [Selected Essays on the 'Pinyin' Spelling System], 146-161. Beijing: Wenzi Gaige Chubanshe, 1985. (in collaboration with A. Lucas).
- 雙賓語結構從漢代至唐代的歷史發展 Shuangbinyu jiegou cong Handai zhi Tangdai de lishi fazhan [The double-object construction: its development from Han to Tang]. 中國語文 *Zhongguo Yuwen* 1986.3:204-216.
- A propos de C. Harbsmeier, Aspects of Classical Chinese Syntax. *Cahiers de linguistique – Asie orientale* 15.2:329-337, 1986.
- The double-object construction in Lunyu and Mengzi. *Wang Li Memorial Volumes*, 331-358. Hong Kong: Joint Publishing Company, 1987.
- L'expression de l'à-peu-près dans la langue chinoise contemporaine. *L'à-peu-près, aspects anciens et modernes de l'approximation*, 243-254. Paris-Ancône: EHESS - Il Lavoro editoriale, 1988. (in collaboration with Shiqi Song).
- Les politiques linguistiques de quatre pays de langue chinoise et les attitudes culturelles des populations concernées: le cas de Taiwan. Rapport ATP du CNRS, 101-123, 1988.
- 早期把字句的幾個問題 Zaoqi 'ba' ziju de jige wenti [Several issues concerning the first *ba* constructions]. 語文研究 *Yuwen Yanjiu* 1989.1:1-9.
- History of the passive construction in Chinese until the 10th century. *Journal of Chinese Linguistics* 17.2:335-372, 1989.
- History of the comparative construction in Chinese from the 5th century B.C. to the 14th century A.D. *Proceedings of the Second International Conference on Sinology: Section on Linguistics and Paleography*, 589-612. Taipei: Academia Sinica, 1989.

- **1990-2000**

- Some diachronic aspects of diglossia/triglossia in Chinese. *Southwest Journal of Linguistics* 10.1:105-124, 1991.
- Syntactic change in Chinese: on grammaticalization. *Bulletin of the Institute of History and Philology: In Memory of Dr. Fang Kuei Li*. Taipei: Academia Sinica 59.3:617-652, 1991.
- 論差不多、幾乎、差點：一個漢語與法語的對比語言學問題 Lun chabuduo, jihu, Chadian: yige Hanyu yu Fayu de duibi yuyanxue wenti [Concerning *chabuduo*, *jihu*, *chadian*: a contrastive linguistic problem in French and Chinese]. *Proceedings of the 2nd International Conference on Teaching Chinese as a Second Language*, 161-176. Taipei: World Chinese Language Association, 1991.
- Some remarks on the history of Chinese classifiers. *Santa Barbara Papers in Linguistics* Volume 3: *Asian Discourse and Grammar*, 106-126, 1991.
- Some remarks on the demonstratives in the Fuzhou dialect with reference to their historical evolution in Medieval and Modern Chinese. *Chinese Languages and Linguistics*, Vol. 1: *Chinese Dialects*, 493-522. Taipei: Academia Sinica, 1992. (in collaboration with Michel Désirat).
- L'iconicité: un nouveau dogme de la syntaxe chinoise? *Faits de langues* 1:69-78, mars 1993. (in collaboration with Marie-Claude Paris).
- Le rôle des classificateurs nominaux en chinois et leur évolution historique: un cas de changement cyclique. *Faits de langues* 2:51-61, septembre 1993. (in collaboration with Thekla Wiebusch).
- Orientalisme et linguistique. *Livre Blanc de l'Orientalisme français*, 101-103. Paris: Société Asiatique, 1993.
- 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: Academia Sinica, 1994.
- History of some coordinative conjunctions in Chinese. *Journal of Chinese Linguistics* 22.2:179-201, 1994. (in collaboration with Jian Liu).
- Is Chinese syntax iconically motivated? *Journal of Linguistics: Working Papers of Hong Kong Baptist University*, 40-48. Hong Kong Baptist University, 1994.
- Nouvelles réflexions sur l'histoire des formes accusatives en 'ba' du chinois. *Cahiers de Linguistique – Asie orientale* 23:265-277, 1994.
- Sur un cas curieux d'ambiguïté en chinois médiéval concernant des formes comparatives. *Faits de langues* 5:73-81, mars 1995. (in collaboration with Thekla Wiebusch).

- Problems relating to the history of different copulas in Ancient Chinese. *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*, ed. by M. Y. Chen & O. J.-L. Tzeng, 383-404. Taipei: Pyramid Press, 1994. (in collaboration with Thekla Wiebusch).
- Recent issues in Chinese historical syntax. *New Horizons in Chinese Linguistics*, ed. by C.-T. J. Huang & Y.-H. A. Li, 161-213. Dordrecht: Kluwer Academic Publishers, 1996.
- Le cantonais est-il du chinois? *Perspectives chinoises* 34:26-29, 1996.
- On word order in Archaic Chinese. *Cahiers de linguistique – Asie orientale* 26.1: 3-20, 1997.
- On the double-object construction and the oblique construction in Cantonese. *Studies in Language* 21.1:105-127, 1997. (in collaboration with Liejiong Xu).
- On word order and word order change in Pre-Archaic Chinese. *Chinese Languages and Linguistics*, Vol. 4: *Typological Studies of Languages in China*, ed. by Chiu-yu Tseng, 105-124. Taipei: Academia Sinica, 1997.
- Cantonese post-verbal adverbs. *In Memory of Mantaro J. Hashimoto*, ed. by A. O. Yue-Hashimoto & M. Endo, 303-313. Tokyo: Uchiyama Shoten, 1997.
- 古代漢語中的動之名結構 Gudai Hanyu zhong de ‘dong-zhi-ming’ jiegou [On the structure ‘verb – zhi – noun’ in Classical Chinese]. 古漢語語法論集 *Gu Hanyu Yufa Lunji* [Anthology of Articles on the Grammar of Classical Chinese], ed. by Guo Xiliang, 392-407. Beijing: Language and Culture Press, 1998.
- On the history of classifiers in Archaic and Medieval Chinese. *Studia Linguistica Serica*, ed. by B. Tsou, 131-145. Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong, 1998.
- Ordre des mots et changement d’ordre des mots en chinois ancien. *Comptes-rendus des séances de l’Académie des Inscriptions et Belles-Lettres*, 2 :533-542. Paris: De Boccard, 1997.
- 二十世紀以前歐洲漢語語法學研究狀況 Ershi shiji yiqian Ouzhou Hanyu yufaxue yanjiu zhuangkuang [Studies in Chinese grammar in Europe before the 20th century]. 中國語文 *Zhongguo Yuwen* 1998.5:346-352.
- 上古、中古漢語量詞的歷史發展 Shanggu, Zhonggu Hanyu liangci de lishi fazhan [Historical evolution of classifiers in Ancient Chinese]. 語言學論叢 *Yuyanxue Luncong* 21:99-122. Beijing: The Commercial Press, 1998.
- On the modal auxiliaries of possibility in Classical Chinese. *Selected Papers from the Fifth International Conference on Chinese Linguistics*, ed. by H. Samuel Wang, Feng-fu Tsao & Chinfa Lien, 27-52. Taipei: Crane, 1999.
- The “verb + zhi + noun” structure in Classical Chinese. *Chinese Languages and Linguistics*, Vol. 5: *Interactions in Language*, ed. by Yuen-mei Yin, I-li Yang & Hui-chan Chan, 593-613. Taipei: Academia Sinica, 1999.

- Sur le changement syntaxique et sémantique: à propos de deux ouvrages récents sur la grammaire historique du chinois. *Cahiers de linguistique – Asie orientale* 28.1:113-124, 1999.
- Les déplacements de constituants en syntaxe historique du chinois. *In Honor of Mei Tsu-Lin: Studies on Chinese Historical Syntax and Morphology*, ed. by A. Peyraube & C. Sun, 131-143. Paris: EHESS, 1999.
- L'origine des langues et du langage. *Sciences humaines* 27:6-10, 1999. (reprinted in the volume edited by J.-F. Dortier, *Le Langage*, 153-163, Editions Sciences Humaines, 2001.).
- Historical change in Chinese grammar. *Cahiers de linguistique – Asie orientale* 28.2:177-226, 1999.
- Sur les sources du Ma shi wen tong. *Histoire, Épistémologie, Langage* 21.2:65-78, 1999.
- **2000-2010**
 - Westernization of Chinese Grammar in the 20th century: myth or reality? *Journal of Chinese Linguistics* 28.1:1-25, 2000.
 - 上古漢語疑問代詞的發展與演變 Shanggu Hanyu yiwen daici de fazhan yu yanbian [Development and evolution of interrogative pronouns in Chinese]. 中國語文 *Zhongguo Yuwen* 2000.4:311-326. (in collaboration with Fuxiang Wu).
 - Idées linguistiques en Chine ancienne. *Corpus représentatif des grammaires et des traditions linguistiques*. Tome II, *Histoire Epistémologie Langage*, hors-série n°3, ed. by Bernard Colombat & Elisabeth Lazcano, 299-324. Paris: Shesl, 2000. (in collaboration with Françoise Bottéro & Barbara Niederer).
 - 二十世紀以前歐洲漢語語法學研究狀況 Ershi shiji yiqian Ouzhou Hanyu yufaxue yanjiu zhuangkuang [Research on Chinese grammar in Europe before the 20th century]. 文化的饋贈 *Wenhua de Kuizeng* [Contribution of Culture], 467-474. Beijing: Peking University Press, 2000. (summarized version).
 - Research on Chinese grammar in Europe before the 20th century. 馬氏文通與漢語語法學—馬氏文通出版百年 (1989-1998) 紀念文集 *Mashi Wentong yu Hanyu Yufaxue*, ed. by Jingyi Hou & Guanjin Shi, 149-161. Beijing: The Commercial Press, 2000.
 - Le rôle du savoir linguistique dans l'éducation et la société chinoise. *History of the Language Sciences*, ed. by S. Aurox, E. F. K. Koertner & H. J. Niederehe, 55-58. Berlin: Walter de Gruyter, 2000.
 - Syntaxe et discours dans la langue chinoise: le mythe de l'occidentalisation. *Mots* 66:19-32, 2001.

- 清代馬氏文通以前的語法知識 Qingdai *Mashi Wentong* yiqian de yufa zhishi [Grammatical knowledge in China of the Qing Dynasty before the *Mashi Wentong*]. *Collected Essays in Ancient Chinese Grammar*, ed. by R. Djamouri, 1-9. Paris: EHESS, CRLAO, 2001.
- Some reflections on the sources of the *Mashi Wentong*. *New Terms for New Ideas: Western Knowledge and Lexical Change in Late Imperial China*, ed. by M. Lackner, I. Amelung & J. Kurtz, 341-355. Leiden: Brill, 2001.
- On the modal auxiliaries of volition in Classical Chinese. *Sinitic Grammar: Synchronic and Diachronic Perspectives*, ed. by H. Chappell, 172-188. Oxford: Oxford University Press, 2001.
- L'évolution des structures grammaticales. *Langages* 146:46-58, 2002.
- Origin and evolution of question-words in Archaic Chinese: a cognitive approach. *Language and Cognition*, ed. by Yuchau E. Hsiao, 258-272. Taipei: National Chengchi University, 2002.
- Ordre des constituants en chinois archaïque. *Cahiers de linguistique de l'INALCO*, 99-110, 2000.
- On moving constituents in Chinese historical syntax. *Historical Development of Chinese Language*, ed. by Dah-an Ho, 141-156. Taipei: Institute of Linguistics, Academia Sinica, 2002.
- Languages at risk. *International Seminar on Nuclear War and Planetary Emergencies: 27th Session*, ed. by A. Zichichi, 274-283. Singapore: World Scientific Publishing Company, 2003.
- On the history of place words and localizers in Chinese: a cognitive approach. *Functional Structure(s), Form and Interpretation: Perspectives from East Asian Languages*, ed. by Y.-H. A. Li & A. Simpson, 180-198. London: RoutledgeCurzon, 2003.
- Some thoughts on cultural relativism. *International Seminar on Nuclear War and Planetary Emergencies: 29th Session*, ed. by R. Ragaini, 62-69. Singapore: World Scientific Publishing Company, 2003.
- Ancient Chinese. *The Cambridge Encyclopaedia of the World's Ancient Languages*, ed. by R. Woodard, 988-1014. Cambridge: Cambridge University Press, 2004.
- Grammaire diachronique et cognition: l'exemple du chinois. *La linguistique cognitive*, ed. by C. Fuchs, 135-154. Paris: Orphys-MSH, 2004.
- Sens et sons en chinois. *Le Monde* 19, mars 2004.
- La traduction — un art ou un produit du marché? *Europa denkt mehrsprachig — L'Europe pense en plusieurs langues*, ed. by F. Nies, 223-228. Tübingen: Gunter Narr Verlag, 2005.

- Origin and evolution of question-words in Archaic Chinese: a cognitive approach. *Cahiers de linguistique – Asie orientale* 34.1:3-24, 2005. (in collaboration with Fuxiang Wu).
- Comment classer les langues? *Aux origines des langues et du langage*, ed. by J.-M. Hombert, 338-345. Paris: Editions Fayard, 2005.
- L’homme aux six mille langues. *Aux origines des langues et du langage*, ed. by J.-M. Hombert, 450-461. Paris: Editions Fayard, 2005.
- Avancées récentes en linguistique et enseignement du chinois. *Enseigner le chinois*, 51-59. Paris: Ministère de l’Education nationale, de l’enseignement supérieur et de la recherche, Direction de l’enseignement scolaire, 2005.
- 漢語的語法演變：論語法化 Hanyu de yufa yanbian: lun yufahua [Grammatical change and grammaticalization in Chinese]. 漢語語法化研究 *Hanyu Yufahua Yanjiu*, ed. by Wu Fuxiang, 44-72. Beijing: The Commercial Press, 2005.
- Motion events in Chinese: a diachronic study of directional complements. *Space in Languages: Linguistic Systems and Cognitive Categories*, ed. by M. Hickmann & S. Robert, 121-138. Amsterdam: John Benjamins, 2006.
- On the history of some adverbs of scope and quantity in Chinese. *Studies in Chinese Language and Culture: Festschrift in Honour of Christoph Harbsmeier in the Occasion of his 60th Birthday*, ed. by C. Anderl & H. Eifring, 269-283. Oslo: Hermes Academic Publishing, 2006.
- The analytic causatives of Early Modern Southern Min in diachronic perspective. *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, 973-1011. Taipei: Institute of Linguistics, Academia Sinica, 2006. (in collaboration with H. Chappell).
- Chinese linguistics and typology: the state of the art. *Linguistic Typology* 11.1: 187-211, 2007. (in collaboration with H. Chappell & Ming Li).
- 近代早期閩南話分析型致使結構的歷史探討 Zaoqi jindai Minnanhua fenxixing zhishi jiegou de lishi tantao [A diachronic exploration of analytic causatives in Early Modern Southern Min]. 方言 *Fangyan* 2007.1:52-59. (in collaboration with H. Chappell).
- Chinese localizers: diachrony and some typological considerations. *Space in Languages of China: Cross-linguistic, Synchronic and Diachronic Perspectives*, ed. by Dan Xu, 15-37. Amsterdam: Springer, 2008. (in collaboration with H. Chappell).
- Languages and genes in China and in East Asia. *Bulletin of Chinese Linguistics* 2.1: 1-16, 2007.
- 語義演變理論與語義演變和句法演變研究 Yuyi yanbian yu yufa yanbian lilun [Theory of semantic and syntactic change]. 當代語言學理論和漢語研究 *Dangdai*

Yuyanxue Lilun he Hanyu Yanjiu [Contemporary Linguistic Theories and Related Studies on Chinese], ed. by Yang Shen, 1-25. Beijing: The Commercial Press, 2008. (in collaboration with Ming Li).

- 漢語趨向補語結構的產生與演變 Hanyu quxiang buyu jiegou de chansheng yu yanbian [Emergence and evolution of the directional verb-complement construction in Chinese]. 歷史語言學研究 *Lishi Yuyanxue Yanjiu* 2008.1:164-181. (in collaboration with Yingfeng Liang & Fuxiang Wu).
- On the history of Chinese directionals. *Proceedings of the Workshop in Chinese Directionals: History and Dialectal Variation, in conjunction with the 6th Cross-Strait Conference on Chinese Historical Grammar*, 174-187. Taipei: Academia Sinica, 2009.
- 漢語意願動詞的歷史演變 Hanyu yiyuan dongci de lishi yanbian [The historical development of volitional verbs in Chinese]. 漢語史學報 *Hanyushi Xuebao* 2009.8:17-35. (in collaboration with Ming Li).
- 漢語的體和時 Hanyu de ti he shi [Tense and aspect in Chinese]. 漢語時體的歷時研究 *Hanyu Shi Ti de Lishi Yanjiu* [Historical Research on Chinese Tense and Aspect], ed. by L. Feng, Y. Yang & C. Zhao, 1-8. Beijing: Language and Culture Press, 2009.
- 漢語歷史語法與類型學 Hanyu lishi yufa yu leixingxue [Diachrony and typology in Chinese grammar]. 歷史語言學研究 *Lishi Yuyanxue Yanjiu* 2009.2:1-10. (in collaboration with Dan Xu).
- 關於連動式的歷史及發展的幾點考慮 Guanyu liandongshi de lishi ji fazhan de ji dian kaolü [Some issues on the historical development of the serial verb constructions]. *Collected Essays for the Centennial Anniversary of Prof. Lü Shuxiang*, 249-258. Beijing: The Commercial Press, 2010. (in collaboration with Hui-ju Hsiung).
- 《老乞大》的個體量詞和語言循環現象之關係 Lao Qida de geti liangci he yuyan xunhuan xianxiang zhi guanxi [The classifiers in *Lao Qida* and the hypothesis of cyclic change]. 漢語史學報 *Hanyushi Xuebao* 10:1-9, 2010. (in collaboration with Jang-Ling Lin).

• **Since 2010**

- Cognition and linguistics for ever together. *The New Waver 4: Cognitive Science*, 56-60. Paris: Editions Autrement, 2011.
- 西漢初期的概數表達 Hanyu de gaishuci [Expressions of approximate numbers in Early Han]. 歷史語言學研究 *Lishi Yuyanxue Yanjiu* 2010.3:114-122. (in collaboration with Dan Xu).

- Grammaticalization in Sinitic languages. *The Oxford Handbook of Grammaticalization*, Chapter 65, ed. by Bernd Heine & Heiko Narrog, 784-794. Oxford: Oxford University Press, 2011. (in collaboration with H. Chappell).
- 上古漢語的語序 Shanggu Hanyu de yuxu [Word order in Archaic Chinese]. Fuxiang Wu (ed.) (forthcoming).
- Introduction. *Dictionnaire des langues*, ed. by E. Bonvini, J. Busuttil & A. Peyraube, XV-XXV. Paris: Presses Universitaire de France, 2011. (in collaboration with Emilio Bonvini).
- Le Cantonais. *Dictionnaire des langues*, ed. by E. Bonvini, J. Busuttil & A. Peyraube, 1009-1015. Paris: Presses Universitaire de France, 2011. (in collaboration with Joëlle Busuttil).
- Les langues sino-tibétaines. *Dictionnaire des langues*, ed. by E. Bonvini, J. Busuttil & A. Peyraube, 977-984. Paris: Presses Universitaire de France, 2011.
- A comitative source for object markers in Sinitic languages: 跟 kai⁵⁵ in Waxiang and 共 kang⁷ in Southern Min. *Journal of East Asian Linguistics* 20.4:291-338, 2011. (in collaboration with H. Chappell & Yunji Wu).
- Les langues d'Asie et d'Océanie. *Dictionnaire des langues*, ed. by E. Bonvini, J. Busuttil & A. Peyraube, 911-918. Paris: Presses Universitaire de France, 2011. (in collaboration with Joëlle Busuttil).
- Mood and modality in Sinitic. *Handbook of Mood and Modality*, ed. by J. Nuyts & J. van der Auwera. Oxford: Oxford University Press. (in preparation). (in collaboration with H. Chappell).
- Individuating classifiers in Early Southern Min (14th-19th centuries). *Quantification and Plurality*, ed. by Dan Xu, 101-120. Berlin: Mouton de Gruyter, 2012. (in collaboration with Jang-Ling Lin).

The Evolutionary Origin of Language Reconsidered

Charles N. Li (李 訥)

University of California, Santa Barbara

There are markers and processes in hominid's evolution and cognitive capability that are relevant to the emergence of language. After a long process of creation and concatenation of symbolic signals in their speech, grammatical structures emerged shortly, in a swift, graded and ongoing manner. This dynamic process defies any attempt to impose on it a non-arbitrary discrete marker as the point of origin of language.

Key words: emergence of human language, language origin, human evolution, symbolic signals, gesturing, animal communication, grammatical principles, *Foxp2*

1. Introduction

Throughout history, humans have been fascinated with origins: The origin of the universe, the origin of the earth, the origin of life, the origin of human, the origin of language. Unlike the origins of the universe, the earth, life and human, which are consigned to scientific investigation, the origin of language has attracted a good amount of speculations. The earliest one on record comes from the Old Testament, where we learn that a vengeful God punished the people of Babylon for their vanity and disobedience by creating a great diversity of languages.

By the 21st century, a new round of speculations on the origin of language and other complex human behavior has prompted the evolutionary biologist, Richard Lewontin, to sound an alarm in an AAAS meeting (Feb., 2008), that research on the evolution of hominid cognition is akin to 'paleofantasy'. There is no fossil record of human cognition, observed Mr. Lewontin.

Even though Lewontin's alarm is salutary, I believe that the combined advances in the study of fossil records (paleoanthropology), neuroscience, human genetics, animal communicative behavior and linguistics have laid the empirical foundation for mapping out the landmarks, the processes, the time frame and an outline of the evolutionary origin of language.

First, several important landmarks in hominid evolution allow us to delimit a period of time during which human language emerged. Secondly, a wide array of research on

animal communicative behavior in the second half of the 20th century has unveiled the fundamental differences between animal communicative behavior and human language, enhancing our ability to infer the transformations undertaken by hominid communicative behavior that led to the emergence of language. Thirdly, archaeological, neurological and human genetic evidence provides a number of clues to the evolutionary processes leading to the emergence of language. Finally, advances in linguistics have elucidated the emergence of grammar, a dynamic process that never ends in a living language.

Let me begin by clarifying the genre of modern language that may give us a hint of language at its origin.

2. Modern language and language origin

Modern human language has multiple genres, similar to a Linnean genus consisting of different species. These genres constitute a spectrum with graded differences in grammar, organization as well as vocabulary. If we compare the two extremes of the spectrum, their differences, instead of being graded and minor, will appear discrete and significant in myriad ways.

At one extreme of the spectrum is the formal written language constructed in sentences and paragraphs, which, at their best, are rendered in logical and coherent arguments or narrative, and composed strictly according to grammatical conventions. In this channel of communication, the linguistic text must bear the entire burden of transmitting information and attaining communicative felicity between the author and the reader. This is the prevalent genre of language in academic research. Educated people favor the formal written language to such an extent that they endeavor, in certain social contexts, to speak in the way they write.

At the other extreme of the spectrum is the casual spoken language of pre-literate cultures, or the casual conversation in contemporary social interaction. This genre of language is composed of words, phrases, intonation units, pauses and various interjections, where grammatical constraints are far weaker than they are in written text. Unlike a piece of written text which must be self-sufficient for the purpose of proper transmission of information to the reader, casual spoken language can draw on extra-linguistic information in the communicative context. In other words, unlike a written text, which stands alone in the transmission of information, the language of casual conversation constitutes only one of many facets of a communicative context. If one compares the language in a high-fidelity transcript of a session of casual conversation with the written language of an academic paper, one will have the impression that the former looks haphazard, anarchistic and often ungrammatical. It is, therefore, not surprising that educated people in a literate society tend to hold the language of casual

conversation in low esteem.¹

To underscore the gulf separating the two opposite ends of the spectrum of language genres, I will point out two pieces of evidence that are relevant to the origin of language.²

First, every human being is genetically disposed toward acquiring a casual spoken or sign language without instruction, whereas no one can make the same claim for the acquisition of a written language.³ To learn to write is an arduous intellectual endeavor that requires long-term schooling. To those who take writing seriously, it is an art and craft to be honed and improved as long as they live. Probably no more than twenty percent of the world's population is fluent in a written language. But every human is fluent in a spoken language.

Secondly, written language was invented five to six thousand years ago. It is a vital invention that has fueled cultural development, because it permits the unlimited preservation and accumulation of knowledge for transmission to future generations. Casual spoken language precedes written language by hundreds of thousands of years. This chronological difference alone is sufficient to compel us to acknowledge that written language is much too far removed from the origin of human language for it to be useful in investigating how the communicative behavior of hominids evolved into language. Casual spoken language is the only conduit through which we may gain some insight into the nature of language at its origin. In particular, we must bear in mind that long before the dawn of civilization, when the behavior of our hominid ancestors was still focused on survival, as it is the case among all undomesticated, non-human animals today, language was the vehicle of communication that elevated humans to the unprecedented position of absolute dominance in the animal kingdom.

I have established an understanding of the linguistic base of what we are looking for as hominid communicative behavior evolved toward language, i.e., a communicative behavior that involves linguistic as well as extra-linguistic information, and a linguistic code that is far less grammatically constrained and far more haphazard than formal written language. I will now proceed to ferret out the major landmarks in hominid evolution that are relevant to the origin of language.

¹ This low esteem has had immense impact on generative linguistic theory which, at least during its early development, uses, as its empirical base, the genre of spoken language that closely approximates formal written language.

² There is a significant literature on the differences between written language and spoken language which does not need to be cited here.

³ There is a significant literature on the acquisition of literacy which is beyond the scope of this article.

3. Hominids evolution and language origin

The family of hominids began approximately six to seven million years ago by a group of great apes who descended from a tree habitat to a savanna habitat in Africa. The cause for this habitat change is assumed to be the geological formation of the Great Rift Valley in East Africa. As East Africa began drifting away from the continent because of tectonic movement, creating the so-called Great Rift Valley, the climate on the East side of the valley turned arid. There, tropical forest gave way to savanna and plains such as the Serengeti and Masai Mara. According to this scenario of geological transformation, our earliest hominid ancestors didn't come down the trees voluntarily. The trees they were living in gradually disappeared.

This is the so-called East Side Story in paleoanthropology. The validity of the East Side Story is now being challenged by recent discoveries of six million year old fossils in Chad and Kenya.⁴ If these fossils turn out to be hominids, the East Side Story will be cast in doubt because both sites of discovery were rain forest habitat six million years ago.

Whether or not the East Side Story proves to be a valid hypothesis for our earliest hominid ancestors to descend from the trees, fossil records confirm that the earliest hominids originated in Africa and genetic studies based on the rate of mutation place the emergence of hominids at around six to seven million years ago.

The second landmark more or less coincides with our hominid ancestor's descent from trees. It is the trait that defines the Linnean family of Hominids, namely, bipedalism. All non-human living great apes knuckle-walk. Humans are the only great ape that has the body and brain structure for walking upright. It took about four million years for hominids to evolve into full-fledged bipedalism and its attendant anatomical structures. This evolutionary transformation involves not just the skeletal structure, but also the muscular composition, the central and the peripheral nervous systems. In other words, the emergence of bipedalism coincides with the start of the evolutionary development of the hominid nervous system to beget a larger and more complex brain.

Lucy, the Australopithecine made famous by Donald Johanson, was a bipedal hominid living approximately 3.3 million years ago. But her skeletal and muscular structure is very different from ours in that while it enabled her to walk upright, it also facilitated tree-climbing.

⁴ The fossil found in Chad is named Toumai, formally classified as *Sahelanthropus tchadensis*. It is dated between 7-6 million years ago. The other fossil, found in Kenya, is the so-called Millennium Man, formally classified as *Orrorin tugenensis*. It is dated around 6 million years ago. Whether or not these two species are hominids remains unresolved at this time.

To date, we still don't have a satisfactory explanation for the emergence of bipedalism. Speculations and theories abound, ranging from freeing the hands, to sexual attraction, to greater visual field. But none can be considered definitive. We do know, however, that bipedalism is swifter and more energy efficient than knuckle-walking. Perhaps it is a combination of all the afore-mentioned factors that drove the evolution toward bipedalism.

The third landmark in hominid evolution is the permanent grounding of hominids. Let me explain what I mean by 'permanent grounding'.

I have mentioned that early hominids such as Lucy, although capable of bipedal locomotion, had an anatomy that also facilitated tree-climbing. Lucy's arms, for example, are longer than her legs — an important anatomical feature for brachiation, whereas modern human legs are much longer than arms; Lucy's fingers are curved — enabling firmer grasp of tree branches, but our fingers are straight; Lucy's feet have longer toes and a significant gap between the big toe and the other toes — allowing a firmer hold on tree trunks, and our feet don't. All of these anatomical features of Lucy's arms, legs, fingers and feet facilitate tree-climbing.

In other words, early hominids such as Lucy might engage in bipedal locomotion on African savannahs, but when pursued by predators, they escaped by climbing trees. Most likely, they also nested in trees during the night for safety, like chimps, bonobos and orangutans do nowadays.

At this point I need to interrupt my explanation of the evolutionary landmark of 'permanent grounding' by injecting a brief description of the ecological background of hominid evolution in Africa.

3.1 Ecological background of hominid evolution in Africa

Recall that when our hominid ancestors descended from the trees six to seven million years ago, they changed their habitat from the canopy of the forests to the savannas. This change of habitat was fraught with danger.

In the canopy of the forest, a good-sized ape faces almost no predatory pressure. The African savanna of six million years ago, however, presented a completely different scenario. There, large and powerful predators such as saber-tooth tigers, lions, hyenas, leopards abound. In addition, these powerful predators had been honing their hunting skills against their preys' defensive capability in an evolutionary arm race for many million years before the arrival of hominids. As new denizens who were not swift, not large and not overwhelmingly powerful, the early hominids which descended from trees became sitting ducks, or choice snacks, for large predators.

In fact, it is remarkable that early hominids were not hunted to extinction by the great predators of the African savannahs.

When we examine fossil records, we can see two very distinct evolutionary trends in hominid development. One trend leads toward a larger and stronger hominid in order to combat predators. The other trend results in a gracile and ultimately more intelligent hominid capable of outwitting faster and more powerful predators. Given that hominids were late-comers to the evolutionary arm race between prey and predators on the Africa savannah, it is not surprising that the large and robust line of hominid evolution headed toward extinction. The gracile line succeeded because those hominids evolved a more effective way of coping with predators by forming large social groups. In other words, the key to survival of the early hominids is collective power.

3.2 Collective power presupposes cooperation and intelligence

In a social group, an individual must calculate incessantly in order to arrive at a reasonable balance between altruistic and selfish behavior in terms of benefit and cost. As hominids went down the path of forming larger and larger social groups, each individual needed greater and greater cognitive power in terms of memory, calculation, assessment, manipulation and decision-making to establish, maintain, terminate or modify his/her social relations with others in order to derive sufficient benefits at a reasonable cost.

An English anthropologist, Robin Dunbar, plotted a graph correlating relative brain size and the size of social groups among modern primates. His data show a direct correlation between the two. That is, the larger the social group, the bigger the brain in proportion to body weight.

Of course, the larger the social group, the greater is the evolutionary pressure for a more efficient means of communication, because establishing, maintaining, modifying and terminating social relations demand an efficient and effective vehicle of communication. For example, the primary communication tool for establishing, repairing and reinforcing social bonds among primates involves mutual grooming, and mutual grooming takes up a significant amount of time of a social primate, approximately one quarter to one third of a social primate's day. Robin Dunbar pointed out, as the size of a social group of hominids grew larger than those of other social primates, a more efficient and effective vehicle of communication other than grooming became imperative.

But a more efficient and effective vehicle of communication requires the evolution of a larger and more sophisticated brain — a mutually reinforcing co-evolution, and a bigger brain begets a bigger skull. A big skull, however, presents an issue in parturition because of the size of the female pelvic opening. From the Australopithecine, Lucy, to modern human, the brain roughly tripled in size. This is an astonishing development in three and half million years of evolution, even if we take into consideration the increase

in brain size attributable to the increase in body size from Lucy to *Homo ergaster*.⁵ But the female pelvic opening did not even double in size during the same period of time.

Thus, in the co-evolution of the hominid brain and communication, the female pelvic opening imposes an upper limit on the size of the skull a fetus can develop. This limit is far below the size of the brain of a modern human. In order to circumvent this limit imposed by the pelvic opening, hominid evolved to give birth prematurely so that the brain can develop after birth.

3.3 The permanent grounding of hominids

Premature birth means the infant is born long before it has developed a brain for motor skills and the ability to integrate a full panel of sensory perception. It also means a prolonged childhood. In contrast to human infants, for example, a new born great ape can hold onto its mother within a few hours after birth.

Premature birth renders a new-born infant helpless and totally dependent on the mother for several years. Specifically, a nursing mother moving from one place to another must carry her infant offspring with her. Even after weaning, a long childhood to allow the development of the large brain ensures a child's total dependency on his parents for travel and survival. Carrying an infant or a young child makes it impossible for a hominid to climb trees. Whereas premature birth may circumvent the limit of the skull size imposed by pelvic opening, it deprives hominids of the possibility of seeking refuge in trees. Thus began the permanent grounding of hominids.

The time line for this landmark in hominid evolution is around 2.2 million years ago, when the size of the hominid brain reached 700-800 cubic centimeters, as indicated by the fossil skull of the Turkana boy, a *Homo ergaster*. The relatively complete fossil skeleton of the Turkana boy also shows that he no longer possesses the anatomical structure that facilitates tree-climbing.

As soon as hominids were permanently grounded, they embarked irrevocably onto the evolutionary path of continuously increasing the size of their social group and enhancing their social organization. It was the only option for them to gain an upper hand in their competition against other animals: predators to defend against and preys to feed on. The need for a larger and better organized social group, in turn, drove the evolution of a more effective and efficient vehicle of communication. Now we have a

⁵ *Homo ergaster* emerged approximately 2.2 million years ago. There is some controversy with respect to the distinction between *Homo ergaster* and *Homo erectus*. But the important point here is that by 2.2 million years ago, hominids reached the physical size of modern humans. Prior to that point in time, all species of the gracile line of Australopithecines are smaller than modern humans. Lucy, for example, weighed no more than 40-50 lb.

powerful feedback loop in hominid evolution, because a more effective and efficient vehicle of communication, in turn, requires a bigger and more complex brain. In other words, ‘permanent grounding’ launches hominid evolution onto the fast lane constituted by two mutually reinforcing co-evolutionary processes: Developing a bigger brain and developing a better vehicle of communication. From that point on, the emergence of language became inevitable.

3.4 Now we come to the fourth landmark in hominid evolution

If we observe animal behavior, even casually, we will notice that animals are, by and large, prisoners of their immediate present in both time and space. By prisoners of the immediate present, I mean an animal in its natural world is bound to respond to a stimulus at the time and the place the stimulus occurs. A trivial example is that if you place food in front of a hungry animal in the absence of any other interfering stimulus, the animal will have to eat it then and there. There can be no waiting, no delay, no consideration of other factors that are not in that place and at that time.

When the brain developed to a point where hominids were completely freed from obligatorily responding to their immediate environment in time and space, they reached a critically important landmark in their evolution, because that freedom makes planning possible, and planning is the all-important prerequisite for the emergence of language and ultimately, civilization.

One cannot make a sophisticated tool without planning deeply ahead in the sense of visualizing and conceiving the tool before producing it. Making tools in multiple steps is, of course, the beginning of hominid culture that has separated hominids from the rest of the animal kingdom forever.

The ability of cogitating beyond the immediate present in time and space is, no doubt, an immense gift from the evolving brain. It opens the door to human creativity.

The part of the brain that mediates displaced thought and behavior resides primarily in the prefrontal lobe, which is also known as the Executive Brain. The non-human great apes have a miniscule prefrontal lobe in comparison to humans. Even Neanderthals didn’t have much of a prefrontal lobe although the overall size of their brain is slightly bigger than the modern human brain. Hominid fossil records indicate that the prefrontal lobe grew disproportionately during the last two million years of hominid evolution.

I should note that while freedom from the immediate present in time and space may be a great evolutionary gift to hominids, it is also a curse.

Once hominids became capable of planning for a complex task and being aware of the future vis-à-vis the present and the past, it was inevitable that they would come to realize that every life ends at some point in time. That awareness probably marks the

beginning of the human anguish which led to the invention of rituals, beliefs and religions in order to ease the gnawing fear of death. We will see later that these rituals, beliefs and religious practices have left us with archaeological evidence that provides a window for inferring the evolutionary development of hominid cognition.

3.5 How do we infer a time frame for the emergence of this important landmark in hominid evolution which I have called freedom from the immediate present?

Hominid stone tools provide us with the evidence for making the inference. The earliest hominid stone tools, called the Olduvie tools, are made about 2.3 million years ago by breaking a good-sized pebble. The sharp edge of the rock splinters allowed the hominids to cut meat and hack at objects. One could argue that the Olduvie tools signal a small step toward planning, similar to the effort of contemporary chimps who strip twigs into a pliable rod to fish for termites. Since Jane Goodall's description of chimp's making termite fishing sticks, primatologists have discovered several other fascinating tool-making behaviors among great apes in nature. But none involves more than one step of shaping or modifying an object that is available in nature. In other words, even if some animals display some ability to plan ahead, the level of this ability is very shallow.

By about 1.8 million years ago, hominid manufactured bi-faced stone tools in what's called the Achulean industry. To make a bi-faced stone tool, the manufacturer has to chip repeatedly and judiciously at the two opposite surfaces of a suitable stone until it becomes a symmetrical, sharp and pointed axe or a cutting tool. To make such an instrument, one has to conceptualize the end product ahead of time and figure out the appropriate knapping process. In other words, it involves multiple step planning.

Since the most primitive hominid tool emerged about 2.3 million years ago, and the oldest bi-faced axes manufactured by *Homo Erectus* is about 1.8 million years old, I place the evolutionary landmark of freedom from the present at about two million years before the present.

The relevance of this landmark in cognitive evolution to the origin of language is obvious. Spoken language is predicated on our ability to think ahead, to weave together different pieces of information in order to achieve communicative felicity at the discourse level. Before hominids were capable of making bi-faced stone tools, it is unlikely that they communicated with language. They might have invented free-standing symbolic signals (which I will clarify later) prior to the appearance of Achulean tools, but they couldn't have strung the signals together to tell a story.

3.6 The mitochondria genome and the Y-chromosome genes

The final landmark of hominid evolution that is relevant to the origin of language comes from the studies of the mitochondria genome and the Y-chromosome genes. The mitochondria genome is transmitted maternally and the Y-chromosome, transmitted paternally. The exclusively uni-sexual lineage of these genes rules out the possibility of random gene mixing in sexual reproduction. By comparing the mutations at each of these two genetic sites among different populations around the world, human geneticists are able to establish that all modern humans descended from a small band of people in east Africa approximately fifty to sixty thousand years ago.⁶ Members of this small band of Africans crossed the outlet of the Red Sea to the Indian Ocean and landed on the Arabian peninsula in several waves of emigration. Their descendants ultimately populated and migrated to every corner of the world. Crossing the open sea, in this case, the Red Sea, requires a high level of technological and organizational skills that could not be achieved without language. We can, therefore, safely assume that language already existed fifty to sixty thousand years ago.

On the basis of the preceding four landmarks in hominid evolution, we can claim that language has been around for at least sixty thousand years, but not more than two million years. However, one may shrink this two million year span by assuming that Anatomically Modern Humans, the species to which all living humans belong, possess language from the day of its appearance.⁷ If this assumption is valid, language will have existed for a minimum of one hundred and sixty thousand years because the oldest fossil record of Anatomically Modern Humans, unearthed in Ethiopia, called the Herto skulls, is about 160,000 years old.⁸

In an earlier paper,⁹ I presented several developments in hominid evolution that are relevant to the origin of language. They are:

- the slow down of the ontogenetic developmental clock
- encephalization
- the shrinking of the gastro-intestinal tract

⁶ A clear and concise summary of decades of research in this arena of human genetics is the article, "Human origins: traces of a distant past" by Gary Stix, *Scientific American* 299.1:56-63, July, 2008.

⁷ This assumption, sensible as it may sound, is not based on any fossil evidence.

⁸ The skulls have a cranial capacity of 1450 cc, larger than that of most modern humans. There is disagreement among paleoanthropologists as to whether Herto should be considered a full-fledged anatomically modern human.

⁹ "On the evolutionary origin of language", C. N. Li & J-M. Hombert (2002), in *Mirror Neurons and the Evolution of Brain and Language*, edited by M. Stamenov & V. Gallese.

- the change from a low caloric and protein-scarce diet to a high caloric and protein-rich one
- the enlargement of the vertebral canal
- the descent of the larynx

Each of these evolutionary processes has been elucidated by trail-blazing scholars during the last few decades. Instead of repeating a lengthy discussion of them here, I will briefly highlight their significance to the origin of language.

Development (1) is connected to (2) because slowing down the ontogenetic developmental clock enables the development of a more complex brain shaped by a prolonged period of experience.¹⁰ In other words, slowing down the developmental clock enhances brain plasticity during youth and lengthens the time for learning, a necessity for hominids because of the cognitive demand of an increasingly complex vehicle of communication and a greater amount of knowledge to be transmitted from one generation to the next.

Developments (2), (3) and (4) form a feedback loop, i.e. dietary improvement enables the shrinking of the gastro-intestinal tract which, in turn, makes it possible for encephalization in terms of energy consumption, and encephalization, which leads to greater cognitive capability, allows hominids to acquire nutrient-rich food and thereby improve their diet.¹¹

Development (5) is concomitant with the thickening of the spinal cord, and thickening of the spinal cord implies better innervation of the chest and abdomen which enhances the control of breathing during speech articulation.

The relevance of development (6), the descent of the larynx, to the origin of language, originally claimed by P. Lieberman and E. S. Crelin,¹² has been challenged by a number of studies.¹³ These studies demonstrate convincingly that the absence of a descended larynx does not morphologically handicap the production of fundamental vowels. The following question then arises: What is the primary evolutionary impetus for the descent of the larynx? J-M. Hombert (personal communication) argues that the

¹⁰ An indication of the length of time required by the completion of brain development is that myelination of the axons in the prefrontal lobe of a modern human brain is not complete until around age twenty.

¹¹ A major step in hominid dietary improvement is the harness of fire. Cooked food not only facilitates the absorption of nutrients, but also provides enriched nutrients in many edible items such as yams, potatoes, etc.

¹² P. Lieberman & E. S. Crelin, "On the speech of the Neanderthal man", *Linguistic Inquiry* 2.2:203-222, 1971.

¹³ See, for example, "The potential Neanderthal vowel space was as large as that of modern humans", *Journal of Phonetics* 30.3:465-484, 2002.

descent of the larynx is primarily driven by sexual selection, i.e. to enable males to have a lower voice, sound larger in size and thereby become sexually preferable.¹⁴ But Hombert's argument begs the question why it did not occur earlier in hominid evolution, say, among the Australopithicines or the Neanderthals,¹⁵ when sexual dimorphism was much more pronounced than what is found among *Homo erectus*.

While the primary evolutionary impetus for the descent of the larynx remains murky, it is non-controversial that the descended larynx has the effect of allowing humans to attain a higher level of articulatory acrobatics and thereby, the ability to produce a greater repertoire of speech sounds.

I have summarized what we can deduce from hominid fossils and provided the reasons for concluding that language emerged within the past two million years. I will now suggest the scenario of the evolutionary process leading to the emergence of language on the basis of our understanding of animal communicative behavior and archaeological evidence.

4. Symbolic behavior and origin of language

Archaeological evidence showing burial rituals, decorative beads and other symbolic behavior provides cues for inferring the origin of language, because language is built on symbolic signals.¹⁶ But the notion, 'symbolic signal', is ill-defined because there are different kinds of symbolic signals with different levels of structural and communicative complexity in language.

The simplest symbolic signal should be defined as a communicative signal referring to a concrete object with the condition that its reference is independent of the communicative function it serves. For example, 'dog' is one of the simplest symbolic signals in English and it refers to the domesticated canid whether the word is used for warning, recruitment, assembly, command, description, etc.

The qualifier, 'concrete', is necessary for the simplest symbolic signal because a concrete object is cognitively easier to grasp than an abstract notion. The ability to

¹⁴ The descent of the larynx is much more pronounced in males than in females.

¹⁵ Fossil records do not allow us to infer whether or not the descent of the larynx occurred among earlier hominids such as *Australopithicus*. However, it is well-established that the Neanderthals do not have a descended larynx. If hominids preceding the Neanderthal had a descended larynx, then it could imply that Neanderthals underwent a reversal in their evolution on this issue.

¹⁶ The term 'symbolic signal' is made popular by Terrence Deacon's 1997 book *The Symbolic Species*, although he did not offer a precise definition.

grasp and manipulate concrete notions must precede the ability to grasp and manipulate abstract notions in evolution.

The constraint, “independent of communicative function”, is important because it secures the free-standing referential status of the signal. Free-standing reference is an all important feature because it gives sense to ‘meaning’. Once meaning is established, communicative signals, each of which has a free-standing and functionally independent reference, can be concatenated to form a larger communicative unit — a critical feature of human language.¹⁷ The meaning of this larger communicative unit will then be based on the amalgamation of the references of its component signal.

4.1 Animal communicative behavior

In animal kingdom, communicative signals do not have meaning. There is no animal communicative signal that has free-standing reference which is independent of the communicative function of that signal. Indeed, the vast majority of animal communicative signals are not referential. A dolphin’s identity whistle, for example, does not have a reference just like the timber and quality of a familiar voice of a close kin has no reference, and a dolphin’s identity whistle has the communicative function of allowing a member of its pod to identify the emitter just like the timber and quality of a familiar voice has the function of allowing a close kin to identify the person behind the voice. The whistle itself, like the timber of a human voice, does not have any meaning. The same can be said of scent-marking, a common communicative signal among most vertebrates.

When someone casually speak of ‘the meaning of an animal communicative signal’, s/he is merely interpreting the communicative function of that signal in terms of human language. Since communicative functions cannot be concatenated to form a new function, animal communicative signals are never concatenated to form a larger signal. Concatenation of communicative signals presupposes that each signal has a free-standing reference independent of the communicative function and context.

¹⁷ Similarly, free-standing referential status of a signal, i.e. reference not inseparably associated with the function of the signal, is a prerequisite for all semantic phenomena. For example, metaphor, which is ubiquitous in language, cannot exist without words that have free-standing references. In fact, it is impossible to speak of ‘the meaning’ of an animal communicative signal, precisely because no animal communicative signal has free-standing referential status. If and when an animal communicative signal has a reference, the reference is always inseparable from the particular function of that signal, be it warning, courtship, recognition, distress, assembly, recruitment, etc.

Since animal communicative behavior does not exhibit the simplest symbolic signal,¹⁸ the first step toward language taken by hominids must be the creation of this uniquely human communicative signal. This significant step of innovation had to occur early during the past two million years because the language-related developments in hominid evolution presented earlier suggest that language did not emerge suddenly. There is no immaculate birth of language and there is no sudden genetic mutation that begets language.¹⁹ The emergence of language is a gradual evolutionary process.

The invention of the first symbolic signal may have its origin in gesturing. The fact that Broca's area, which plays an important role in mediating the production of grammatical utterance, is located in the pre-motor cortex lends support to the hypothesis that gesturing might have led to the invention of verbal symbolic signals. Addition support for the hypothesis derives from the observation that higher primates, especially chimps, employ a variety of gestures for communicative purposes, for example, in order to beg for meat, a chimp often extends an open palm toward an ally which has successfully captured a prey, typically a colobus monkey.

4.2 Gesturing

Gesturing as a component of a communicative behavior is one step closer to pointing at an object in order to bring it to the attention of another party, and pointing at an object to bring it to the attention of another party represents a major cognitive leap toward creating a signal symbolizing that object because the action mimics symbolizing the object. Our cousins in the primate world do not have symbolic signals because they have not yet taken the stride of exhibiting the behavior of pointing at an object in order

¹⁸ Vervet monkey warning calls are not symbolic signals because their references to different predators are inseparable from their warning function. In other words, vervet monkeys do not have a communicative signal that is not a warning and refers to a predator.

¹⁹ The *Foxp2* gene has been misrepresented as a "language gene". Discovered in the study of the speech pathology among members of the KE family, *Foxp2* is a Transcription Factor. A Transcription Factor interacts with the regulatory regions of genes and modulates their level of transcription. A Transcription Factor, therefore, plays a significant role in determining the relative abundance of other proteins in a cell, which is critical in the normal development of an organ. The *Foxp2* gene is expressed in the central nervous system, the lungs and other parts of the body. Among the afflicted members of the KE family, the DNA binding domain of the *Foxp2* gene is mutated. This mutation adversely affects the development of the central nervous system in multiple ways. Language disorder in the afflicted ones represents one of the pervasive consequences of this mutation. Thus, *Foxp2* is not any more of a 'language gene' than the homeobox gene, *Lim-1*, whose deletion causes the entire head anterior to the third rhombomere fail to develop, while the rest of the body remains intact.

to bring it to the attention of another primate.²⁰ The primate brain has not evolved the cognitive capability to execute this behavior.

It is my hypotheses that an overwhelming portion of the gradual evolutionary process leading toward the emergence of language during the past two million years involves the creation of symbolic signals, one by one on an evolutionary time scale. Cultural transmission, then, insures the retention of every symbolic communicative signal because of its immense adaptive values. Imagine a band of hominids who share among them the symbolic signal meaning ‘food’. That signal alone would confer a significant advantage in fitness to the band over another one whose members do not yet have that communicative signal, because it would allow the first band to mobilize more speedily to access food.

5. Emergence of grammatical principles

During the first stage of the long evolutionary process leading toward the emergence of language, the symbolic signals are mono-morphemic words each of which refers to a concrete, observable object.

The second stage begins with the quantum leap of creating the first mono-morphemic word referring to an abstract concept such as an action, an activity, an event, or an experience. Continuous encephalization and therefore, improved cognitive capability in hominid evolution makes this quantum leap possible.

When the repertoire of words reaches beyond a thousand, the third stage of the evolutionary process leading toward the emergence of language is set. This stage begins with the concatenation of mono-morphemic words, creating new linguistic units. It is during this stage that grammar emerges.

Extant literature on the origin of language in the linguistic discipline emphasizes grammar as the critical feature of language. Linguists have also produced a significant body of literature elucidating the processes and diachronic pathways through which grammatical structures appear. For our purpose, however, it is important to view grammar within the scheme of evolution and, especially, within the scheme of hominid communication. The following observations help to frame such a perspective:

(1) Grammatical principles are akin to social conventions in the sense that their emergence does not require time on a geological scale, once the concatenation of vocabulary has begun. Here we are talking about time in terms of a few generations, NOT thousands of generation. The vicissitude of articulation and production of speech

²⁰ “Two hypotheses about primate cognition” by Michael Tomasello (2000), in *The Evolution of Cognition*, edited by C. Heyes & L. Huber.

when words are strung together serve as a platform for content words to metamorphose into grammatical markers.

From an evolutionary perspective, the first grammatical principles to emerge are those which govern the concatenation of symbolic signals, i.e. word order. The earliest word order, (Consider, for instance, whether the modifier precedes or follows the modified), emerges serendipitously. Its canonization and spread within a population is accomplished in a few generations, similar to the establishment of a social convention.

The appearance of grammatical particles follows the establishment of the earliest word orders. This is the process of grammaticalization which has been the focus of illuminating research in diachronic linguistics during the past few decades.

Like social conventions, the earliest grammatical principles are not rigid and inviolate. This is confirmed by the study of modern casual spoken language, the genre of language that also draws on extra-linguistic, contextual information to maintain communicative felicity.

(2) A reasonable corpus of vocabulary in the low thousands together with a few word order principles, in the absence of inflection, agreement, derivation, declension and grammatical markers or functor words, can go a long way as a vehicle of communication serving humans in a social group which existed more than one hundred thousand years ago, long before the dawn of agriculture, sedentary society, complex material culture and a burgeoning population. Even though such a vehicle of communication is distinct from and infinitely superior to what we know as animal communication today, we may wonder, as a number of linguists writing on the origin of language have contemplated, “Does it qualify as language?”

It is my contention that beyond a corpus of vocabulary consisting of one or two thousand single words and concatenated words, it is difficult, if not impossible, to ascertain how much grammar is present in a spoken language when anatomically modern human first appeared on the scene approximately two hundred thousand years ago. In the course of the past two hundred thousand years, human beings have evolved to become more diverse in their physical characteristics,²¹ and cultures have grown to become increasingly complex at an ever accelerating pace. There is every reason to believe that human cognition and language, the primary human vehicle of communication, have also evolved during this same span of time. Indeed, archaeological evidence has unveiled a continuous development starting more than two hundred thousand years ago that supports a scenario of incremental evolution of hominid cognitive capability in terms of abstract thinking, deep planning and symbolic behavior. The archaeological evidence of this incremental evolutionary development involves increased ritual, artifact diversity,

²¹ Consider, for example, skin colors and other physical characteristics associated with different races.

standardization of artifact types, blade technology, worked bone and other organic materials, personal ornaments, exploitation of aquatic resources, etc.²²

I have argued that the mainstay of the evolutionary development of language involves the development of a vocabulary during the past two million years, and that, beyond the evolutionary development of a good sized vocabulary,²³ a set of grammatical principles and grammatical markers probably emerged within a few generations during the past two hundred thousand years.

Let us now return to the critical question raised at the beginning of (2): Does a communication system consisting of a few thousands of vocabulary items and a few word order conventions qualify as language?

Consider a basic grammatical phenomenon: parts of speech. From the study of language universals and language typology, we know every existing language has, at a minimum, two parts of speech: noun and verb. Does this language universal oblige us to assume that language, at its origin, has at least two parts of speech: 'Noun' vis-à-vis 'Verb'?

The answer is affirmative if we consider a noun as a word referring to a concrete object and a verb as a word referring to an action or activity. We may call such nouns and verbs, 'proto-nouns' and 'proto-verbs'. However, if we take into account the syntactic criteria for defining parts of speech in terms of distribution, which is the proper grammatical definition of 'noun' and 'verb', the answer is far from clear. For sure, in the evolution of language, the parts of speech, 'noun' and 'verb' with their full-fledged syntactic definitions emerged at some point in time, most likely before the emergence of any other parts of speech. The question then becomes this: Is there any non-arbitrary reason for claiming that prior to the existence of nouns and verbs as distributionally-defined grammatical categories, the communicative behavior of hominids does not qualify as language?

This question can be further generalized into a challenge questioning how much grammar is required of a language from the perspective of evolution:

Is it reasonable to consider a human communicative system consisting of one or two thousand words denoting concrete objects, abstract notions, activities, events, experiences and states, and a few word order principles 'LANGUAGE'?

²² S. McBrearty & A. Brooks, "The revolution that wasn't: a new interpretation of the origin of modern human behavior", *Journal of Human Evolution* 39.5:453-563, 2000.

²³ The exact size of the vocabulary at the time of the emergence of the first grammatical principle remains indeterminate. If nineteenth century pidgin provides a hint, the size will be in the low thousands. A native speaker of a modern language has an estimated vocabulary of forty to fifty thousands items. A vocabulary is a set of expressions a native speaker has committed to memory. An expression, of course, may be a single word or a composite of several words.

This question has its parallel in the ontogenetic studies of language acquisition. During the long process of first language acquisition, is it possible to arrive at a non-arbitrary definition of language so that we can assert that only after, but not before, achieving what is specified in the definition, a child is deemed to have acquired a spoken language? Normally a child is considered fluent in his/her native language by the age of three or four. But such an assertion merely begs another question:

What does 'fluency' entail grammatically?

For instance, does it imply that the child in an English-speaking community must be capable of relativizing the subject, the object and the head noun of a prepositional phrase, or, does it mean that the child must be capable of distinguishing "I'm easy to please" from "I'm eager to please"?

I submit that we do not have a non-arbitrary criterion to carve out a set of grammatical pre-requisites for defining language because the development of grammar, in both phylogeny and ontogeny, is graded, continuous, and NOT discrete. It is not possible to choose a point in this dynamic process and claim that before that point, the communication system is not a language, but only after that point, it qualifies as a language. Indeed, the grammar of a language never ceases to change, contrary to a misconception commonly held among people who are not familiar with linguistics. In other words, at the origin of language, as I have already stated, grammar begins to emerge and evolve, and that process of evolving has never stopped even after the rise of thousands of mutually unintelligible languages with dramatically different grammatical structures.

6. Conclusion

To sum up, I have pointed out the futility of searching for a specific demarcation point in hominid evolution as the origin of language. There are markers, beyond the cranial capacity, such as permanent grounding, freedom from the immediate present and others, which indicate stages of the expansion of human cognitive capability that is relevant to the emergence of language. There are also processes in hominid evolution, such as the lengthening of the developmental clock, the expansion of the vertebral canal and others, which facilitate the emergence of language. I have inferred from fossil evidence that the emergence of a new vehicle of communication, distinct from animal communication, began a little less than two million years ago when *Homo erectus* started to invent symbolic signals with free-standing references independent of the communicative function of the signal. I have argued that the lengthy evolutionary process of a little more than one and half million years leading to the emergence of language is taken up primarily by the creation of a sufficiently large repertoire of

symbolic signals, and that grammatical structures emerges shortly, within a few generations, after hominids begin to concatenate symbolic signals in their speech. Finally, I have shown the futility of searching for an answer to the question: How much grammar does a language need in order to qualify as a language? From an evolutionary perspective, the emergence of grammar is swift, graded and ongoing. This dynamic process defies any attempt to impose on it a non-arbitrary discrete marker as the point of origin of language.

References

- Boë, Louis-Jean, Jean-Louis Heim, Kiyoshi Honda, and Shinji Maeda. 2002. The potential Neandertal vowel space was as large as that of modern humans. *Journal of Phonetics* 30.3:465-484.
- Deacon, Terrence William. 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W. W. Norton.
- Li, Charles N., and Jean-Marie Hombert. 2002. On the evolutionary origin of language. *Mirror Neurons and the Evolution of Brain and Language*, ed. by Maxim I. Stamenov & Vittorio Gallese, 175-205. Amsterdam & Philadelphia: John Benjamins.
- Lieberman, Philip, and Edmund S. Crelin. 1971. On the speech of Neanderthal man. *Linguistic Inquiry* 2.2:203-222.
- McBrearty, Sally, and Alison S. Brooks. 2000. The revolution that wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution* 39.5:453-563.
- Stix, Gary. 2008. Human origins: traces of a distant past. *Scientific American* 299.1:56-63.
- Tomasello, Michael. 2000. Two hypotheses about primate cognition. *The Evolution of Cognition*, ed. by Cecilia M. Heyes & Ludwig Huber, 165-183. Cambridge: MIT Press.

Language Learning and the Brain: An Evolutionary Perspective^{*}

William S-Y. Wang (王士元)
The Chinese University of Hong Kong

In this essay I will discuss some aspects of the phylogeny and the ontogeny of language within an evolutionary perspective. The first part of the essay stresses the heterogeneity of language, a view proposed by Weinreich et al. (1968). In contrast to Generative Grammar, which represents linguistic structures in terms of deep derivations, I submit that these structures can be better studied in terms of shallow memorial processes, cf. Bolinger (1961), and Tomasello (2003). This view of shallow memorial processes is consistent with an observation von Neumann (1958) made when he compared the computer with the brain, and with the recent proposals toward flat structure by Culicover & Jackendoff (2005, Chapter 4). The second part begins with a very brief review of recent developments in cognitive neuroscience, and especially of EEG methods. Several experiments are then discussed which reveal the remarkably rich biological resources for learning that the infant brings to the challenge of language acquisition. These include its ability to imitate facial expressions almost immediately at birth, and to distinguish the phonetic features of its native language from non-native languages, etc. These developments are demystifying language acquisition, and steadily laying a solid empirical foundation upon which our understanding of language ontogeny can build.

Key words: evolution, brain, cognition

1. Phylogeny and ontogeny

In a panoramic survey of the origins of life, two eminent biologists divided the billions of years of life on earth by eight major transitions (Maynard Smith & Szathmáry 1999). They started at the very beginning, with the appearance of isolated molecules that could replicate themselves — the first condition for life surely being the ability to reproduce. After several billion years of evolution, during which life forms on earth made several momentous changes, including the formation of DNA, sexual reproduction, etc.,

^{*} This essay is offered in celebration of Alain Peyraube, a dear friend of several decades, in appreciation for his numerous important contributions to Chinese linguistics. I hope he enjoys the position taken here.

the final stage of life came with the emergence of human language. Indeed, more than anything else, language is the mental instrument that contributes to the unique achievements of our species.

Compared with the other seven biological transitions in evolutionary time, the emergence of language is really an extremely recent affair. According to the best estimates from anthropology and population genetics, modern humans left Africa to colonize the world around 100,000 years ago (Cavalli-Sforza et al. 1988, and Cavalli-Sforza & Feldman 2003). Many behaviors which we consider human-like that clearly require language, such as ritual burials, coordinated hunting, art and music, and sea crossings, came dozens of millennia later.

Yet 100,000 years is much too short a time span for biological evolution to endow us with anything like a “language organ” (Anderson & Lightfoot 2002), when we consider the millions of years it took Nature to fashion much less demanding body parts, such as the heart and the eye. The central issue for understanding language, therefore, can be cast as two related questions on its emergence. Phylogenetically, the question is how our species made the transition in the first place, from having no language to having language; that is, how did language emerge? Ontogenetically, the question is how does the child manage to learn something as complex as a language, so reliably, and seemingly so without effort?

Pinker (1994) advocated the idea of a “language instinct”, which he used as the title of his highly popular book. But as Tomasello (1995) rightly points out in his review of the book, Pinker’s “instinct” is quite distinct from the way the word is usually used. An instinct typically refers to a behavior that is highly stereotyped; yet it is evident that a major hallmark of language is its great diversity in time and in space. Also, an instinct typically refers to a behavior which does not require cultivation. Yet, it is clear that without a linguistic environment, no language can be acquired.

In the same review, Tomasello raises a deeper objection to the book. These are his words:

Some of the materials in this essay are extracted from keynote addresses given at the International Conference on Computer Processing of Oriental Languages (Hong Kong, March 26, 2009), the Conference in Evolutionary Linguistics (Guangzhou, March 30, 2009), and the International Conference on Applied Linguistics and Language Teaching (Taipei, April 16, 2009). I thank the organizers of these conferences for their gracious invitations.

Research at the Language Engineering Laboratory is supported by the Research Grants Council of Hong Kong SAR, and by the Shun Hing Institute of Advanced Engineering of the Chinese University of Hong Kong.

“The problem with the book is that Generative Grammar and its concomitant nativism are presented to the reader as established scientific facts, without even the hint of a hint that there are fierce theoretical and empirical debates currently raging over almost every issue discussed. That many linguists, indeed the majority of linguists, do not believe in a Chomsky-like Universal Grammar is not acknowledged anywhere in the 430 pages of the book.”

Indeed, in its early days the movement of Generative Grammar rolled over the landscape of linguistics in juggernaut fashion. Some attribute this movement to the personality, some to the political times — in any case we must wait for a historian of our field to eventually sort out the actual socio-dynamic elements involved.¹ Nonetheless, the dissenting voices from highly respected sources can no longer be simply ignored, as they have been for several decades.

One of the purposes of this essay is to bring together some of the other voices from both within linguistics and from scholars in other fields who comment cogently on language. Language is not the deeply derivational, computational instrument that some believe it to be, but a rich heterogeneity of massively stored patterns. I will also try to review some of the recent research on the powerful resources for learning that infants bring to the acquisition of language. That these numerous patterns can be acquired so uniformly and so effortlessly is largely due to the availability of these powerful resources which each infant is endowed with even at birth.

Returning for now to the Generative Grammar movement, many criticize it for its sectarian stance, insulating linguistics from healthy interactions with other disciplines, including even the data-rich area of language pedagogy. For example, Philip Lieberman decried the “hermetic disjuncture” the movement created, and started his discussion with this assessment: “*The major ‘contribution’ of generative grammar to cognitive science is negative.*” (2005:289).

From another academic discipline, the well-known neuroscientist Ramachandran compared the ‘nativism’ of Generative Grammar with the divine intervention of Alfred Russell Wallace, which both shocked and saddened Charles Darwin:

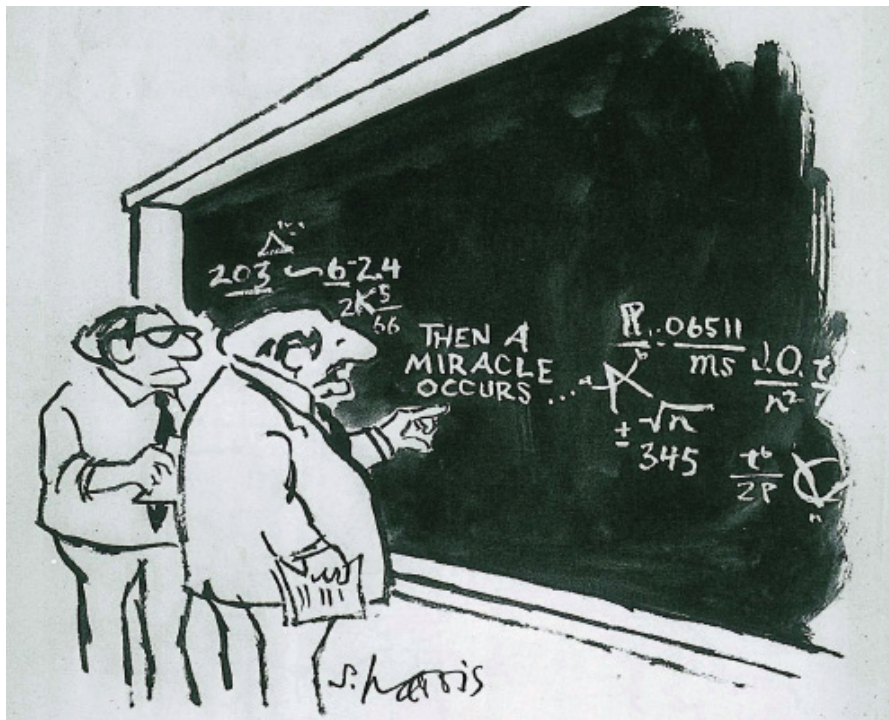
“Alfred Russell Wallace said the mechanism is so complicated it couldn’t have evolved through natural selection at all and must have resulted from divine intervention.... Chomsky said something quite similar, although he didn’t

¹ Sven Öhman, for instance, wrote that “Chomsky is not a scientist at all, but a political ideologue...” (2007:5). Indeed, Chomsky’s meteoric rise in public visibility in the 1960s was clearly due in large part to his political involvement with the Vietnam War.

invoke God.... He almost says it's a miracle. Unfortunately, neither Wallace's nor Chomsky's theory can be tested." V. S. Ramachandran (2004:75).

A cartoon by the humorist Sydney Harris makes the point well; it is reproduced below as Figure 1. In a spirit similar to Ramachandran's observation, my commentary of 1984 on Derek Bickerton's "bioprogram" was titled *organum ex machina* by analogy with the *deus ex machina* used in Greek dramas. Indeed, attributing the human ability to learn language to some implausible 'organ', 'instinct', or 'bioprogram' is not really facing the challenge directly to look for an explanation.

The most detailed analysis of Generative Grammar from the inside is that by Paul Postal, who devotes the better half of a volume to dissect both the reports and the behavior of the reporters, calling his exercise *Studies of Junk Linguistics* (2004:205-338). He recently recapitulated some of his more trenchant arguments in the latest issue of the journal *Biolinguistics* (2009), in an article entitled *The Incoherence of Chomsky's 'Biolinguistic' Ontology*.



No doubt a major weakness of Generative Grammar can be traced to its early ethnocentric bias toward language structures of the European type, which carried over

to the Universal Grammar hypothesis. As late as 1980, given the tremendous wealth of linguistic diversity that has been reported from different parts of the world, it is incredible that anyone can still take a position of ethnocentric hubris expressed by the following words:² “*I have not hesitated to propose a general principle of linguistic structure on the basis of observation of a single language.*” Chomsky (1980:48).

The recent contribution by Evans & Levinson (2009), surveying a wide spectrum of far flung languages and demolishing one putative universal after another, should go a distance toward stemming such hubris. Although some of their conclusions may need to be revised with future work,³ it rightly stresses an empirical aspect of language research that has been slighted too long. Furthermore, even if we had full records of all the 6,000 languages spoken in the world today, these authors remind us, they are still but a small fraction of the totality of languages that have ever been used in the millennia since language emerged. Considering the innumerable languages that have not survived to this day, the total diversity must be much greater than we could ever discover.

Moving past misleading metaphors like ‘organ’, ‘bioprogram’, and ‘instinct’, linguists are now much more open to interaction with other disciplines, realizing what an enormously complex challenge language is. Ray Jackendoff recently criticized the excessively narrow perspective of Generative Grammar, especially for its counter-productive isolation from other disciplines, and for its excessive emphasis on syntax. He spoke for many when he concluded his huge synthetic volume with these final words:

“But linguistics alone cannot sustain the weight of the inquiry. We need all the help we can get from every possible quarter. Above all, my hope for the present work is that it can help to encourage the necessary culture of collaboration.”
Jackendoff (2002:429)

In fact, in a later work reviewing various versions of Generative Grammar over several decades, Culicover & Jackendoff (2005:88) rightly point to several principal failings of this movement:

² Quoted in Evans, Nicholas & Stephen Levinson (2009).

³ The study of linguistic diversity has typology as one of its main tools at present. However, Ernst Mayr urged the replacement of typological thinking with population thinking: “Typological (essentialistic) thinking is misleading when applied to organisms. What must be used instead is population thinking which realizes that in a biological population every individual is unique and differs from all others. The statistical mean value of a population is merely an abstraction. Dual causality as well as the uniqueness of every individual of a biopopulation characterize the world of living beings and are therefore characteristic for biology” (excerpted from his Walter Arndt Lecture). These remarks concerning research methodology certainly merit our reflection.

“First, it does not take the task of the learner to be central; rather, it develops out of a conception of human language as a ‘perfect’ system. Second, it assumes without argument that the correct characterization of linguistic knowledge takes the form of a derivation. Third, it imposes certain notions of economy on its formulations that appear to have little if any empirical motivation. And fourth, it relies heavily on the Uniformity methodology of earlier work.”

Indeed, many linguists have been handicapped too long by the artificial and counter-productive stance of thinking of their discipline as “autonomous,” taking refuge in such simple dichotomies as Saussure’s *langue:parole*, or the *competence:performance* of Generative Grammar. How can language learning, whether by child or by adult, be anything but “central” to a discipline dedicated to language?

The child gathers its language samples from a diversity of sources — from its primary caregiver as well as from other members of the household, and later on from playmates as well as from other adults in school or in the neighborhood. As Gauchat (1905) demonstrated over a century ago, these language samples display a high degree of variation, even within a single family in an isolated Alpine village. We have longitudinal data here as well, thanks to the follow-up study of Hermann (1929). These works and many other similar studies have been put in a modern perspective on linguistic variation by Labov (1994). In order to construct language from these samples, the child does its best to impose some sort of order on this heterogeneity, recalling these terms from the insightful paper by Weinreich, Labov & Herzog (1968). But the end product of this process of construction is very far from “perfect”.

The variation can be readily seen not only in the spoken and written language samples that people produce, but also in their judgments on grammaticality and acceptability. J. R. Ross reports on some experiments carried out with professional linguists, who were asked to make such judgments. The extensive variation he found among these subjects led him to remark:

“The view of language that seems most plausible to me is that the sentences of a language are points in an n-space... An idiolect is a vector in this n-space... And each speaker’s vector, or path, through the space will, I expect, be as individual as his or her face — a linguistic fingerprint.” (Ross 1979:160)

In a related vein, Edward Sapir once said that “all grammars leak.” Natural languages by their very nature are basically different from the clean homogeneity of formal systems, and Procrustean efforts to model them as such have steered us in

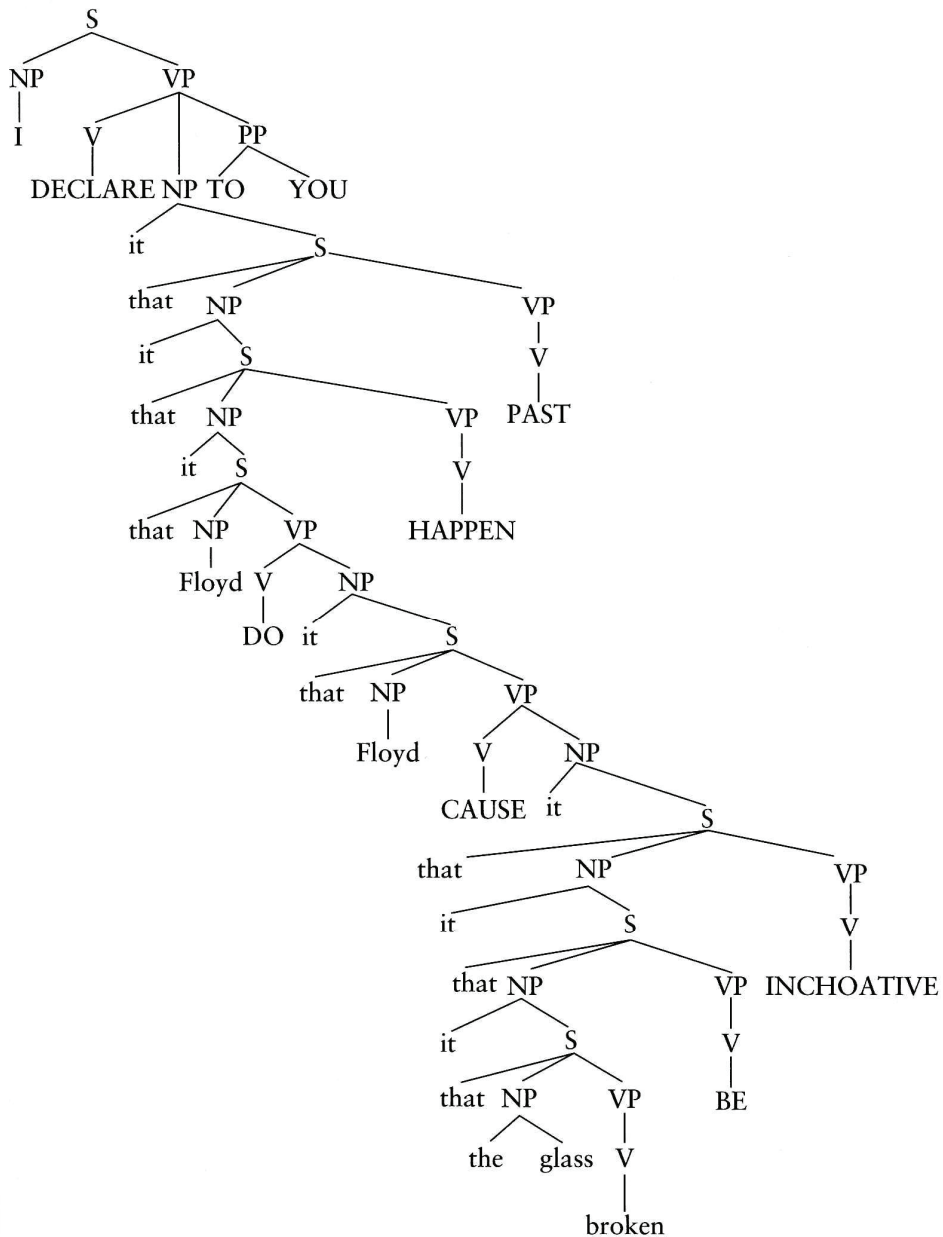
counter-productive directions. In an insightful discussion of evolution, the geneticist François Jacob (1977) wrote:

“Living organisms are historical structures, literally creations of history. They represent not a perfect product of engineering, but a patchwork of odd sets pieced together when and where opportunities arose.”

Languages too are clearly historical structures, both phylogenetically and ontogenetically (Schoenemann & Wang 1996). Their evolution is opportunistic in much the same way, at both the level of our species from generation to generation, and at the level of the child as it pieces together the odd samples of language in the ambient environment to construct a language of its own. What Jacob calls ‘patchwork’ above I have elsewhere called ‘mosaic’, with the same scenario in mind. The situation is similar when an adult adds new words or constructions to his language, driven by fortuitous social needs.

Regarding the second point on derivational complexity made by Culicover & Jackendoff above, it is instructive to examine a sentence that has been discussed for many years in Generative Grammar, reproduced here from Culicover & Jackendoff (2005:99). The sentence is actually a very simple one: *Floyd broke the glass*. However, as can be seen in Figure 2, it is derived by a set of eight sentences marking the tense, aspect, transitivity, inchoativeness, etc., each successively adding to the depth of the derivation.

As a result, a very simple sentence is represented by a monstrously complex structure, reminiscent of Rube Goldberg machines, created for satire rather than function. It is hard to imagine that such a heavy syntactic structure ever comes into play in the mind, either when the child learns the sentence or when anyone uses it in a real communicative situation. The derivational theory which produces such structures has clearly drifted too far away from our goal of understanding language in a realistic way.



In addition to the various arguments advanced by Culicover & Jackendoff for flat structures in syntax, a general point to be made here has to do with an important distinction between how brains and computers process information. A comparison of

the nerve cell with the electronic transistor was first made by John von Neumann (1958), whose pioneering design for computer architecture remains centrally relevant today.⁴ His book has been recently reprinted (2000), with a substantive foreword by Paul and Patricia Churchland.

In terms of current knowledge in these fields, transistors are both much faster and much more accurate than neurons. This fact was dramatically revealed by the mass media in 1997, when IBM's Deep Blue Supercomputer won the world chess championship beating Grandmaster Garry Kasparov. On the other hand, the immense power of the brain resides in its massive number of synapses, several orders of magnitude greater than the number of transistors in any supercomputer. It is the parallel and distributed activities of these innumerable synapses that enable the brain to outperform the computer in many tasks. As the Churchlands put it:

“Conjointly, these two severe limitations — one on speed, and the other on accuracy — drive von Neumann to the conclusion that whatever computational regime the brain is using, it must be one that somehow involves a minimum of what he calls ‘logical depth’. That is, whatever the brain is doing, it cannot be sequentially performing thousands upon thousands of sequentially computational steps... Given the slowness of its neuronal activities, there isn’t enough time for the brain to complete any but the most trivial of computations. And given the low accuracy of its typical representations, it would be computationally incompetent even if it did have enough time.”

Taking the above into consideration, it would seem that the assumptions behind the syntactic analysis in Figure 2, leading to excessive derivational depth, are not justified. Just as natural languages are basically different from formal systems, the parallel processing human brain works very differently from the serial processing computer. In fact, the derivational approach of Generative Grammar, challenged by Culicover & Jackendoff, was questioned much earlier by Dwight Bolinger (1961), who asked:

“Is grammar something where speakers ‘produce’ (i.e., originate) constructions, or where they ‘reach for’ them, from a pre-established inventory, when the occasion presents itself? If the latter, then... constructions are not produced one from another or from a stock of abstract components, but filed side by side, and their interrelationships are not derivative but mnemonic.”

⁴ A more recent comparison of real and artificial neural networks has been made by Crick (1989), who discusses some of the more important differences between them.

Looking at Bolinger's question⁵ from the viewpoint of language acquisition, Tomasello (2003) similarly opts for a memorial approach, as stated in these words:

"This means that in many cases children's comprehension and production of relatively complex utterances are based on a simple retrieval of stored expressions, whereas in other cases they are based on the cutting and pasting together of stored linguistic schemas and constructions of various kinds and degrees of abstraction."

2. Evolutionary perspective

To move forward on our inquiry, we must put language back into an evolutionary framework, connecting it securely to its biological and social roots. The geneticist T. Dobzhansky (1973) once famously said that "*Nothing in biology makes sense except in the light of evolution.*" The basic concepts of evolution, variation and selection, apply no less to linguistics, except for the fact that whereas organisms typically evolve along one tract, language evolves along two tracts — the biological and the cultural.

The major course of how an organism will develop is determined at the instant when the sperm and the egg fuse to produce a new cell, which contains the entire genetic blueprint necessary for the organism to grow biologically. In contrast, there is no such blueprint for language. The child must depend on its biological resources, constantly sample the surrounding world of sights and sounds, people and objects, comforts and frustrations, and construct a language on its own. Since each child has its own unique history of development, the language it constructs will be no less than a *linguistic fingerprint*, in the sense of Ross quoted earlier.

If this ability to construct a language seems miraculous and mysterious, it is because such biological resources have been evolving for millions of years, the power of which is something we are only now beginning to appreciate. This perspective is described vividly in Jacob's famous metaphor that evolution works like a *tinkerer* (1977), giving old parts new uses.⁶ The mechanism is so central to evolution that Gould & Vrba (1982) proposed a new word to highlight it, which is '*exaptation*'. Recently, Ramachandran (2004:76) put language squarely within such an exaptive perspective:

⁵ Bolinger's question has been recently addressed in neuroscience with respect to arithmetic, cf. Grabner et al. (2009). Perhaps similar methods can be used for language.

⁶ Some readers will be familiar with the Tinkertoy, very popular with children before electronic toys. It is discussed by Dewdney (1993).

“I suggest... that it is the fortuitous synergistic combination of a number of mechanisms which evolved for other purposes initially that later became assimilated into the mechanism that we call language.”

The tinkerer metaphor is descriptive and vivid, though I came to know Jacob’s 1977 paper only many years later.⁷ In reaching for a metaphor toward explicating language evolution in a series of lectures I gave in India in 1978, I opted for the word “mosaic”, as in the paragraph below, and also in Wang (2007).

“Many of these abilities are present to different degrees in other animals (witness the instances of tool making and problem solving in chimpanzees). Most of them probably emerged much earlier than language in hominid evolution. Gradually and piece by piece, these abilities were increasingly made accessible for use in the elaboration of language, much as adding pieces to a **mosaic**. In parallel fashion, these abilities have also been made accessible to several other elaborate human institutions, most notably mathematics and music.”

Independently of my usage, Jim Hurford also settled on the mosaic metaphor in a recent paper (2003). The theme that is recurrent across these metaphors, *mosaic*, *tinkerer*, *patchwork*, *cut-and-paste*, is that language emerged fortuitously as a complex adaptive system, not as a new instinct or a new organ (Tzeng & Wang 1983b). Language grew gradually as various domain general abilities were tinkered together.⁸ Clearly, many regions of the brain must participate to make language possible. But in the words of Dick et al. (2001:760),

“... it is no more appropriate to refer to a participating region as a language zone or grammar zone than it would be to refer to the elbow as a ‘tennis organ’.”

It is only natural that we share these domain general abilities to varying extents with other species, especially with the chimpanzee, our closest relative, separated from us some six million years ago. Recent research in primate behavior has revealed a whole series of interesting discoveries here, shedding much light on our homological development. However, as Bolhuis & Wynne (2009) remind us, a lot can be learned as

⁷ I thank Tom Schoenemann for calling my attention to Jacob’s paper, as well as to many other important works in evolution and in neuroscience over the years.

⁸ Over the last two decades there have been numerous computer simulations of language emergence through the random interactions of populations of artificial agents. Gong (2009) provides an overview of these efforts as well as a model of his own.

well from more distant species; birds in particular have evolved remarkable vocal skills by convergent evolution.

Chimpanzees have a complex social structure that permits them to hunt in cooperative groups, mothers teaching skills to the young, such as how to crack nuts, as well as the transmission of various cultural traditions which are specific to regional populations. Comparative studies of the two species of chimpanzees, *Pan paniscus* and *Pan troglodytes* reveal very interesting differences between them, both in group structure and in social behaviors, which are radically different from each other. F. De Waal has produced a series of useful reports of our knowledge in this area (1998, 2005a, and 2005b). Evidence is also accumulating that chimpanzees can plan ahead, such as gathering stones for future throwing, cf. Osvath (2009).

Here is a succinct description by the biologist Christophe Boesch (1991) of an instance of a chimpanzee mother teaching her son the proper way to crack a nut. The observation is all the more valuable since it took place in the wild rather than in a zoo or primate center, where the animal's behavior may be influenced by humans.

“... After successfully opening a nut, Sartre replaced it haphazardly on the anvil in order to gain access to the second kernel. But before he could strike it, Salome took the piece of nut in her hand, cleaned the anvil, and replaced the piece carefully in the correct position. Then, with Salome observing him, he successfully opened it and ate the second kernel.”

Given our close genetic relations with the chimpanzee, such studies provide an important though approximate baseline toward understanding the biological resources the child can have at his disposal toward language acquisition. The recent experiments by Herrmann et al. (2007) are especially instructive in comparing the human child with the chimpanzee and the orangutan, exploring what they call the Cultural Intelligence Hypothesis. They find that whereas all three apes have similar cognitive skills dealing with the physical domain, the human child, at age two and a half, performs much better with those skills dealing with the social domain. We are much better endowed genetically for social cognition than the other apes.

It is important to remember, of course, that six million years of independent development separating us from the chimpanzee can bring about major changes, even given the slow pace at which biological evolution operates. As noted earlier, two million years of separation between the two species of chimpanzees has already produced radical differences in their societies. Our understanding here has increased significantly now that the genomes of both the human and the chimpanzee have been completely sequenced, and powerful algorithms have become available to compare these masses of data to hunt for significant differences between these genomes.

Hidden among these billions of nucleotides are short stretches of human DNA which have mutated at much faster rates, which presumably have been favored by natural selection. One such stretch discovered early in this research, consisting of 118 nucleotides, has been called HAR-1 — human accelerated region 1. This stretch appears to underlie some general abilities which have special implications for the emergence of motor movements which we have adapted to produce speech. An accessible discussion of these exciting new developments is Pollard (2009). Rapid progress is being made in identifying more HARs which greatly enhance our understanding of the uniqueness of our species.

However, in interpreting these findings, it is important to keep in mind that just as there is no ‘language organ’ *per se*, it makes even less sense to speak of a ‘language gene.’ Rather, a balanced statement of where we stand at present in our understanding of the problem is the following by Dick et al. (2004:226):

“The most parsimonious account of language evolution is one where incremental, quantitative changes in primates’ vocal tract, fiber pathways, and neural anatomy converge with social and cultural developments. From this convergence arises the framework upon which the complex language skills could build.”

3. Language in the brain

Although language disorders have been observed since antiquity, cumulative interest in the relation between language and the brain began in the 18th century, when various researchers observed the special role the left hemisphere of the brain seems to play. Injured people who are paralyzed on the left side of the body often keep their language intact, while paralysis on the right side of the body typically is accompanied by language impairment. Since it was known that the right side of the body is controlled by the left hemisphere, the inference was made that an injury there could be responsible for the language impairment as well.

Systematic study of language and the injured brain began in the middle of the 19th century, with the French scholar, Paul Broca (1824-1880), who has given his name to a type of language disorder, Broca’s Aphasia. Patients with this disorder typically exhibit difficulties in producing fluent speech. Broca identified a region in the frontal lobe in the left hemisphere which is correlated with this syndrome. Broca’s discovery was soon followed by a German scholar, Carl Wernicke (1848-1904), who identified a language region in the temporal lobe in the left hemisphere. Patients with Wernicke’s Aphasia can speak fluently with normal rhythm and intonation. However, the speech they

produce is often semantically meaningless, with non-words mixed in with words, resulting in what has been called ‘verbal salad.’

Although these two discoveries were of fundamental importance, the science that they needed for a foundation was not yet available. The realization by the great Spanish anatomist, Santiago R. Cajal (1852-1934), that the human brain was made up of many billions of unconnected neurons was to come only at the very end of the 19th century. Nonetheless, their influence has been pervasive, especially in clinical work in helping aphasics.

The Harvard neurologist, Norman Geschwind (1926-1984), enhanced their influence by high-lighting a fiber tract connecting these two areas in the left hemisphere, called the Arcuate Fasciculus because of its bend around the lateral sulcus that separates the frontal lobe from the temporal lobe. Although nowadays neuroscientists are much more aware of many other parts of the brain involved in the use of language, cortical as well as subcortical, there is little disagreement that these three regions constitute a ‘core’ area.

The last quarter of the 20th century saw an explosive growth of interest in the brain, together with revolutionary advances in the technology of studying it non-invasively in the normal brain. No doubt this progress was much facilitated by George Bush, in a Presidential Proclamation (#6158) on July 17, 1990, which named the period 1990-2000, Decade of the Brain to stimulate research funding. We can now measure the movement of blood and changes in electromagnetism in the head while the subject is doing various tasks, as well as trace the fiber pathways from region to region by a method called Diffusion Tensor Imaging. A recent investigation by Marco Catani et al. (2005) using the latter method shows the nerve fibers of this core area of language in vivid colors (see their Figure 3), thus giving the Broca-Wernicke-Geschwind model a solid empirical interpretation.

Amid this wealth of technology for imaging the brain’s activities, including MRI, PET, MEG, and EEG, some are better at spatial resolution while others are better at temporal resolution. There is now a sizeable literature on the application of each of these methods in language research.

Let us first consider an early MRI study on bilingualism by Kim and his group in New York (1997). MRI experiments can yield very precise spatial information on which area of the brain is active during a specified event. Theirs is a functional MRI study on 12 bilingual subjects, involving a spectrum of languages. Six of the subjects grew up bilingual, and six subjects learned their foreign language as young adults. While in the MRI machine, each subject was asked to describe some events silently by internal speech, in one of his two languages.

Two interesting and closely related results emerged from their experiment that can be roughly summarized as follows. One is that approximately the same posterior region

of the left hemisphere is active regardless of the language involved, and regardless of the age of language acquisition. This region is assumed to correspond to Wernicke's area. The other result has to do with an anterior region of the left hemisphere, assumed to correspond to Broca's area. It turns out the age of acquisition makes a critical difference here. The subjects who are bilinguals from childhood also activate approximately the same region, regardless of the language in which they were describing the events while in the machine. On the other hand, the bilinguals who learned their other language as young adults activated separate and distinct portions of the anterior region for the two languages.

These results may be interpreted as saying that the semantic structures have a lot in common across various languages, and they draw upon the same neural resource. On the other hand, the phonetic forms of the words across languages involve motor skills the acquisition of which is much more age sensitive. This reminds us of authors who can write eloquently no matter how late they have mastered a foreign language. But they give away their native language the minute they start to speak.

Perhaps the best known illustration of this dissociation is the famous Polish author Joseph Conrad, who did not speak English until his twenties. He wrote elegant and perceptive English, as in his novel *Heart of Darkness*, but always spoke with a strong Polish accent. In my personal experience, the most vivid example of this dissociation of semantics from phonetics was the great linguist Roman Jakobson, who was once introduced at the University of Michigan as being able to speak twelve languages fluently, "all in Russian."⁹

A more recent MRI study I will review here, by Siok et al. (2009), concerns the famous Sapir-Whorf hypothesis, which speculates that our behavior is influenced by the language we speak, even when language is not overtly involved, cf. Kay & Kempton (1984). In a series of studies by Paul Kay and his associates over the last several years, this hypothesis has been examined from the viewpoint of color perception. These studies make use of the fact that our visual field can be divided into the left visual field, which projects more strongly to the right hemisphere of the brain (RH), and the right visual field, which projects more strongly to the left hemisphere (LH). By appropriate experimental conditions, such as centering a fixation point to minimize eye movement and presenting visual stimuli very briefly on the screen, we can control which hemisphere will receive the stimuli first.

The subjects are presented with 12 color patches, arranged in a ring around the fixation point. Eleven of these patches are identical in color; they are called 'distracters';

⁹ The story of Jakobson's linguistic prowess and of his strong Russian accent is recorded for posterity in Lamb (1999:41). Lamb's volume also provides a good discussion of the cognitive neuroscience of language from the viewpoint of a linguist.

one patch has a different color and may appear in any position in the ring, it is called the ‘target.’ The task is for the subject to focus on the fixation point and report as quickly as possible whether the target appeared on the left side of the ring, or on the right side of the ring. A key element in this experiment lies in how the target color differs from the distracter color — whether the difference crosses a linguistic boundary. Physically, the two colors in the experiment are always the same distance apart in terms of their wave lengths. However, they may both fall within the same linguistic category, for instance ‘*green*’. Or, they may fall in different linguistic categories, for instance ‘*green*’ for one and ‘*blue*’ for the other.¹⁰ The purpose is to examine how linguistic categories, not overly involved in the task, influence the perceptual judgment, and to find the brain regions which reflect this influence.

Brain activities were monitored by event-related functional magnetic resonance imaging (fMRI) during the experiment. Consistent with earlier studies on color perception, it is found that LH activities are stronger and initiate earlier when the subject responds to color differences if they fall across linguistic categories, i.e., when different color names are involved. More interestingly, activities in the V2/3 area of the brain, which is known for color perception, coincide in time with those of the temporoparietal language region in the LH.

This synchronization between two separate brain regions provides the empirical basis for confirming the Sapir-Whorf hypothesis in color perception, where linguistic categories play a facilitating role even when language is not overtly involved while subjects perform their assigned task. We may hope that in the years to come, more and more distinct brain regions will be identified in their operation to support various linguistic processes. These are the distinct pieces that evolution has tinkered together to make the language mosaic.

4. Language and brain waves

I will now turn to another technology, the EEG, or **E**lectro-**E**ncephalo-**G**raphy. Specifically I will focus on the recent use of the EEG to probe into various distinctive aspects of language acquisition in the case of very young children.

One important advantage of EEG is its fine temporal resolution — on the order of just milliseconds following the brain’s activity. The brain is monitored simply by an

¹⁰ It is well-known that languages differ in how they divide up the color spectrum. For example, in contrast to English, Japanese does not use separate lexical labels for ‘*green*’ and ‘*blue*’, whereas Russian has two separate lexical labels for different shades of ‘*blue*’. Winawer et al. (2007) report on color perception in the Russian case.

ensemble of electrodes attached to the scalp; hence another advantage is that the method is completely non-invasive and free from problems of radiation associated with some other methods. The subject can move about relatively freely during an experiment, and can carry the EEG machinery constantly with him.

Such is the case, for instance, in the growing field of research on Brain Computer Interface - BCI, where the paraplegic patient controls the movement of his wheelchair by his brain waves. The method is also especially useful for young children who will not stay still for any length of time. Also, the subject is not required to perform any prescribed judgments or motor responses, which are often not possible for prelinguistic children or various types of apes.

Brain waves were first noted by a German psychiatrist, Hans Berger (1873-1941), some two decades after Cajal discovered the neuron. His 1929 paper showed the alpha wave, with a frequency within the range of 8 to 12 Hz; many more brain waves have been discovered since, covering a wide range of frequencies. We now know that these oscillations in electric potential picked up at the scalp are due to the electro-chemical signals the neurons send to each other when they communicate across synapses. In particular, when certain groups of neurons are aligned in parallel and fire synchronously, they will produce a more stable EEG measurement.

There are trillions of synapses in the human brain, with a numerous quantity active at any given time, reporting different sensory stimuli, processing a multitude of diverse information, giving commands to various muscles, etc. The waves we pick up with the electrodes attached to the scalp are a jumbled up summation of all these neural activities, filtered through the intervening membranes covering the brain and the bony skull which encase them. It is a monumental challenge to decode these jumbled waves, and interpret them according to the particular linguistic task the subject is performing at that time. With the help of powerful computational equipment and sophisticated statistical methods, we are beginning to make some headway in face of this challenge.

A standard method of research in EEG is to mark the time of an event with respect to the brain waves which are associated with it. As this event is repeated a large number of times, the waves of electric potential associated with the event will summate, while other incidental unrelated waves will tend to cancel each other out. The result is a wave form with distinctive properties of polarity, latency, and shape, which are presumed to be specifically related to the event. Such waves are called ERPs, or Event Related Potentials. Thus the P300 is an ERP that represents a positive potential that peaks around 300 ms after the event, and the N400 is an ERP which is negative and peaks around 400 ms after the event.¹¹ These two are among the handful of ERPs which have

¹¹ The abbreviation 'ms' stands for 'millisecond(s)'.

attracted a great deal of attention in the past decades. A recent analysis of the P300 is Polich (2007), and of the N400 is Lau et al. (2008). An overview of EEG research in language is Kutas et al. (2007).

I was first exposed to ERP research on language early in the 1980s, through collaboration with the lab members of Emanuel Donchin, a pioneer in electrophysiological studies, who was then at the University of Illinois. See his historical note on EEG in Donchin (2006). We were inspired by exceptional irregularity of spelling in English, caused largely by the large influx of French words due to the Norman Conquest. Our stimuli consisted of four lists of word pairs. The Ro words rhyme, and are written alike, such as *plea* and *flea*. The R words also rhyme, but their spelling is quite different, such as *make* and *ache*. The Wo words are written as though they should rhyme but they do not, such as *said* and *paid*; as in this example, the Wo words differ only in the initial consonant. Lastly, there is the control list of words which are not related either in rhyme or in written form, such as *mind* and *wall*. See Polich et al. (1983).

These lists of word pairs were presented visually under two conditions. In one condition the pair of words were separated by a very brief ISI (**I**nter**S**timulus **I**nterval) of 50 ms. In the other condition, they were separated by an ISI of 600 ms. The subjects performed two tasks. Both the reaction times of their responses and ERP data were taken during the experiment. Since the reaction time requires a motor response, it involves a longer latency time. The reaction times ranges from 600 to 1,000 ms, while the ERP latency ranges narrowly from 530 to 620 ms.

In the rhyme task, respondents judge whether words rhyme or not. To make this judgment, the subject must retrieve the presented visual words from his mental lexicon, access their phonetic forms, and compare them. The Ro words encourage a fast ‘yes’ response since the visual forms already suggest that the words rhyme. For R words, like *make* and *ache*, the visual form is actually interference, urging a ‘no’ response when they do rhyme, a judgment that is called a ‘false negative’. For Wo words, like *said* and *paid* the interference is in the other direction, urging a wrong ‘yes’ response, which is a ‘false positive’. In fact, these are the words that produced the greatest latency in the rhyme task.

In the other task, the visual task, the subjects judge whether or not words have similar spellings. Like in the rhyme task, the fastest latency is with the Ro words; even though no reference to the phonetics was required. The fact that the words rhymed facilitated the judgment of their visual similarity, and the judgment comes earlier than the Wo words. Since the words in both these lists differ only in their first consonant, speech recoding apparently kicked in automatically to speed up the judgment for Ro words.

Another interesting observation can be made on these data with respect to the two settings of the ISI, i.e., 50 ms and 600 ms. This has to do with the consistent greater

latency for the shorter ISI across both tasks and all four lists, on average about 25 ms. The subject can only make a comparative judgment, of course, after the second word is presented. In the case of the 600 ms ISI, we can assume that the first word has been fully retrieved and analyzed from the mental lexicon, and the comparison can be accomplished as soon as the second word is retrieved and analyzed. On the other hand, an ISI of 50 ms may be too brief for the full analysis of the first word, and this process continues into later stages, thus delaying the judgment. Whether this is a correct interpretation must await future experiments.

The brain must process languages written with alphabets differently from those written with other orthographic principles. In written language as well as in spoken language, we must resist the easy temptations of ethnocentrism. In particular, investigations on how the brain processes Chinese writing,¹² with its complex shapes and dual construction of semantic and phonetic components, can tell us a great deal about language in the brain. An overview of research in this area is available in the comprehensive anthology compiled by Ping Li et al. (2006), where a third of the volume is devoted to language and the brain. Rapid progress is being made in this area through the ERP experiments of Chiaying Lee and her group in Taiwan (Lee et al. 2006, 2007, and Lee 2008), and the fMRI experiments of Lihai Tan and his group in Hong Kong (Siok et al. 2004, 2008, 2009, and Tan et al. 2005, 2008). Hopefully, before long we will have a more complete understanding of how the brain supports reading and writing in our species; (cf. Tzeng & Wang 1983a).

Another exciting frontier in applying ERP methods is to provide information on children in assessing their neural development for language by comparing their brain waves with those generated by adults. The working hypothesis is that when the children's brain waves become the same as those of the adult, they will have developed the biological resources for that linguistic task. Valerie Shafer (2007) and her group have been investigating this issue, and provide some rough landmarks on this issue.

“Synaptogenesis is creation of synaptic sites on which axons can connect. Subcortical and primary cortical areas are relatively mature in newborns and reach a peak in synaptogenesis by 3 months of age. Association cortex, such as prefrontal regions, does not reach peak synaptogenesis until around 3 years of age. These peaks are followed by loss of synaptic sites and loss of neurons related to absent or weak connections....” (Shafer et al. 2007:32)

¹² A historical account of literacy in Chinese is Wang et al. (2009).

The general principle here is approximately ontogeny recapitulating phylogeny. We share many subcortical parts of the nervous system with other species, and construction of these parts is relatively complete at birth. It is our cortex that has evolved much more elaborately than that in other species; this takes place much later phylogenetically, as well as ontogenetically.

The primary cortical areas which are responsible for sensori-motor systems are constructed comparatively early. The association cortex, which is active in linking and integrating sensory modalities, is completed much later. This is another region of the cortex that Norman Geschwind focused on, especially the Angular Gyrus, situated close to the junction of three hemispheric lobes — temporal, parietal and occipital. It is of special importance for language, since language is the associator *par excellence* of the various modalities. All languages appear to have words that are shared across modalities, e.g., soft light, soft sound; sharp taste, sharp edge. Primates without language have much greater difficulty making cross-modal associations.

An early discovery in this area is the ability of newborn infants to imitate adults' facial expressions, the youngest surveyed being less than one hour old (Meltzoff & Moore 1977). The vision of infants that young is still quite underdeveloped, and of course they have yet to see themselves in a mirror to understand their body parts. Yet they are able to imitate a gesture made by an intruding face. This issue of imitation has been recently discussed (Meltzoff & Decety 2003) in the context of mirror neurons (Rizzolatti & Craighero 2004, and Arbib 2006) and the theory of mind that every child will develop. On the whole, this is a very exciting and active area of research, and in the years ahead we may expect significant advances in our understanding here.

Closer to speech, the very early ability to structure perception to serve language on the basis of communicative sounds is certainly an important part of the biological resource that a child brings to his task. Similar abilities have been reported for other species as well for species-specific vocalizations, though in humans they are considerably more refined. Kuhl (2004) provides a valuable overview of our knowledge here, including a developmental time-table of how perception relates to production, as well as the child's use of statistical information in locating word boundaries.

Several ERP studies have been reported on the question of when infants begin to detect phonetic contrasts which are found in their native language, as opposed to those found in a foreign language. For one study in Seattle, Rivera-Gaxiola et al. (2005) prepared three sets of syllables which differed in their Voice Onset Times (VOT). Spanish speaking infants have a pre-voiced /d/ with a VOT of -24 ms in contrast with a /t/ with a VOT of +12 ms. On the other hand, a stop consonant with a +12 ms VOT would be heard as a /d/ by English speaking infants. For this latter group, a /t/ is typically strongly aspirated, with a VOT of +46 ms.

The ERP finding of these researchers is that both the English-speaking infants and the Spanish-speaking infants can readily distinguish native contrasts from non-native contrasts at as early as seven-months old. Furthermore, in the same experiment they found an intriguing difference in the 11-month old infants they tested, which were divided into two groups. While both groups maintained the ability to distinguish native and non-native contrasts, the polarity of the ERP in one group was positive, while it was negative in the other group. Such a finding would be obscured if an average ERP were taken across all 11-month olds, as Kuhl (2004) also observes. Given this finding however, it would be of great interest to follow them longitudinally to see how the difference in polarity affected their linguistic abilities.

The Seattle study was followed by an ERP study on a similar issue, reported by a collaborative team consisting of Angela Friederici and her colleagues based in Leipzig and in Paris (2007). The stimuli they constructed were the disyllables which they transcribed as /ba:ba/ and /baba:/, where the colon indicates a long vowel. Their subjects were four-month old infants, still younger than the Seattle study. The German-speaking infants were assumed to feel native with the German pattern of word initial stress, suggested by the long vowel in the first syllable. The same went for the French infants with the French pattern of word final stress, by the long vowel in the second syllable of the stimuli. On the basis of the ERP patterns these stimuli elicited in the two groups of infants, they suggest that neural representations of word forms are already in place in the infant brain at as early as four-months of age.

Working with infants still younger, Chao He et al. (2007) studied the EEG of three groups. They monitored two types of mismatch responses in two-month-olds, three-month-olds, and four-month-olds. One type of response, a left-lateralized positive wave (MMP) is strong in two-month olds, weaker in three-month olds, and insignificant in four-month olds. In contrast, the other type of response, a faster, adult-like negativity (MMN), lateralized to the right hemisphere begins to appear in the two-month olds and becomes earlier and stronger as age increases. The dissociation between these two ERP components is a very interesting finding on the maturational time-tables in infant perception in this early age range. The stimuli these investigators used are changing pitches in piano tones. Hopefully, future studies will reveal brain responses to other age ranges, as well as to other forms of auditory stimuli more closely related to speech sounds.

5. Concluding remarks

As the studies reviewed above show, we are just at the beginning of laying an empirical foundation toward understanding how language learning takes place. Building

upon the legacy of pioneers like Broca, Wernicke and Geschwind, we now have a much more quantitative knowledge of brain functions which make language possible. The experiments are getting increasingly focused on the biological resources which make language possible.

The coverage in this essay is neither balanced nor comprehensive, leaving out many topics that are directly relevant to our concern. Among these are the area of children with Selective Language Impairment (SLI) compared with those of Typical Language Development (TLD), discussed by Shafer et al. (2005), and the area of the language of neurologically intact adults compared with aphasic patients, discussed by Dick et al. (2001). Nonetheless, it is easy to see the impressive progress that has been made over the last decade or so. In large part, this new empirical foundation is due to the convergence of expertise and interest from many disciplines, ranging from developmental psychology, to genetics, to animal behavior, as well as the explosive advances in the cognitive neurosciences.

Paradoxically, linguistics has not been at center stage in the progress that has been made, as one might expect of a discipline whose central focus is on language. This is in part due to the unfortunate “hermetic disjuncture”, in Philip Lieberman’s words, that Generative Grammar has created for several decades between linguistics and other disciplines. A great deal of time and energy has been spent by many linguists in deriving very complex syntactic trees, with little regard for the relevance these diagrams may have for the language user, despite the early doubts raised by eminent scholars, like the late Dwight Bolinger (1907-1992).

We can only hope that in the decades to come, more and more linguists will take the evolutionary perspective to heart, and commit more positively to the inter-disciplinary research on language. Language is a gift that natural selection has endowed our entire species with, and it is the business of linguists to know how it varies among the numerous diverse populations in space and time. Knowledge of such variation is of course an indispensable ingredient for inter-disciplinary research if we are ever to truly understand the nature of this precious gift that is uniquely our own.

References

- Anderson, Stephen R., and David W. Lightfoot. 2002. *The Language Organ: Linguistics as Cognitive Physiology*. Cambridge & New York: Cambridge University Press.
- Arbib, Michael A. (ed.) 2006. *Action to Language via the Mirror Neuron System*. Cambridge & New York: Cambridge University Press.
- Boesch, Christophe. 1991. Teaching among wild chimpanzees. *Animal Behaviour* 41.3: 530-532.
- Boesch, Christophe. 2005. Joint cooperative hunting among wild chimpanzees: taking natural observations seriously. *Behavioral and Brain Sciences* 28.5:692-693.
- Bolhuis, Johan J., and Clive D. L. Wynne. 2009. Can evolution explain how minds work? *Nature* 458:832-833.
- Bolinger, Dwight L. 1961. Syntactic blends and other matters. *Language* 37.3:366-381.
- Call, Josep, and Michael Tomasello. 2008. Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences* 12.5:187-192.
- Catani, Marco, Derek K. Jones, and Dominic H. ffytche. 2005. Perisylvian language networks of the human brain. *Annals of Neurology* 57.1:8-16.
- Cavalli-Sforza, Luigi Luca, and Marcus W. Feldman. 2003. The application of molecular genetic approaches to the study of human evolution. *Nature Genetics Supplement* 33:266-275.
- Cavalli-Sforza, Luigi Luca, Alberto Piazza, Paolo Menozzi, and Joanna L. Mountain. 1988. Reconstruction of human evolution: bringing together genetic, archeological and linguistic data. *Proceedings of the National Academy of Sciences* 85.16:6002-6006.
- Chomsky, Noam. 1980. On cognitive structures and their development: a reply to Piaget. *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*, ed. by Massimo Piattelli-Palmarini, 35-52. Cambridge: Harvard University Press.
- Churchland, Paul M., and Patricia S. Churchland. 2000. Foreword to 2nd edition of *The Computer and the Brain*. New Haven: Yale Nota Bene.
- Crick, Francis. 1989. The recent excitement about neural networks. *Nature* 337:129-132.
- Culicover, Peter W., and Ray S. Jackendoff. 2005. *Simpler Syntax*. Oxford & New York: Oxford University Press.
- De Waal, F. B. M. 1998. *Chimpanzee Politics: Power and Sex among Apes*. Baltimore: Johns Hopkins University Press.
- De Waal, F. B. M. 2005a. *Our Inner Ape: A Leading Primatologist Explains Why We Are Who We Are*. New York: Riverhead Books.
- De Waal, F. B. M. 2005b. A century of getting to know the chimpanzee. *Nature* 437:56-59.

- Dewdney, Alexander Keewatin. 1993. *The Tinkertoy Computer and Other Machinations*. New York: W. H. Freeman and Company.
- Dick, Frederic, Elizabeth Bates, Beverly Wulfeck, Jennifer Aydelott Utman, Nina Dronkers, and Morton Ann Gernsbacher. 2001. Language deficits, localization, and grammar: evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review* 108.4:759-788.
- Dick, Frederic, Ayse Pinar Saygin, Suzanne Moineau, Jennifer Aydelott, and Elizabeth Bates. 2004. Language in an embodied brain: the role of animal models. *Cortex* 40.1:226-227.
- Dobzhansky, Theodosius. 1973. Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher* 35.3:125-129.
- Donchin, Emanuel. 2006. fMRI: Not the only way to look at the human brain in action. *APS Observer* 19.11.
- Evans, Nicholas, and Stephen C. Levinson. 2009. The myth of language universals: language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.5:429-492.
- Friederici, Angela D., Manuela Friedrich, and Anne Christophe. 2007. Brain responses in 4-month-old infants are already language specific. *Current Biology* 17.14:1208-1211.
- Gauchat, Louis. 1905. L'unité phonétique dans le patois d'une commune. *Aus Romanischen Sprachen und Literaturen: Festschrift Heinrich Morf*, 175-232. Halle: Niemeyer.
- Gong, Tao. 2009. *Computational Simulation in Evolutionary Linguistics: A Study on Language Emergence*. Taipei: Institute of Linguistics, Academia Sinica.
- Gould, Stephen Jay, and Elisabeth S. Vrba. 1982. Exaptation—a missing term in the science of form. *Paleobiology* 8.1:4-15.
- Grabner, Roland H., Daniel Ansari, Karl Koschutnig, Gernot Reishofer, Franz Ebner, and Christa Neuper. 2009. To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia* 47.2: 604-608.
- He, Chao, Lisa Hotson, and Laurel J. Trainor. 2007. Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience* 19.5:878-892.
- Hermann, Eduard. 1929. Lautveränderungen in der Individualsprache einer Mundart. *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen* 11:195-214.
- Herrmann, Esther, Josep Call, María Victoria Hernández-Lloreda, Brian Hare, and Michael Tomasello. 2007. Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317:1360-1366.

- Hurford, James R. 2003. The language mosaic and its evolution. *Language Evolution*, ed. by Morten Christiansen & Simon Kirby, 38-57. Oxford & New York: Oxford University Press.
- Jackendoff, Ray S. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford & New York: Oxford University Press.
- Jacob, François. 1977. Evolution and tinkering. *Science* 196:1161-1166.
- Jacob, François. 1982. *The Possible and the Actual*. Seattle: University of Washington Press.
- Johnson, Jacqueline S., and Elissa L. Newport. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21:60-99.
- Kay, Paul, and Willett Kempton. 1984. What is the Sapir-Whorf hypothesis? *American Anthropologist* 86:65-79.
- Kim, Karl H. S., Norman A. Relkin, Kyoung-Min Lee, and Joy Hirsch. 1997. Distinct cortical areas associated with native and second languages. *Nature* 388:171-174.
- Kuhl, Patricia K. 2004. Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience* 5.11:831-843.
- Kutas, Marta, Kara D. Federmeier, Jenny Staab, and Robert Kluender. 2007. Language. *Handbook of Psychophysiology* (3rd edition), ed. by John T. Cacioppo, Louis G. Tassinary & Gary G. Berntson, 555-580. Cambridge & New York: Cambridge University Press.
- Labov, William. 1994. *Principles of Linguistic Change*, Vol. 1: *Internal Factors*. Oxford: Blackwell.
- Lamb, Sydney M. 1999. *Pathways of the Brain: The Neurocognitive Basis of Language*. Amsterdam & Philadelphia: John Benjamins.
- Lau, Ellen F., Colin Phillips, and David Poeppel. 2008. A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience* 9.12:920-933.
- Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: Wiley.
- Lee, Chia-Ying, Jie-Li Tsai, Yu-Chin Chiu, Ovid J. L. Tzeng, and Daisy L. Hung. 2006. The early extraction of sublexical phonology in reading Chinese pseudocharacters: an event-related potentials study. *Language and Linguistics* 7.3:619-635.
- Lee, Chia-Ying, Jie-Li Tsai, Wen-Hsuan Chan, Chun-Hsien Hsu, Daisy L. Hung, and Ovid J. L. Tzeng. 2007. Temporal dynamics of the consistency effect in reading Chinese: an event-related potentials study. *Neuroreport* 18.2:147-151.
- Lee, Chia-Ying. 2008. Rethinking of the regularity and consistency effects in reading. *Language and Linguistics* 9.1:177-186.
- Li, Ping, Li-Hai Tan, Elizabeth Bates, and Ovid J. L. Tzeng. (eds.) 2006. *The Handbook of East Asian Psycholinguistics*, Vol. 1: *Chinese*. Cambridge & New York: Cambridge University Press.

- Lieberman, Philip. 2005. The pied piper of Cambridge. *The Linguistic Review* 22.2-4: 289-301.
- Maynard Smith, John, and Eörs Szathmáry. 1999. *The Origins of Life: From the Birth of Life to the Origin of Language*. Oxford & New York: Oxford University Press.
- Meltzoff, Andrew N., and Jean Decety. 2003. What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences* 358:491-500.
- Meltzoff, Andrew N., and M. Keith Moore. 1977. Imitation of facial and manual gestures by human neonates. *Science* 198:75-78.
- Neumann, John von. 1958. *The Computer and the Brain*. New Haven: Yale University Press.
- Öhman, Sven. 2007. *The Essence of Language: A Philosophical Problem: Why Noam Chomsky was never a linguist*. Nora, Sweden: Nya Doxa.
- Osvath, Mathias. 2009. Spontaneous planning for future stone throwing by a male chimpanzee. *Current Biology* 19.5:R190-R191.
- Pinker, Steven. 1994. *The Language Instinct*. New York: Morrow and Company.
- Polich, John. 2007. Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology* 118.10:2128-2148.
- Polich, John, Gregory McCarthy, William S-Y. Wang, and Emanuel Donchin. 1983. When words collide: orthographic and phonological interference during word processing. *Biological Psychology* 16.3-4:155-180.
- Pollard, Katherine S. 2009. What makes us human? *Scientific American* 300:44-49.
- Postal, Paul M. 2004. *Skeptical Linguistic Essays*. Oxford & New York: Oxford University Press.
- Postal, Paul M. 2009. The incoherence of Chomsky's 'Biolinguistic' ontology. *Biolinguistics* 3.1:104-123.
- Ramachandran, Vilayanur S. 2004. *A Brief Tour of Human Consciousness: From Imposter Poodles to Purple Numbers*. New York: Pi Press.
- Rivera-Gaxiola, Maritza, Juan Silva-Pereyra, and Patricia K. Kuhl. 2005. Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science* 8.2:162-172.
- Rizzolatti, Giacomo, and Laila Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27:169-192.
- Ross, John Robert. 1979. Where's English? *Individual Differences in Language Ability and Language Behavior*, ed. by Charles J. Fillmore, Daniel Kempler & William S-Y. Wang, 127-163. New York: Academic Press.
- Sapir, Edward. 1929. The status of linguistics as a science. *Language* 5.4:207-214.

- Schoenemann, P. Thomas, and William S-Y. Wang. 1996. Evolutionary principles and the emergence of syntax—Commentary on Müller: innateness, autonomy, universality. *Behavioral and Brain Sciences* 19.4:646-647.
- Shafer, Valerie L., Mara L. Morr, Hia Datta, Diane Kurtzberg, and Richard G. Schwartz. 2005. Neurophysiological indexes of speech processing deficits in children with specific language impairment. *Journal of Cognitive Neuroscience* 17.7:1168-1180.
- Shafer, Valerie L., and Karen Garrido-Nag. 2007. The neurodevelopmental bases of language. *Blackwell Handbook of Language Development*, ed. by Erika Hoff & Marilyn Shatz, 21-45. Malden & Oxford: Blackwell.
- Siok, Wai Ting, Charles A. Perfetti, Zhen Jin, and Li-Hai Tan. 2004. Biological abnormality of impaired reading is constrained by culture. *Nature* 431:71-76.
- Siok, Wai Ting, Zhendong Niu, Zhen Jin, Charles A. Perfetti, and Li-Hai Tan. 2008. A structural-functional basis for dyslexia in the cortex of Chinese readers. *Proceedings of the National Academy of Sciences* 105.14:5561-5566.
- Siok, Wai Ting, Paul Kay, William S-Y. Wang, Alice H. D. Chan, Lin Chen, Kang-Kwong Luke, and Li-Hai Tan. 2009. Language regions of brain are operative in color perception. *Proceedings of the National Academy of Sciences* 106.20:8140-8145.
- Tan, Li-Hai, Alice H. D. Chan, Paul Kay, Pek-Lan Khong, Lawrance K. C. Yip, and Kang-Kwong Luke. 2008. Language affects patterns of brain activation associated with perceptual decision. *Proceedings of the National Academy of Sciences* 105.10: 4004-4009.
- Tan, Li-Hai, John A. Spinks, Guinevere F. Eden, Charles A. Perfetti, and Wai Ting Siok. 2005. Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences* 102.24:8781-8785.
- Tomasello, Michael. 1995. Language is not an instinct. *Cognitive Development* 10.1: 131-156.
- Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.
- Tzeng, Ovid J. L., and William S-Y. Wang. 1983a. The first two R's: the way different languages reduce speech to script affects how visual information is processed in the brain. *American Scientist* 71.3:238-243.
- Tzeng, Ovid J. L., and William S-Y. Wang. 1983b. Search for a common neuro-cognitive mechanism for language and movements. *American Journal of Physiology* 246: R904-R911.
- Wang, Feng, Yaching Tsai, and William S-Y. Wang. 2009. Chinese literacy. *The Cambridge Handbook of Literacy*, ed. by David R. Olson & Nancy Torrance, 386-417. Cambridge & New York: Cambridge University Press.

- Wang, William S-Y. 1978. Explorations in language evolution. Diamond Jubilee Lecture. Osmania University. First published in 1982 as Supplement to Vol. 8 of *Osmania Papers in Linguistics*.
- Wang, William S-Y. 1984. Organum ex machina. *Behavioral and Brain Sciences* 7.2: 210-211.
- Wang, William S-Y. 2007. The language mosaic and its biological bases. *Journal of Bio-Education* 2.1:8-16.
- Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. *Directions for Historical Linguistics: A Symposium*, ed. by Winfred P. Lehmann & Yakov Malkiel, 95-188. Austin: University of Texas Press.
- Winawer, Jonathan, Nathan Witthoft, Michael C. Frank, Lisa Wu, Alex R. Wade, and Lera Boroditsky. 2007. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences* 104.19:7780-7785.

On Moving in Conceptual Space: Readers' Responses to a Parallel Text in Chinese and English*

Gisela Bruche-Schulz
Hong Kong Baptist University

The question addressed in this paper is how a reader's mind receives a narrative text, in other words, how a reader conceptualizes the experience that is mediated through the semantics of the text. It is suggested that the experimental design used in this paper allows tapping into the conceptual activity accompanying reading and viewing in novel ways. The design differs from, but also shares points of view with approaches used in eye-tracking research and read-aloud studies (Holšánová 2008, Rayner 1998, Pressley & Afflerbach 1995). Two different groups of Hong Kong students have been given an excerpt of Saint-Exupéry's *Le Petit Prince* in its English and Chinese translation, respectively. They were asked to 'glance through the text, underline the words for which something comes to mind, and jot down what it is that comes to mind in the margin of the same page'. They were given 10 minutes for this task. No restrictions in terms of content or quantity were imposed on their responses.

It turned out that the experimental design favours the production of local-view responses, i.e. responses to an element of a clause segment whose core is a verb. Global view, i.e. attention to larger portions of text, occurs, but to a distinctly lesser degree. Responses are of two kinds: overt and covert. They are overt when perceived as the jottings in the margin of a page, they are covert when observed by the pattern of their quantitative distribution across the segments of the text. The overt responses, consisting of verbal and pictorial expressions map the elements cued by the text onto those elements of conceptual space which are of momentary significance to the reader (after Fauconnier 1997). This explanatory background is assumed to apply, but the process is not elaborated in the present paper in any detail.

* The elicitation of the responses that represent the data of the present study was undertaken by the author in January 2004 (in English) and, in the context of a faculty research grant of Hong Kong Baptist University, by Maggie L. F. Tsung in June 2004 (in Chinese). In this very context, further elicitation events were conducted during the summer of 2004, with not only narrative, but also expository and hortatory text types in focus as well. I am grateful for the support of the following research assistants: Clara C. Y. Hui, Heather Y. C. Lei, and Maggie L. F. Tsung (Bruche-Schulz 2005). — The final version of the present paper was accepted for publication in June 2009.

In this paper, the main findings are about the covert display of sensitivity to textual structure (Hoey 2001). The latter is revealed through the quantitative patterning of readers' local-view responses that reflect a discerning distribution of attention to the textual means. The distribution of attention to local foci reveals constructions of a *want-to-know* figure within the textual progression. When a *want-to-know* impulse is sensed, readers' attention is kept in suspense, and response numbers are generally low. When suspense is followed by a temporary resolution ('knowing is granted'), response numbers go up. Both in Chinese and English, the quantitative patterning of readers' responses to textual cues allow us to see how the text evokes a thread of subsequent expectations towards the question of 'what comes next'. The phenomenon occurs across two different languages, with different groups of participants. This means that the elicitation procedure should be replicable and yield comparable results across other languages as well.

Key words: covert and overt conceptualization, local and global-view perspectives on a written text, *want-to-know* figure, construal in conceptual space

1. Introduction

The present study has settled for a design that does not ask for any particular kind, type, or quantity of responses. There are two groups of student-readers (in Hong Kong), one at a university, the other one at a secondary school, who are requested by their teachers to read quickly through a text on an A-4 page in front of them, and then to 'underline the words for which something comes to mind, and jot down what it is that comes to mind in the margin of the same page'. They are not told why they are asked to do this, are not asked to reveal their names either, but are promised a full explanation once their responses have been reviewed.

There are no constraints in regard to 'what' and 'how' to respond. (The design of the study borrows this open-endedness from the participant-observer method in ethnographic interviews.) The constraints which do exist come from three directions: (a) from the medium in which the students are asked to respond, namely, with a pen on paper, (b) from the restriction on the time allowed for completing the task, and (c) from the specifics of the task. The latter makes use of the main purpose that motivates the reading of narrative texts. Narrative texts invite the mind's wandering through "fictional woods".¹

¹ Other possible purposes: to check the text for spelling errors; to use it as a piece of evidence; to judge it for its style; to use it for the teaching of a language and so forth. The present paper does not deal with such purposes. — Neither do I discuss the question of the narrative being a 'mode' of discourse rather than a genre.

Even when there are no well-trodden paths in a wood, everyone can trace his or her own path, deciding to go to the left or the right of a certain tree and making a choice at every tree encountered. (Eco 1994:6)

Every lexical trigger may cause a reader's mind to enter conceptual space, and place attentional focus on those regions of the experiential frames therein that are of current relevance within the concrete spatio-temporal moment of reading. The mind, in an unending flow of attention, moves from a momentary starting point (base space) to a next point (focus space), the latter turning into a new starting point which again prepares for a new focus, and so forth (Fauconnier 1997, Fauconnier & Turner 2002). To be sure, the mind's wandering through 'fictional woods' is not without guidance. It follows its own purpose, namely, to get to know 'what comes next'. A base space may convey information about a situation, leaving open an aspect that gives rise to a [*W*] *ant-to-know* impulse. The focus space will respond to the [*W*] *ant* by some temporarily satisfying [TS] information, and will normally give way to a new question. In the words of Hoey:

Whenever we read a sentence it sets up expectations in our mind and those expectations shape our interpretation of what comes next. We do not of course formulate these expectations as [open] questions [...]. But that we do have expectations is demonstrated every time we are surprised or disconcerted by the direction a text takes. (2001:22-23)

The sequence of micro-thought figures [W-TS] is not lexically signalled. The reader construes the [W-TS] sequence through the support of the semantic constellations that effect the signposting in conceptual space. When the semantics of the text prompts the setting up of a new expectation, the readers are taken aback, as it were, and react with low response values. High response values are observed when readers are 'happy' to find temporary answers to a *want-to-know* situation.

In §2, three approaches to reading, each using a different experimental design, are briefly surveyed. In §3, the elicitation technique used in the present study will be discussed as a basically qualitative instrument supported by quantitative information. In this same §3, it will also be demonstrated which response formats the technique allows for. In §4, I will show how the thread of the text develops along [W-TS] sequences of the thought figure that, both in Chinese and English, are accompanied by a distinct quantitative patterning of the readers' responses. It will be shown, how responses are held back at [W] segments, and go up at [TS] segments.

2. Approaches to reading

Approaches to reading, relevant to this paper, have come from three directions. One is represented by the well-established tradition of eye-tracking research (for an overview account, see Rayner 1998), the second by the investigation of reading processes in think-aloud protocols (Pressley & Afflerbach 1995), and the last combines insights into the simultaneity of verbal conceptualizations and attentional focus, expressed by the movements of the eye during picture description (Holšánová 2008). The present paper shares with eye-tracking research the assumption that the inner processes of meaning-making by the mind should have outward expressions, and therefore be open to an investigation. It shares with think-aloud protocols the belief that relevant aspects of the reader's thought process can be self-reported, and it draws on Holšánová's work for organizing the discursive reactions to the text used in the elicitation events by their expressions of global and local foci.

Eye-tracking research explores eye movements that can reliably be tracked by measuring and tracking devices. The latter devices register oculomotor processes that are taken to represent evidence of what is believed to truly express the movement of attention. While such work follows the external traces of mental activity closely, it does not touch upon the conceptually shaped relevance of what the eyes (of readers in laboratories) meet when they move in saccades and engage in fixations. It is true that readers in laboratories are not "interrupted with requests for metacognitive reports" (Richardson et al. 2007:324). But it seems that none of the tasks mentioned in Rayner's survey "of 20 years of research" (1998) explores the discursive-conceptual impact of the text presented to readers in eye-tracking laboratories.

The second direction of research relevant to this paper, think-aloud protocols, request verbal (spoken) reports of what a text 'means' in regard to a meta-perspective assumed to be routinely adopted by competent readers. While this method may have the benefit of producing self-reports that are maximally 'on-line', the interruption of the reading process "with requests for metacognitive reports" (Richardson et al. 2007:324) compromise this potential benefit by putting pressure on the spontaneous attentional work and distorting it through additional attentional burdens, caused by distracting demands that tax the respondents abilities to, e.g. judge style or content of a text, report on strategies of comprehension, and the like.

The third approach avoids the pitfalls of the first two, as I see it. Instead of merely measuring oculomotor processes, or imposing a metacognitive task, it lets people describe a picture verbally, thus staying close to the link of a visually and conceptually processed object in front of a respondent. It maintains a direct connection between perception and conceptualization. Before I present my own experimental design, I will

briefly dwell on the way Holšánová (2008) organizes a picture description.

She combines the viewing of a picture and a verbal report thereof when monitoring the movement of the eyes. She aims at concretely capturing both the activity of the moving eye, and the conceptual activity that is captured by the spoken picture description. True, in Holšánová's experimental design, there *is* the 'bottleneck' of phonetic articulation, which holds back the system of cognitive processing that can "run about four times faster" (Levinson 2000:6). However, it seems that, among the three approaches that are relevant to my own attempt at the matter, hers comes closest to a glimpse into the conceptual activity accompanying viewing — with two components described, the contribution of the eye and of verbal conceptualizations, organized by the inter-relation of global and local foci of attention.

By introducing a combined "verbal focus and visual focus [...]" as 'two windows to the mind'" (2008:6), Holšánová has found ways of describing the mind's ability to handle both verbal and visual, and global and local foci in a given moment together. When words are spoken that describe regions of the picture, the verbal and visual foci may either be synchronized, delayed, or redirected to a super-focus (88-93). The mind holds in consciousness a holistic view on a situation and, at the same time, may focus attention on a smaller part of it. What is more, her findings suggest that eye movement patterns "are not restricted to the process of planning and structuring verbal description of the scene [viewed in the picture] but are rather connected to the scene semantics" (141), observed by distinct synchronized, intertwined, or delayed mappings of visual and verbal activity (106-113). Holšánová demonstrates how eye movements accompany the apperception of semantic constellations. However, she does not deal with the reading of written texts.

The present paper deals with written language, a tool which uses a channel obviously different from that of speaking, but is closely related to viewing. Reading is an activity that is sustained by the deployment of attention like vision and speaking. Reading shares with the two activities described by Holšánová (viewing a scene and describing it verbally) the same cognitive mechanism for the deployment of attention: *a text (space), like a picture (space), may be held in consciousness as a whole, while local foci are marked as active* at a particular moment, with the 'particular moment' determined to occur within the 10 minutes allowed for the response to the request 'to jot down what comes to mind'.

3. Eliciting of responses to a written text

At each elicitation event, students are asked to glance through the text, underline the words for which something comes to mind, and jot down what it is that comes to

mind in the margin of the same page. It is pointed out explicitly that topics, subject matters, and forms of what is jotted down are entirely up to the respondents. No upper or lower limit is set for the response numbers. Responses elicited in this manner are enacted silently by the movements of a pen on paper. The time allowed for the procedure is 10 minutes.

The text is an excerpt of Saint-Exupéry's *Le Petit Prince*, in two parallel versions (Chinese and English).² The English version comprises 316 words, the Chinese version 489 characters. The excerpt is taken from the introductory paragraph of Chapter twenty-five. Saint-Exupéry's *Le Petit Prince* narrates the story of a pilot who meets the little prince in the Sahara after the pilot's plane has crashed there. The text evolves around narrative episodes which are interspersed with reflections on life.

The English-language version of the excerpt was presented to a group of 21 undergraduates of a university in Hong Kong in their third year of studies. The Chinese-language version was read by a group of 17 secondary-school students in Hong Kong, school-leavers who were coming to school for the very last days of the school year. The groups vary by educational level and the usual language(s) spoken by the students. The language in which the excerpt is read is also the (dominant) language of instruction.

3.1 Questions of segmentation

The clauses of the text form the basis for its segmentation. There are altogether 44 segments (s[1-44]), shown in §4.2 underneath Figures 1-6 in full. In a few instances, the sequence of the clauses is different between English and Chinese. There are two differences which should be mentioned.

In Chinese, there is the need for a preposed speech introducer, while it may be postposed in English. This leads to a reversed numbering of the segments in the Chinese version at s[14-15], s[22-23], s[27-28], and s[34-35]. With English here treated as providing the default sequence,³ the difference of sequence is rendered as in example (1) below.

(1) Reversed sequence of clauses

14. "It is strange,	15. 我對小王子說：
15. I said to the little prince	14. 「好奇怪，

² The excerpt of the English version was extracted from the translation by Katherine Woods (1945[2002:76-77]). The Chinese translation from a trilingual edition published in Taiwan, translated by 張譯 Zhang Yi (1999:260-262).

³ By doing so, I do not intend to argue for a particular view of the iconic motivation of the preferred clause sequences.

The second difference concerns the possibility of moving constituents of the direct-speech clause to a position in front of the speech introducing clause in English. Here, Chinese is taken to provide the default order, and the numbering of the segments is arranged accordingly, as in example (2).

(2) Preposing of direct-speech elements

2.a) "Men", 1.) said the little prince,	1. 小王子說：
2.b) "set out on their way in express trains,	2. 「人們只顧著匆匆忙忙擠到快車裡，

The rationale for choosing this format of segmentation is as follows: When processed statistically, the segments in both languages have to be arranged in a way that allows the schematic representation of response numbers together with their correct match of semantically equivalent segments in the other language; see the diagrams in §4.

In two cases, translators have added extra clauses that do not appear in the other language version, and are also not motivated by the French original. The extra clauses have been identified by an asterisk (*), but not been given a numbered segment (as shown in Figures 4 and 5 in §4.2).⁴

The core of a segment is its verb that enunciates an occurrence of an event or state. Since verbs may enunciate occurrences of situations not only in main clauses, but also in modifying relative clauses and non-finite adverbial clauses, the latter are treated as separate segments. Likewise, speech and thought introducing clauses are treated as separate segments since the act of introducing brings about an event of its own.⁵

The responses to the text consist of verbal and pictorial expressions that, given the material means used during the elicitation, are jotted down with a pen on the page placed on a desk in front of the student-participants. Jotted-down responses, from now on sometimes also simply called 'jottings', are then visible to the researcher as words, elaborate drawings, or simple sketches.

⁴ An additional speech-introducing clause is inserted before s[29] in Chinese and, in English, an adverbial clause is added after s[31].

⁵ Support for this view comes from Smith, as I see it. Smith discusses the role of complement clauses as contributing to "narrative advancement". She touches only lightly on direct quotations though (Smith 2003:104-105). I believe, however, that a similar view should be taken on direct quotations.

3.2 About the respondent groups and the format of the jottings

The respondents, who will be referred to only by number, were assured anonymity so as to not jeopardize their readiness to participate in the elicitation events. The numbering of the response sheets follows the order in which the pages were handed back during the events. Group E[nglish] initially consisted of P[ersons](1-21), aged 19-21 on average. Since P18 remarked that s/he could not connect to the story characters, group E was processed with only 20 persons. Group CH[inese] consisted of P[ersons] (1-17), aged 17-19 on average.

Table 1: Summary of information about respondent groups

Groups/ Characteristics	English (E)	Chinese
Number of respondents	20	17
Social profile	Undergraduates in 3 rd year of studies, Hong Kong, January 2004	Secondary-school students, in their final year, Hong Kong, June 2004
Setting	Classroom: Beginning of second semester (Textlinguistics course)	Classroom: End of school year, close to last days of school (subject: Remedial English)
Date of elicitation event	15 January 2004	21 June 2004
Average age	20 – 22	17 – 19

The students used the right margin of the text for jotting down notes, drawings, and sketches. In most instances, they did underline the places in which ‘something came up in the mind’, but they also used brackets, circles or other forms of text marking. The majority of jottings are clearly related to text marked off in the ways just mentioned. I consider these items as instances of a ‘local view’ of the text.

Text has, at times, also been marked globally. The instances of global marking signal that more than one clause, possibly a whole paragraph, or several paragraphs, are kept in view. Such marking may be done by brackets or just by the positioning of a jotting beside the paragraph in view. Jottings may also be placed in the margin without affirming either a local or a global view. I have listed the latter kind of positioning of responses as an UnAffirmedView or UAV marking. Below, examples (3)-(5) show each type of marking.

(3) Local-view text marking (P2E, at s[3b])

"Men," said the little prince, "set out on their way in express trains, but they do not know what they are looking for. Then they rush about, and get excited, and turn round and round . . ."

men look for sth

P2E secures understanding by a paraphrase of what is underlined. I have regarded such responses as instances of an *expert* role that adopts a bird's eye view on the frame that, according to the respondent, has been evoked by the marked-off places of the text. Other instances in this category present commentary containing cultural models ('men are that way, always rushing'), moral judgements ('men are aimless, have forgotten their soul'), or display technical knowledge ('fast trains are equipped with [xy-type] engines').

(4) Global-view text marking (P7E, scope over s[5-33])

And he added :
" It is not worth the trouble . . . "

The well that we had come to was not like the wells of the Sahara. The wells of the Sahara are mere holes dug in the sand. This one was like a well in a village. But there was no village here, and I thought I must be dreaming . . .

" It is strange," I said to the little prince. " Everything is ready for use : the pulley, the bucket, the rope . . . "

He laughed, touched the rope, and set the pulley to working. And the pulley moaned, like an old weathervane which the wind has long since forgotten.

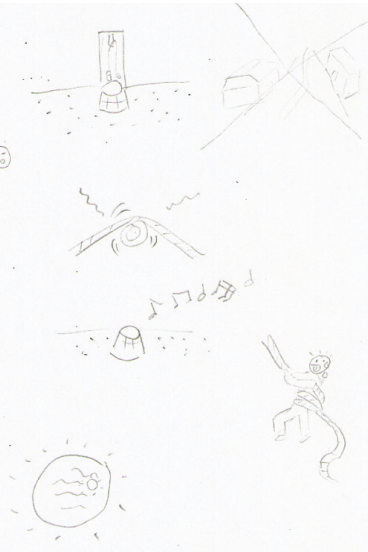
" Do you hear ? " said the little prince. " We have wakened the well, and it is singing . . . "

I did not want him to tire himself with the rope.

" Leave it to me," I said. " It is too heavy for you. "

I hoisted the bucket slowly to the edge of the well and set it there—happy, tired as I was, over my achievement. The song of the pulley was still in my ears, and I could see the sunlight shimmer in the still trembling water.

" I am thirsty for this water," said the little prince. " Give me some of it to drink . . . "



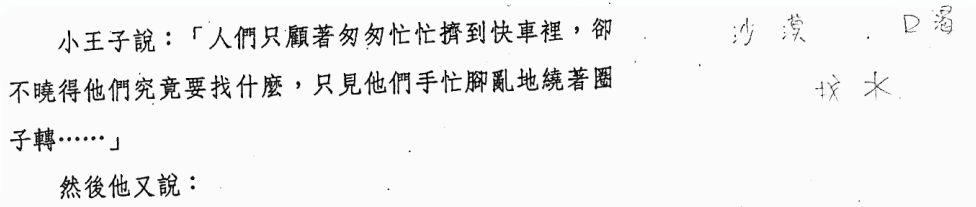
In the example (4), P7E does not mark off text locally. S/he places sketches beside the paragraphs for which something 'comes up in the mind', and the reference to globally marked portions of the text is performed by the position of the jottings.

I have labelled the role adopted by P7E as the *viewer* role. When the authors of such jottings describe objects and spatial spaces of experiential frames by pictorial means they make the observer 'see' what is in a scene or scenario that is triggered by a lexical cue. Even if the pictorial form is minimal, only adumbrating a shape or a space,

such drawings and sketches invite a “symbol formation process in reverse” adding in “large amounts” of sensory information to objects, even if filtered out substantially before being schematized as a graphic symbol (Barsalou 1999:593).

Example (5) below serves two purposes. It exemplifies a *viewer* role that is not expressed pictorially, but verbally, and it illustrates a UAV (UnAffirmedView) jotting that does not in any way allow relating the jotting(s) with a locally or globally marked-off portion of text. Here, the jottings are placed beside a paragraph that describes how people are moving around, aimlessly, in crowded trains (at s[1-5]), but the jotted-down remarks refer forward to the paragraphs that talk about wells in the desert, being thirsty and so on (at s[6-34]). No attempt is made at aligning the jottings with particular local portions of the text.

(5) UAV jotting (P14CH, at s[1-5])



Jotted down notes and drawings are taken to be outward signs of an “inner world” that generates the “spotlight of attention” (Holšánová 2008:83).⁶ Attention may keep the ‘spotlight’ on foregrounded elements for a moment, and then background them as semi-active elements of the next. The design of the elicitation procedure appears to favour the production of local-view responses.

3.3 Methodological considerations

The elicitation technique I have developed is of a novel kind, and the question arises as to whether it is capable of yielding valid results when considering the objective of tapping into the conceptual activity of a reader’s mind. The technique rests on the principle that responses be maximally unconstrained, both in regard to their content and their number. It is thus set up as a qualitative method that resembles a conversational interview rather than a survey or a test. However, when response numbers are correlated with the semantic constellations found in the text, a quantitative parameter emerges as well. Therefore, we deal with a mixed-methods elicitation technique: The presence of a

⁶ Traced by verbal and visual foci during picture viewing in Holšánová (2008).

causal (independent) variable in the form of semantic constellations is met by distinctive correlations between this causal variable and response numbers. I will elaborate the quantitative side of the findings first.

With reference to Table 2 below, I will first argue that the participation rates in the experiments show that participation is comparable across the two different groups. In addition, with reference to Table 2 and Table 3, I argue, that the design of the elicitation procedure not only seems to favour local-view responses across the two different groups, but also yields comparable distributions across the two groups. Therefore, it seems reasonable to assume that external validity holds.

Table 2: Response formats in absolute numbers and average per person

Formats Groups	Local-view (Loc)		Global-view (GI)		Unaffirmed view (UAV)		Total (Loc, GI, UAV)	
	Average per person	Total n°	Average per person	Total n°	Average per person	Total n°	Average per person	Total n°
E (20 part.)	7.4	148	0.3	6	0.6	12	8.3	166
CH (17 part.)	8.4	144	0	0	0.1	2	8.6	146

* Note: all decimal values have been rounded to the nearest tenth.

Table 2 gives evidence of a comparable response rate. There is one difference which warrants mentioning. The university students who responded to the English translation, have used more non-local-view markings than the secondary-school students. However, I do not see this as a difficulty, but rather as a signal for what one will find when replicating the experiment, namely, differences of response behaviour that are motivated by social background, possibly also age, and the corresponding different daily routines that come with those factors and influence what is 'in or on one's mind'.

Table 3 presents the statistical measures that inform us about the spread of local-view responses among the participants.

Table 3: The spread of local-view responses

Response activity Groups	Mean	Median	Standard deviation (SD)	Maximum n° of responses per person	Minimum n° of responses per person
E	7.6	2	3.6	12	0
CH	8.4	2	3.6	13	0

A median of two indicates that about half of the respondents provided between zero and two local-view responses, whereas the other half of the respondents gave between two and as many as 12(E) or 13(CH) responses.

With the bulk of responses being given to local segments instead of to larger portions of the text, it seems safe to say that—in both groups—about half of the respondents were inclined to participate actively, whereas the other half were not inclined to do so. It is interesting to note that this pattern is consistent across the two groups.

As for the validity of the observations yielded by the experimental design, it can be stated that the design allows for a participation rate which is not unknown in other experimental procedures, e.g. in marketing research, where ‘subjects’ are allowed the freedom to decide whether or not to participate.⁷ In the following §4, I suggest that it is not only *external* validity that holds, but also what is called construct validity, namely, that the method employed indeed taps into ‘something’.⁸ Looking at the actual correlations of numbers and semantic constellations, it seems clear to me that they provide a glimpse of the mind’s conceptual activity when reading a text.

The quantitative correlations between the text semantics and response numbers yield patterns that carry all the imprints of semantic-pragmatic relevance. (In the meantime, responses have also been elicited from groups of readers who were given the German, Russian, and Turkish translations respectively.⁹) In §4, I will show how the local-view responses pattern onto segments that evoke semantically distinct meanings. The patterning comes about on the basis of a moment-by-moment movement, resulting in a series of tiny ‘moves’ driven by the purpose of wanting to know what is ‘up’ for ‘uptake’, resting at places where the wanting to know is temporarily satisfied, then forming another expectation towards “what happens next” (Hoey 2001:25).

⁷ I have learned from various elicitation events during the summer of 2004 (Bruche-Schulz 2005), that two factors should be observed when replicating the elicitation procedure: (a) the number of participants in an event should be between around 15 and 20 (maybe even more); (b) the event is best planned with a homogeneous group which is tied together by some common activity.

⁸ For the criteria relevant to quantitative methods, see, e.g. Trochim (2006) and a discussion of the differences between quantitative and qualitative methods, see Creswell (1994).

⁹ The outcome of the elicitation event based on the German translation of the excerpt has been reported in Bruche-Schulz (2008). The examination of the responses yielded during two more elicitation events, using the Russian and Turkish translations of the excerpt, is near completion.

4. Conceptual structure: thought figures in texts

In this section, I deal only with local-view responses since global-view and UAV responses cannot be quantified and correlated with marked-off portions of text in relation to its segments. Local-focus, or local-view, responses to linguistic cues reveal a pattern that oscillates between suspense (expressed by zero and low response numbers), and the sensation of temporary satisfaction when getting to know how an expectation is met (expressed by higher response numbers).

In §4.1, I will point out what the features of the semantic constellations are where readers *want to know* what comes next (signalled by zero or low response numbers), and where their expectation is temporarily satisfied (signalled by high response numbers). In §4.2, I will provide graphical presentations of the readers' response behaviour in Chinese and English.

4.1 Expectations tied to linguistic cues

As mentioned already, text progresses along expectations. At the core of an expectation is a textual cue that triggers a *want-to-know* impulse, in other words, the sensing of the desire to get to know 'what comes next'. Positioning, and lexical elements may work together for developing the thread of expectations. Initial positions of a text start out with a *want-to-know* signal [W], since the 'meaning' of an initial position is that of a stage setter. If something is marked as a beginning, it makes the reader wonder 'what the following will be about?'. After the stage is 'set', with the initial query being resolved by some preliminary information that temporarily satisfies the reader's expectation [TS], new queries open up by *want-to-know* signals [W], each then followed by a stage at which a temporary satisfaction of the *wanting-to-know* impulse is achieved [TS], thus bringing about the textual progression. — In short, a reader moves along a sequence of micro-acts that consists of a string of [W]-[TS] units.

When a *want-to-know* impulse [W] is sensed, readers' attention is kept in suspense, and response numbers are generally low. When the state of the world appears to be temporarily known, i.e. when the wish of wanting to know is temporarily satisfied [TS], response numbers go up. Table 4 summarizes the semantic situations found in the segments that motivate either low or high response numbers. (The meanings of 'high' and 'low' are respectively orientated towards the maximum responses, 12 in group E, and 13 in group CH, and the minimum, 0[zero], in both groups. Response numbers are perceived to be low when in the range of 1-3 response instances.)

The forming of expectations and the perceived meeting of such expectations, immediate or delayed, is not expressed explicitly by the readers. The analyst may infer

them from the response numbers in relation to the lexical cues they correlate with, but the reader herself probably rather ‘senses’ them at the moment of reading. It is suggested to understand such ‘sensing’ along the line of the assumptions hypothesized to accompany all conceptualizations (following Barsalou 2003, Cienki 2008). Such kind of sensing is not passively receptive, but actively giving a conceptual form to the object or event perceived whereby, in the case of a textual pattern, the conceptual form is not a symbolic unit represented by a linguistic sign but a “constructional schema” (Taylor 2002:167).

Table 4: Movement of attention

Motivations (semantic constellations)	ZERO, and LOW response numbers: <i>Want-to-know signals [W]</i>	HIGH response numbers <i>Temporary satisfaction of wanting to know [TS]</i>
(a) Modalities. Lasting events and states	Want to know what ‘is/will be real?’. <i>Triggered by:</i> Expressions of possibility, future, intention, obligation, [degrees of] ability, attempt, readiness, hope, look for, seem.	Descriptions of positive and lasting emotive states (love, like, other pleasant sensations) accompanying situations; assertions of positive knowing.
(b) Negations. Tapping into presuppositions	Want to know ‘what is the (congruous, positive) reality?’. <i>Triggered by:</i> Negated assertions, overt questions, and other means that partially overlap with (a).	Contrastive negations; evoking the sense of knowing something already, with the named part of the contrast triggering an inference as to the nature of the unnamed contrasting element.
(c) Perspectives on bounded and unbounded events; perfective and imperfective aspect; speech and thought introducers	Want to know ‘what next?’. <i>Triggered by:</i> Actions that bring about changes to a situation (as when moving something to a temporal or spatial bound). ¹⁰ Speech and thought introducers that trigger expectations of a ‘follow-up’ event of saying and thinking.	Ongoing activities filled with sensory detail, with no indication of a terminal bound, with or without explicitly formulated or implicated evaluations, both positive and negative. (For related observations, see Madden & Zwaan 2003.)

¹⁰ I adopt the term ‘bound’ in this analysis, in preference to ‘boundary’, following Kearns (2000).

What follows are six diagrams (Figures 1-6) which are meant to visualize the observations. The diagrams show different actual response numbers for each group of respondents on the one hand, but highly comparable patterns on the other. Most often, if there are *high* values in one group for particular segments, they are mostly so in the other group as well. If there are *low* values in one group, they are mostly low in the other group also.

There are six places where the proportions of the response numbers of group E and CH do not 'match' (namely, at s[1], s[10], s[21], s[24], s[25], and s[40]). Lack of space forbids discussing in very great detail all these cases which may be hypothesized to pertain both to socio-cultural differences of perspective, for example at s[10], and also to language-specific characteristics of the linguistic cues, for example, right at the start of the excerpt at s[1].

4.2 Movement of attention as visible from response numbers

Response numbers can be read off the scale on the left of each diagram. The values marked in red indicate the number of responses given by group E[nglish], those in green indicate the number of responses given by group CH[inese]. The scale to the left of the diagram *adds* the responses of group CH onto those of group E in terms of absolute numbers. There are six diagrams altogether (Figures 1-6). Each diagram shows the quantities of the responses to portions of the text. For ease of reference, the full text of the respective segments to which the responses belong is presented underneath each diagram. (Figure 1: s[1-6], Figure 2: s[7-13], Figure 3: s[14-21], Figure 4: [s22-29], Figure 5: s[30-38], Figure 6: s[39-44].)

4.2.1 Setting the stage

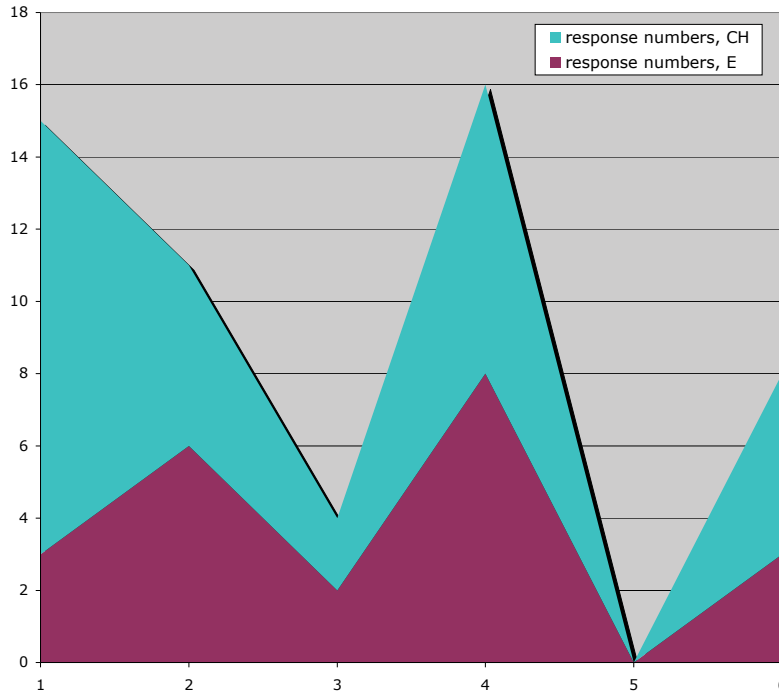


Figure 1: Segments 1-6

2.a) 1. “Men”, 1. said the little prince,	1. 小王子 說 ：‘syut’, ‘said’
2.b) “set out on their way in express trains,	2. 「人們只顧著匆匆忙忙擠到快車裡，
3.a) but they do not know	3.a) 卻不曉得
3.b) what they are looking for.	3.b) 他們究竟要找什麼，
4. Then they rush about, and get excited and turn round and round ...”	4. 只見他們手忙腳亂地繞著圈子轉……」
5. And he added:	5. 然後他又說：
6. “It is not worth the trouble ...”	6. 「何苦這樣嘛……」

Segment (1) sets the stage. It cues both a [W] and a [TS] perception. Readers ‘know’ that a story starts [TS] but, at the same time, expect more elaboration to come [W]. The very first segment at s[1] draws in three responses in English, and twelve in Chinese. This is noteworthy since, in the remaining text, speech-introducing clauses have, in both respondent groups, received only 0 (zero) or low responses. Without exception, all responses to s[1] deal with the main character, the ‘little prince’ — who is

thus marked as the most prominent lexical cue, introducing a main character and his discourse world to the reader. Interestingly, the preposed word 'men', the subject of s[2], invited only one global response, not a single local response in English.¹¹

The Chinese version of s[1] appears to direct attention towards the conventional format of story-telling. Although s[1] reads '小王子 說 ...', it is easily inferred to stand for the (ellipted) formula '[從前] 小王子 說 ...' ('Once upon a time...'), since it is placed at the very beginning of the excerpt. We hear the formula, we sit down to listen! Thus s[1] works both ways: as [W] since it sets the scene, and as [T]emporary [S]atisfaction of the wish of being prepared for what is going to come [TS_{story telling}]. After commencement of the storytelling that sets off the pattern with deviating response numbers at s[1], there is then a correspondence of response numbers observed in both language versions. Numbers oscillate between [W] (low) and [TS] (high numbers) at particular segments in a quite regular fashion.

Attention is undecided at s[2] by an activity that seems to be ongoing (the crowding in trains and the traveling), but normally should find an end; s[2] results in a cautious [TS₁]. At s[3a-b], a negative assertion ('but they do not know ...'), triggers a want-to-know impulse (want-to-know what? 'they don't know') [W₂]. What follows in s[4] is an evaluation couched in the form of an authoritative cultural model that describes an ever ongoing fruitless activity [TS₂]. Readers respond with a high number of responses to s[4].

A speech verb at s[5] prepares a new stage of expectation (W₃), letting the response numbers drop to zero. The following negative evaluation at s[6] is cloaked in the form of a negative assertion in English. According to the observations concerning crucial semantic constellations, summarized in Table (4), s[6] should therefore receive low response numbers. However s[6] continues the gist of the cultural model at [s4] and, instead of being read as a negative assertion, it is probably more likely read as a negative evaluation — a situation which is even clearer in Chinese where the translation gives us a rhetorical question (「何苦這樣嘛……」). Consequently, s[6] has its responses increase again [TS₃]. (In English, the increase is, however, only slight.)

The discussion of the diagram of Figure 1 has gone to some length in order for the main line of argument to be sufficiently explicated. It is trusted that a briefer commentary on the diagrams in Figures 2-6 will serve the purpose from this point onwards.

¹¹ The preposed word 'men' in s[1] of the English text seems to be perceived as a "light subject" (Chafe 1994:88-90) which counts neither towards [W] nor [TS].

4.2.2 What wells are like

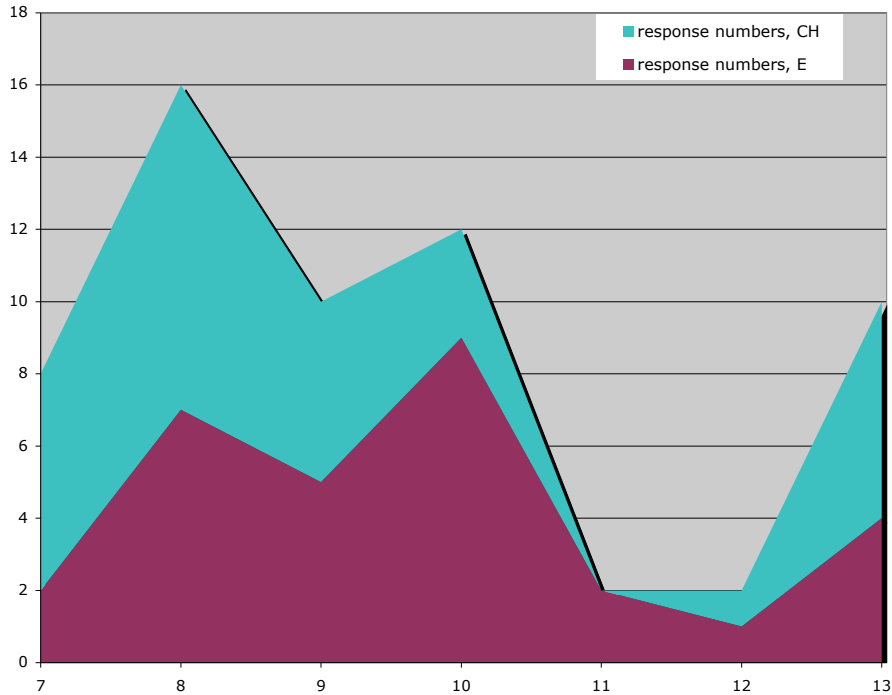


Figure 2: Segments 7-13

7. The well that we had come to	7. 我們找到的那口井
8. was not like the wells of the Sahara	8. 不像是撒哈拉沙漠地帶的井。
9. The wells of the Sahara are mere holes dug in the sand.	9. 撒哈拉沙漠地帶的井通常只是一些在沙地裡挖個簡單的洞。
10. This one was like a well in a village.	10. 這口井卻很像一般村裡的井，
11. But there was no village here,	11. 但是那裡的周遭沒有任何村莊，
12. and I thought	12. 於是我開始懷疑
13. I must be dreaming ...	13. 自己是否在作夢。

The *wanting-to-know* figure shown in the diagram of Figure 2 starts out with s[7] that informs us about a new location, achieved through reaching a spatial bound ([‘the well] that we had come to’) [W1]. The description of the new location begins with a contrastive negation at s[8] (‘not like the wells of the Sahara’) that motivates higher response rates since an inference is triggered as to what the ‘other’ wells will be like. The contrastive negation presupposes knowledge of the first part of the contrast (‘the

wells of the Sahara': *are* such and such); so s[8] temporarily satisfies the *wanting-to-know* impulse sensed at s[7] as [TS₁].

The clause at s[9] further confirms what the wells of the Sahara are like, but includes a relativizing evaluation ('*mere* holes'). Can holes be wells? Readers hesitate and responses are held back. At [10], a contrastive assertion that conveys more positive knowledge about those wells lets response numbers go up again. However, the negative assertion at s[11] triggers a new expectation. Readers want to know how to solve the puzzle of a situation where there is a well, but no village [W₂]. The *want-to-know* move is prolonged by s[12] with a thought verb that promises a perspective that has yet to be clarified [W₃]. Thus, s[11] and s[12] draw in low values. At s[13] the depiction of an ongoing activity of the mind ('dreaming') triggers again more response numbers [TS₂].

4.2.3 The little prince, the pulley, and the rope

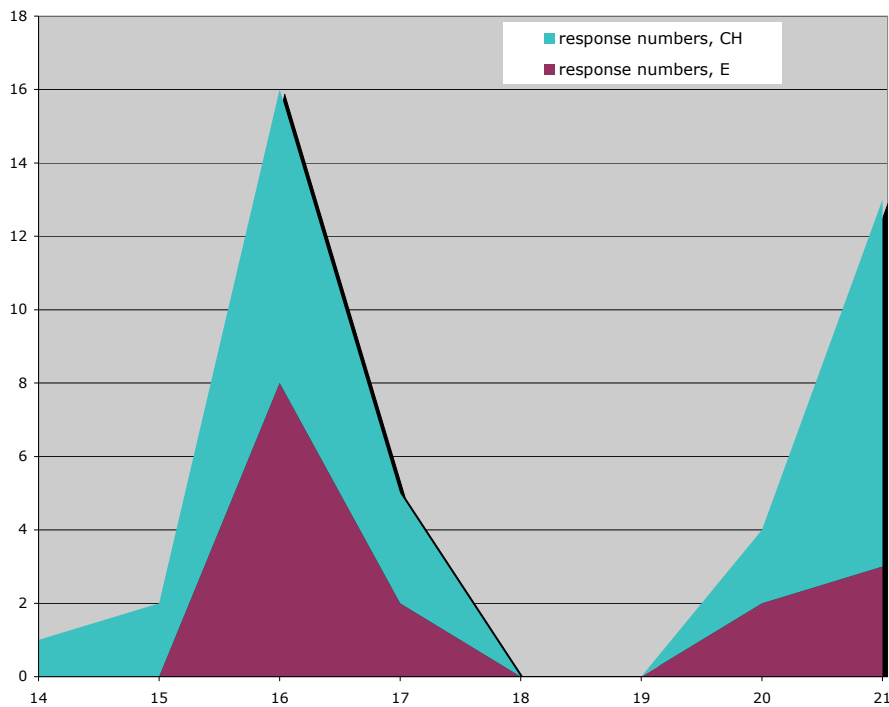


Figure 3: Segments 14-21

14. “It is strange,	15. 我對小王子說：
15. I said to the little prince	14. 「好奇怪，
16. Everything is ready for use: the pulley, the bucket, the rope...”	16. 什麼都準備好了，滑輪，水桶和繩子……」
17. He laughed,	17. 他笑笑地碰了，[碰繩子]
18. touched the rope	18. 碰繩子
19. and set the pulley to working	19. 玩了玩滑輪，
20. And the pulley moaned,	20. 滑輪喀吱喀吱地發出聲響
21. like an old weathervane which the wind has long forgotten.	21. 好像沒風的日子裡，一具老舊的風信雞在微哼著單調小曲一樣。

With the keyword *strange* at s[14], [W₁] is opened up in English, and supported by the speech introducer [W_{supp}] in s[15]. The Chinese segments are presented in reverse order, s[15-14]. In Chinese, the speech introducer triggers [W₁], which is then strengthened by 好奇怪 [W_{supp}]. The diagram shows very few responses at s[14], and s[15], in both languages. Then, the promise of some solution (‘everything is ready’) triggers eight responses in Chinese and eight in English [TS₁]. After that, a new riddle [W₂] is on its way at s[17-19]. It is being prepared through the cue ‘laughing’, and continued by two bounded events at s[17], and s[18] which keep up suspense regarding new events that are expected to follow. At [s18-19], the response numbers die down to zero. But, then, at s[20], the pulley sends out, or ‘utters’, negatively evaluating noises, it ‘moans’ [W₃] — why? Readers wonder how the moaning will be resolved, and at s[21] they receive an explanation in the form of an event frame that is filled with a state of the wind’s ‘forgetting’ to apply its causative force for an untold, seemingly unlimited period of time [TS₃].

4.2.4 We have wakened the well

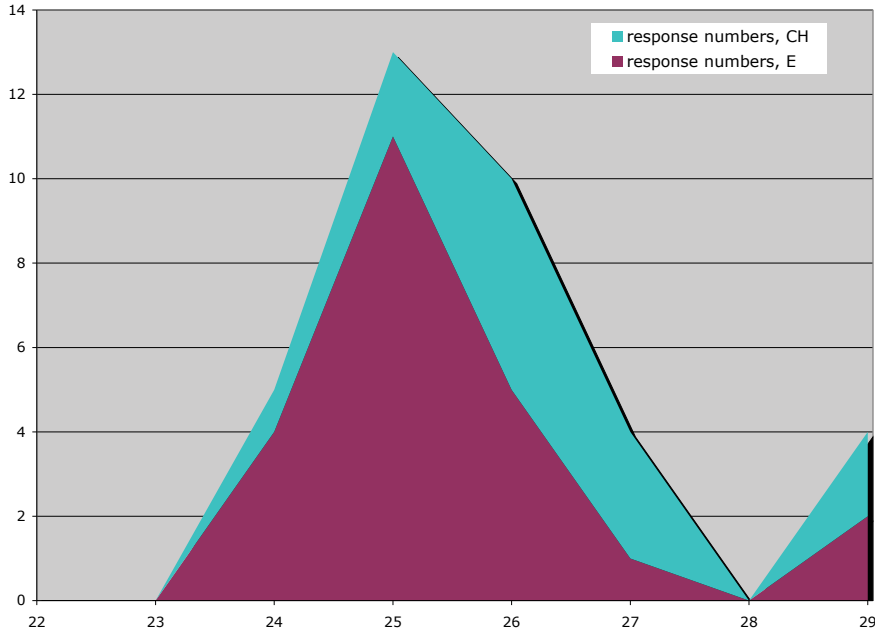


Figure 4: Segments 22-29

22. "Do you hear?"	23. 小王子說：
23. said the little prince	22. 「你聽到沒，
24. "We have wakened the well,	24. 我們叫醒了這口井，
25. and it is singing"	25. 而它正在唱歌哩……」
26. I did not want him to tire himself with the rope.	26. 我不希望他太用力，
27. "Leave it to me,"	28. 於是說：
28. I said.	27. 「讓我來拉吧！」
[no translation equivalent in English]	*我對他說：
29. "It is too heavy for you."	29. 「這個對你來講太重了。」

Semantic elements that build up the *want-to-know* figure in the sequence of segments s[22-29] are: the question expressing a *want-to-know* [W₁] at s[22] together with the speech introducer at s[23]. The stative present perfect in English (resultative in CH) at s[24] is obviously not only read as an answer to the question 'do you hear?', but also as conveying the sense of 'the well is awake now', i.e. as a starting point for s[25]. Segment s[25] depicts an event frame with a state reading, filled with a positively evaluative ongoing activity, with no indication of a terminal bound [TS₁].

The negated *want*-clause at s[26] carries a contrastive thrust. But the negation is yet to be decided as either working towards a negative assertion ('pilot will not accept thing xy'), or a contrastive negation ('pilot does not want little prince do something, but wants himself doing the something'). At s[27], an imperative determines the contrast as a contrastive negation in English (not 'him' is to tire himself, but 'me', the pilot). While the sense of a contrastive negation seems to prevail in the English imperative ('leave it to me', implying 'not to you'), the effect is shading into an optative element of a request in the Chinese text at s[27] ('let *me* do it'). In both ways, an expectation [W₂] is invited as to whether the little prince would oblige or not. So response numbers are low. At s[28], a speech introducer makes responses drop to zero. Responses do not rise significantly upon the start of a new problem at s[29]. If something is labelled as 'too heavy' [W₃], the question arises as to what to do, or what to think about it.

4.2.5 Thirsty for this water

Segments s[30-31] describe bounded events which 'bring' the bucket 'to the edge of the well' and 'set it there', thus creating a new situation causing the reader wanting to know what is next [W₁]. Next comes the description of a world filled with positively evaluative sensory experience at s[32-33], triggering high response numbers [TS₁]. What follows is then a string of [W]-elements that do not find an immediate [TS]-complement. [W]-elements are "carried forward", they pile up so to speak (Hoey 2001:23).

At s[34] a new [W]ant-to-know sequence [W₂] begins by describing a sensation felt to be a problem (being 'thirsty'). At s[35], an event of 'saying', i.e. an event that construes the text world by introducing speech, follows [W₃]. (Speech introducers are greeted with suspense since they trigger the expectation of what someone will be saying or thinking.) At s[36], an imperative construction triggers a further expectation [W₄] as to whether the other story character, the pilot, will oblige or not.¹² The order is reversed in Chinese: [W₃] – [W₂] – [W₄], but the effect is the same, or very similar at least.

¹² Due to lack of space, I disregard the semantic-pragmatic differences between the different types of [W].

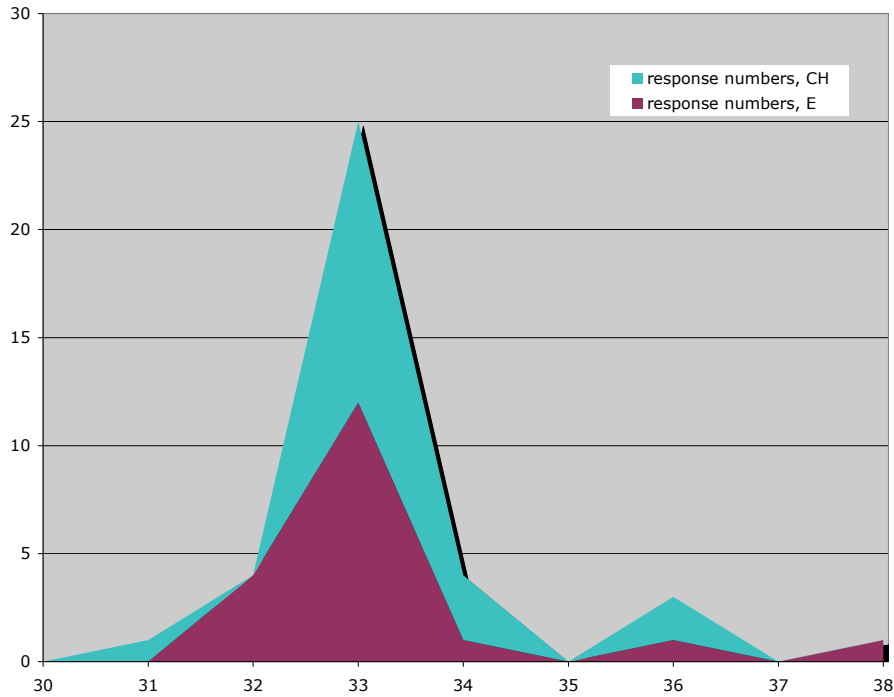


Figure 5: Segments 30-38

30. I hoisted the bucket slowly to the edge of the well	30. 我慢慢地把水桶拉到井口上面，
31. and set it there.	31. 然後放在一旁。
* - happy, tired as I was, over my achievement	* [no translation equivalent in Chinese]
32. The song of the pulley was still in my ears,	32. 我的耳裡一直迴響著滑輪喀吱喀吱的聲響，
33. and I could see the sunlight shimmer in the still trembling water.	33. 而陽光正在波動的水面上粼粼發光著。
34. "I am thirsty for this water,"	35. 小王子說：
35. said the little prince	34. 「我好想喝這水，
36. "Give me some of it to drink..."	36. 給我喝一點嘛……」
37. And I understood	37. 我這才明白
38. what he had been looking for.	38. 他所要尋找的！

The thought-introducing verb at s[37] is joined by a complement clause at s[38] acknowledging that the little prince ‘had been looking for something’, thus hinting at some solution. But instead of becoming an intermediate [TS₅], s[38] immediately triggers a new *want-to-know* impulse directed towards the new question as to what exactly he had been looking for. Meanwhile, [W₂] is being carried forward, all along from s[34]: Thirst demands quenching — which is not resolved from s[35/36] through and up to s[38] inclusively.

It is only in the stretch from s[39] through s[44] where the piled up expectations of s[34-37/38] receive a [T]emporary [S]atisfaction of the *want-to-know* impulse starting at s[34].

4.2.6 這些水不僅是用來喝的

‘This water was indeed a different thing from ordinary nourishment.’

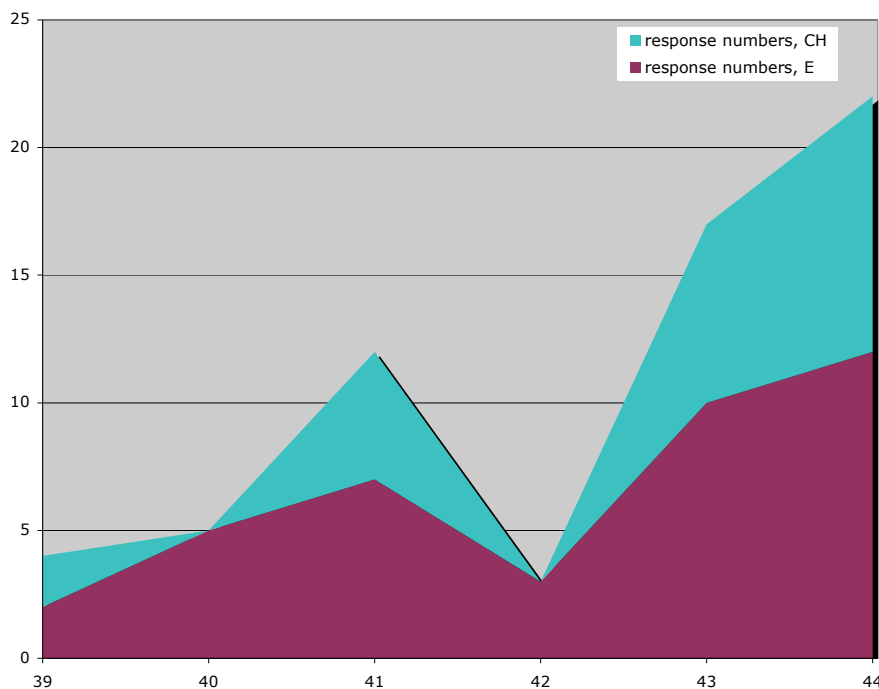


Figure 6: Segments 39-44

39. I raised the bucket to his lips.	39. 我把水桶一直提到他的嘴邊，
40. He drank, his eyes closed.	40. 他閉著眼睛喝，
41. It was as sweet as some special festival treat.	41. 那感覺就嘉年華一樣地甜蜜。

42. This water was indeed a different thing from ordinary nourishment.	42. 這些水不僅是用來喝的，
43. Its sweetness was born of the walk under the stars, the song of the pulley, the effort of my arms.	43.a) 它還是來自我們星光底下的長途跋涉，[comes from a long way under the stars]
	43.b) 來自滑輪單調有致的歌唱，[comes from the pulley's singing]
	43.c) 來自我臂膀出力的結果。[comes from the efforts of my arm] ¹³
44. It was good for the heart like a present.	44. 它就像天賜的禮物一樣，好喝又有益心靈。

The text that triggers the flow of attention reflected in the diagram of Figure 6 contains practically one single [W]-element at s[39] only. The pilot raises a 'bucket' full of water, and lets the little prince drink. Thus, a bounded event makes room for the expectation of a further event to come next. However, s[39] works in two ways: it works first as a major [TS] that responds to the expectations built up from s[34] through s[38] and second it also forms a [W] element that subsumes the build-up of [TS] from s[40] through s[44].

At s[40], an event follows which is only incrementally accomplishing a bound¹⁴ ('drank, his eyes closed'). While the numbers of the responses to s[40] rise somewhat in the English-reading group, those in the Chinese group remain at zero level. After all, we do not know whether the little prince drank all the water, or rather continued to enjoy the drinking.

Far from introducing a new event, s[40] rather conveys a positive evaluation of the situation described in the stretch of text from s[30-38] and [39], thus temporarily satisfying a *want-to-know* need that has been maintained across a few segments (s[30-39]).

A further positive evaluation completes this aspect at s[41]. At s[42], the respondents to the English text appear to treat the semantics of the clause as a contrastive focus that promises more good news, the respondents to the Chinese text seem to draw from s[42] the meaning of a negative assertion that pushes responses to zero. Positive evaluations, using words conveying explicitly positive values, are then found at s[43] and s[44] in both languages, and the respondents in both groups react with high response numbers. In those final evaluations that celebrate the precious, and mythical qualities of the water from the desert well, both groups unite in high response numbers again.

¹³ Segment s[43] may be seen to contain three different clauses in Chinese. However, it seems to me that this situation presents a potentially new question which I do not pursue in this paper.

¹⁴ See Kearns (2000:246) about incremental themes.

The diagram in Figure 6 concludes the thread of expectations found in the excerpt selected from a longer narrative. The thread of expectations can be shown to evolve along a template consisting of [W]-[TS] units. It has been suggested understanding the template as a sequence of *want-to-know* figures which are enacted as micro-acts. Such micro-acts are performed by a core, a *want-to-know* impulse, and a subsequent temporarily satisfying [TS] rejoinder. The resulting units may come as plain serial strings, or may also have several [W] elements embedded in a ‘master’ [W], i.e. in a *want-to-know* impulse that may stretch over several impulses related by their semantic-pragmatic connections.

5. Concluding remarks

The experimental design used in the present study allows for reaching out to readers’ conceptualizations along two tracks: the overt expression of ‘what comes to mind’, and the covert sensing of the textual progression along semantic-pragmatic meaning constellations, thus reflecting, enactment, and thereby construing a template of a text schema. I have suggested that mental-space theory (see Fauconnier 1997, among others) currently looks best equipped to provide an adequate explanatory framework for overtly expressed conceptualizations. However it is the covert sensing of textual structure, that has been the main concern of the present investigation.

As for the nature of the readers’ covert sensing of textual structure, I showed how this covert sensing leaves its traces in response patterns that correlate with semantic-pragmatic constellations found in the text. What is more, I suggest that the match of these patterns across two language groups adds relevance to the observations.

The readers themselves have been unaware of the process. That they provided a pattern gives credibility and soundness to the observations on the one hand, but raises more questions than answers on the other; the most urgent question of all appear to be what kind of consciousness is involved when referring to ‘overt’ and ‘covert’ processes. Availability in memory is, at it seems, only part of the story.

Conceptualization sits at the core of the black-box problem of the human mind [...]: it never reveals itself directly at the observable surface of human behaviour; it only ‘appears’ indirectly, in disguise, coded or filtered through the ‘structural principles’ of the many types of human behavioural systems, linguistic and otherwise. (Pederson & Nuyts 1997:3)

To get at the heart of the ‘many types of human behavioural systems’ requires finding ways of getting to know about these systems. In addition, we would need to

know how human consciousness registers the activity of such systems in relation to intentional symbolic action. An explanation should answer queries such as (a) what the site of overt and covert responses could be within the activity of the mind, (b) what exactly the instrumental processes within this activity are, and (c) what the quality of the experience is felt to be when engaging in overt and covert conceptualizations.

I follow Barsalou in assuming that, at each level of conceptualization, we deal with perceptual processes that always remain grounded in sensory-motor systems (1999:583), and which simultaneously bring in all the perceptual systems¹⁵ at once. I would like to believe that covert textual gestalts, such as the *want-to-know* figure described here, may be sensed—apperceived and thereby construed—by unconscious, or even preconscious, processes of enactment.

When the text is talking about the wells of the Sahara, and a reader jots down that her or his mind sets off to 'look for some water' (找水 [P14CH]), the act of jotting down words and the sensing of the scene is probably fully conscious. However, the very impulse of jotting down words in a particular place on the margin of a page must surely be channelled through a different pathway. Just at which level of consciousness the different kinds of sensing and enactment of conceptualizations are operating is a question that challenges cognitive scientists and linguists alike.

¹⁵ Introspective and proprioceptive experience belongs here as well (577, and *passim*).

References

- Barsalou, Lawrence W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22.4:577-660.
- Barsalou, Lawrence W. 2003. Situated simulation in the human conceptual system. *Language and Cognitive Processes* 18.5-6:513-562.
- Bruche-Schulz, Gisela. 2005. The construction of meaning by readers. *Research Grant Completion Report*. Hong Kong: Hong Kong Baptist University.
- Bruche-Schulz, Gisela. 2008. The hidden cognitive process. Paper presented at the 4th Lodz Symposium: New Developments in Linguistic Pragmatics, May 15-18, 2008. Lodz: Department of Pragmatics, University of Lodz.
- Chafe, Wallace L. 1994. *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press.
- Cienki, Alan. 2008. Looking at analyses of mental spaces and blending / Looking at and experiencing discourse in interaction. *Mental Spaces in Discourse and Interaction*, ed. by Todd Oakley & Anders Hougaard, 235-245. Amsterdam & Philadelphia: John Benjamins.
- Creswell, John W. 1994. *Research Design*. London: SAGE.
- Eco, Umberto. 1994. *Six Walks in the Fictional Woods*. Cambridge: Harvard University Press.
- Fauconnier, Gilles. 1997. *Mappings in Thought and Language*. Cambridge & New York: Cambridge University Press.
- Fauconnier, Gilles, and Mark Turner. 2002. *The Way We Think*. New York: Basic Books.
- Hoey, Michael. 2001. *Textual Interaction*. London: Routledge.
- Holšánová, Jana. 2008. *Discourse, Vision, and Cognition*. Amsterdam & Philadelphia: John Benjamins.
- Hutchins, Edwin. 2005. Material anchors for conceptual blends. *Journal of Pragmatics* 37.10:1555-1577.
- Kearns, Kate. 2000. *Semantics*. Basingstoke: Macmillan.
- Khanina, Olesya. 2008. How universal is *wanting*? *Studies in Language* 32.4:818-865.
- Levinson, Stephen C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge: MIT Press.
- Madden, Carol J., and Rolf A. Zwaan. 2003. How does verb aspect constraint event representations? *Memory & Cognition* 31.5:663-672.
- Pederson, Eric, and Jan Nuyts. 1997. Overview: on the relationship between language and conceptualization. *Language and Conceptualization*, ed. by Jan Nuyts & Eric Pederson, 1-12. Cambridge & New York: Cambridge University Press.
- Pressley, Michael, and Peter Afflerbach. 1995. *Verbal Protocols of Reading*. Hillsdale: Lawrence Erlbaum Associates.

- Rayner, Keith. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124.3:372-422.
- Richardson, Daniel C., Rick Dale, and Michael J. Spivey. 2007. Eye movements in language and cognition. *Methods in Cognitive Linguistics*, ed. by Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson & Michael J. Spivey, 323-344. Amsterdam & Philadelphia: John Benjamins.
- Smith, Carlota S. 2003. *Modes of Discourse: The Local Structure of Texts*. Cambridge & New York: Cambridge University Press.
- Taylor, John R. 2002. *Cognitive Grammar*. Oxford & New York: Oxford University Press.
- Trochim, William M. K. 2006. Introduction to Validity. <http://www.socialresearchmethods.net/kb/introval.php>. Last revised: 20-10-2006. Accessed on 27-02-2009.

The sources of the excerpt

- Antoine de Saint-Exupéry. 1938. *Le Petit Prince*. Paris: Editions Gallimard 1946.
- The Little Prince*. Translated by Katherine Woods. First published in 1945. Current edition: Egmont Books. London 2002. Pages 76-77 (= excerpt, English translation).
- 小王子 – *Le Petit Prince – The Little Prince*. Translated by 張譯 'Zhang Yi'. 希代 'Xi Dai' Publishers. Taipei 1999. Pages 260-262 (= excerpt, Chinese translation)

(According to information from <www.petit-prince.at/pp-chines.htm> [accessed in 2007], the translation by 張譯 was the first of eleven subsequent translations into Chinese.)

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)
National Chung Cheng University

Encoding patterns of motion events in Frog Story narratives of Mandarin Chinese and Taiwan Sign Language are identified with reference to Talmy's framework of conceptual analysis. In terms of Slobin's tripartite distinction, both languages can be regarded as complex verb-framed languages, though with different distribution in encoding patterns. Their typological similarity comes from conflation of conceptual elements of Manner, Path, and Motion. However, in spoken languages, these elements occur in the linear serialization, whereas in signed languages, they are simultaneously conflated in classifier predicates. A revised tripartite distinction among satellite framed, verb-framed, complex verb-framed languages is therefore proposed across the audio-vocal and the visual-gestural modalities of human languages. Co-speech gestures in spoken languages are capable of encoding more refined manners as are facial expressions in sign languages. Therefore, a more comprehensive language typology must take into consideration grammatical structures in both modalities, co-speech gestures in spoken languages, and facial expressions in sign languages.

Key words: typology, spoken languages, signed languages, serial verb languages, classifier predicates, sequentiality, simultaneity

1. Introduction

Talmy (1985, 2000) proposed a conceptual analysis of motion events which consists of six cognitive components, namely Figure, Ground, Path, Motion, Manner and Cause. He further proposed a typological distinction between 'satellite-framed' languages and 'verb-framed' languages in terms of different encoding patterns of motion events. In

* This paper was presented at Second International Conference of the French Association for Cognitive Linguistics, University of Lille, France, May 10-12, 2007. We thank Dan Slobin, Adam Schembri, Sonja Erlenkamp, and others for their valuable comments and suggestions at the Conference. We also thank a traveling grant from the Ministry of Education which enabled Su Shiou-fen to present the paper at the Conference.

‘satellite-framed’ languages such as English and German, the verb incorporates Motion and Manner. In contrast, in ‘verb-framed’ languages such as French and Spanish, the verb incorporates Motion and Path. According to Talmy, Chinese is a ‘satellite-framed’ language. However, Tai (2003) argued that Chinese cannot properly be treated as a ‘satellite-framed’ or as a ‘verb-framed’ language because of abundance of resultative verb compounds, and thereby suggesting a tripartite typological distinction. In fact, Slobin (2004) proposed ‘complex verb-framed’ languages as a third type to include Chinese and American Sign Language (hereafter ASL).¹ Thus, instead of Talmy’s bipartite typology, with Slobin, we adopt a tripartite typology of ‘satellite-framed’, ‘verb-framed’ and ‘complex verb-framed’ languages.

Mandarin Chinese is neither a ‘satellite-framed’ language nor a ‘verb-framed’ language for three reasons. First, Path can be expressed as an independent verb alone. Second, it serves as the center of predication in resultative verb compounds. Third, Manner is not syntactically subordinated to Path as in ‘verb-framed’ languages (Huang & Tanangkingsing 2005, Slobin 2004, and Tai 2003). In this paper, we propose that Mandarin Chinese is a ‘complex verb-framed’ language.

Sign languages, in visual-gestural modality, are capable of simultaneously encoding different cognitive components of motion events. How does this simultaneity in sign languages fit into the typological framework initiated by Talmy? Supalla (1990) treated ASL as a satellite-framed language. However, Slobin & Hoiting (1994) proposed that ASL is a kind of complex verb-framed languages in which the main verb conflates Path and Motion, and Manner can be encoded either by a verb before the path verb or by an affix incorporating into the path verb. In this paper, we also observe that Taiwan Sign Language (hereafter TSL) can be treated as a complex verb-framed language, just like ASL and other sign languages.²

Following Slobin & Hoiting (1994), Taub & Galvan (2001) and Galvan & Taub (2004) investigated how typologically different languages, such as English and ASL, encode information about motion events by using Frog Story narratives as materials. Frog Story (*Frog, where are you?*) is a story picture book without written text drawn by Mayer (1969) about a boy and a dog’s searching for a frog.

¹ Slobin (2004) also extended this type of languages as equipollently-framed languages. The reason that we adopt the term ‘complex verb-framed’ languages is to emphasize the encoding of motion events with obligatory path verb but with optional manner verb or Manner incorporated into path verb.

² Taiwan Sign Language (TSL) is a language naturally developed in Taiwan Deaf Community. For an introduction, see TSL browser (<http://tsl.ccu.edu.tw>) compiled by Jane Tsay and James Tai.

As a part of sign language project at National Chung Cheng University, we have collected Frog Story narratives of Mandarin Chinese and TSL for various research issues. These narratives provide our data base for the present study. The materials we investigated include narratives by 20 subjects in Mandarin Chinese and another 20 subjects in TSL.³ The purpose of this paper is two-fold. One is to report different encoding patterns of motion events between TSL and Mandarin Chinese as spoken in Taiwan. The other is to discuss their typological implications with respect to previous studies.

The organization of this paper is as follows: Section 2 reviews the typology of motion events and identifies patterns of Frog Story narratives in Mandarin Chinese. Section 3 briefly discusses previous studies of motion events in sign languages and identifies patterns of Frog Story narratives in TSL. In §4, based on our observations in Mandarin and TSL Frog Story narratives, we examine the manner-salient property across modalities, with special reference to co-speech gestures in general. Section 5 concludes the paper with typological implications.

2. Typology of motion events and their expression

Talmy (1985, 2000) proposed a bipartite typology of construction types according to whether the core schema PATH is mapped onto a verb or a satellite: (1) *verb-framed languages (V-languages)*, such as Spanish, map the core schema on the verb, and (2) *satellite-framed languages (S-languages)*, such as English, map the core schema on the satellite. In addition, Slobin & Hoiting (1994) and Slobin (2004) proposed that ASL is a kind of *complex verb-framed languages* in which Path is obligatorily encoded as a main verb and Manner can be optionally encoded as a verb sequentially before a path verb, or as an affix simultaneously incorporating into a path verb.

2.1 Satellite-framed and verb-framed languages

In Talmy's (1985, 1991, 2000) analysis of motion events, a basic motion event involves one object (the Figure) moving or locating with respect to another object (the reference object or Ground). There are four major cognitive components in the conceptual structure of a motion event: Figure, Ground, Motion, and Path, while Manner or Cause will sometimes be encoded as co-event. Different strategies to encode these conceptual

³ Forty narratives were collected as a part of James Tai's sign language project sponsored by National Science Council in Taiwan (2001-2004). The project title is "A Study of Taiwanese Sign Language: Phonology, Morphology, Syntax, and A Digital Graphic Dictionary".

elements are used among different languages. Talmy (2000:26) illustrated the encoding of these two co-events with sentences (1) and (2). In sentence (1), Manner of the verb ‘roll’ is incorporated into Motion, while in sentence (2), Cause of the verb ‘blow’ is conflated with Motion. Both sentences use a preposition ‘off’ as the satellite to represent Path.

- | | | | | |
|-----|-------------------|---------------|------------|-------------------|
| (1) | <i>The pencil</i> | <u>rolled</u> | <u>off</u> | <i>the table.</i> |
| | Figure | Motion | Path | Ground |
| | | Manner | | |
| (2) | <i>The pencil</i> | <u>blew</u> | <u>off</u> | <i>the table.</i> |
| | Figure | Motion | Path | Ground |
| | | Cause | | |

As illustrated in sentences (1)-(2) above and (3) below, in *satellite-framed languages*, such as English, the core schema ‘Path’ is encoded as the satellite and the main verb conflates Manner or Cause with Motion. In contrast, as illustrated in sentence (4) below, in *verb-framed languages*, such as Spanish, the main verb conflates Motion and Path with Manner expressed optionally as an adverbial or a gerundive constituent (Talmy 2000:49).

- | | | | | |
|-----|-------------------------------------|----------------|-------------------|--------------------|
| (3) | English: | | | |
| | <i>The bottle</i> | <u>floated</u> | <u>into</u> | <i>the cave.</i> |
| | | Motion.Manner | Path | |
| (4) | Spanish: | | | |
| | <i>La botella</i> | <u>entro</u> | <i>a la cueva</i> | <i>(flotando).</i> |
| | The bottle | moved-in | to the cave | (floating) |
| | | Motion.Path | | (Manner) |
| | ‘The bottle floated into the cave.’ | | | |

2.2 Typological category of Mandarin Chinese and its expression

In addition to *verb-framed* languages and *satellite-framed* languages proposed by Talmy, there is a third type, *complex verb-framed* languages proposed by Slobin. Mandarin is a highly verb serializing language where the most preferred strategy to encode motion events is to combine manner verb, path verb and deictic motion verb (Huang & Tanangkingsing 2005). The question then arises: which kind of language does Mandarin Chinese belong to? In this section, we will review the previous studies

on this issue in §2.2.1 and examine the narratives we have elicited from Frog Story in §2.2.2.

2.2.1 Chinese as a complex verb-framed language

Talmy analyzed Path element in Mandarin Chinese as a satellite and therefore treated Mandarin as a satellite-framed language (1985, 2000). However, Tai (2003) pointed out that Path element in Mandarin Chinese can be used independently as a main verb and in V-V compound, Path element is often the center of predication.

At first glance, it appears that Mandarin Chinese should belong to the *verb-framed* language group, since the ‘core schema’ is expressed in path verb compounds such as *chu^lai*⁴ ‘exit or come out of’ and *jin^qu* ‘enter or go into’. However, a closer inspection shows that these path verbs can be preceded by manner verbs and there is no syntactic or semantic evidence that either verb is subordinate to the other. Slobin & Hoiting (1994) and Slobin (2004) proposed that spoken serial-verb languages and sign languages belong to the type of *complex verb-framed languages*.

It seems to be more appropriate to treat Mandarin Chinese as a complex verb-framed language. First, as illustrated in sentence (5), path verb or path verb compound can occur independently as a main verb. Second, Manner can be optionally encoded within a verb compound, in which Manner is encoded with V₁ and Path, V₂, such as ‘*pao^jin(^qu)*’,

⁴ Notational conventions:

<i>word^word</i>	Compound words in Chinese, e.g. <i>chu^lai</i> ‘exit or come out of’, <i>jin^qu</i> ‘enter or go into’, <i>fei^chu^lai</i> ‘fly out (toward the speaker)’ etc. The symbol # is for word boundary.
SIGN	English glosses for manual signs are written in capitals, such as JAR, FROG, DOG, TAKE, etc.
SIGN>SIGN	The change of signing from one stage to another stage.
SIGN-SIGN	Multiword glosses connected by hyphen dash are used when more than one English word is required to translate a single sign (e.g. GO-AWAY).
SIGN(SIGN^SIGN)	Compound sign. For example: ESCAPE(ESCAPE^RUN).
Gpro+Fpro-move	Classifier predicates to express motion events are typically composed of a proform for Ground (Gpro) and a proform for Figure (Fpro), the later being integrated with Motion and Path (Fpro-move). The marker ‘+’ means that Gpro and Fpro are signed simultaneously by both hands, with Gpro signed by the non-moving hand, and Fpro by the moving hand. The marker ‘-’ means that the conceptual components ‘Figure proform’, ‘Motion’ and ‘Path’ are indivisible and are expressed as one unit by the moving hand, e.g. <i>JARpro+FROGpro-move.out.of</i> . Sometimes classifier predicates consist of two Figures moving at the same time. Parentheses are used to group them together, as in <i>(DEERpro + DOGpro) – move</i> .

as illustrated in sentence (6). Third, the manner verb itself can be a replicated verb preceding the main verb compound, such as ‘*paopaopao pao^jin(^qu)*’, in sentence (7). Fourth, there is no evidence that manner verb and path verb in both sentences (6) and (7) are syntactically subordinate to each other. Fifth, the ungrammaticality of sentence (8) clearly shows that the manner verb alone cannot encode a motion event in Mandarin Chinese. In contrast, in sentence (5) path verb alone without manner verb *pao* ‘run’ is grammatical.

- (5) *Gou jin(^qu) fangjian li.*
 Dog enter(^go) room inside
 ‘The dog entered the room.’
- (6) *Gou pao^jin(^qu) fangjian li.*
 Dog run^enter(^go) room inside
 ‘The dog ran and entered the room.’
- (7) *Gou paopaopao pao^jin(^qu) fangjian li.*
 Dog run++ run^enter(^go) room inside
 ‘The dog ran and ran and ran and entered the room.’
- (8) **Gou pao fangjian li.*
 Dog run room inside

In short, Chinese is a serial verb language, and like other serial verb languages, it belongs to the third type, namely, *complex-verb framed* languages.

2.2.2 Mandarin patterns of motion event in Frog Story narratives

Manner verb cannot occur alone at the sentence level as illustrated in sentence (8) above. No motion event encoding manner verbs alone is observed in Frog Story narratives either. In other words, we didn’t find sentences such as (9) and (10) in the narratives. This kind of sentences can only be used for contrastive purpose at the sentence level. Therefore, it is not surprising that they are absent in Frog Story narratives.

- (9) ?*Mifong fei.*
 bee fly
- (10) ?*Xiaogou pao.*
 dog run

We have identified several encoding patterns in Mandarin Chinese narratives of Frog Story. These patterns can be grouped into two categories. They are (a) motion events that encode Path, and (b) motion events that encode both Manner and Path.

In (a) category, Path in Mandarin motion events can be encoded by path verbs alone, such as *jin* ‘enter’, *chu* ‘exit’, *dao* ‘arrive’, *likai* ‘leave’, etc. It can also be encoded by a path compound that is composed of a path verb and a deictic motion verb such as *lai* ‘come’ and *qu* ‘go’ in *jin^qu* (enter^go) ‘enter’, and *chu^lai* (exit^come) ‘exit’, etc., as illustrated in sentences (11) and (12). The verb *lai* ‘come’ represents the perspective of ‘toward the speaker’; the other verb *qu* ‘go’ encodes the perspective of ‘away from the speaker’.

(11) Path verb alone (P)

Tamen yiqi dao chuang shang shuijiao.
they together arrive bed top sleep
‘They went to bed together to sleep.’

(12) Path compound composed of path verb and a deictic motion verb (P^D)

Mifong tongtong chu^lai le.
bee all exit^come ASP
‘The bees all came out.’

In category (b), there are three patterns of motion events that encode both Manner and Path. The first pattern uses V-V compound that is composed of a manner verb and a path verb, as exemplified by *pa^dao* (crawl^arrive) ‘crawl (on)to’ in (13). The second pattern uses V-V-V compound that is composed of a manner verb, a path verb, and a deictic verb, as exemplified by *fei^chu^lai* (fly^exit^come) ‘fly out (toward the speaker)’ in (14). The third pattern uses V-V compound that is composed of a manner verb and a path verb, with a deictic verb occurring at the end of a sentence. For example, *pao^dao...qu* ‘run^arrive...go’ in (15).

(13) V-V compound composed of a manner verb and a path verb (M^P)

Xiaohai pa^dao da shu shangmian.
child crawl^arrive big tree top
‘The boy crawled onto the top of the tree.’

(14) V-V-V compound composed of a manner verb, a path verb, and a deictic verb (M^P^D)

Mifong fei^chu^lai le.
bee fly^exit ASP
‘The bees flew out.’

- (15) V-V compound composed of a manner verb and a path verb plus a separated deictic verb (M[^]P#D)

Xiaopengyou you *pa^dao* *shu* *shang* *qu*.
 child again crawl^arrive tree above go
 'The boy crawled onto the top of the tree again.'

Sometimes the aspect marker 'le' is inserted in between a manner verb and a path-deictic verb compound, as illustrated in sentence (16). Furthermore, to emphasize Manner, the manner verb alone in a reduplicated or even triplicated form occurs before a V-V compound composed of a manner verb and a path verb. This is shown in (17).

- (16) Manner verb and a path-deictic verb compound with an aspect marker in between (M-asp.-P[^]D)

Xiao qingwa *cong* *pingzi* *li* *tiao* *le* *chu^lai*.
 little frog from jar inside jump ASP exit^come
 'The little frog jumped out of the jar.'

- (17) Reduplicated or triplicated manner verb before a compound composed of a manner verb and a path verb (M#M[^]P)

Xiaohai *zouzou(zou)* *zou^dao* *shenlin* *limian*.
 boy walk++ walk^arrive forest inside
 'The boy walked, walked, walked, and walked into the forest.'

Path can also be encoded by an adverbial or a prepositional phrase, as shown in (18) and (19).

- (18) Manner verb with an adverbial phrase represented as Path

Mifong *sichude* *fei*.
 bee everywhere fly
 'The bees flew around.'

- (19) Manner verb with a prepositional phrase represented as Path

Lu *wang xuanyai fangxiang* *pao* *zhe*.
 deer toward cliff direction run ASP
 'The deer ran toward the cliff.'

3. Motion events in sign languages

Cross-linguistically, motion events in sign languages necessarily involves classifier predicates. A full sentence expressing motion event can be analyzed as Figure-Ground-classifier predicate or Ground-Figure-classifier predicate. A classifier predicate is typically composed of a proform for Ground and a proform for Figure with Motion and Path conflated (Chang, Su & Tai 2005). In addition, Manner can be simultaneously incorporated into the classifier predicate or sequentially expressed as a verb occurring before the classifier predicate. Therefore, Supalla (1990) and Slobin & Hoiting (1994) proposed a serial-verb analysis of predicates in ASL. A serial-verb construction in ASL is composed of a manner verb and a path classifier predicate in sequence. Supalla (1990) analyzed the manner verb as the main verb, and thus treated ASL as a satellite-framed language.⁵ However, Slobin & Hoiting (1994) analyzed the path classifier predicate as the main verb because the manner verb is only optionally encoded. In their analysis, ASL belongs to a complex verb-framed language akin to spoken serial verb languages. In this paper, based on TSL Frog Story narratives we have collected, we will show that TSL, like ASL and other documented sign languages, can be more properly treated as a complex verb-framed language. In the following section, we report various encoding patterns in TSL Frog Story narratives.

3.1 Various encoding patterns in TSL Frog Story narratives

Both classifier predicates and lexicalized verbs can be utilized to encode motion events. We will illustrate patterns of classifier predicates first, and then those of lexicalized verbs.

3.1.1 Classifier predicates

Sign languages typically use classifier predicates to represent motion events. Each classifier predicate is composed of a proform for Ground and a proform for Figure conflated with Path and Motion (Chang, Su & Tai 2005).

In the examples below, we present full sentences expressing motion events as consisting of Figure-Ground-classifier predicate or Ground-Figure-classifier predicate. The underlined parts of sentences (20)-(23) below are classifier predicates as shown in Figures (1)-(4). The ‘proform’, represented by ‘pro’ is a replaced form for the preceding argument(s), i.e., Figure or Ground. The marker ‘+’ means two components are signed simultaneously by both hands. The non-moving hand represents the proform for Ground

⁵ Supalla (1990) referred to this kind of Manner verb as body classifier.

and the moving hand represents the proform for Figure. The marker ‘–’ means that the components ‘proform for Figure’ and ‘Motion/Path’ are simultaneously expressed as a unit by the moving hand.

- (20) *JAR* *FROG* *JARpro + FROGpro – move.out.*
 Ground Figure Ground pro + Figure pro – Motion/Path
 ‘The frog moved out of the jar.’



Figure 1: JARpro + FROGpro – move.out

A proform for Ground can encode either source or goal for Path. In TSL sentence (20) above, the motion event is about Figure (a frog) moving out of Ground (a jar). In TSL sentence (21) below, the motion event is about Figure (a dog) moving into Ground (a jar). Both sentences (20) and (21) represent a motion event by a classifier predicate composed of the proform for Ground and the proform for Figure with Motion and Path conflated. The difference is that the proform for Ground of sentence (20) represents the source for Path, while that of sentence (21), the goal for Path.

- (21) *DOG* *JAR* *JARpro + DOGpro – move.into.*
 Figure Ground Ground pro + Figure pro – Motion/Path
 ‘The dog moved into the jar.’



Figure 2: JARpro + DOGpro – move.into

Sentence (20) above describes a motion event that Figure (a frog) is moving out of Ground (a jar). Sentence (22) represents the same motion event by a similar classifier predicate with Manner incorporated. In this case, Manner is encoded by the turning movement of the wrist. Such local movement of Manner can be interpreted as ‘the manner of movement along the path’. Accompanied by facial expressions and contexts, the local movement can represent different manner of moving, such as stealthily or quietly crawling/climbing, etc.

- (22) *FROG* *JARpro + FROGpro – move.out.with wrist turning.*
 Figure Ground pro + Figure pro – Motion/Path/Manner
 ‘The frog moved out of the jar by crawling/climbing stealthily or quietly.’



Figure 3: *JARpro + FROGpro – move.out.with wrist turning*

Let's turn to a more complex representation of a motion event. Sentence (23) describes a motion event in which one Figure (a dog) is running after the other Figure (a deer), and the classifier predicate is composed of two proforms for two Figures conflated with Motion, Path and Manner. We use parentheses to group the two proforms together in (23) and Figure 4. In this case, Manner is encoded by the wiggling of the fingers. What's complex here is the proform of the deer plays the role of both Figure and Ground at the same time. As Figure, it represents the running deer, and as Ground, it represents the deer that dog is running after. The meaning that the dog is running after the deer is encoded by the arrangement of both hands.

- (23) *DOG* *(DEERpro + DOGpro) – move.with finger wiggling.*
 Figure (Figure pro + Figure pro) – Motion/Path/Manner
 ‘The dog ran after the running deer.’



Figure 4: (DEERpro + DOGpro) – move.with finger wiggling

From the above examples, we can see that the handshape, movement, location and arrangement of both hands in the classifier predicate of TSL are all meaningful. These meaningful elements are important for sign languages to conflate motion events.

3.1.2 Lexicalized verbs

In addition to classifier predicates, motion events can be encoded by lexicalized verbs, as shown in sentences (24)-(27).

The motion event in sentence (24) is expressed with a lexicalized manner verb only. A large part of upper body is involved when signing this kind of manner verbs. Supalla (1990:136) regarded such verbs as ‘locomotion verbs involving body classifiers’, noting that they refer to ‘manner of locomotion of the agent’, not ‘manner along the path’. Slobin & Hoiting (1994) analyzed them as ‘manner verbs’ referring to different types of movement of the agent, such as run, crawl, fly, etc. In this paper, we adopt Slobin & Hoiting’s analysis. The data we collected show that manner verbs in TSL express Path as well by the orientation and movement of the body.

- (24) Motion event with a lexicalized manner verb in TSL (M)

EAGLE TREE EAGLE FLY.

Figure Ground Figure Motion/Manner/(Path)
‘An eagle flew out of the tree.’

Motion events can also be expressed by lexicalized path verbs, as illustrated in sentence (25).

- (25) Motion event with two lexicalized path verbs in TSL (P)

I GO-AWAY COME HOME.

Figure Motion/Path Motion/Path Ground
‘I left and came home.’

Furthermore, manner verbs can be signed sequentially with path classifier predicates in motion events, as illustrated in sentence (26). The motion event is about Figure (a frog) crawling and jumping out of Ground (a jar). The lexicalized manner verb expresses movement of Figure (a frog) in a ‘crawling’ manner. It precedes the path classifier predicate which incorporates another Manner element expressing the frog’s movement out of the jar by ‘jumping’.

- (26) Motion event with a lexicalized manner verb and a path classifier predicate represented sequentially in TSL (M#CL)

FROG JAR CRAWL JARpro + FROGpro – be.in > jump.out.

Figure Ground Motion/Manner Ground pro + Figure pro – Motion/Path/Manner

‘The frog crawled, and it was in the jar first and jumped out of the jar.’



Figure 5: a. CRAWL



b. JARpro + FROGpro – be.in > jump.out

In our TSL Frog Story narratives, we also found several sentences like (27), in which Path verb precedes Manner verb. We feel they might not represent genuine pattern of TSL. It appears that ‘ESCAPE RUN’ is adopted from Chinese verb compound *tao^pao* ‘escape^run’ meaning ‘escape’.

- (27) *FROG ESCAPE(ESCAPE^RUN) LE.*
 Figure Motion/Path ASP
 ‘The frog escaped.’

3.2 Typological category of TSL

As mentioned in the previous section, in ASL, Manner can be encoded either as a verb or as a local movement incorporated in a path classifier predicate, and when both occur, the manner verb and the path classifier predicate are represented sequentially, as in sentence (26) above. Both Supalla (1990) and Slobin & Hoiting (1994) proposed a serial-verb analysis of ASL motion events. However, they differ with regards to whether

ASL is a ‘satellite-framed’ or ‘verb-framed’ language.

Supalla (1990) treated the fuller and more elaborated manner verb as a main verb and the path classifier predicate as a reduced form, like a satellite, and thus claimed that ASL is a satellite-framed language. In contrast, Slobin & Hoiting (1994) treated ASL as a kind of verb-framed languages because the path classifier predicate is the main verb and the manner verb is optionally encoded. Furthermore, they proposed that ASL as a serial verb language belongs to a third typological type, and referred to sign language in general as ‘complex verb-framed’ languages. Supalla (1990) and Slobin & Hoiting (1994) made their different analysis on sentence level. Later, Taub & Galvan (2001) and Galvan & Taub (2004) compared motion events of Frog Story narratives between English and ASL, and concluded that ASL is a verb-framed language because of the high percentage use of path classifier predicates.

In §3.1, we have identified various encoding patterns in TSL Frog Story narratives. The most abundant pattern involves classifier predicates of which Path is the necessary element with Manner optionally incorporated into Path. That means Path can be encoded without Manner (i.e., (20) and (21)) or Manner can be simultaneously encoded with Path (i.e., (22) and (23)). Therefore, it is reasonable to conclude that TSL is a verb-framed language. Furthermore, Manner in some cases can also be encoded as a verb occurring sequentially before a path verb (i.e., (26)). Therefore, TSL can be regarded as a complex verb-framed language.

In §2, we have shown that Mandarin Chinese is a complex verb-framed language on both sentence level and narrative discourse level. In §3, we have shown that TSL is also a complex verb-framed language. Although both Mandarin Chinese and TSL can be regarded as the same type of language, the frequencies of different encoding patterns of motion events are quite different, as shown in Table 1 below.

In Table 1, M stands for Manner; P for Path; CL for bare classifier predicate; CLm for classifier predicate with Manner incorporated; M[^]P(^D) for compound composed of Manner and Path (and Deictic); M#CLm for manner verb and classifier predicate encoded sequentially; and M#M[^]P(^D) for manner verb and manner-path compound encoded sequentially.

Table 1

	M	P	CL	CLm	M [^] P(^D)	M#CLm	M#M [^] P(^D)	Total
Chinese	2%	32%			59%		7%	100%
TSL	9%	13%	25%	44%		9%		100%

4. Is TSL a manner-salient language

As observed by Supalla (1990) and Slobin & Hoiting (1994), in ASL, Manner can be encoded either simultaneously with a path classifier predicate or with a manner verb sequentially represented before the path classifier predicate. We found it is also true for TSL. Sentence (28)-(29) illustrates the simultaneously incorporated Manner in a classifier predicate with the form of finger wiggling. The interpretation of this local movement incorporated in classifier predicate can be ambiguous, meaning ‘crawling, running, flying, etc’. The specific meaning of this local movement depends on facial expressions as well as context. Sentence (28) can be interpreted as ‘crawling’, while sentence (29), ‘running’. In contrast, in sentence (30), a manner verb is sequentially represented before a path classifier predicate and its meaning can be more clearly specified as ‘crawling’.

- (28) Classifier predicate with Manner simultaneously incorporated
FROG JARpro + FROGpro – move.out.with finger wiggling.
 Figure Ground pro + Figure pro – Motion/Path/Manner
 ‘The frog moved out of the jar by crawling.’
- (29) Classifier predicate with Manner simultaneously incorporated
DOG (DEERpro + DOGpro) – move.with finger wiggling.
 Figure (Figure pro + Figure pro) – Motion/Path/Manner
 ‘The dog moved after the running deer by running.’
- (30) Classifier predicate with a manner verb sequentially represented
FROG CRAWL JARpro + FROGpro – move.out.
 Figure Motion.Manner Ground pro + Figure pro – Motion.Path
 ‘The frog crawled out of the jar.’

Compare sentence (28) in TSL to sentence (31) below in Mandarin Chinese, they are similar in encoding both Manner and Path. The difference is that Manner in TSL is simultaneously encoded, while in Mandarin Chinese, sequentially. Furthermore, the meaning that Manner encodes in TSL is not as clearly specified as in Mandarin Chinese.

- (31) *Qingwa pa^chu boli ping.*
 Frog crawl^exit glass jar
 ‘The frog crawled out of the jar.’

It is worth of noting that Manner in spoken Spanish can be encoded on path verbs with co-speech gestures accompanied simultaneously (McNeill 1992, 2000).⁶ We illustrate this with a schema in (32). This co-speech gesture express Manner like the manner verb in (30) in TSL to yield more specific meaning for Manner. The difference is that Manner is encoded simultaneously with the co-speech gesture in Spanish, but sequentially with the manner verb before the classifier predicate in TSL. The co-speech gesture in (32) can also be compared with the finger wiggling in (28). Both are encoded simultaneously with Path. The difference is that the gesture expresses more clearly specified meaning than the finger wiggling. It appears that co-speech gesture is capable of representing a more specific meaning than Manner incorporated into path classifier predicate.

- (32) Spanish:
 Figure Motion.Path Ground
 (with co-speech gesture)
 (run, crawl, fly, etc.)

In addition to the distinction between verb-framed and satellite-framed languages, languages can also be categorized into manner-salient or non-manner-salient languages (Slobin 2004, Huang & Tanangkingsing 2005). Satellite-framed languages are all manner-salient languages. Although Mandarin Chinese is a complex verb-framed language, it is also a manner-salient language. In manner-salient languages, Manner element is expressed by both manner verbs and manner adjuncts (Slobin 2004). We illustrate the idea with sentences (33)-(34) of Frog Story narratives. In Mandarin Chinese, Manner is encoded by manner verbs to express Manner of the motion, i.e., *pao* ‘run’ in sentence (33). But it can also be encoded by adverbial adjuncts to express the inner state of the agent, i.e., *jinzhangdi* ‘nervously’ in sentence (34).

- (33) *Lu* *pao^dao* *xuanyaibian*.
 deer run^arrive cliff.side
 ‘The deer ran and arrived the side of the cliff.’
- (34) *Xiaohai* *jinzhangdi* *pa^chu* *chuang* *wai*.
 child nervously crawl^exit window outside
 ‘The child crawled and exited the window nervously.’

⁶ Relating to this, gestures can be co-expressed as Manner in Spanish (Kita & Özyürek 2003, McNeill 1992, 2000, Senghas, Kita & Özyürek 2004) and in Basque (Ibarretxe-Antuñano 2004).

We have seen that in TSL Manner can be made salient with manner verbs sequentially preceding classifier predicates. However, this pattern is not frequent, only 7% in our Frog Story narratives (see Table 1). In addition, Manner can also be simultaneously encoded with Path in classifier predicates, with high percentage, 44% (see Table 1). However, since we do not think that the typological distinction by Slobin (2004) between manner-salient and non-manner-salient sign languages is necessary, we do not want to conclude here that TSL is a manner-salient sign language.

5. Conclusion

The result of our study shows that Mandarin Chinese speakers use at least four patterns to encode motion events: (a) Path verb only, (b) a compound composed of Manner verb and Path verb, (c) Manner verb with Path adverbial and (d) Manner expressed sequentially as a verb before the compound composed of Manner and Path verb. Deictic verbs can optionally occur after the verb or verb compound. TSL signers use five patterns to encode motion events: (a) Path verb only, (b) Path classifier predicate, (c) Manner verb only, (d) Manner inflected into Path classifier predicate, and (e) Manner expressed sequentially as a verb before Path classifier predicate. One important finding in this study is that the linear-order representation of Manner-Path-Deictic verb is the most abundant pattern in Mandarin Chinese, while the simultaneous conflation of Manner-Path classifier predicate is the most abundant pattern in TSL. The other is that the sequential arrangement of Manner and Path in TSL is small in percentage, a natural consequence of the simultaneous conflation of Manner-Path classifier predicate in sign language in general. These two findings should not be surprising. For Mandarin Chinese, as a spoken language, encodes conceptual elements of motion events through audio-vocal modality, while TSL, as a signed language, encodes conceptual elements of motion events through visual-gestural modality. In audio-vocal modality, spoken languages are limited to one-dimensional linear order of representation. In contrast, in visual-gestural modality, signed languages are capable of conflating conceptual elements simultaneously.

In this study, we observe that the most prevalent encoding patterns of motion events in Mandarin Chinese and in TSL as well are verbs conflating Path and Motion and with Manner encoded optionally. Both languages are therefore treated as a kind of complex verb-framed languages. Extending previous typological studies, we propose a tripartite framework of language typology (as shown in Table 2) which reflects modality effects at the same time.

Table 2

Typology Modality	Satellite-framed	Verb-framed	Complex verb-framed
Visual-gestural			ASL, TSL, and other sign languages
Audio-oral	English, German	French, Spanish	Mandarin Chinese and other serial-verb languages ⁷

Independent of the tripartite distinction, Slobin (2004) also proposed a distinction between manner-salient or non-manner-salient languages. The visual-gestural modality of sign languages allows conceptual elements to be conflated simultaneously. The simultaneous encoding explains why classifier predicates are more abundant than lexicalized verbs in TSL, and perhaps in sign language in general. However, even though classifier predicates in sign languages are more iconic than verbs or satellites in spoken languages, without facial expressions or contexts, the meaning of the simultaneously incorporated Manner tends to be ambiguous. This observation propels us to examine further the manner-salient encoding across modalities. Speakers of languages with abundance of manner verbs (i.e., English) or manner-path verb compounds (i.e. Mandarin Chinese) in their grammar, they naturally tend to encode Manner in their narratives. In contrast, in sign languages, it is not common for manner verbs to occur alone. Therefore, we are tempted to think that without facial expressions and contexts, spoken languages like English and Mandarin Chinese are more capable than signed languages of expressing various kinds of Manner more clearly.

In the grammar of verb-framed languages like Spanish, Manner is encoded as an adjunct or a gerundive phrase. As observed in a series of works by Slobin (1996, 1997, 2000), Spanish speakers tend to less frequently and less differentiatedly express Manner of movement than English Speakers do. However, as pointed out by McNeill (1992, 2000), co-speech gestures expressing manners co-occur with path verbs. Co-speech gestures in spoken languages enable speakers to express Manner in more detail. In the same vein, facial expressions and perhaps co-sign gestures in sign languages serve the same purpose. Thus, further studies of co-speech gestures in spoken languages and facial expressions and co-sign gestures in sign languages would shed a new light to the typology of motion events within and across modalities.

⁷ Such as Creole languages and serial verb languages in Africa, Indonesia, and Papua New Guinea.

References

- Chang, Jung-hsing, Shiou-fen Su, and James H-Y. Tai. 2005. Classifier predicates reanalyzed, with special reference to Taiwan Sign Language. *Language and Linguistics* 6.2:247-278.
- Galvan, Dennis, and Sarah Taub. 2004. The encoding of motion information in American Sign Language. *Relating Events in Narrative*, Vol. 2: *Typological and Contextual Perspectives*, ed. by Sven Strömquist & Ludo Verhoeven, 191-218. Mahwah: Lawrence Erlbaum Associates.
- Huang, Shuanfan, and Michael Tanangkingsing. 2005. Reference to motion events in six western Austronesian languages: toward a semantic typology. *Oceanic Linguistics* 44.2:307-340.
- Ibarretxe-Antuñano, Iraide. 2004. Motion events in Basque Narratives. *Relating Events in Narrative*, Vol. 2: *Typological and Contextual Perspectives*, ed. by Sven Strömquist & Ludo Verhoeven, 89-112. Mahwah: Lawrence Erlbaum Associates.
- Kita, Sotaro, and Asli Özyürek. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48.1:16-32.
- Mayer, Mercer. 1969. *Frog, Where Are You? Sequel to a Boy, a Dog, and a Frog*. New York: Puffin Pied Piper.
- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about the Thought*. Chicago: University of Chicago Press.
- McNeill, David. 2000. Analogic/analytic representations and cross-linguistic differences in thinking for speaking. *Cognitive Linguistics* 11.1-2:43-60.
- Senghas, Ann, Sotaro Kita, and Asli Özyürek. 2004. Children creating core properties of language: evidence from an emerging Sign Language in Nicaragua. *Science* 305: 1779-1782.
- Slobin, Dan I. 1996. Two ways to travel: verbs of motion in English and Spanish. *Grammatical Constructions: Their Form and Meaning*, ed. by Masayoshi Shibatani & Sandra A. Thompson, 195-220. Oxford: Clarendon Press.
- Slobin, Dan I. 1997. Mind, code, and text. *Essays on Language Function and Language Type: Dedicated to T. Givón*, ed. by Joan L. Bybee, John Haiman & Sandra A. Thompson, 437-467. Amsterdam & Philadelphia: John Benjamins.
- Slobin, Dan I. 2000. Verbalized events: a dynamic approach to linguistic relativity and determinism. *Evidence for Linguistic Relativity*, ed. by Susanne Niemeier & René Dirven, 107-138. Amsterdam & Philadelphia: John Benjamins.
- Slobin, Dan. I. 2004. The many ways to search for a frog: linguistic typology and the expression of motion events. *Relating Events in Narrative*, Vol. 2: *Typological and*

- Contextual Perspectives*, ed. by Sven Strömquist & Ludo Verhoeven, 219-257. Mahwah: Lawrence Erlbaum Associates.
- Slobin, Dan I., and Nini Hoiting. 1994. Reference to movement in spoken and signed languages: typological considerations. *Berkeley Linguistics Society (BLS)* 20:487-505. Berkeley: Berkeley Linguistics Society.
- Supalla, Ted. 1990. Serial verbs of motion in ASL. *Theoretical Issues in Sign Language Research*, Vol. 1: *Linguistics*, ed. by Susan D. Fischer & Patricia Siple, 127-152. Chicago: University of Chicago Press.
- Tai, James H.-Y. 2003. Cognitive relativism: resultative construction in Chinese. *Language and Linguistics* 4.2:301-316.
- Talmy, Leonard. 1985. Lexicalization patterns: semantic structure in lexical forms. *Language Typology and Syntactic Description*, Vol. 3: *Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 57-149. Cambridge & New York: Cambridge University Press.
- Talmy, Leonard. 1991. Path to realization: a typology of event conflation. *Berkeley Linguistics Society (BLS)* 17:480-519. Berkeley: Berkeley Linguistics Society.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*, Vol. 2: *Typology and Process in Concept Structuring*. Cambridge: MIT Press.
- Taub, Sarah, and Dennis Galvan. 2001. Patterns of conceptual encoding in ASL motion descriptions. *Sign Language Studies* 1.2:175-200.

Psycholinguistique comparative du traitement des phrases: référence spéciale au chinois

Michèle Kail

CNRS et Université Paris 8

Ce chapitre est consacré aux apports du chinois dans les recherches inter-langues concernant le traitement des phrases. On examine les effets de l'absence de morphologie flexionnelle et de la présence de marqueurs spécifiques dans le traitement de monolingues chinois analysé dans le cadre du Modèle de Compétition. On souligne l'intérêt du recours au chinois dans l'étude de la dynamique des transferts L1 L2 et dans celle des jugements de grammaticalité chez des sujets atypiques (SLI). La vitalité de la psycholinguistique chinoise est clairement attestée tant empiriquement que théoriquement.

Mots clés: traitement des marqueurs syntaxiques du chinois, traitement on-line des phrases, transfert, jugements de grammaticalité, Modèle de Compétition, psycholinguistique chinoise

1. Introduction

Cet anniversaire me fournit l'occasion d'évaluer l'apport de la langue chinoise dans la psycholinguistique qui s'intéresse au traitement des phrases. Ce faisant, c'est un clin d'œil que j'adresse à l'éminent spécialiste de syntaxe chinoise qu'est Alain Peyraube. Dans cet article, après avoir rappelé les propriétés du chinois qui en font une langue cruciale pour nombre de domaines de la psycholinguistique, je concentrerai mon propos sur trois domaines dans lesquels j'ai réalisé des recherches comparatives: le traitement off-line et online des phrases chez les sujets monolingues, la dynamique des transferts chez les sujets bilingues et enfin les jugements de grammaticalité chez divers sujets présentant des dysfonctionnements du langage (SLI, aphasiques). Cet article ne constitue pas une synthèse exhaustive: il s'agit plutôt de présenter certaines recherches emblématiques qui permettront d'illustrer tout l'apport de ces travaux réalisés sur la langue chinoise dans une perspective comparative.

Durant les trois dernières décennies, un ensemble considérable de recherches a été réalisé sur les processus cognitifs et les mécanismes cérébraux sous-jacents au traitement, à l'acquisition et aux dysfonctionnements du langage. Le savoir ainsi accumulé provient massivement des langues indo-européennes (en particulier l'anglais). L'examen des

relations entre unités de traitement supposées universelles et l'impact des variations liées à la diversité des langues est progressivement devenu un thème central de la psycholinguistique générale et développementale. Cette démarche comparative inter-langues (Kail 2002) issue de courants linguistiques distincts est actualisée par deux options principales: l'une qui, dans la tradition chomskyenne, privilégie les invariants et propose une variabilité restreinte; l'autre qui, informée par les théories fonctionnalistes, place la variabilité linguistique au centre de ses entreprises. Le Modèle qui sert de cadre à nos travaux s'inscrit dans ce courant. Enfin, plus récemment, se développe une troisième direction de recherche, qui, remettant en cause l'existence de bases cognitives universelles au langage (Bowerman & Levinson 2001) admise par les deux autres courants, renoue avec la tradition relativiste de la linguistique whorfienne.

A la faveur des nombreux débats et controverses qui alimentent les avancées théoriques et empiriques, le recours à des langues non indo-européennes s'est récemment imposé non seulement comme une source précieuse d'informations à verser aux débats mais aussi comme une nouvelle fenêtre sur les processus cognitifs et une opportunité de nouvelles investigations psycholinguistiques.

Cet article considère la langue chinoise dans une acception large de l'ensemble sino-tibétain majeur, incluant le chinois mandarin et le cantonais. Parlé par un cinquième de la population mondiale, le chinois diffère notablement de la plupart des langues indo-européennes et présente des traits tout à fait spécifiques dans ses structures phonologiques, orthographiques, lexicales et syntaxiques. Ce sont les propriétés morphosyntaxiques qui feront ici l'objet de notre réflexion.

Il semble néanmoins utile de mentionner très brièvement les particularités orthographiques du chinois très éloignées de celles des langues alphabétiques ainsi que l'existence d'une phonologie tonale. Ces dimensions ont été exploitées dans les controverses actuelles qu'il s'agisse du rôle de l'orthographe et de la phonologie dans le traitement lexical ou des spécificités neuroanatomiques liées au traitement et à l'acquisition du chinois. Pour une synthèse on pourra consulter le *Handbook of East Asian Psycholinguistics*, Vol. 1: *Chinese* (Li, Tan, Bates & Tzeng 2006).

Au niveau orthographique, le chinois utilise des caractères et non des lettres alphabétiques comme unité de base de la langue écrite, lesquels entrent dans des configurations qui s'apparentent à des morphèmes plutôt qu'à des phonèmes de la langue orale. Ainsi, un caractère chinois présente une relation plus directe avec sa signification qu'un mot écrit ne le fait dans d'autres langues. Par ailleurs, la majorité des caractères chinois sont composés de multiples radicaux phonologiques et sémantiques qui, dans la plupart des cas, sont eux-mêmes des caractères. On entrevoit donc que cette organisation contraint les systèmes de traitement d'une manière différente de celle induite par les systèmes alphabétiques. Nombre de recherches ont été conduites pour

savoir si la lecture du chinois implique un accès automatique à la phonologie et si oui, comment les codes phonologiques sont activés par rapport aux traits orthographiques des caractères et à l'information sémantique des noms. Des travaux très récents indiquent que la connaissance et l'aisance dans le maniement par les enfants de la structure phonologique des caractères contribuerait de manière importante à la maîtrise de la lecture (Leong 2006, McBride-Chang & Zhong 2006).

Au niveau phonologique, le chinois fait usage d'un système tonal comportant jusqu'à neuf tons dans certains dialectes qui permet de désambigüiser les items lexicaux. Le chinois présente un très grand nombre d'homophones. Selon le *Dictionary of Frequency of Modern Chinese Words* (Beijing Language Institute 1985) 80% des monosyllabes différenciés par les tons sont sémantiquement ambigus et 55% ont 5 ou 6 homophones. Les composantes temporelles de l'accès au lexique durant la compréhension ont été étudiées à la lumière de cette importante ambiguïté des mots. Une recherche très récente (Liu, Shu & Li 2007) présente la première étude normative de 2423 mots chinois à un caractère. Pour chaque mot 15 variables ont été analysées: la fréquence, la fréquence cumulée, la densité des homophones, la fréquence phonologique, l'âge d'apprentissage, l'âge d'acquisition, le nombre de mots formés par un caractère, le nombre de significations, le nombre de composants, le nombre d'attaques, la familiarité, le caractère concret, l'iconicité, la régularité et la nature du phonème initial. La latence de dénomination du mot est affectée par la fréquence, la sémantique, les traits visuels et la consistance des caractères chinois mais pas par la phonologie. Ces résultats distinguent donc le chinois des langues alphabétiques pour lesquelles les facteurs phonologiques et orthographiques jouent un rôle beaucoup plus important que les facteurs sémantiques.

2. Propriétés particulièrement pertinentes de la morphosyntaxe chinoise

Notre article étant centré sur les apports du chinois dans le domaine du traitement des phrases, nous mentionnons brièvement les principales caractéristiques retenues dans les recherches.

En chinois, il n'y a pas morphologie flexionnelle sur les noms ou les verbes, mais certaines procédures servent d'indices pour l'assignation des fonctions syntaxiques: l'ordre des mots, le marqueur de l'objet, *ba*, le marqueur du passif, *bei*, le marqueur du datif, *gei*, ainsi que les marqueurs aspectuels, *le*, *zhe* et *zai* (Li & Thompson 1981, Li 1990).

2.1 L'ordre des mots

L'ordre de base en chinois est SVO, mais dans le langage oral on rencontre les ordres OSV, SOV et VOS qui sont contraints sémantiquement ou pragmatiquement. C'est ainsi que OSV met l'emphasis sur le caractère topicalisé du sujet de la phrase. Les phrases SOV se distinguent selon la présence ou non du marqueur *ba* précédant l'objet: sa présence réfère à des événements hautement transitifs et résultatifs; son absence est une indication pragmatique et concerne des situations dans lesquelles le locuteur apporte une information contraire aux attentes de l'auditeur. Ce dernier type d'usage de SOV est particulièrement marqué dans la mesure où une suite NNV sans marqueur *ba*, est plutôt interprétée comme OSV que SOV en chinois. Enfin, les phrases VOS sont restreintes à un usage spécifique, celui où S exprime une réflexion après coup. De l'existence de SVO, OSV et SOV, il découle que la position préverbale n'est pas en chinois associée à une fonction spécifique comme elle l'est en anglais. Par ailleurs, le fait que OVS et VSO n'existent pas en chinois signifie qu'un syntagme en position postverbale est objet. De sorte qu'en chinois, c'est la position postverbale qui constitue un indice fiable pour l'assignation de l'objet. En revanche, pour l'assignation du sujet, la position préverbale n'est pas un indice fiable. Il faut ajouter qu'il existe en chinois l'omission du sujet quand le contexte est clair, voire même l'omission d'autres constituants. Ainsi, le langage informel peut être restreint à un seul nom ou verbe et ces omissions ont pour conséquence une réduction de la fiabilité de l'ordre des mots.

2.2 Le marqueur d'objet BA

Ce marqueur concerne essentiellement les phrases SOV. En lien avec son sens originel, *ba* requiert un objet fortement affecté par l'activité dénotée par le verbe. Les verbes causatifs et résultatifs exigent des constructions avec *ba* qui obéissent à d'autres contraintes. Ainsi par exemple, *ba* concerne généralement un objet défini et un verbe structurellement complexe (les verbes monosyllabiques ne pouvant figurer seuls avec *ba*). Par ailleurs, *ba* est partiellement homophone avec le marqueur de question *ba* et le marqueur d'hésitation *ba*. Au total, les contraintes sémantiques, syntaxiques et phonologiques qui pèsent sur *ba* en réduisent la validité comme marqueur spécifique de l'objet de la phrase.

2.3 Le marqueur du passif BEI

Ce marqueur indique que c'est le second N et non le premier qui est l'agent de la phrase N *bei* NV. Comme *ba*, *bei* requiert un verbe hautement transitif, une sémantique

causative et un syntagme verbal structurellement complexe. Néanmoins les deux marqueurs diffèrent puisque *bei* indique une structure OSV alors que *ba* renvoie à une structure SOV. Peu fréquent dans le langage parlé, *bei* constitue néanmoins un indice très fiable puisque tout nom qui le suit est toujours l'agent de l'action. Deux raisons semblent concourir à sa faible fréquence. Tout d'abord, le chinois fait usage de constructions topicalisées là où d'autres langues ont recours à une construction passive. Ensuite, la construction *bei* demeure associée à un contenu adversatif même si cette association est plus ténue en chinois moderne.

2.4 La définitude et YI

En Chinois, il n'y a ni article défini ni indéfini. Les noms qui apparaissent en position préverbale sont plus volontiers considérés comme définis alors que ceux en position postverbale sont plutôt indéfinis. Dans une phrase canonique SVO, le sujet est préverbal, défini et l'objet postverbal indéfini, de sorte qu'il y a une association position, définitude et fonction. Toutefois, en raison d'ordres non canoniques, cette association n'est pas absolue. Il existe des marquages explicites du caractère indéfini: le marqueur *yi* plus un classificateur tel que *zhi* placés devant le syntagme nominal préverbal affaiblissent l'agentivité alors que placés devant un nom postverbal, ils renforcent le caractère d'objet. L'utilisation de *yi* non accentué comme marqueur de l'indéfini l'apparente au fonctionnement de l'article indéfini dans d'autres langues.

2.5 Les classificateurs nominaux

L'ensemble que représentent les classificateurs est intéressant car il se situe à l'interface de la grammaire et de la sémantique. Les noms sont assignés à des classificateurs sur la base de certaines dimensions sémantiques abstraites et particulièrement des traits relatifs à la forme physique.

3. Traiter les phrases dans le Modèle de Compétition

Alors que la recherche d'universaux constitue une entreprise classique en linguistique, ce n'est qu'à partir des années 1980, sous l'influence pionnière de Slobin (1973, 1982, 1985), que la psycholinguistique s'est engagée dans de nouvelles problématiques comparatives, exploitant judicieusement la diversité des langues pour tester des hypothèses qui ne peuvent l'être à partir d'une langue unique.

Le Modèle de Compétition (Bates & MacWhinney 1982, MacWhinney 1987b, MacWhinney & Bates 1989) se donne pour objectif d'expliquer et de prédire les

spécificités du traitement des phrases en conférant à la variabilité linguistique un rôle central. Après avoir souligné l'intérêt de ce modèle fonctionnel et probabiliste pour l'acquisition du langage (Kail 1983a, 1983b), nos propres travaux ainsi que ceux d'autres chercheurs ont contribué à une version enrichie de ce modèle qui a suscité et continue de susciter de nombreuses investigations dans différents domaines: acquisition, traitement chez des adultes normaux, des sujets aphasiques, des apprenants de langues secondes, des sujets bilingues, simulations connexionnistes et explorations neuronales. Les travaux réalisés dans ce cadre théorique ont fait l'objet de diverses synthèses (Kail 1995, 1999, 2000, 2004b) auxquelles nous nous référons pour l'essentiel.

En invoquant les grammaires fonctionnelles (Lakoff & Thompson 1975, Li & Thompson 1976, Givón 1979, 1984), le Modèle de Compétition définit la connaissance linguistique non pas comme un ensemble de règles grammaticales mais comme un réseau complexe de correspondances pondérées entre formes et fonctions. Toute langue fournit des indices lexicaux, syntaxiques, morphologiques ou prosodiques qui signalent la présence de telle ou telle fonction, par exemple la fonction sémantique d'agent. Toutefois, la force des correspondances entre les différents indices et les différentes fonctions varie aussi bien d'une langue à l'autre qu'au sein d'une même langue.

L'assignation des rôles fonctionnels dans la phrase résulte d'un appariement entre le niveau formel de la langue, où les formes de surface sont disponibles, et le niveau fonctionnel, où les significations de l'énoncé se trouvent représentées. Cet appariement est envisagé comme aussi direct que possible sans être pour autant un appariement terme à terme, une même forme pouvant être multi-fonctionnelle et une même fonction pouvant être exprimée par différentes formes. Le principe de l'appariement direct entre les formes et les fonctions signifie en réalité la possibilité pour l'analyseur d'intégrer, au même niveau, différents types d'indices dans les processus de traitement: il traite ainsi des combinaisons d'indices lexico-sémantiques (caractère animé des noms), d'indices morphologiques (morphologie nominale, verbale, marquage d'accord), d'indices relatifs à l'ordre des mots (par exemple la position préverbale) et d'indices prosodiques (par exemple l'accent contrastif).

L'appariement entre les formes et les fonctions implique des corrélations verticales alors que des corrélations horizontales sont à l'œuvre dans des appariements possibles au niveau formel (corrélations formes-formes) et au niveau fonctionnel (corrélations fonctions-fonctions). Ainsi, la position préverbale en anglais est hautement corrélée avec la fonction syntaxique de sujet (corrélation forme-fonction). D'autre part, la fonction syntaxique de sujet renvoie au niveau formel à un ensemble de caractéristiques fortement corrélées entre elles comme la position préverbale, la position initiale dans la phrase, le marquage casuel nominatif, le caractère défini, l'absence d'accentuation, etc. (corrélations formes-formes). Enfin, au niveau fonctionnel, le rôle syntaxique de sujet

est souvent corrélé avec le rôle sémantique d'agent, étant donné que ces rôles sont assignés habituellement à un même référent présentant un ensemble de caractéristiques telles que le caractère humain, le caractère animé, l'intentionnalité, etc., caractéristiques qui le désignent pour assumer ces fonctions dans la phrase (corrélations fonctions-fonctions).

Selon le Modèle de Compétition, ces trois types d'appariement ne se réalisent pas de manière indépendante mais tendent au contraire à s'organiser en coalitions. Une telle organisation caractérise toutes les langues en raison des limitations des capacités du système de traitement et des contraintes temporelles strictes auxquelles celui-ci est soumis lors de l'utilisation du langage. Comme dans la théorie des prototypes de Rosch (1978), le modèle suppose que les catégories grammaticales sont structurées sous forme de prototypes. Une illustration de sujet prototypique est le sujet en anglais qui actualise des coalitions aussi bien au niveau formel (position préverbale, accord avec le verbe) qu'au niveau fonctionnel (agent d'une action transitive, thème dans le discours). En anglais, la thématization d'un constituant de la phrase autre que le sujet nécessite la rupture de ces coalitions dont la passivation est un exemple.

Dans le cas où, consécutivement à l'activation parallèle du niveau fonctionnel et du niveau formel, diverses sources d'information revendiquent la même fonction, une compétition peut s'instaurer. Une co-évaluation de ces sources devient alors nécessaire, fortement conditionnée par la validité relative des indices concernés dans une langue donnée.

Le concept de *validité des indices* est central dans le Modèle de Compétition. Il renvoie à la valeur informative d'une source d'information donnée (par exemple la position préverbale), eu égard à une fonction ou signification particulières (par exemple, le rôle d'agent). La validité des indices est définie en termes de deux paramètres: leur disponibilité et leur fiabilité. Si un indice est toujours disponible quand on en a besoin, il a une disponibilité maximale. Si un indice n'est jamais ambigu et si sa prise en compte donne toujours lieu à une interprétation correcte, il possède une fiabilité maximale. Un exemple d'indice pourvu d'une haute validité pour la fonction d'agent est la position préverbale en anglais, qui est à la fois très disponible et très fiable. En revanche, ce n'est pas le cas ni en italien, ni en espagnol.

Une autre notion fondamentale est celle de poids des indices. A la différence de la validité, qui est une propriété objective et quantifiable de l'indice lui-même (McDonald & MacWhinney 1989), à partir d'échantillons d'inputs fournis à l'apprenant d'une langue donnée, le poids est envisagé comme une propriété subjective de l'utilisateur; c'est une construction psychologique inférée à partir du comportement des sujets dans certaines conditions expérimentales. Le poids d'un indice est la probabilité que le sujet traitant affecte à une information donnée pour assigner telle ou telle fonction. Selon ce

modèle intégratif et probabiliste, le poids d'un indice dans l'interprétation d'un énoncé est fonction de la validité de cet indice dans la langue donnée. Chez l'adulte, poids et validité des indices sont isomorphes quand les correspondances entre formes-fonctions sont «ajustées» de façon optimale. Chez l'enfant, l'ordre d'émergence des indices est censé refléter leur validité chez l'adulte.

3.1 Le traitement off-line des phrases: comparer pour généraliser

La valeur heuristique de la notion de validité a été testée dans un large éventail de recherches de compréhension portant sur des langues très variées. Dans ces études, on demande aux enfants ou adultes d'identifier l'agent ou le patient des phrases transitives actives au moyen de procédures d'investigation «off-line»: mime de l'action exprimée dans l'énoncé (enfants) et temps de décision du choix de l'agent (adultes). Un des résultats stables de ces recherches est l'existence d'une corrélation étroite entre la validité des indices et leur poids dans le traitement. Ce poids rend compte du degré d'activation d'une interprétation particulière issue de la convergence ou de la compétition des indices. Pour les locuteurs anglais, l'ordre des mots est un indice très dominant de même que l'est la préposition accusative (pour l'accès au patient de l'action) en espagnol et en hébreu. Les italiens se fondent principalement sur l'accord verbal même si dans les cas d'ambiguïté, ils ont recours à certaines combinaisons du caractère animé, de l'ordre et de l'accentuation. Les adultes français établissent leurs stratégies sur le contraste animé/non animé des noms et l'accord verbal même lorsqu'il y a compétition entre ces indices et l'ordre des mots. En allemand, l'indice le plus valide est le marquage casuel qui, bien que disponible s'avère parfois ambigu et donc assez peu fiable, de sorte que les locuteurs allemands se fondent aussi sur le contraste animé/non animé et à un moindre degré sur l'accord verbal. [Pour une synthèse des travaux (cf. Tableau 1)].

Tableau 1: Hiérarchie des indices pour assigner l'agent dans différentes langues
(adapté de MacWhinney & Bates 1989)

Arabe*		Italien	
Adultes:	Cas > VSO	Adultes:	Accord verbal > Clitique > Animé > SVO > accent
Chinois		Avant 7:	Animé > SVO > Accord verbal > Clitique > SOV, VSO
Adultes:	Animé > SVO	Japonais	
Hollandais		Adultes:	Cas > Animé > SOV
Adultes:	Cas > SVO > Animé	Polonais*	
Avant 10:	SVO > Cas > Animé	Adultes:	Cas / Genre > SVO
Anglais		Portugais*	
Adultes:	SVO > VOS, OSV > Animé	Adultes:	Animé > accord verbal > (S)VO
Accord verbal > Accent		Serbo-Croate	
5-7:	SVO > Animé > Accord verbal > SOV, VSO > Accent	Adultes:	Cas > Accord verbal > Animé > SVO, VSO, SOV
Avant 5:	SVO > Animé > Accent > SOV, VSO > Accord verbal	Avant 5:	Animé > Cas > SVO, VSO, SOV > Accord verbal
Français*		Espagnol*	
Adultes:	Accord verbal > Clitique > Animé > SVO > Accent	Adultes:	Préposition Accusative > Accord verbal > clitique > SVO, VSO
Avant 6:	SVO > Animé > VSO, SOV	Avant 10:	Préposition Accusative > SVO, VSO
Allemand		Suédois*	
Adultes:	Cas > Accord verbal > Animé > SOV, VSO, SVO	Adultes:	Animé > SVO
Grec*		Turc	
Adultes:	Cas > Animé > SVO	Adultes:	Cas > Animé > SVO
Avant 3:	Animé > Cas > SVO	Avant 2:	Cas > SVO
Hébreu		Warlpiri	
Adultes:	Cas > Accord verbal > SVO	Adultes:	Cas > Animé > SVO, VSO
Avant 10:	Cas > SVO > Accord verbal	Avant 5:	Animé > Cas > SVO, VSO
Hongrois			
Adultes:	Cas > Accord verbal > SVO		
Avant 3:	Animé > Cas > SVO > Accent		

* Recherches de M. Kail et collègues

En ce qui concerne l'acquisition, les recherches interlinguistiques indiquent une évolution générale qui recouvre trois phases distinctes: une phase initiale où les stratégies de traitement sont essentiellement contrôlées par la disponibilité des indices suivie par une phase d'apprentissage placée sous le contrôle de la validité globale des indices et une phase finale régie par la dominance des indices dans des configurations de compétition, ce que MacDonald (1987) appelle la «validité conflictuelle».

S'il est vrai que l'ordre d'émergence des indices au cours du développement est largement prédictible à partir de la validité des indices chez l'adulte, les enfants doivent parfois procéder à un réajustement de leurs stratégies de compréhension. Ainsi, en français par exemple, la notion de validité ne permet pas d'expliquer pourquoi les jeunes enfants se placent d'abord sous la dominance de l'ordre des mots (comme les enfants anglais) puis privilégient, à partir de 6 ans, le contraste animé des noms (comme les enfants italiens) dans l'assignation de la fonction d'agent (Kail 1986, 1987, Kail & Charvillat 1986). Pour prendre en compte ces phénomènes, nous avons proposé d'intégrer dans le Modèle de Compétition, un nouveau concept celui de *coût du traitement des indices*, envisagé en termes de procédures d'accès et de temps de traitement des indices. L'articulation des concepts de validité et de coût a été examinée dans une série de recherches comparatives français/espagnol (Charvillat 1988, Kail & Charvillat 1988, Kail 1989). On a ainsi montré que d'une part, un indice est d'autant plus valide pour l'attribution des fonctions et de traitement moins coûteux qu'il est plus local (par exemple les marques casuelles) et que d'autre part, avec le développement, les traitements locaux sont privilégiés aux dépens des traitements topologiques ou globaux (par exemple l'ordre des mots). La hiérarchie des indices dans les différentes langues suit un principe qui peut être résumé en termes de principe de localité.

Par ailleurs, pour chaque langue, il existe une configuration d'indices susceptible d'optimiser le traitement, la multiplication d'indices convergents pour une interprétation donnée n'entraînant pas nécessairement une réduction du temps de traitement, ainsi que l'ont montré des travaux en espagnol et en français (Kail & Charvillat 1988), en serbo-croate (Mimica, Sullivan & Smith 1994) et en hongrois (MacWhinney & Osmán-Sági 1991).

3.2 Le traitement on-line des phrases: un domaine en pleine expansion

Les recherches sur le traitement en temps réel d'unités plus étendues tels les énoncés, réalisées dans le cadre du Modèle de Compétition, partent de l'idée que le système de traitement de l'information linguistique s'engage dans le réajustement continu des assignations des rôles thématiques de la phrase. Ce système tend ainsi à combiner les différentes sources d'information linguistique en assignant le plus rapidement possible une signification aux indices, intégrant les fragments linguistiques dans des structures plus larges, compatibles avec l'information déjà traitée. Ce mode de traitement en parallèle optimise les rattachements locaux entre unités, assurant ainsi une moindre charge pour le processeur. Au contraire des relations entre formes et fonctions (corrélations verticales) qui caractérisent le traitement off-line, les relations des formes entre elles (corrélations horizontales) constituent des informations cruciales dans cette

conception du traitement en temps réel. Les langues à morphologie riche sont de bonnes candidates pour déterminer les procédures d'accès aux indices pertinents et les modalités de leur intégration dans l'énoncé en cours de traitement.

Quelques travaux conduits chez l'adulte ont commencé à mettre ces hypothèses à l'épreuve dans différentes langues: Wulfeck, Bates & Capasso (1991), pour l'anglais et l'italien; Wulfeck (1993) pour l'anglais; Devescovi, D'Amico & Gentile (1999) pour l'italien; Li, Bates & MacWhinney (1993), pour le chinois; Mimica et al. (1994), pour le serbo-croate; Kail & Bassano (1997); Lambert & Kail (2001); Kail (2004a, 2004b) pour le français; Kail & Diakogiorgi (1998), pour le grec; Kempe & MacWhinney (1999), pour le russe et l'allemand, Staroń, Bokus & Kail (2005) pour le polonais et Costa, Hub Faria & Kail (2004); Kail, Costa & Hub Faria (2008) pour le portugais.

Dans le domaine du traitement des phrases en temps réel, domaine émergent chez l'enfant, les comparaisons interlangues systématiques sont encore rares. On pourra néanmoins consulter quelques synthèses récentes de ce champ en pleine expansion (Bates, Devescovi & D'Amico 1999, Bates, Devescovi & Wulfeck 2001, Devescovi & D'Amico 2005, Kail 1999, 2000, Kempe & MacWhinney 1999). Sans doute la rareté de ces études est-elle liée aux difficultés méthodologiques mais depuis fort longtemps, Tyler & Marslen-Wilson (1981) ont souligné l'intérêt théorique de travaux développementaux portant sur le traitement en temps réel des énoncés dans une recherche pionnière sur l'anglais.

Des paramètres tels que la *perceptibilité* des indices ou leur *assignabilité* ont fait l'objet de certains de nos travaux notamment dans des langues à morphologie riche, arabe, grec, espagnol, polonais. L'assignabilité d'un indice recouvre sa capacité à être intégré à l'information déjà traitée.

Dans une langue comme le hongrois l'assignabilité des suffixes casuels est optimale: le rôle sémantique d'un nom est attribué dès que le suffixe attaché à ce nom peut être classifié.

C'est ce qu'Ammon & Slobin (1979) ont appelé «indice local». En revanche, certains indices comportent des fragments ou des éléments discontinus, ce que nous avons appelé des indices «topologiques», lesquels ont une assignabilité plus faible puisque le sujet doit attendre d'avoir l'ensemble des informations éparses avant d'effectuer l'attribution d'une fonction sémantique (les pronoms clitiques par exemple, cf. Weissenborn, Kail & Friederici 1990). L'assignabilité d'un indice dépend de certains facteurs parmi lesquels: le degré de contrastivité du marquage (flexions verbales non contrastées en français dans les verbes du 1^{er} groupe vs marquage verbal transparent en italien et en espagnol); la continuité vs discontinuité du marquage morphologique (cas de l'accord verbal en arabe marocain); la position des affixes: flexions pré vs post-verbales; la régularité vs irrégularité du marquage morphologique.

L'étude de certains des facteurs de l'assignabilité a été réalisée dans une série de recherches portant sur le traitement chez les adultes de phrases clivées en arabe marocain (Bamhamed & Kail 1996). L'intérêt de l'arabe marocain réside dans le fait que cette langue à morphologie riche possède des indices grammaticaux permettant de tester le rôle de la continuité *vs* discontinuité du marquage casuel ainsi que celui de la position des affixes (flexions pré- et post-verbales). Le verbe est en effet le pivot de la phrase où l'ordre dominant est VSO et l'assignation de la fonction d'agent se fait par la prise en compte des flexions verbales (genre, nombre, cas). Ainsi, en arabe marocain, on a montré que la position des affixes verbaux (préfixes *vs* suffixes), la continuité *vs* discontinuité et la longueur (phonème *vs* syllabe) des marqueurs morphologiques déterminent étroitement les temps d'accès à la fonction de sujet de la phrase.

Si le caractère de localité des marques, par exemple le marquage nominal casuel, constitue une bonne condition d'assignabilité, nos travaux sur le grec moderne (Diakogiorgi & Kail 1994, 1996) ont montré que dans des phrases où varient l'ordre des mots (NVN, NNV, VNN) et la morphologie casuelle, les temps d'attribution de l'agent dépendent de l'organisation séquentielle des noms. Ainsi, la juxtaposition des noms portant les marques casuelles dans un ordre NNV permet une assignation beaucoup plus rapide des rôles thématiques que lorsque les marques sont distribuées (NVN) ou placées en fin de séquence. Des résultats convergents ont été mis en évidence en italien (Bates, communication personnelle) où l'ordre VNN, qui donne le plus rapidement accès à l'accord verbal qui est l'indice dominant dans cette langue, entraîne l'assignation la plus rapide. Il est intéressant de noter qu'en italien cet ordre ne correspond pas à l'ordre dominant dans la langue (S)VO, pas plus que l'ordre NNV ne correspond à l'ordre dominant en grec. Ces résultats semblent indiquer qu'en temps réel, les contraintes du traitement mettent l'ordre des mots dans son aspect fonctionnel au service de la morphologie.

Nous avons récemment développé un programme impliquant des langues choisies en raison des contrastes pertinents eu égard aux spécificités du traitement en temps réel (français, anglais, suédois, grec, portugais, polonais). Parmi les paradigmes expérimentaux que nous avons retenus, l'un des plus productifs concerne les jugements de grammaticalité en temps réel (Blackwell, Bates & Fisher 1996, Kail & Bassano 1997). Des violations morphologiques d'accord ou d'ordre des mots sont introduites dans les phrases et doivent être détectées par les sujets le plus rapidement possible. L'hypothèse générale est que plus un indice est valide dans une langue, plus sa violation devrait être détectée facilement et rapidement. Par ailleurs ces violations interviennent au début ou à la fin de la phrase (contexte maximal). Enfin, l'empan de la violation contraste des violations intrasyntagmes et intersyntagmes. Ces deux derniers facteurs sont supposés davantage liés à des contraintes contextuelles et structurales considérées

comme universelles que par les spécificités des langues reflétées par la validité.

Les résultats très brièvement résumés indiquent que conformément à nos prédictions, dans les cinq langues, et pour chaque niveau d'âge, les violations tardives sont détectées significativement plus rapidement que les violations précoces. Cet effet qui vaut aussi bien pour les violations d'ordre que les violations morphologiques d'accord, est interprété comme l'indication que les locuteurs utilisent dès 6 ans leurs connaissances grammaticales pour construire des hypothèses tout au long de leur traitement, utilisant à profit le contexte morphosyntaxique pour anticiper la suite à venir.

Dans les différentes langues, les violations intrasyntagmatiques sont détectées significativement plus rapidement que les violations intersyntagmatiques. Un tel résultat souligne que le traitement en temps réel est dépendant des capacités de la mémoire de travail et il indique que le coût du traitement de l'indice, peut devenir plus déterminant que sa validité. L'influence globale de ce facteur est modulée selon les langues par des interactions avec les autres facteurs.

Enfin, la prédiction selon laquelle plus un indice est valide dans une langue, plus sa violation devrait être détectée rapidement, n'est que partiellement vérifiée. En français, en portugais, en grec, les résultats sont conformes à cette prédiction: pour chaque âge, les violations morphologiques d'accord sont plus rapidement détectées que les violations d'ordre des mots et cet effet augmente avec l'âge. En anglais, contrairement aux prédictions issues de la validité des indices, ce ne sont pas les violations de l'ordre des mots qui sont les plus rapidement détectées mais les violations morphologiques. En suédois, les deux types de violations entraînent des temps comparables. Ces résultats indiquent de manière claire que, contrairement à ce qui était jusque là admis, la validité des indices ne conduit pas nécessairement à de bonnes prédictions pour le traitement en temps réel.

Des analyses plus détaillées ont été conduites qui indiquent des similitudes et des différences entre les langues et soulignent des spécificités développementales. La conclusion essentielle de ces travaux peut être ainsi résumée: le concept de validité des indices linguistiques, central dans le Modèle de Compétition et très heuristique pour le traitement off-line, ne permet pas de faire à tout coup de bonnes prédictions pour le traitement on-line à tout le moins pour une langue où l'ordre des mots est très dominant comme l'anglais ou à un moindre degré le suédois. Ces résultats, plaident en faveur de l'introduction dans le Modèle de Compétition d'un nouveau concept que nous proposons, "*cue localness*" qui permet de préciser le concept global de "*cue cost*" et qui contraint les performances dès l'âge de 6 ans (Kail 2004a, Kail 2012a, Kail, Costa & Hub Faria 2008, Kail, Kihlstedt & Bonnet 2012).

De manière convaincante, Kempe & MacWhinney (1999) ont récemment montré que des langues telles que l'allemand et le russe qui présentent un parallélisme quant à

la validité des indices morphologiques peuvent susciter des traitements en temps réel relativement différents. A l'inverse, nous avons montré que des langues très contrastées (anglais et langues romanes par exemple) peuvent susciter des traitements en temps réel comparables.

Ces premiers travaux pionniers devraient encourager les recherches futures à mettre en œuvre des comparaisons interlangues systématiques dans des tâches variées, susceptibles de fournir des données cruciales sur le développement du traitement du langage examiné sous l'angle de ses contraintes temporelles. Certains principes sont actuellement proposés comme organisant la microstructure temporelle du traitement tels l'assignabilité, la prédictibilité (Lambert & Kail 2001) et la localité des indices linguistiques, mais l'investigation sur un éventail plus large de langues doit être poursuivie.

3.3 L'interprétation des phrases en chinois: quels indices pour quels traitements?

Les travaux réalisés sur le traitement des phrases en chinois dans le cadre du Modèle Compétition sont encore peu nombreux, mais ils constituent d'ores et déjà un apport irremplaçable pour illustrer le fonctionnement d'un système de traitement reposant sur l'interaction des indices. Nous rapportons ici les résultats de Li, Bates, Liu & MacWhinney (1992) et Li, Bates & MacWhinney (1993). Dans la première de ces études sont comparés les traitements off-line et on-line alors que la seconde est consacrée à l'examen des temps de décision pour choisir l'agent de la phrase. Les phrases présentées aux locuteurs chinois intègrent les dimensions que nous avons évoquées en 1. Les résultats des expériences off-line et on-line sont très consistants et mettent en évidence un effet très marqué du contraste animé des noms, un effet notable de l'ordre des mots et une interaction entre les deux. Quand l'ordre des mots et le contraste animé entrent en compétition, le dernier l'emporte. La hiérarchie des indices est la suivante: marqueur du passif *bei* si présent > contraste animé des noms > ordre des mots > marqueur de l'objet «ba» > marqueur de l'indéfini *yi*. Cette hiérarchie dans laquelle les marqueurs «semi-morphologiques» encadrent le contraste animé et l'ordre des mots est très spécifique et semble indiquer que le poids des indices en chinois ne suit pas une progression univoque (par exemple morphologie > ordre des mots) mais résulte d'interactions complexes entre les divers types d'indices.

Etant donné que le passif *bei* est rare dans la langue informelle, c'est le contraste animé qui est l'indice le plus valide en chinois.

L'importance de l'ordre des mots ne se traduit pas comme en anglais. En effet, dans la compétition ordre des mots et caractère animé, ce dernier l'emporte en chinois,

mais l'issue de la compétition est inverse en anglais où l'ordre des mots l'emporte systématiquement. De plus, si les locuteurs anglais se fondent sur la position préverbale pour déterminer le sujet de la phrase, les locuteurs chinois prennent la position postverbale comme indice pour déterminer l'objet de la phrase. Les locuteurs chinois présentent une stratégie du second nom comme agent dans les suites non canoniques VNN et NNV. C'est ainsi que VNN est interprété comme VOS et l'usage plus fréquent de l'ordre OSV par rapport à l'ordre marqué SOV dans les phrases simples tend à promouvoir le 2^{ème} nom comme sujet dans NNV.

Alors que le marqueur du passif *bei* quand il est présent joue un rôle prédominant dans l'identification de l'agent, ces recherches on-line n'ont pas mis en évidence un effet du marqueur de l'objet *ba* dans l'identification du patient. Les auteurs avancent plusieurs raisons pour expliquer ce résultat inattendu. Tout d'abord, ils notent que *ba* a pour fonction de marquer l'objet préverbal dans des phrases SOV. Contrairement au marquage de l'accusatif dans les langues flexionnelles, il ne marque pas l'objet dans des positions postverbales. Par ailleurs, les constructions *ba* véhiculent une signification fortement transitive et causative liée à l'origine du verbe *ba* (Sun 1991); enfin, *ba* marque un objet plutôt défini qu'indéfini. Une telle plurifonctionnalité rendrait plus opaque son rôle d'indicateur de fonction syntaxique dans la phrase. Contrairement à *ba*, *bei* n'est pas associé à la définitude et remplit une fonction uniforme.

Le marquage de l'indéfini affecte les phrases à ordre NVN interprétées plus rapidement comme SVO quand «*yi*» marque le second nom. En revanche, il n'a pas d'impact sur les suites non canoniques VNN et NNV qui sont des constructions déjà marquées. Selon les auteurs, le marquage indéfini pourrait relever d'une catégorie plus large d'indices qui n'auraient pas d'effet absolu mais un effet contingent de l'ordre dans lequel ils sont impliqués. Une interprétation de ce type a été avancée pour expliquer les effets de l'accentuation contrastive ou ceux des pronoms clitiques dans des langues comme l'italien, l'espagnol ou le français (Kail 1989, Bates & MacWhinney 1989).

L'utilisation des indices en chinois apparaît comme plus complexe que dans d'autres langues. En l'absence de morphologie flexionnelle, les locuteurs chinois utilisent tous les indices disponibles en les intégrant d'une manière interactive. Dans une langue comme le chinois, les locuteurs ne peuvent s'appuyer sur un seul type d'information car certaines marques sont ambiguës ou peu disponibles. Le fait que le contraste animé des noms soit l'indice le plus important en chinois est tout à fait compatible avec le caractère de langue orientée par le contexte dans laquelle les locuteurs utilisent l'information sémantique et pragmatique pour déterminer l'agent d'une action. Contrairement à beaucoup d'autres langues, l'information syntaxique n'est pas déterminante dans le traitement des phrases.

Il n'est donc pas étonnant que la première recherche consacrée au rôle du contexte menée dans le cadre du Modèle de Compétition concerne une comparaison entre le chinois et l'anglais. Dans une première recherche, Su (2001) analyse comment, dans le traitement des phrases, les locuteurs chinois et anglais intègrent des informations contextuelles et quelle est leur incidence par rapport aux indices dominants que sont l'ordre des mots en anglais et le contraste animé des noms en chinois. Les résultats indiquent que les monolingues anglais ne tiennent pas compte des indices contextuels apportés par une proposition précédant la phrase et biaisant l'interprétation de N1 ou N2 comme agent. Les monolingues chinois quant à eux font usage du contexte, bien que l'effet de celui-ci demeure limité, expliquant 29% de la variance. Ces résultats s'apparentent aux résultats antérieurs montrant le rôle d'appoint de certains indices (les clitiques dans les langues romanes ou l'accent en allemand) dans le traitement de langues où l'ordre des mots est assez flexible. Pour les apprenants de l'anglais et du chinois comme L2, les effets sont comparables en chinois à ceux des monolingues par contre, les anglais qui ont le chinois comme L2 font un peu plus usage du contexte dans le traitement des phrases non canoniques. Dans une seconde recherche, Su (2004) a proposé aux sujets des contextes discursifs plus étendus. Dans ces conditions, les locuteurs anglais font un plus grand usage des indices contextuels que ne l'a montré la première recherche et les locuteurs chinois font toujours plus d'usage du contexte que les locuteurs anglais. Ces résultats illustrent bien la nécessité de recours à d'autres langues que l'anglais pour comprendre les déterminants de l'interprétation des phrases en temps réel et d'autres recherches interlangues sont encore nécessaires avant de pouvoir évaluer les effets complexes du contexte.

4. Dynamique des transferts dans le traitement des phrases par des sujets bilingues

En raison de sa capacité à mettre à jour des différences qualitatives et quantitatives dans le traitement et l'acquisition des sujets monolingues, le Modèle de Compétition est un outil performant pour aborder les spécificités de l'acquisition bilingue, notamment sous l'angle des transferts (Kilborn 1994, MacWhinney 1987b). La connaissance précise des indices dominants et de leur dynamique développementale dans la langue source et la langue cible permet de formuler des hypothèses précises sur la nature des transferts attendus.

Différents types de patterns peuvent être avancés pour rendre compte du traitement des phrases par les sujets bilingues. *La Différenciation*: un ensemble de stratégies est exclusivement appliqué dans les contextes où L2 est utilisée, de telle sorte que le sujet se comporte comme un monolingue dans chacune des langues. Ce pattern est considéré

comme révélant une excellente maîtrise dans les deux langues. Le *transfert proactif*: les stratégies de traitement de L1 sont appliquées aux deux langues. Un tel pattern est attendu chez les bilingues tardifs. Le *transfert rétroactif*: par feed-back, les stratégies de L2 sont appliquées à L1. Ce pattern est attendu chez les bilingues précoces. Enfin, l'*amalgame*: des stratégies nouvelles sont adoptées au cours de l'acquisition de L2 et sont assimilées au sein d'un ensemble de stratégies appliquées aux deux langues. Ces hypothèses ne sont pas mutuellement exclusives; chacune de ces possibilités peut s'avérer juste à tel ou tel moment de l'acquisition d'une langue seconde (L2).

Les premiers travaux réalisés dans le cadre du Modèle de Compétition ont surtout mis en évidence le transfert proactif et l'amalgame chez des sujets dont la L2 est par exemple: *l'italien* (Gass 1987), *l'anglais* (Gass 1987, Harrington 1987, Hernandez, Bates & Avila 1994, Kilborn 1989, Kilborn & Cooreman 1987), *le hollandais* (McDonald 1987), *le français* (Heilenman & McDonald 1993, McDonald & Heilenman 1991) ou encore *le japonais* (Sasaki 1991, 1994) ou *l'arabe* (Bamhamed & Kail 1996). En revanche, il faut noter que la différenciation et le transfert rétroactif ont été peu mis en évidence dans les paires de langues étudiées dans la littérature sur l'acquisition de L2, où le plus souvent, l'anglais est L2 et où par conséquent, la langue première des sujets est une langue à morphologie plus riche que l'anglais.

C'est tout l'intérêt du recours au chinois comme L1 ou L2 pour l'analyse de certains déterminants des types de transfert dans une vision dynamique des habiletés acquises au cours de l'apprentissage de L1 et L2. Une recherche particulièrement intéressante (Liu, Bates & Li 1992) concerne les stratégies de compréhension de phrases chez des locuteurs bilingues chinois-anglais et anglais-chinois en examinant le rôle de l'âge d'exposition à la L2. Les auteurs inscrivent leurs travaux dans le cadre d'une réflexion relative à l'hypothèse de période critique avancée pour expliquer que la réussite dans l'acquisition d'une langue seconde dépend de l'âge auquel le sujet est confronté à cette langue. Les sujets bilingues ont été exposés à L2 (anglais) avant 4 ans, entre 6 et 10 ans, entre 12 et 16 ans et après 20 ans (bilingues tardifs qui inclut aussi un groupe d'anglais exposés à L2 (chinois)). Pour ces groupes de bilingues, on dispose des données monolingues sur les mêmes tâches expérimentales de choix de l'agent dans des phrases où varient l'ordre des mots (NVN, VNN, NNV) indice dominant des monolingues anglais et le contraste animé des noms de la phrase (AA, AI, IA), indice dominant des monolingues chinois. Les résultats révèlent des effets relativement complexes. Comme attendu, les bilingues tardifs montrent des transferts de stratégies L1 vers L2, en d'autres termes appliquent à l'anglais (L2) les stratégies sémantiques du chinois (L1) ou au chinois (L2) les stratégies d'ordre des mots de l'anglais. Les bilingues exposés à l'anglais entre 6 et 10 ans d'âge présentent un pattern de différenciation, appliquant à chacune des langues les stratégies qui lui sont spécifiques.

Les sujets exposés avant 4 ans à l'anglais, présentent un pattern très clair de transfert rétroactif, utilisant les stratégies de l'anglais pour les deux langues. Ce même pattern est retrouvé chez des sujets arrivés aux USA entre 12 et 16 ans dont les questionnaires semblent indiquer un relatif rejet de la langue maternelle, entraînant une perte de la pratique de L1 à l'adolescence. De sorte que les patterns de transferts observés obéissent à une dynamique complexe d'activation/inhibition qui régit l'acquisition de L2 mais aussi le maintien de L1, dynamique dans laquelle, l'âge d'exposition mais aussi la pratique quotidienne des deux langues sont en interaction.

Nombre d'études consacrées aux divers types de transfert n'ont envisagé qu'un niveau de maîtrise de L2 et par suite il est difficile de savoir si les patterns de transfert changent en fonction du niveau atteint en L2. C'est cette relation qu'aborde l'étude d'I-Ru Su (2001) impliquant des sujets chinois apprenant l'anglais (E-FL) et des sujets anglais apprenant le chinois (C-FL). Ces apprenants sont répartis en trois groupes selon la durée de leur apprentissage de L2: débutants (2-3 ans); intermédiaires (6 ans) et avancés (10 ans). Les phrases présentées aux sujets, comme dans la recherche précédente, varient quant au contraste animé/non animé des noms et à l'ordre des mots. Les E-FL montrent une dépendance croissante vers la syntaxe, et un abandon progressif des stratégies sémantiques du chinois pour traiter l'anglais, en fonction de leur niveau de maîtrise (transfert proactif). Par ailleurs, avec l'augmentation de la maîtrise de l'anglais, les chinois utilisent davantage les stratégies d'ordre pour traiter leur langue (L1) (transfert rétroactif). Avec l'expérience, les sujets anglais prennent conscience de l'importance du caractère animé des noms pour traiter les phrases chinoises. Mais, même chez les sujets avancés ce facteur est loin d'être aussi important que l'ordre des mots. Par ailleurs, le transfert rétroactif est beaucoup moins important pour les C-FL que pour les E-FL et n'est pratiquement mis en évidence que chez les sujets avancés. Il en résulte une dissymétrie surtout notable dans l'importance de l'influence de L2 sur L1. La question de savoir pourquoi ce transfert est plus important quand on apprend une langue L2 fondée sur l'ordre des mots plutôt que fondée sur la sémantique reste posée. Il conviendrait ici d'évaluer sur d'autres paires de langues (italien/anglais) la robustesse de cette dissymétrie du transfert rétroactif. Il faudrait aussi disposer de données sur le traitement en temps réel chez les bilingues puisque les recherches précédentes concernent la compréhension off-line.

En effet, très peu de travaux ont été conduits sur le traitement en temps réel des énoncés chez les sujets bilingues, lesquels semblent indiquer une prévalence des stratégies d'amalgame chez les sujets adultes (Hernandez, Bates & Avila 1994, Bamhamed & Kail 1996, Kempe & MacWhinney 1999, Reyes & Hernández 2006).

Nous avons entrepris d'examiner, dans un programme consacré au coût du traitement en temps réel, la flexibilité des stratégies de compréhension de phrases et la dynamique

des transferts chez des sujets bilingues en maintenant constants certains facteurs comme l'âge d'acquisition et la nature de l'exposition à L2. Notre programme implique plusieurs équipes (à Lund, à Lisbonne, à Paris, à Londres). Les sujets bilingues choisis sont des bilingues simultanés de divers niveaux d'âge, de 6 ans, 8 ans, 10 ans et adultes. Ces sujets ont été dès leur naissance confrontés à deux langues dans une configuration un parent, une langue. Toutefois, en raison de leur résidence dans un pays, et malgré une scolarité dans des lycées internationaux, ces sujets ont une langue dominante. Le français sert de pivot aux autres langues, anglais, suédois, portugais, chacune des langues étant tantôt forte, tantôt faible. Portugais et français sont des langues romanes à morphologie riche, suédois et anglais sont des langues à moindre morphologie et l'anglais notamment à forte contrainte d'ordre des mots. Les données bilingues sont confrontées à celles des monolingues de même âge évoqués précédemment et soumis à la même tâche de jugements de grammaticalité en temps réel.

Les résultats globaux sont complexes et dans cet article nous ne mentionnons que certains d'entre eux (pour une synthèse, cf. Kail 2012b). Dans leur langue dominante, les bilingues font preuve d'une capacité de détection des violations grammaticales comparable à celle des monolingues. En revanche, dans leur langue non dominante, jusqu'à 10 ans, les violations sont moins bien détectées que chez les monolingues correspondants. En ce qui concerne les temps de détection, les bilingues tendent à être plus lents que les monolingues aussi bien dans la langue dominante que dans la langue non dominante. D'autres travaux utilisant des données chronométriques et d'imagerie ont également fait mention de cette tendance dont une interprétation possible serait la persistance de l'activation de l'autre langue qui ralentirait l'accès à la langue testée (Grosjean 1998, Hernandez, Martinez & Kohnert 2000, Reyes & Hernández 2006).

Le coût du traitement est examiné comme chez les monolingues à travers le poids des facteurs généraux du traitement: le contexte (P) qui contraste les violations précoces et tardives; les contraintes structurales (S) qui contrastent les violations intrasyntagmes et intersyntagmes. De plus, les spécificités des langues contrastent les violations de la morphologie et de l'ordre des mots (T). Le tableau 2 est un exemple qui concerne les bilingues français-anglais qui ont le français comme langue dominante au regard des monolingues correspondants. Dans ce tableau, sont exprimés les % de la variance des temps en fonction des âges pour les trois facteurs P, S et T.

Tableau 2: Facteurs du coût (% de variance des temps) pour les monolingues bilingues français-anglais en fonction de l'âge.

	Age 6-7			Age 8-9			Age 10-11			Adultes		
	T	S	P	T	S	P	T	S	P	T	S	P
Monolingues français			85			87	53	38		60	16	14
Bilingues français		6	17 TxP 68		7	29 TxP 55		11	37 TxP 28	25	30 TxS 28	16
Monolingues anglais			89			71			73		71	
Bilingues anglais		TxS 8	87		TxS 6	87		TxS 15	79	10	23	27 TxP 36

En premier lieu, on constate que le coût du traitement tant chez les monolingues français qu'anglais est dominé jusqu'à 10 ans par le facteur général de contexte (P). A partir de 10 ans, la morphologie (T) en français explique 63% de la variance. Chez les adultes anglais, c'est la structure qui devient le facteur dominant (71% de la variance). Dans les deux langues, le poids du contexte diminue au profit de facteurs internes à la langue. En second lieu, dans les deux langues des bilingues, on constate d'importantes interactions entre les facteurs, absentes chez les monolingues interactions d'autant plus importantes que la langue (ici le français) est dominante. En particulier, les bilingues testés en français montrent une forte interaction (TP) entre le contexte et la morphologie et chez les adultes, une interaction entre la morphologie et la structure (TS). Ces interactions indiquent que les contraintes qui pèsent sur les monolingues dans le traitement en temps réel sont pour partie qualitativement différentes de celles qui régissent les monolingues.

Ces données en temps réel indiquent à leur façon que la comparaison avec les monolingues n'est pas dépourvue d'intérêt même s'il est vrai comme Grosjean (1989) l'a affirmé que les bilingues ne sont pas deux monolingues dans le même cerveau.

Par ailleurs, il est intéressant de noter que le poids des facteurs structuraux chez les adultes monolingues et bilingues dans les langues que nous avons étudiées trouve un écho dans une recherche très récente utilisant les potentiels évoqués recueillis au cours du traitement en chinois de divers niveaux de la structure hiérarchique des constituants dans les constructions *ba* (Jiang & Zhou 2009). En utilisant le paradigme des violations des contraintes structurales, les auteurs mettent en évidence un engagement différentiel des structures neuronales lié au niveau (bas ou élevé) de la structure hiérarchique de la

phrase, résultat compatible avec les travaux de Friederici (2004), Friederici & Meyer (2004).

5. Jugements de grammaticalité et troubles du langage: le cas particulier du chinois

Les jugements de grammaticalité en temps réel ont été utilisés pour comprendre certaines spécificités du traitement de sujets présentant des pathologies du langage, enfants SLI (Specific Language Impairment) ou dysphasiques et sujets aphasiques. En dépit d'une audition normale, de scores également normaux dans les tests d'intelligence non verbale et d'absence de lésions neurologiques, les enfants SLI se caractérisent par d'importantes difficultés dans le traitement des composantes morphosyntaxiques du langage. Pour certains auteurs, le déficit relève d'un trouble spécifique des représentations grammaticales, pour d'autres, il résulte de difficultés perceptives ou mnémoniques qui affectent les procédures de traitement de l'input linguistique.

Le Modèle de Compétition nous semble pertinent pour mettre à l'épreuve certaines hypothèses. L'une d'elles est que les enfants SLI pourraient rencontrer des difficultés quand le coût des indices excède les ressources disponibles ou lorsque les structures à traiter comportent des compétitions d'indices.

Une série de recherches a été réalisée avec des enfants français SLI de 12 ans soumis à une tâche de détection de violations grammaticales prenant en compte les facteurs déjà présentés (Kail & Plaza 2004). En premier lieu, ces enfants ont des scores élevés de non détection des violations (44%). Le facteur qui explique massivement ces résultats est le facteur de structure: les erreurs concernent les violations intersyntaxes, comme chez les enfants normaux plus jeunes. Lorsque ces enfants SLI détectent correctement les violations, leur temps moyen de détection est trois fois plus élevé que les temps des enfants normaux de 10 ans. L'analyse des différents facteurs (contexte, structure et type de violation) qui contrôlent le coût du traitement en temps réel indique que les enfants SLI ont un profil qui les apparente aux enfants de 6;8 ans pour lesquels le facteur dominant expliquant les différences de temps est le contexte (85% de la variance) alors que pour les enfants français de 10 ans, 65% de la variance des temps est expliquée par la plus grande facilité de traitement de la morphologie. Ces travaux semblent plaider en faveur d'une vision qui interroge la spécificité des SLI en tant que c'est principalement un fort ralentissement dans l'accès aux informations qui les caractérisent. Une conclusion voisine est proposée par Dick, Wulfeck, Krupa-Kwiatkowski & Bates (2004) dans une recherche avec des sujets SLI anglais.

En raison des effets d'interaction de la variabilité de l'ordre des mots, du caractère optionnel des mots fonctionnels, de l'omission de certains constituants, il est très

difficile de créer des violations grammaticales non ambiguës en chinois. D'ailleurs la langue chinoise n'a pas de terme correspondant à «agrammatical», ce qui implique de longues explications aux sujets soumis aux jugements de grammaticalité. Dans ces conditions, l'étude des jugements de grammaticalité chez les patients aphasiques chinois vient élargir le débat sur la nature des déficits grammaticaux dont nombre de recherches sur les langues indo-européennes ont discuté l'alternative suivante: déficit des connaissances ou déficit de l'accès aux connaissances linguistiques.

Dans une série de recherches, Lu, Bates, Li, Tzeng, Hung, Tsai, Lee & Chung (2000) ont utilisé un sous ensemble de constructions qui permettent de réaliser des phrases agrammaticales (marquage de l'objet et du passif; classificateurs de noms; marquage de l'aspect) et ont comparé les jugements de sujets normaux, d'aphasiques de Broca et de Wernicke. Les résultats de ces recherches indiquent tout d'abord que les aphasiques chinois demeurent sensibles aux caractéristiques grammaticales des constructions choisies au sein de leur langue, suggérant qu'un certain savoir est préservé dans l'aphasie en dépit des déficits constatés dans l'accès à ce savoir en temps réel. En second lieu, contrairement aux résultats obtenus dans de nombreuses recherches sur d'autres langues montrant que chez les adultes normaux, le pourcentage de détections correctes est très élevé, les résultats des adultes chinois sont beaucoup plus contrastés. C'est ainsi que leur acceptation des phrases grammaticales est élevée (95%) alors que le pourcentage de détections correctes oscille entre 91% pour les violations d'ordre sur les items *ba/bei* et 66% pour les violations concernant les omissions de classificateurs. On peut en déduire que les jugements de grammaticalité sont en chinois un phénomène probabiliste pour les locuteurs très directement lié au caractère de langue dépendante du contexte comme nous l'avons précédemment mentionné. Il est donc d'autant plus intéressant de mettre en évidence une sensibilité des aphasiques à la grammaticalité des phrases en chinois même si les jugements y revêtent un caractère beaucoup plus fragile que dans d'autres langues.

6. Pour ne pas conclure

Il serait paradoxal de conclure alors que les recherches sur le traitement en temps réel sont en pleine expansion. Ces recherches contribuent largement à la compréhension conjointe des phénomènes de plasticité comportementale liée aux contraintes cognitives et linguistiques et de plasticité neuronale. De ce dernier point de vue, on assiste à une véritable explosion de travaux sur les bases neurales du traitement du langage qui ont recours à des langues diverses et notamment au chinois. Ces travaux sont encore le plus souvent restreints aux composants phonologiques ou lexicaux mais c'est un challenge des années futures que d'aborder les phénomènes d'interaction des composants et

d'accroître les travaux en direction des dimensions phrastiques et discursives. Un autre challenge important serait une approche plus informée ou dotée de critères opérationnels pour lier les traitements morphosyntaxiques à l'idée de distance entre langues, en d'autres termes de revisiter les propositions typologiques.

Références

- Ammon, Mary Sue, and Dan I. Slobin. 1979. A cross-linguistic study of the processing of causative sentences. *Cognition* 7.1:3-17.
- Bamhamed, Mohamed, and Michèle Kail. 1996. Mode d'appropriation des langues et traitement psycholinguistique: cas de l'arabe marocain et du français. *Diálogos Hispánicos* 17:45-66 (N° Spécial: les langues dans l'Europe communautaire).
- Bates, Elizabeth, and Brian MacWhinney. 1982. Functionalist approaches to grammar. *Language Acquisition: The State of the Art*, ed. by Eric Wanner & Lila R. Gleitman, 173-218. Cambridge & New York: Cambridge University Press.
- Bates, Elizabeth, and Brian MacWhinney. 1989. Functionalism and the competition model. *The Crosslinguistic Study of Sentence Processing*, ed. by Brian MacWhinney & Elizabeth Bates, 3-73. Cambridge & New York: Cambridge University Press.
- Bates, Elizabeth, Antonella Devescovi, and Simona D'Amico. 1999. Processing complex sentences: a cross-linguistic study. *Language and Cognitive Processes* 14.1:69-123.
- Bates, Elizabeth, Antonella Devescovi, and Beverly Wulfeck. 2001. Psycholinguistics: a cross-language perspective. *Annual Review of Psychology* 52:369-398.
- Blackwell, Arshavir, Elizabeth Bates, and Daniel Fisher. 1996. The time course of grammaticality judgment. *Language and Cognitive Processes* 11.4:337-406.
- Bowerman, Melissa, and Stephen C. Levinson. (eds.) 2001. *Language Acquisition and Conceptual Development*. Cambridge & New York: Cambridge University Press.
- Charvillat, Agnès. 1988. *Etude développementale de la compréhension et du traitement en temps réel des phrases pronominales en français et en espagnol*. Thèse de Doctorat de Psychologie non publiée, Université de Paris V.
- Costa, Maria Armanda, Isabel Hub Faria, and Michèle Kail. 2004. Semantic and syntactic cues' interaction on pronoun resolution in European Portuguese. *Proceedings of 5th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2004)*, ed. by Antonio Branco, Tony McEnery & Ruslan Mitkov, 45-50. Lisboa: Edições Colibri.
- Devescovi, Antonella, and Simona D'Amico. 2005. The competition model: cross-linguistic studies of online processing. *Beyond Nature-Nurture: Essays in Honor of Elizabeth Bates*, ed. by Michael Tomasello & Dan I. Slobin, 165-191. Mahwah: Lawrence Erlbaum Associates.

- Devescovi, Antonella, Simona D'Amico, and Paola Gentile. 1999. The development of sentence comprehension in Italian: a reaction time study. *First Language* 19:129-163.
- Diakogiorgi, Kleopatra, and Michèle Kail. 1996. De l'intérêt psycholinguistique d'une langue minoritaire: cas du grec. *Diálogos Hispánicos* 17:111-135.
- Dick, Frederic, Beverly Wulfeck, Magda Krupa-Kwiatkowski, and Elizabeth Bates. 2004. The development of complex sentence interpretation in typically developing children compared with children with specific language impairments or early unilateral focal lesions. *Developmental Science* 7.3:360-377.
- Friederici, Angela D. 2004. The neural basis of syntactic processes. *The Cognitive Neurosciences III*, ed. by Michael S. Gazzaniga, 789-801. Cambridge: MIT Press.
- Friederici, Angela D., and Martin Meyer. 2004. The brain knows the difference: two types of grammatical violations. *Brain Research* 1000.1-2:72-77.
- Gass, Susan M. 1987. The resolution of conflicts among competing systems: a bidirectional perspective. *Applied Psycholinguistics* 8.4:329-350.
- Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.
- Givón, Talmy. 1984. *Syntax: A Functional-typological Introduction*, Vol. 1. Amsterdam & Philadelphia: John Benjamins.
- Grosjean, François. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36.1:3-15.
- Grosjean, François. 1998. Transfer and language mode. *Bilingualism: Language and Cognition* 1.3:175-176.
- Harrington, Michael. 1987. Processing transfer: language-specific processing strategies as a source of interlanguage variation. *Applied Psycholinguistics* 8.4:351-377.
- Heilenman, L. Kathy, and Janet L. McDonald. 1993. Processing strategies in L2 learners of French: the role of transfer. *Language Learning* 43.4:507-557.
- Hernandez, Arturo E., Elizabeth Bates, and Luis X. Avila. 1994. On-line sentence interpretation in Spanish-English bilinguals: What does it mean to be in between? *Applied Psycholinguistics* 15.4:417-446.
- Hernandez, Arturo E., Antígona Martínez, and Kathryn Kohnert. 2000. In search of the language switch: an fMRI study of picture naming in Spanish-English bilinguals. *Brain and Language* 73.3:421-431.
- Jiang, Xiaoming, and Xiaolin Zhou. 2009. Processing different levels of syntactic hierarchy: an ERP study on Chinese. *Neuropsychologia* 47.5:1282-1293.
- Kail, Michèle. 1983a. L'acquisition du langage repensée: les recherches interlangues. Partie I. Principales propositions théoriques. *L'Année Psychologique* 83:225-258.
- Kail, Michèle. 1983b. L'acquisition du langage repensée: les recherches interlangues. Partie II. Spécificités méthodologiques et recherches empiriques. *L'Année Psychologique* 83:561-596.

- Kail, Michèle. 1986. Validité et coût des indices linguistiques dans la compréhension des phrases. *Bulletin de Psychologie* (Numéro spécial “Jugement et Langage”) 39:387-397.
- Kail, Michèle. 1987. The development of sentence interpretation strategies from a cross-linguistic perspective. *First and Second Language Acquisition Processes*, ed. by Carol Wollman Pfaff, 28-55. Cambridge: Newbury House.
- Kail, Michèle. 1989. Cue validity, cue cost, and processing types in sentence comprehension in French and Spanish. *The Crosslinguistic Study of Sentence Processing*, ed. by Brian MacWhinney & Elizabeth Bates, 77-117. Cambridge & New York: Cambridge University Press.
- Kail, Michèle. 1995. Variations interlinguistiques et développement du langage. *Revue de Phonétique Appliquée* 112-113:239-261.
- Kail, Michèle. 1999. Linguistic variations and cognitive constraints in the processing and the acquisition of language. *Language Diversity and Cognitive Representations*, ed. by Catherine Fuchs & Stéphane Robert, 179-194. Amsterdam & Philadelphia: John Benjamins.
- Kail, Michèle. 2000. Acquisition syntaxique et diversité linguistique. *L’acquisition du langage*, Vol. 2, ed. by Michèle Kail & Michel Fayol, 9-44. Paris: Presses Universitaires de France.
- Kail, Michèle. 2002. Variabilités en psycholinguistique développementale. *Invariants et variabilités dans les sciences cognitives*, ed. by Jacques Lautrey, Bernard Mazoyer & Paul van Geert, 209-224. Paris: Editions de la Maison des Sciences de l’Homme.
- Kail, Michèle. 2004a. On-line grammaticality judgments in French children and adults: a cross linguistic perspective. *Journal of Child Language* 31.3:713-737.
- Kail, Michèle. 2004b. Le développement morphosyntaxique du langage. *Développement cognitif et troubles des apprentissages*, ed. by Marie-Noëlle Metz-Lutz, 123-150, Marseille: Editions Solal.
- Kail, Michèle. 2005. Le Modèle de Compétition: un modèle en développement. *Le Langage et l’Homme*, N° Spécial, 40.2:41-60.
- Kail, Michèle. 2012a. Online sentence processing in children and adults: general and specific constraints — a crosslinguistic study in four languages. *Comparative Perspectives to Language Acquisition: A Tribute to Clive Perdue*, ed. by Marzena Watorek, Sandra Benazzo & Maya Hickmann, 586-612. Bristol: Multilingual Matters.
- Kail, Michèle. 2012b. Coût du traitement chez les sujets bilingues: perspective développementale et comparative. *Nada na linguagem ihe estranho*, ed. by Armanda Costa et al., 633-650. Porto: Edições Afrontamento.
- Kail, Michèle, and Dominique Bassano. 1997. Verb agreement processing in French: a study of on-line grammaticality judgments. *Language and Speech* 40.1:25-46.

- Kail, Michèle, and Agnès Charvillat. 1986. Linguistic cues and processing types in French sentence comprehension. *Knowledge and Language*, ed. by Ida Kurcz, Grace W. Shugar & Joseph H. Danks, 349-357. Amsterdam: North Holland.
- Kail, Michèle, and Agnès Charvillat. 1988. Local and topological processing in sentence comprehension by French and Spanish children. *Journal of Child Language* 15.3: 637-662.
- Kail, Michèle, Armanda Costa, and Isabel Hub Faria. 2008. Jugements de grammaticalité en temps réel: analyse comparative du français et du portugais. *Apprentissage des langues*, ed. by Michèle Kail, Michel Fayol & Maya Hickmann, 199-222. Paris: CNRS Editions.
- Kail, Michèle, and Kleopatra Diakogiorgi. 1994. Morphology and word order in the processing of Greek sentences: a crosslinguistic and developmental study. *Themes in Greek Linguistics*, ed. by Irene Philippaki-Warbuton, Katerina Nicolaidis & Maria Sifianou, 325-332. Amsterdam & Philadelphia: John Benjamins.
- Kail, Michèle, and Kleopatra Diakogiorgi. 1998. On-line integration of morphosyntactic cues by Greek children and adults: a crosslinguistic perspective. *Issues in the Theory of Language Acquisition*, ed. by Norbert Dittmar & Zvi Penner, 177-201. Berlin: Peter Lang.
- Kail, Michèle, Maria Kihlstedt, and Philippe Bonnet. 2012. On-line sentence processing in Swedish: cross-linguistic developmental comparison with French. *Journal of Child Language* 39.1:28-60.
- Kail, Michèle, and Monique Plaza. 2004. How specific are Specific Language Impaired children in off-line and on-line sentence comprehension. Paper presented at the Second Lisbon Meeting on Language Acquisition with Special Reference on Romance Languages, June 1-4, 2004. Lisbon, Portugal.
- Kempe, Vera, and Brian MacWhinney. 1999. Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes* 14.2:129-171.
- Kilborn, Kerry. 1989. Sentence processing in a second language: the timing of transfer. *Language and Speech* 32.1:1-23.
- Kilborn, Kerry. 1994. Learning a language late: second language acquisition in adults. *The Handbook of Psycholinguistics*, ed. by Morton Ann Gernsbacher, 917-944. San Diego: Academic Press.
- Kilborn, Kerry, and Ann Cooreman. 1987. Sentence interpretation strategies in adult Dutch-English bilinguals. *Applied Psycholinguistics* 8.4:415-431.
- Lakoff, George, and Henry Thompson. 1975. Introducing cognitive grammar. *Berkeley Linguistics Society (BLS)* 1:295-313. Berkeley: Berkeley Linguistics Society.
- Lambert, Ludovic, and Michèle Kail. 2001. Le traitement en temps réel des marques morphologiques d'accord dans les phrases en français. *L'Année Psychologique* 101.4:561-592.

- Leong, Che Kan. 2006. Making explicit children's implicit epilanguage in learning to read Chinese. *The Handbook of East Asian Psycholinguistics*, Vol. 1: *Chinese*, ed. by Ping Li, Li-Hai Tan, Elizabeth Bates & Ovid J. L. Tzeng, 70-80. Cambridge & New York: Cambridge University Press.
- Li, Charles N., and Sandra A. Thompson. 1976. Subject and topic: a new typology of language. *Subject and Topic*, ed. by Charles N. Li, 457-489. New York: Academic Press.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Li, Ping. 1990. *Aspect and Aktionsart in Child Mandarin*. Leiden: Leiden University dissertation.
- Li, Ping, Elizabeth Bates, and Brian MacWhinney. 1993. Processing a language without inflections: a reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language* 32.2:169-192.
- Li, Ping, Elizabeth Bates, Hua Liu, and Brian MacWhinney. 1992. Cues as functional constraints on sentence processing in Chinese. *Language Processing in Chinese*, ed. by Hsuan-Chih Chen & Ovid J. L. Tzeng, 207-234. Amsterdam: North-Holland.
- Li, Ping, Li-Hai Tan, Elizabeth Bates, and Ovid J. L. Tzeng. (eds.) 2006. *The Handbook of East Asian Psycholinguistics*, Vol. 1: *Chinese*. Cambridge & New York: Cambridge University Press.
- Liu, Hua, Elizabeth Bates, and Ping Li. 1992. Sentence interpretation in bilingual speakers of English and Chinese. *Applied Psycholinguistics* 13.4:451-484.
- Liu, Youyi, Hua Shu, and Ping Li. 2007. Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods* 39.2:192-198.
- Lu, Ching-Ching, Elizabeth Bates, Ping Li, Ovid J. L. Tzeng, Daisy L. Hung, Chih-Hao Tsai, Shu-Er Lee, and Yu-Mei Chung. 2000. Judgments of grammaticality in aphasia: the special case of Chinese. *Aphasiology* 14.10:1021-1054.
- MacWhinney, Brian. 1987a. Applying the Competition Model to bilingualism. *Applied Psycholinguistics* 8.4:315-327.
- MacWhinney, Brian. 1987b. The competition model. *Mechanisms of Language Acquisition*, ed. by Brian MacWhinney, 249-308. Hillsdale: Lawrence Erlbaum Associates.
- MacWhinney, Brian, and Elizabeth Bates. 1989. *The Crosslinguistic Studies of Sentence Processing*. Cambridge & New York: Cambridge University Press.
- MacWhinney, Brian, and Judit Osmán-Sági. 1991. Inflectional marking in Hungarian aphasics. *Brain and Language* 41.2:165-183.
- McBride-Chang, Catherine, and Yiping Zhong. 2006. Emergent literacy skills in Chinese. *The Handbook of East Asian Psycholinguistics*, Vol. 1: *Chinese*. ed. by Ping Li,

- Li-Hai Tan, Elizabeth Bates & Ovid J. L. Tzeng, 81-89. Cambridge & New York: Cambridge University Press.
- McDonald, Janet L. 1987. Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics* 8.4:379-413.
- McDonald, Janet L., and L. Kathy Heilenman. 1991. Determinants of cue strength in adult first and second language speakers of French. *Applied Psycholinguistics* 12.3: 313-348.
- McDonald, Janet L., and Brian MacWhinney. 1989. Maximum likelihood models for sentence processing. *The Crosslinguistic Study of Sentence Processing*, ed. by Brian MacWhinney & Elizabeth Bates, 397-421. Cambridge & New York: Cambridge University Press.
- Mimica, Ivo, Michael Sullivan, and Stan Smith. 1994. An on-line study of sentence interpretation in native Croatian speakers. *Applied Psycholinguistics* 15.2:237-261.
- Reyes, Iliana, and Arturo E. Hernández. 2006. Sentence interpretation strategies in emergent bilingual children and adults. *Bilingualism: Language and Cognition* 9.1: 51-69.
- Rosch, Eleanor. 1978. Principles of categorization. *Cognition and Categorization*, ed. by Eleanor Rosch & Barbara B. Lloyd, 18-49. Hillsdale: Lawrence Erlbaum Associates.
- Sasaki, Yoshinori. 1991. English and Japanese interlanguage comprehension strategies: an analysis based on the competition model. *Applied Psycholinguistics* 12.1:47-73.
- Sasaki, Yoshinori. 1994. Paths of processing strategy transfers in learning Japanese and English as foreign languages. *Studies in Second Language Acquisition* 16.1:43-72.
- Slobin, Dan I. 1973. Cognitive prerequisites for the development of grammar. *Studies in Child Language Development*, ed. by Charles A. Ferguson & Dan I. Slobin, 175-208. New York: Holt, Rinehart and Winston.
- Slobin, Dan I. 1982. Universal and particular in the acquisition of language. *Language Acquisition: The State of the Art*, ed. by Eric Wanner & Lila R. Gleitman, 128-170. Cambridge & New York: Cambridge University Press.
- Slobin, Dan I. 1985. Introduction: Why study acquisition crosslinguistically? *The Crosslinguistic Study of Language Acquisition*, Vol. 1: *The Data*, ed. by Dan I. Slobin, 3-24. Hillsdale: Lawrence Erlbaum Associates.
- Staroń, Małgorzata, Barbara Bokus, and Michèle Kail. 2005. On-line sentence processing in Polish children and adults. *Studies in the Psychology of Child Language*, ed. by Barbara Bokus, 227-245. Warszawa: Matrix.
- Su, I-Ru. 2001. Transfer of sentence processing strategies: a comparison of L2 learners of Chinese and English. *Applied Psycholinguistics* 22.1:83-112.
- Su, I-Ru. 2004. The effects of discourse processing with regard to syntactic and semantic cues: a competition model study. *Applied Psycholinguistics* 25.4:587-601.

- Sun, Chaofen. 1991. Transitivity and the *ba* construction. Paper presented at the 3rd North American Conference on Chinese Linguistics (NACCL-3). Ithaca: Cornell University.
- Tyler, Lorraine Komisarjevsky, and William D. Marslen-Wilson. 1981. Children processing of spoken language. *Journal of Verbal Learning and Verbal Behavior* 20.4:400-416.
- Weissenborn, Jurgen, Michèle Kail, and Angela D. Friederici. 1990. Language particular or language independent factors in acquisition? *First Language* 10:141-166.
- Wulfeck, Beverly B. 1993. A reaction time study of grammaticality judgments in children. *Journal of Speech, Language and Hearing Research* 36.6:1208-1215.
- Wulfeck, Beverly B., Elizabeth Bates, and Rita Capasso. 1991. A crosslinguistic study of grammaticality judgments in Broca's aphasia. *Brain and Language* 41.2:311-336.

Comparative Psycholinguistics of Sentence Processing with a Special Reference to Chinese

Michèle Kail

CNRS and University of Paris 8

During the last decades of psycholinguistic research, cross-linguistic comparisons have become a very powerful method in the field of language acquisition and language processing. Because of the diversity of the world's languages, comparisons across languages allow to disentangle variables that are likely to be confounded in a single language. However, only Indo-European languages have been examined so far, all of which have to various degrees inflectional devices marking number, gender, or case relations. In contrast to these languages, other languages such as Chinese make virtually no use of inflectional morphology (no marking for case, agreement, tense). Thus, as shown in the first part of my paper, Chinese offers very different structural properties as compared to Indo-European languages, thereby providing a particularly interesting case study for theories of sentence processing.

In the following parts of the paper I present an overview of cross-linguistic research conducted in three domains with a special focus on my own studies: off-line and on-line sentence processing in monolingual children and adults, in bilingual children and adults, and in atypical populations (SLI) in various languages: French, English, Spanish, Portuguese, Swedish, Polish, Greek. My theoretical background is the Competition Model, an interactive activation model initially proposed by MacWhinney & Bates (1989), which I contributed to develop with colleagues from the beginning. In each of the three domains, I report representative Chinese studies conducted within the same theoretical background. I emphasize how they bring a new and unique window onto various topics and results, particularly towards a better understanding of the time course of sentence processing across languages.

Key words: processing syntactic markers in Chinese, on-line sentence processing, transfer, grammaticality judgments, the Competition Model, Chinese psycholinguistics

Une des formes d'expression de la subjectivité en chinois mandarin: L'auxiliaire modal 要 *yao*

Viviane Alleton (艾樂桐)

EHESS

On peut se demander pourquoi *yao* est le moins étudié des verbes modaux chinois, alors qu'il représente un cas particulièrement intéressant d'expression de la subjectivité. Cela tient peut-être à la complexité de ses usages et plus probablement au fait que ses valeurs modales, absentes du chinois classique, ne sont attestées que tardivement. Jusqu'aux Song il nomme certes des actions impliquant des relations humaines telles que «demander, inviter, revendiquer», sans jamais, semble-t-il, les modaliser spécifiquement.

Mots clés: subjectivité, modalités, pouvoir/devoir, prévision

1. Introduction

Il faut rappeler ce qu'on entend par «subjectivité» et par «modalité» ainsi que les questions posées par l'articulation de ces deux notions.

Les auteurs qui se sont intéressés à la théorie de la subjectivité ont été, de longue date, relativement nombreux; cependant on est bien loin de disposer d'une théorie unifiée. Émile Benveniste et John Lyons, qui lui donnent une place centrale dans leur système, l'articulent à la notion de deixis, c'est-à-dire au fait que tout discours suppose un «je» énonciateur, et au-delà, un lieu, «ici», et un moment de l'énonciation, «maintenant».¹ Benveniste définit la subjectivité comme «la capacité du locuteur à se poser comme sujet» et précise que le discours «est la langue en tant qu'assumée par l'homme qui parle, dans la condition d'intersubjectivité, qui rend seule possible la communication linguistique».²

Dans les années récentes, la question de la subjectivité semble avoir été traitée surtout par des fonctionnalistes. Pour opposer subjectif à non subjectif, Halliday, par

¹ Benveniste, Emile. [1958]1966. «De la subjectivité dans le langage». *Problèmes de linguistique générale I*, 259-276; Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press; «Deixis and Subjectivity: *Loquor ergo sum* ?» in Jarvella & Klein, *Speech, Place and Action*, John Wiley & Sons, 1982.

² *Op. cit.* 259-260.

exemple, distingue un composant «idéationnel» et un composant interpersonnel.³ Austin et Searle s'appuient sur la distinction de base entre locutionnaire et illocutionnaire. Une illustration de ces positions théoriques est donnée par John Myhill et Laura Smith, dont l'article «*The Discourse and Interactive Functions of Obligation Expressions*» traite des marques de l'obligation en anglais, chinois, hopi et hébreu biblique.⁴ Leur analyse du chinois,⁵ fondée sur les dialogues de deux pièces de Cao Yu, *Beijingren* (1954) et *Leiyu* (1961), prend en compte essentiellement l'effet négatif du contenu propositionnel sur le locuteur ou l'interlocuteur, opposant la «réprimande» à «ce qu'on peut conseiller» (*advisable action*). Cette analyse quelque peu réductrice semble arbitraire dans son application.

À strictement parler, la subjectivité prise comme liée au nécessaire repérage d'un énonciateur (par la deixis) étant consubstantielle au langage, on pourrait en conclure que cette dimension n'est pas pertinente dans l'analyse des langues particulières. Effectivement, puisque le simple fait de «dire» implique un «je», il semble préférable de laisser l'objectivité de côté. Néanmoins, il importe d'étudier les formes d'expression et les degrés d'explicitation de la subjectivité.

Quant à la modalité, on peut la définir, à la suite de Benveniste, comme une «assertion complémentaire portant sur l'énoncé d'une relation»,⁶ autrement dit l'ensemble des opérations qui situent l'énonciateur par rapport à son énoncé.

Comment articuler ces deux notions ? La subjectivité est une dimension générale du langage, centré sur l'énonciateur. Sont modales toutes les opérations qui, mises en œuvre différemment dans chaque langue, explicitent la position de cet énonciateur par rapport à son énoncé. Admettant que toute énonciation implique un degré plus ou moins grand de subjectivité, on ne peut pas suivre des auteurs, comme ceux réunis dans le recueil de Bybee et Fleischman, qui opposent modalités «subjectives» et «objectives» — tout en reconnaissant que cette «objectivité» est difficile à prouver.

Je préfère poser que les modalités étant des formes privilégiées d'expression de la subjectivité, leur étude permet d'en apprécier le degré d'explicitation.

Si l'on reprend les classifications des modaux par les différents auteurs, les expressions de la capacité et de la volonté seraient «moins» subjectives que les déontiques, lesquels seraient «moins» subjectives que les épistémiques. C'est ainsi que, sur le plan diachronique, Elisabeth Traugott a intitulé l'un de ses articles «*On the rise of*

³ Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*. London: Arnold.

⁴ Bybee, Joan & Susanne Fleischman. 1995. *Modality in Grammar and Discourse*. Amsterdam & Philadelphia: John Benjamins.

⁵ *Op. cit.* 257-271.

⁶ Benveniste, Emile. [1965]1974. «Structure des relations d'auxiliarité». *Problèmes de linguistique générale II*, 187.

epistemic meanings in English: an example of subjectification in semantic change.⁷

Il faut rappeler enfin que l'énoncé modalisé ne garantit pas la réalité de l'action ou la qualité indiquée par le prédicat. Autrement dit, il échappe au test de non contradiction: [x mais non-x] est une forme contradictoire, inacceptable, cependant [devoir/vouloir x mais non-x] est acceptable.

2. Les modalités en chinois

On repère différents types de formes modales. Les modalités énonciatives (assertion, interrogation, exclamation, injonction) sont bien identifiées par des mots grammaticaux, non seulement sous leurs formes directes mais aussi dans leurs dérivations, comme l'interrogation rhétorique. Celle-ci est formellement marquée par les positions respectives non standard de la négation et de certains adverbess, de la négation et de certains verbes auxiliaires. Cette construction marque une assertion forte.⁸ Dans les marges du système adverbial, on observe également des opérateurs argumentatifs tels que *you* «de nouveau, encore», qui implique un jugement distancié, — en contraste avec *zai* «de nouveau», qui est neutre. On peut citer aussi d'autres adverbess comme *ye* «aussi», qui peut parfois signifier «tous comptes faits», ou *dou* «tout», qui, peut exprimer un sentiment d'excès, «au point de», «déjà».

Cependant, il existe en chinois une classe de mots caractérisée *toute entière* par ses emplois modaux, celle des verbes auxiliaires. En cela, le chinois diffère de langues comme le français ou l'anglais dans lesquelles les modaux partagent la classe des auxiliaires avec les formes de la temporalité et de la diathèse.

Parmi les verbes auxiliaires de mode du chinois contemporain, 要 *yao* est remarquable par sa polyvalence. L'analyse de ses divers emplois permet d'illustrer non seulement les oppositions entre les types de modalité, mais aussi les ambiguïtés, les superpositions possibles. Ce mot a plusieurs emplois verbaux.⁹ Sa valeur de base, «vouloir», est commune au verbe transitif («vouloir [ceci]»), au verbe régissant une proposition («vouloir [que]») et au verbe auxiliaire («vouloir [+ V]»).¹⁰ Une telle

⁷ Traugott, Elizabeth C. 1989. «On the rise of epistemic meanings in English: an example of subjectivisation in semantic change». *Language* 65.1:31-55.

⁸ Cf. Alleton, Viviane. 1988. «The so-called 'Rhetorical Interrogation' in Mandarin Chinese», *Journal of Chinese Linguistics* 16.2:278-296.

⁹ En outre il s'emploie comme conjonction (forme courte de *yaoshi* «si»).

¹⁰ Par «valeur de base», je me réfère à la définition de J. Kurylowicz, selon qui «la valeur de base d'un mot est celle qui a le plus large éventail de fonction et le moins de contraintes contextuelles». Cf. Kurylowicz, Jerzy. 1966. «Dérivation lexicale et dérivation linguistique», in Hamp, *Readings in Linguistics*, vol. 2, 42-50.

plasticité ne nous étonnera pas si l'on se souvient que le verbe français *vouloir* peut être qualifié de «modalisant par occasion».¹¹

Quand *yao* est en fonction de verbe auxiliaire modal, il peut exprimer, outre sa valeur volitive, le déontique et l'épistémique. Dans ces deux derniers cas, il s'oppose à 應該 *yinggai*, dont la valeur de base est «devoir». Pour l'épistémique, il existe un troisième terme, 會 *hui*, marque d'éventualité. Ces différents verbes auxiliaires peuvent pour une partie de leurs emplois, figurer dans des contextes identiques, ils commutent. Ce qui les différencie est, polairement, le renvoi à une norme extérieure avec *yinggai*, à la volonté du sujet énonciateur avec *yao*, à la possibilité pour *hui*. Dans le cas de *yao*, il ne s'agit pas de polysémie mais plutôt d'invariance: ses différentes valeurs ne sont pas entièrement disjointes.¹²

Sa caractéristique majeure, en contraste avec *yinggai*, est une forte composante subjective. L'énonciateur prend en charge l'obligation énoncée, alors qu'avec *yinggai* il peut la mentionner tout en acceptant qu'elle ne soit pas remplie.

- (1) 你應該買這本書，但不買也沒關係。

Ni yinggai mai zhei ben shu, dan bu mai ye
2SG *yinggai* acheter ce CL livre mais NEG acheter aussi
mei guanxi.
ne.pas.avoir conséquence
«Tu **dois** acheter ce livre, mais si tu ne l'achètes pas, ça ne fait rien.»

Yao est exclu dans un tel énoncé.

On notera que les cas dans lesquels l'énonciateur présente sa propre volonté comme étant le devoir de celui auquel il s'adresse supposent qu'on donne une place majeure à l'autorité dans la morale sociale. Autrement dit, la formule «tu dois faire ce que je veux» vient assez naturellement, sans toutefois impliquer un quelconque automatisme.

La troisième valeur de *yao* illustre une dérivation souvent observée dans les langues, du déontique à l'épistémique. Dans ce cas, il est fréquent que *yao* revoie à un événement situé dans le futur tel qu'il est prévu par le sujet énonciateur. Cette valeur est bien rendue en français par les formes auxiliaires du verbe «aller». Il peut exprimer aussi l'appréciation d'une qualité. Le terme chinois 估計 *guji* semble correspondre à ces deux notions. *Yao* commute dans une partie de ses emplois épistémiques avec *hui*, dont le caractère subjectif est moins marqué.

¹¹ Cf. Benveniste, *op. cit.* note 6, 188.

¹² Cela est sans doute vrai également pour les autres verbes auxiliaires, mais j'examine ici seulement le cas de *yao*.

3. Analyse syntaxique

Il ne convient pas d'étudier les cas où *yao* est modal (structure d'auxiliation) en les classant par catégories sémantiques. En effet, dans un même énoncé un auxiliaire peut avoir plusieurs significations. Celles-ci se différencient éventuellement en fonction du contexte large, cependant, et cela est important, dans l'usage il n'est pas rare que l'ambiguïté subsiste.

On examinera les types de sujet grammatical et de verbe auxilié (V2), les marques aspectuelles et la négation.

3.1 Sujets en position de thème¹³ et verbe d'action en V2

3.1.1 Le sujet grammatical et l'énonciateur sont confondus («je»)

Yao peut être volitif, déontique ou épistémique.

- (2) a. 我要向他學習。
Wo yao xiang ta xuexi.
1SG *yao* vers 3SG étudier
«[Il travaille efficacement], je **veux/dois** prendre modèle sur lui.»

On ne sait pas s'il s'agit d'une volonté personnelle de l'énonciateur — une résolution un peu floue — ou d'une obligation.

- (2) b. 我要向他學習了。
Wo yao xiang ta xuexi le.
1SG *yao* vers 3SG étudier *le*
«Je **vais** prendre modèle sur lui.»

La présence d'un *le* final indique qu'il s'agit d'une résolution que l'énonciateur a l'intention suivre dans un futur proche (valeur épistémique).

- (3) 我要洗澡了。
Wo yao xizao (le).
1SG *yao* se.laver *(le)*
«Je **vais** prendre une douche.»

¹³ On sait que le sujet grammatical n'est pas nécessairement exprimé en chinois. Pour la commodité de l'analyse, je considère ici des cas où il l'est.

La valeur volitive forte de *yao* n'est pas exclue ici, cependant elle est rare en contexte quotidien (on emploie plutôt 想 *xiang* «penser à, avoir l'intention de»). Elle est attestée seulement s'il y a un obstacle. La valeur déontique semble rare. Pour dire «Je dois prendre une douche», on emploiera plutôt *yinggai* (c'est la règle) ou *xuyao* (c'est une nécessité: je suis sale).

- (4) 我一定要蓋一所房子。

Wo yiding yao gai yi-suo fangzi.
 1SG certainement *yao* construire un-CL maison
 «Je **veux/dois** absolument construire une maison.»

En présence d'un adverbe exprimant la certitude (*yiding*), l'interprétation épistémique de *yao* est exclue. En l'absence de contexte justificatif, le volitif est plausible; s'il y a un contexte tel que «mes enfants sont à la rue», on peut supposer que le volitif et le déontique se conjuguent.

3.1.2 Le sujet en position de thème est l'interlocuteur («tu»)

Énonciateur et sujet grammatical sont dissociés. Dans le cas où ce sujet est «tu», on pourrait penser que le volitif est exclu parce qu'on peut difficilement asserter la volonté d'autrui en s'adressant à lui. Néanmoins dans la mesure où l'énonciateur rappelle l'existence d'un devoir sans préciser ce qui le fonde, il se peut qu'il le prenne en charge. Autrement dit, ces énoncés déontiques impliquent une part de volitif — qui serait inhérente à l'auxiliaire *yao* dans ses emplois actuels.

- (5) 你要向他學習。

Ni yao xiang ta xuexi.
 2SG *yao* vers 3SG étudier
 «Tu dois prendre modèle sur lui.»

Le locuteur, qui énonce une obligation incombant à son interlocuteur, cherche à la lui imposer. À la forme écrite, on peut se demander s'il s'agit d'une assertion ou d'un ordre. Dans le cas où ce serait impératif, le pronom *ni* «toi» ne serait pas employé (sauf pour identifier la personne en cause dans un groupe).

- (6) a. 你要早起。

Ni yao zao qi.
 2SG *yao* tôt se.lever
 «tu **dois** te lever tôt.»

- b. 我看，你要早起。

Wo kan, ni yao zao qi.
 1SG considérer 2SG yao tôt se.lever
 «À mon avis tu **dois** te lever tôt.»

Il s'agit d'un conseil fondé sur le jugement du locuteur; celui-ci peut disposer d'une autorité de fait ou morale sur son interlocuteur.

- (7) a. 按計畫，你應該早起。

An jihua, ni yinggai zao qi.
 Selon règlement 2SG yinggai tôt se.lever
 «Selon le règlement, tu **dois** te lever tôt.»

Yinggai renvoie à une source d'obligation extérieure. L'énoncé a les apparences de l'objectivité. Il est purement déontique.

- (7) b. 按新計畫，你要早起了。

An xin jihua, ni yao zao qi le.
 Selon nouveau règlement 2SG yao tôt se.lever *le*
 «Selon le nouveau règlement, tu **dois** te lever tôt.»

Le indique ici une nouvelle situation créée ici par le changement de règlement. *Yao* est déontique.

Bien que la valeur volitive de *yao* soit exclue en ce qui concerne le sujet de la phrase (le destinataire de l'énoncé), elle semble être la motivation implicite de l'énonciateur: cela ressort du contraste avec *yinggai* en (7a), qui ne comporte pas cette nuance.

3.1.3 Le sujet est une ou plusieurs tierce(s) personne(s)

Si le sujet est une personne autre que l'énonciateur ou le destinataire, les trois interprétations de *yao* sont possibles, alors qu'avec *yinggai* c'est nécessairement un déontique.

- (8) a. 老王要向他學習。

Lao Wang yao xiang ta xuexi.
 Lao Wang yao vers 3SG étudier
 «Lao Wang **veut/doit** prendre modèle sur lui.»

L'interprétation de *yao* comme «vouloir» implique que l'énonciateur a fait une hypothèse relative à la volonté d'autrui. Pour interpréter *yao* comme déontique, il faut supposer que l'énonciateur a autorité ou qualité pour apprécier la situation.

- (8) b. 老王應該向他學習。

Lao Wang yinggai xiang ta xuexi.

Lao Wang *yinggai* vers 3SG étudier

«Lao Wang **doit** prendre modèle sur lui.»

Avec *yinggai*, l'énonciateur exprime une implication minimale, non nulle mais plus faible qu'en (8a). C'est la simple assertion du devoir qui incombe à Lao Wang.

- (9) a. 警察要管流氓。

Jingcha yao guan liumang.

police *yao* gérer voyous (après un incident sur la voie publique)

«la police **devrait** s'occuper des voyous.» (déontique)

«la police **va** s'occuper des voyous.» (prévision – rare)

- b. 警察應該管流氓。

Jingcha yinggai guan liumang.

police *yinggai* gérer voyous (discussion entre amis sur le rôle de la police)

«La police **doit** s'occuper des voyous.»

En (9a) et (9b) l'énonciateur affirme l'obligation pour la police de s'occuper des voyous. La différence est qu'en (9a) il suggère que telle est aussi sa propre volonté (*yao*) alors qu'en (9b) il constate seulement que telle est l'obligation professionnelle des agents de police.

- (10) a. 大家都要去吃飯。

Dajia dou yao qu chifan.

Tout.le.monde tout *yao* aller manger

«Tout le monde **veut** aller manger.»

Ici, *dajia* désigne un groupe de personnes données dans le contexte. Exemple de situation: on a travaillé dur, il est tard.

- b. 大家都要吃飯，富人窮人都一樣。

Dajia dou yao chifan, furen qiongren

tout.le.monde tout *yao* manger riche pauvre

dou yiyang.
 tout être.semblable
 «Tout le monde **doit** manger, que l'on soit riche ou pauvre.»

L'énonciateur atteste la légitimité de la volonté des sujets.

- c. 大家應該吃飯。
Dajia yinggai chifan.
 tout.le.monde yinggai manger
 «Tout le monde **doit** (aller) manger.»

La présence de *yinggai* suffit à rappeler qu'il s'agit d'une norme.

3.1.4 Le mot en position de sujet désigne un inanimé

C'est le seul cas excluant radicalement une interprétation volitive.

- (11) a. 屋頂要塌下來。
Wuding yao ta xia lai.
 plafond yao s'effondrer descendre venir
 «Le plafond **va** s'écrouler !»
- b. 屋頂會塌下來。
Wuding hui ta xia lai.
 plafond hui s'effondrer descendre venir
 «Le plafond **risque** de s'écrouler.»

Ces deux énoncés ont une valeur épistémique: il s'agit de l'évaluation d'un risque qui se situe dans le futur proche. En (11a) *yao* marque l'implication du locuteur qui donne l'alerte. En (11b) avec *hui*, il n'y a pas la même nuance d'urgence, c'est une appréciation des données.

- (12) 白菜要煮五分鐘。
Baicaí yao zhǔ wú fēnzhōng.
 chou yao cuire cinq minute
 «Le chou **doit** cuire cinq minutes» (à mon avis).

- (13) 這衣裳要洗了。

*Zhe yishang yao xi le.*¹⁴

Ce vêtement yao laver le

«Ces vêtements sont à laver.»

Les énoncés (12) et (13) comportent les notions d'évaluation et de prescription, caractéristiques des emplois épistémiques de *yao*.

3.2 Le verbe auxilié (V2) est un verbe de qualité

Le partage entre l'estimation (épistémique) et l'énoncé d'un devoir (déontique) est le même que le prédicat soit un verbe d'action (動詞 *dongci*) comme dans les exemples précédemment cités ou un verbe de qualité (形容詞 *xingrongci*) comme ci-après.

- (14) a. 在戰爭中，一切行動都應該迅速果斷。

Zai zhanzheng zhong, yiqie xingdong dou yinggai

être.à guerre dans tous mouvement tout yinggai

xunsu guoduan.

être.rapide être.résolu

«À la guerre tous les mouvements **doivent** être rapides et résolus.»

- b. 在戰爭中，一切行動都要迅速果斷。

Zai zhanzheng zhong, yiqie xingdong dou yao

être à guerre dans tous mouvement tout yao

xunsu guoduan.

être.rapide être.résolu

«À la guerre tous les mouvements **doivent** être rapides et résolus.»

Le message déontique est identique que l'on emploie *yinggai* (14a) ou *yao* (14b), mais l'énonciateur choisira plutôt *yinggai* dans une conférence et *yao* sur le terrain, où l'on suppose son implication personnelle plus grande.

Dans le cas d'une évaluation (épistémique) *yao* est privilégié en présence d'un terme de comparaison:

- (15) 這兩張照片前一張要清楚些。

Zhei liang-zhang zhaopian qian yi zhang yao qingchu xie.

ce deux-CL photo devant un-CL yao être clair quelque peu

«La première de ces deux photos **doit** être plus nette.»

¹⁴ Chao Yuen Ren (1966:736): «These clothes must be washed».

Quand l'estimation est présentée comme fondée sur des données objectives, on préfère *yinggai*:

- (16) 這是鋼材的，應該比較結實。
Zhei shi gangcai de, yinggai bijiao jieshi.
Ce être acier *de yinggai* plutôt être.solide
«C'est en acier, ce **doit** être plutôt solide.»

Ce n'est pas le cas dans l'exemple suivant qui a la forme d'une assertion, sans justification:

- (17) 比墨還要墨。
Bi mo hai yao mo.
comparé à encre encore *yao* être.noir
«C'est encore plus noir que l'encre.»¹⁵

Dans les énoncés où *yao* a une valeur épistémique, la particule *le* est fréquente. L'énonciateur considère que le changement de situation impliqué par *le* est virtuellement acquis.

- (18) 天要黑了，還不走？
Tian yao hei le, hai bu zou?
ciel *yao* être noir *le* encore NEG partir
«Il **va** faire nuit, pourquoi ne t'en vas-tu pas ?»
- (19) 你再也不快點兒，就要遲到了。
Ni zai bu kuai dianr, jiu yao chi dao le.
2SG encore NEG être.rapide un.peu alors *yao* en.retard arriver *le*
«Si tu ne te dépêches pas un peu plus, tu **vas** être en retard.»

3.3 Phrases complexes

Quand le prédicat constitué de *yao* +V2 est précédé d'une subordonnée (marquée ou non comme telle), il est rarement volitif. Il est le plus souvent épistémique.

¹⁵ Chao Yuen Ren (1966:737): «(...) even blacker than ink».

- (20) 你不開窗，大家都要頭疼。

Ni bu kai chuang, dajia dou yao touteng.
 2SG NEG ouvrir fenêtre tout.le.monde tout *yao* avoir.mal.à.la.tête
 «Si tu n’ouvres pas la fenêtre, tout le monde **va** avoir mal à la tête.»

- (21) 空氣太壞，大家都會頭疼。

Kongqi tai huai, dajia dou hui touteng.
 air trop être.mauvais tout.le.monde tout *hui* avoir.mal.à.la.tête
 «L’air est trop vicié, on (tout le monde) **va** avoir mal à la tête.»

Dans ces deux exemples, l’énonciateur fait une prévision fondée sur son évaluation de la situation (épistémique). Avec *hui* (21), il se contente de prévoir, avec *yao* (20), il fait en outre pression sur son interlocuteur.

- (22) 一沾火就要炸！

Yi zhan huo jiu yao zha!
 dès.que toucher feu alors *yao* exploser
 «Dès que cela touchera le feu, ça explosera !»

L’énonciateur suggère une relation nécessaire, qui peut être contestée [*nei ge meiyou weixian* «mais non, cela est sans danger»]. Il insiste sur l’imminence.

- (23) 如果沾火，就要炸。

Ruguo zhan huo, jiu (yao) zha.
 si toucher feu alors (*yao*) exploser
 «Si cela touche le feu, ça explose.»

En (23) *yao* est également épistémique, mais alors qu’en (22), où le danger est imminent, il est indispensable, en (23), qui peut figurer dans un mode d’emploi, il n’est pas nécessaire.

Dans les exemples précédents, il s’agissait de ce qui peut se produire. Dans les suivants, on a l’évaluation d’une quantité suffisante.

- (24) 這項任務要十天才能完成。

Zhei-xiang renwu yao shi tian cai neng wancheng.
 ce-CL tâche *yao* dix jour alors.seulement pouvoir achever
 «Cette tâche, **il faut** dix jours pour l’achever.»

(25) 這項任務應該十天才能完成。

Zhei-xiang renwu yinggai shi tian cai neng wancheng.
ce-CL tâche *yinggai* dix jour alors.seulement pouvoir achever
«Cette tâche, **il faut** dix jours pour l'achever.»

Ici, l'enjeu n'est pas le même selon qu'on ait *yao* ou *yinggai*: *yao* est bien venu dans la discussion d'un devis (l'énonciateur souhaite ce délai) alors que *yinggai* s'emploie pour une simple estimation.

4. Conclusion

Cette esquisse d'analyse porte uniquement sur les emplois de *yao* comme verbe auxiliaire de mode en chinois contemporain. On sait que les fonctions décrites ici ne sont pas très anciennes: il faudrait situer plus précisément et décrire les conditions de leur apparition. Par ailleurs, il serait bon de faire l'inventaire des formes dialectales remplissant des fonctions identiques ou proches.

On peut cependant affirmer que dans la langue commune contemporaine, ce qui oppose *yao* aux autres verbes auxiliaires avec lesquels il peut commuter est l'intensité de l'élément subjectif. Il reste à intégrer les phénomènes décrits dans le cadre d'une théorie d'ensemble des marques de la subjectivité — théorie qui reste à élaborer.

Références

- Alleton, Viviane. 1984. *Les auxiliaires de mode en chinois contemporain*. Paris: Éditions de la Maison des Sciences de l'Homme.
- Austin, John L. 1962. *How to do Things with Words*. Oxford: Clarendon Press.
- Benveniste, Emile. [1958]1966. De la subjectivité dans le langage. *Problèmes de linguistique générale* I, 259-276. Paris: Gallimard.
- Benveniste, Emile. [1965]1974. Structure des relations d'auxiliarité. *Problèmes de linguistique générale* II, 177-193. Paris: Gallimard.
- Bybee, Joan L., and Suzanne Fleischman. 1995. *Modality in Grammar and Discourse*. Amsterdam & Philadelphia: John Benjamins.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Kerbrat-Orecchioni, Catherine. 1980. *L'énonciation: De la subjectivité dans le langage*. Paris: Armand Colin.
- Lyons, John. 1977. *Semantics*. Cambridge & New York: Cambridge University Press.
- Lyons, John. 1982. Deixis and subjectivity: Loquor ergo sum? *Speech, Place and Action: Studies in Deixis and Related Topics*, ed. by Robert J. Jarvella & Wolfgang Klein, 101-124. Chichester: John Wiley & Sons.
- Myhill, John, and Laura A. Smith. 1995. The discourse and interactive functions of obligation expressions. *Modality in Grammar and Discourse*, ed. by Joan L. Bybee & Suzanne Fleischman, 239-291. Amsterdam & Philadelphia: John Benjamins.
- Palmer, Frank R. 1979. *Mood and Modality*. Cambridge & New York: Cambridge University Press.
- Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge & New York: Cambridge University Press.
- Traugott, Elisabeth Closs. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65.1:31-55.
- Verstraete, Jean-Christophe. 2001. Subjective and objective modality: interpersonal and ideational functions in the English modal auxiliary system. *Journal of Pragmatics* 33.10:1505-1528.

An Expression of Subjectivity in Mandarin Chinese: The Modal Auxiliary 要 *yao*

Viviane Alleton

EHESS

One might wonder why *yao* is the least studied of all the modal verbs in Chinese, even though it represents a particularly interesting case with regard to the expression of subjectivity. This may be due to the complexity of its usages but more likely to the fact that its modal values, which are absent in Classical Chinese, are only documented at a late period. Before the Song period, even when associated with actions involving human relationships such as “to ask, to invite, to demand”, it does not appear to be used to add a specific modal value to these actions.

Key words: subjectivity, modality, possibility, obligation, prediction

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

Zhitang Yang-Drocourt (楊志棠)

INALCO-CRLAO

En chinois standard contemporain, la reduplication est un procédé commun à la formation lexicale et à la syntaxe. Nous discuterons ici du statut grammatical des adjectifs formés par reduplication. A la différence de plusieurs auteurs qui, partant des formes en *input*, considèrent une partie des adjectifs redupliqués comme relevant de la réalisation morphologique de la syntaxe, nous nous proposons de les replacer dans le domaine de la formation lexicale. Pour nous, qu'il existe ou non en *input* un mot attesté, que celui-ci soit lui-même un adjectif ou un mot d'une autre catégorie, la reduplication permet de dériver une sous-catégorie de mots, distincts de leur forme simple par des propriétés spécifiques. Cette analyse sur le statut des adjectifs redupliqués nous permet également de mieux cerner leur sémantisme.

Mots clé: chinois moderne, lexique, adjectif, reduplication, morphologie, formation lexicale

1. Introduction

Traditionnellement, lorsqu'on aborde les formes redupliquées du chinois moderne, toutes catégories confondues, on les regroupe le plus souvent en un grand ensemble, puis, on les range dans différentes sous-catégories. L'analyse de Li & Thompson (1981:28) illustre bien cette position. Dans le chapitre «Word Structure», les auteurs traitent l'affixation et la reduplication comme des procédés morphologiques de la dérivation,¹ les distinguant par là même de la composition, en précisant que «as a morphological process, *reduplication* means that a morpheme is repeated so that the original morpheme together with its repetition form a new word». Ces auteurs retiennent cinq types de reduplication en chinois moderne:

¹ Les auteurs n'emploient pas ce terme dans le texte, mais, dans leur index, on trouve une entrée «dérivation», qui renvoie aux seules sections «affixation» et «reduplication».

- A. «Volitional Verbs»: *chang²-chang⁰* ‘taste a little’, *tao³lun⁰-tao³lun⁰* ‘discuss a little’, *shui⁴-shui⁰-jiao⁴* ‘sleep a little’
- B. «Adjectives»: *hong²hong²* ‘really red’ (more vivid description), *shu¹shu⁰fu⁴fu⁴* ‘comfortable’
- C. «Measure Words»: *ge⁴-ge⁴* ‘every person’, *tian¹-tian⁰* ‘every day’
- D. «Kinship Terms»: *ba⁴ba⁰* ‘father’, *po²po⁰* ‘husband’s mother’
- E. «Miscellaneous Reduplicated Terms»: *mao²mao²yu³* ‘drizzle’, *beng¹beng⁰ cui⁴* ‘crackling crisp’, *gang¹-gang⁰* ‘just now’, *ding¹dang⁰-ding¹dang⁰* ‘sound of ringing bells’²

Ce faisant, Li & Thompson impliquent que toutes ces formes rédupliquées occupent une seule et même place dans la langue, celle des mots dérivés, et qu’il s’agit donc d’un fait lexical.

Depuis le début des années 1980, on marque de plus en plus la distinction entre *gou⁴xing²* 構形, littéralement «réalisation morphologique», et *gou⁴ci²* 構詞 «formation lexicale».³ Ce faisant, les linguistes affirment bel et bien l’existence en chinois d’une réalisation morphologique de la syntaxe, qui jouit d’un statut grammatical similaire à celui de la flexion, en même temps, ils rangent tous les éléments issus de la «dérivation» (*pai⁴sheng¹* 派生) dans le domaine de la «lexicologie» (*ci²hui⁴xue²* 詞彙學).

Cette tendance reflète sans doute l’influence exercée par la théorie dominante dans la linguistique occidentale de l’époque. En effet, c’est précisément au cours des années 1970-80 que la «morphologie lexicale» (Matthews 1974[1991:37]) a pu être réexaminée avec une certaine autonomie par rapport à la syntaxe, à travers notamment le principe des «Word Formation Rules» qui, selon Aronoff (1976:22), «are completely separate from the syntactic and phonological rules of the grammar».

Néanmoins, il nous semble bien qu’un décalage terminologique s’est opéré entre les termes de la linguistique générale et ceux de la linguistique chinoise: pour la première, «morphologie» prend plutôt le sens étroit que lui donne Matthews (*op. cit.*): elle est «lexicale» et étudie la structure interne des mots complexes et les régularités dans la formation des mots. Le terme chinois *gou⁴xing²* «réalisation morphologique» renvoie, lui, expressément aux faits syntaxiques. C’est sans doute la raison pour laquelle

² Nous avons repris tels quels, d’après le texte original, les tons et les traits d’union qui relient les séquences.

³ Nous savons que, dans les années 1950, plusieurs linguistes, notamment Yu & Lu (1954), Hu & Wen (1954), Wang (1956) et Cen (1957), ont déjà discuté des «modifications morphologiques» *xing²tai⁴bian⁴hua⁴* 形態變化, et selon Wang (1956:48), «La morphologie au sens étroit s’occupe seulement de *gou⁴xing²fa³* 構形法 ‘réalisation morphologique’, alors que la morphologie au sens large s’occupe également de *gou⁴ci²fa³* 構詞法 ‘formation lexicale’».

les linguistes chinois préfèrent au terme *xing²tai⁴xue²* 形態學 «morphologie» celui de *ci²hui⁴xue²* «lexicologie». ⁴

Certes, il s'agit là de termes empruntés. Nous n'avons pas l'intention de discuter ici du bien-fondé épistémologique de cette terminologie et de son emploi en chinois; nous voulons seulement mettre le doigt sur les problèmes que soulève l'application de ces notions au phénomène de la reduplication en chinois. En effet, on peut très bien employer d'autres termes que «réalisation morphologique» en parlant, par exemple, de «fait syntaxique» *ju⁴fa³xian⁴xiang⁴* 句法現象 (Guo 2002:40), et considérer le processus dérivationnel comme un «fait lexical». Or, la question n'en reste pas moins entière et la distinction pertinente. Pour nous, il existe bel et bien en chinois deux types de «réalisation morphologique», l'une relevant de la syntaxe, l'autre de la formation lexicale. Comme le fait remarquer Zwicky (1990:230), «it might seem that inflection is a topic of rather limited interest to the theory of grammar, given that so many languages have little or no inflectional morphology. However there are special relationships between inflection and a type of lexemes that is exemplified in all the world's languages, so that inflectional morphology cannot be so easily dismissed as a collection of exotica; these lexemes labelled grammatical words, particle words, non lexical items, or particles, I will call them particle lexemes».

Ce qui nous importe ici est de définir le statut grammatical des adjectifs formés par la reduplication en répondant notamment aux questions suivantes: relèvent-ils de la réalisation morphologique de la syntaxe ou de la formation lexicale? Faut-il les analyser cas par cas en fonction de leur forme en input, comme le proposent de nombreux linguistes?

2. Les redupliques entre réalisation syntaxique et dérivation

Matthews (1974[1991:24-36]) propose de distinguer d'abord trois types de «mots»: «mots-forme» pour les mots phonologiques, «lexèmes» pour les unités lexicales abstraites hors de tout contexte syntaxique et «mots grammaticaux» pour les unités lexicales telles qu'elles apparaissent dans un syntagme ou dans une phrase. À partir de cette distinction, il considère que la morphologie étudie deux domaines distincts: l'un concerne la forme et la réalisation des mots grammaticaux; c'est la morphologie «flexionnelle», l'autre, la structure et la formation des lexèmes; c'est la morphologie «lexicale».

Mel'čuk (1993:308) introduit, pour sa part, les notions de «grammèmes» et de «dérivatèmes». Les premiers (c'est le cas des conjugaisons d'un verbe français par

⁴ Ce n'est pas non plus tout à fait exact, puisque la «lexicologie» est concernée à la fois par la morphologie lexicale et par la sémantique lexicale.

exemple), nécessairement fondés sur une identité catégorielle entre input et output, permettent de réaliser des formes nouvelles d'un même lexème sur le plan strictement synchronique. Ces formes sont nouvelles dans la mesure où elles ne sont pas préalablement emmagasinées dans l'esprit du locuteur, mais actualisées «ici» et «maintenant» par celui-ci selon les règles grammaticales standard qui lui sont imposées par la langue. Par définition, les grammèmes obéissent à des automatismes et des régularités; leur emploi est obligatoire.

Ainsi, les linguistes qui acceptent la notion de «réalisation morphologique»⁵ s'accordent en général pour dire que certaines formes de mots chinois sont précisément marquées de ce caractère grammatical et automatique. Par exemple, un locuteur peut exprimer, à l'aide du suffixe *-men⁰* 們, le pluriel défini de n'importe quel nom humain, [X]_N-*men⁰*, et dire *fen³si¹-men⁰* 粉絲們 'les fans', sans que cette réalisation soit inscrite dans les dictionnaires ou préalablement stockée dans sa mémoire. Néanmoins, cette règle n'est pas absolue en chinois, car la pluralité peut aussi être exprimée par la présence d'un autre morphème, à savoir, l'adverbe *dou¹* 都: *ta¹men⁰ dou¹ shi⁴ ta¹ de⁰ fen³si¹* 她們都是他的粉絲 (3PL tous être 3SG SUB fan) 'Ce sont toutes ses fans'.

Voyons aussi l'exemple de la répétition⁶ des verbes volitifs. Dans (1)-(2), la répétition (a) se réalise chaque fois sur la base d'un verbe et sa fonction syntaxique est d'exprimer l'aspect «diminutif». Dans (b-d), ce sont les suffixes verbaux qui permettent de réaliser d'autres expressions aspectuelles:⁷

- | | |
|---|--|
| (1) <i>kan⁴</i> 'voir, regarder' | > a. <i>kan⁴ kan⁰</i> (ASP dimin.) 'jeter un coup d'œil' |
| | > b. <i>kan⁴-le⁰</i> (ASP accompli) 'avoir vu' |
| | > c. <i>kan⁴-zhe⁰</i> (ASP duratif) 'en train de voir' |
| | > d. <i>kan⁴-guo⁰</i> (ASP expérience) 'avoir vu' |

⁵ Ren (1981), Zhang (1981), Liu (1990[2005]), Shi (1996), Ge (2001), entre autres.

⁶ Nous utilisons expressément le terme «répétition» lorsqu'il s'agit des verbes pour marquer son statut syntaxique, en opposition à «réduplication», que nous réservons à la morphologie lexicale.

⁷ Le terme est de Li & Thompson (1981). Chao (1968) emploie celui de *tentative*. Selon Paris (2007:61), La reduplication verbale «code des variations de point de vue sur une situation (atténuation du degré ou de la force illocutoire)». En effet, «diminution» ou «atténuation» doivent être pris au sens large: cette expression aspectuelle peut signifier, selon le sens lexical des verbes et/ou le contexte, (i) la courte durée ou le peu d'occurrences répétitives du procès; (ii) le peu d'ampleur ou la légèreté du mouvement ou du geste; (iii) l'adoucissement de ton d'un ordre, d'une demande ou d'une incitation; (iv) l'essai ou la tentative d'un procès; (v) la répétition d'un procès perçue comme une routine dépourvue d'importance, etc.

- (2) *kao³lǔ⁴* ‘réfléchir’ > a. *kao³lǔ⁰ kao³lǔ⁰* (ASP dimin.) ‘essayer d’y réfléchir’
 > b. *kao³lǔ⁴-le⁰* (ASP accompli) ‘avoir réfléchi’
 > c. *kao³lǔ⁴-zhe⁰* (ASP duratif) ‘en train de réfléchir’
 > d. *kao³lǔ⁰-guo⁰* (ASP d’exp) ‘avoir réfléchi’

L’analyse apparemment anodine, qui consiste à traiter ces formes répétitives comme des réalisations syntaxiques,⁸ effectue en fait une rupture avec la tradition,⁹ telle qu’elle est présentée dans Li & Thompson (*op. cit.*). Pour nous, la distinction entre les domaines syntaxique et lexical est essentielle si on veut parvenir à une meilleure description du chinois et améliorer son enseignement. La conséquence est double:

Premièrement, en parlant de réalisation syntaxique, on montre que l’on est en présence d’expressions de régularité grammaticale permettant de formuler des règles. On pourrait tenter de dire, par exemple, que tous les verbes volitifs capables d’exprimer un procès inscrit dans la durée sont susceptibles d’être répétés et ce, selon des schémas phonologiques fixes et réguliers.¹⁰ De plus, l’interprétation sémantique de cette forme étant mécaniquement déductible, le locuteur n’a rien à mémoriser.

Deuxièmement, on accorde ainsi une certaine spécificité à d’autres formes similaires qui, étant issues de la dérivation ou de la formation lexicale, ne possèdent pas les mêmes caractères systématiques.

En effet, la dérivation doit d’abord être vue comme un processus. D’une base «primitive» dérive un autre lexème, qui n’a pas nécessairement la même identité catégorielle que sa base: du radical *vend-* dérive *vendeur*, [X]V > [Xeur]N, ou bien, du verbe *voler* dérive *vol*, [Xer]V > [X]N. Autrement dit, la dérivation permet de construire des mots différents de ceux dont elle part, qu’ils soient de la même catégorie ou non. Ici, la relation entre la base et sa forme dérivée relève de la formation lexicale, procédé productif qui permet d’enrichir le lexique de nouveaux éléments. Par exemple, une fois créés, les néologismes français *chatcher*, *chatcheur*, *chatcheuse*, qui dérivent du mot anglais *chat* ‘discuter en ligne’, sont intégrés dans le lexique virtuel et sont susceptibles de se voir un jour inclus dans les dictionnaires, donc officiellement reconnus.

Ces mots-là, le locuteur ne les ressent pas comme les réalisations d’une règle de grammaire, mais comme des éléments de son vocabulaire, qui auraient été emmagasinés

⁸ Cela permet de prédire et d’expliquer l’acceptabilité de *kan⁴ yi⁰ kan⁴* 看一看 «jeter un coup d’œil» et de *kan⁴-le⁰ kan⁴* 看了看 «avoir jeté un coup d’œil»: on répète le verbe tel quel, ce qui permet de lui adjoindre un suffixe aspectuel.

⁹ Rappelons néanmoins que Chao (1968:205) proposait déjà de considérer la répétition du verbe comme «a borderline case between morphological and syntactic constructions».

¹⁰ Soit [A^T] > [A^T] [A⁰], soit [A^TB^T] > [A^TB⁰] [A^TB⁰] (T: ton que portent les morphèmes avant la répétition).

au fur et à mesure dans sa mémoire. Si les procédés de la formation lexicale, dont fait partie la dérivation, obéissent aussi à des régularités imposées par la langue, et c'est ce que le montrent de nombreuses études morphologiques, ces régularités ne se plient en revanche à aucune clause grammaticale, comme celle que nous venons de formuler pour la répétition des verbes. Car, de telles réalisations ne sont souvent ni systématiques ni automatiques ni standard, leurs mécanismes et leurs résultats ne sont pas toujours prédictibles. C'est ainsi que, dans le même sens: 'discuter en ligne', 'celui/celle qui discute en ligne', d'aucuns écrivent *tchatcher*, *tchatcheur*, *tchatcheuse*, les faisant dériver de *tchache* 'bavarder', mot lui-même originaire de l'espagnol *chacharear*.

En chinois, la dérivation est un procédé morphologique plutôt marginal, la majorité des mots étant formée par composition. Mais, c'est en ces termes que les linguistes s'accordent à dire que les exemples (3)-(4) sont dérivés des bases primitives que constituent les adjectifs *tian*² 甜 et *tian*²*mi*⁴ 甜蜜:

- (3) *tian*² 甜 'sucré, doux'
 > *tian*²*tian*² (*de*⁰) 甜甜(的) 'sucré, doux' (+ traits sémantiques) [A]_{ADJ} > [AA]_{ADJ}¹¹
- (4) *tian*²*mi*⁴ 甜蜜 (sucré miel) 'doux, heureux'
 > a. *tian*²*mi*⁴*mi*⁴ 甜蜜蜜 (sucré miel+DUP) [AB]_{ADJ} > [ABB]_{ADJ}
 > b. *tian*²*tian*²*mi*⁴*mi*⁴ 甜甜蜜蜜 (sucré+DUP miel+DUP) [AB]_{ADJ} > [AABB]_{ADJ}

Certes, si on compare (3) avec (1), on constate qu'il y a, en surface, une forte similitude entre la reduplication des adjectifs et la répétition des verbes. En effet, la réalisation syntaxique d'un lexème et la dérivation ayant pour cadre commun le niveau lexical, il est malaisé d'établir la frontière pour les séparer. Même dans une langue dite flexionnelle, il existe de nombreux cas intermédiaires.¹² Cependant, cette similitude apparente est trompeuse. Car, un procédé dérivationnel et un procédé syntaxique sont deux phénomènes différents, appartenant à des niveaux d'analyse différents. C'est ce que nous allons essayer de montrer à travers le cas de la reduplication en chinois.

¹¹ Les adjectifs redupliqués prennent de manière systématique et automatique une valeur sémantique supplémentaire (*Cf. infra*). Pour simplifier, nous n'intégrons pas cette valeur dans notre traduction.

¹² Par exemple, le participe présent du français (*convaincant, surprenant...*) qui se trouve à mi-chemin entre le verbe et l'adjectif.

3. Le phénomène de la reduplication en chinois

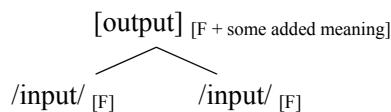
Depuis une vingtaine d'années, on a vu paraître nombre d'études, notamment en chinois,¹³ sur le phénomène appelé *chong²die²* 重疊, terme le plus employé pour désigner la reduplication. Ces travaux portent aussi bien sur le chinois standard que sur les dialectes ou d'autres langues sinitiques, et traitent autant de leurs fonctions syntaxiques que de leurs formes phonologiques et de leurs interprétations sémantiques. A tel point que Shi, en 1996 déjà, qualifiait la reduplication de «domaine le plus étudié de la grammaire chinoise».

Néanmoins, cet auteur fait remarquer que, faute sans doute d'une vision «typologique», il existe des confusions sur la vraie nature de la reduplication et sur la place qu'elle occupe dans le système grammatical du chinois. Car, selon lui, susceptibles d'exprimer une quantification, la plupart des formes redupliquées (verbes, adjectifs, classificateurs, etc.) doivent être considérée comme relevant de la «reduplication syntaxique» *ju⁴fa³ chong²die²* 句法重疊. Autrement dit, à l'opposé du point de vue traditionnel, Shi fait basculer les redupliqués de ces trois catégories du domaine lexical où on les rangeait dans celui de la syntaxe.

Pour nous, ce traitement n'est pas fondé. Concernant le statut grammatical de la reduplication, la confusion règne toujours et ce, en dépit des descriptions plus fines et plus approfondies qu'on a pu proposer dans des études récentes.

Dans la présente étude, le terme «reduplication» est employé dans un sens plus large, tel qu'il est communément accepté dans la littérature linguistique: il désigne toute réitération de MORPHEME en tant que procédé morphologique ou morphosyntaxique.

Pour définir ce qu'elles appellent la reduplication morphologique (morphological reduplication) et la distinguer de la «copie phonologique» (phonological copying), Inkelas & Zoll (2005) présentent le schéma suivant:



where [F] = semantic feature bundle

Ainsi, on exclut les phénomènes suivants:

¹³ En effet, il existe peu d'études en langues occidentales portant sur la reduplication en chinois. On peut citer ici Paul (2004, à paraître) et Paris (2007). Dans Packard (2000), en revanche, la question de la reduplication n'est guère mentionnée.

- (i) La «copie phonologique», telle qu'elle est définie par Inkelas et Zoll (2005: 197), et qui opère sur des segments plus petits que la syllabe, comme dans *you²yu⁴* 猶豫 'hésiter' (allitération) ou *mei²gui⁴* 玫瑰 'rose' (rime);
- (ii) La «réitération iconique» qui relève du moyen discursif de renforcement communicatif, comme *fei¹chang²fei¹chang²da⁴* 非常非常大 'très très grand', *gong¹xi³gong¹xi³* 恭喜！恭喜！ 'Félicitations! Félicitations!' ou *kuai⁴lai², kuai⁴lai²!* 快來！快來！ 'Viens vite! Viens vite!';
- (iii) La reprise d'un morphème dans les expressions phraséologiques, comme *you³tiao²you³li³* 有條有理 (avoir ordre avoir raison) 'bien organisé';
- (iv) La reduplication des onomatopées qui se présente de manière particulière.

Deux cas particuliers qui méritent des études plus approfondies:

1. La reduplication des adjectifs dissyllabiques [AB]_{ADJ} sur le modèle morphologique de la répétition des verbes, [[AB] [AB]], et ce, tant sur le plan phonologique que sémantique. Par exemple, *wo³xiang³an¹jing⁰-an¹jing⁰* 我想安靜安靜 (1SG vouloir être.tranquille être.tranquille) 'J'ai envie d'être un peu tranquille'. Le sujet parlant exprime un état qu'il souhaite atteindre, à savoir, «(qqn) devenir ainsi» ou «rendre qqn ainsi». En fait, très peu d'adjectifs se prêtent à ce type de «prédication causative [+contrôle]» (Paris 2007:71). Pour Li (1996), il s'agit d'«adjectifs provisoirement employés comme des verbes», tandis que pour Paul & Yang-Drocourt (2009), *anjing* 'être tranquille' est ici un verbe, homonyme de l'adjectif 'calme'.

2. La reduplication des verbes [A]_V et [B]_V empruntant à la la reduplication des adjectifs sa forme morphologique et phonologique, [AABB]_V. Par exemple, *ke⁴ren⁰-men⁰shuo¹shuo¹xiao⁴xiao⁴(de⁰), dou¹hen³kai¹xin¹* 客人們說說笑笑(的), 都很開心 (invité-PL parler+Dup rire+Dup (de), tous très content) 'Les invités parlaient et riaient, ils étaient tous très contents'. Grâce à cette forme, les verbes acquièrent des propriétés syntaxiques et sémantiques que, séparées, leurs bases ne possèdent pas: ils deviennent plus descriptifs et statifs. En exprimant le caractère répétitif et ininterrompu des procès s'étalant sur une certaine durée, ils atténuent leur dynamique et leur ponctualité. A l'instar des adjectifs, ils assument non seulement la fonction de prédicat, mais surtout celle de modifieur et de circonstant.

Après avoir ainsi délimité le terrain et selon les critères que nous venons d'exposer, nous proposerons un premier classement de tous les redupliques en trois grands groupes, en fonction de leurs statuts de «mots grammaticaux» ou de «mots construits», les premiers étant, rappelons-le, des réalisations syntaxique, les seconds des résultats de la dérivation.

A. La répétition ou le redoublement des lexèmes au niveau de la syntaxe. Cela concerne:

- (i) Les verbes volitifs, $[_{SV}[A]_V [A]_V]$ ou $[_{SV}[AB]_V [AB]_V]$, dont il a été question plus haut, ainsi que la structure $[_{SV}[A]_V [A]_V NP]$, comme *san san bu* 散散步 ‘faire une petite promenade’.
- (ii) Les classificateurs et les «noms d’unité» — à distinguer des «noms communs» ci-dessous. Ils se répètent selon un modèle unique: $[_{CL/SN}[A]_{CL}[A]_{CL}]$ *ge⁴ge⁴* 個個 (CL+DUP) ‘tous, chaque’ et $[_{SN}[A]_{NU}[A]_{NU}]$ *chu⁴chu⁴* 處處 (endroit+DUP) ‘partout’. Comme le signale Lü (1954), très peu de noms communs sont susceptibles de se prêter à ce genre de variation. Il s’agit en fait de noms utilisés comme des classificateurs, sorte de «vestige du chinois classique». Quant aux classificateurs, tous ceux qui peuvent être quantifiés par n’importe quel numéral sont susceptibles d’être ainsi redoublés¹⁴ (Shi 1996:4) et la forme produite crée automatiquement le sens grammatical de distributivité, ‘tous, chaque’.

B. Mots construits par redoublement dans le processus diachronique de dissyllabisation du chinois. Ce sont:

- (i) Les noms communs dissyllabiques. Ils sont issus de la réduplication complète, $[AA]_N$ tels que *qu¹qur⁰* 蚰蚰兒 ‘grillon’, *ma¹ma⁰* 媽媽 ‘maman’ ou *xing¹xing⁰* 星星 ‘étoile’.
- (ii) Tous les adverbes $[AA]_{ADV}$. Là aussi, le redoublement se fait sur un modèle unique: *chang²chang²* 常常 ‘souvent’ ou *pian¹pian¹* 偏偏 ‘comme un fait exprès’. Pour la plupart des membres de ce groupe, le procédé a permis de créer une unité lexicale à partir d’un morphème non libre ou d’un «morphème unique» vide de sens. Mais, ce type de dissyllabisation n’est plus productif sur les noms communs et les adverbes. En chinois moderne, on a donc deux petites listes fermées. En outre, la différence sémantique entre la base et la forme dérivée est soit inexistante, soit à peine perceptible.
- (iii) Les noms communs trisyllabiques peuvent éventuellement être inclus dans ce groupe. Mais, il faut signaler que leur formation est différente: il s’agit de la composition d’une forme rédupliquée et d’un monosyllabe, $[ABB]_N$ ou $[AAB]_N$. Cette forme rédupliquée peut être elle-même un dissyllabe existant et

¹⁴ Par exemple, on dit *liang²-ma³ shi⁴* 兩碼事 (deux-CL chose) ‘deux choses (différentes)’, mais **san¹-ma³ shi⁴* *三碼事 (trois-CL chose) et **ma³-ma³ shi⁴* *碼碼事 (CL-DUP chose) sont impossibles.

fait donc partie des deux groupes ci-dessous, comme *lao³tai⁴tai⁰* 老太太 (vieux dame) ‘vieille dame’, *mao²mao²yu³* 毛毛雨 (poil.fin pluie) ‘bruine’ ou *qiao¹qiao¹hua⁴* 悄悄話 (à.voix. basse parole) ‘bavardage à voix basse’, mais elle peut aussi être un morphème lié qui ne figure jamais ailleurs que dans ce trisyllabe, comme *la¹la¹dui⁴* 啦啦隊 (*la*_{Interj}+Dup équipe) ‘équipe de supporters’.

C. Mots construits par reduplication tout en créant une valeur supplémentaire. Ce sont:

- (i) Les adjectifs redupliques, tous modèles confondus, [AA]_{ADJ}, [AABB]_{ADJ}, [ABB]_{ADJ}, [BABA]_{ADJ} et [AliAB]_{ADJ}.
- (ii) Les noms quadrisyllabiques formés par la reduplication de deux morphèmes monosyllabiques, [AABB]_N. Ceux-ci sont obligatoirement des co-hyponymes appartenant au même champ lexical: *tan²tan²guan⁴guan⁴* 罐罐罐罐 (jarre+Dup pot+Dup) ‘babioles’. Ainsi, [AABB]_N permet de créer le sens supplémentaire [+hétérogène], [+pluriel].

Malgré leur diversité, tous les modèles du groupe C peuvent être marqués d’au moins deux grandes caractéristiques de la morphologie lexicale: premièrement, ils sont tous productifs et permettent au locuteur des créations nouvelles, voire idiosyncrasiques; deuxièmement, il n’existe pas nécessairement en input des mots attestés permettant de construire une forme plus complexe.

4. Différentes analyses concernant les adjectifs redupliques

Parmi tous les éléments du groupe C, les adjectifs sont numériquement les plus importants et morphologiquement les plus complexes. Leur analyse diverge considérablement selon les linguistes. Cela se traduit essentiellement par un flottement notionnel entre «réalisation syntaxique» et «dérivation»: devant les mêmes formes d’un même modèle, on parle tantôt de l’une, tantôt de l’autre; ou, au contraire, on met dans le même panier tous types de reduplication et tout cela, selon des critères sensiblement différents. En bref, nous nous trouvons devant un «état de l’art» assez surprenant.

Pour résumer les raisons de cette forte divergence, voyons quels sont ces critères ainsi que les principaux arguments avancés en leur faveur par les uns et les autres:

I. Critère fondé sur l’existence ou non d’un mot attesté en input. Selon Zhang (1981), Liu (1990[2005]) et Ge (2001), la reduplication doit être considérée comme une

réalisation syntaxique quand elle s'opère sur la base d'un mot (*ci*² 詞) du lexique pour aboutir à une autre forme du même mot. Dans le cadre de la formation lexicale, au contraire, on échappe à cette contrainte et l'on peut, par le même procédé, former des mots nouveaux à partir d'éléments de nature diverse, que ce soient des syllabes non significatives ou des morphèmes liés.

Selon ce critère, la majorité des adjectifs redupliqués relève du fait syntaxique, seulement une petite partie étant issue de la formation lexicale. Par exemple, si *hong*²*hong*² 紅紅 'rouge' [AA]_{ADJ} n'est qu'une réalisation de *hong*² 紅 'rouge' [A]_{ADJ}, *jiu*¹*jiu*¹ 赳赳 '(l'air) gaillard' [AA]_{ADJ} est, en revanche, une création lexicale, puisque *jiu*¹ 赳 /A/ seul n'est qu'une syllabe vide de sens. Pour la même raison, *gan*¹*gan*¹*jing*⁴*jing*⁴ 乾乾淨淨 'propre' [AABB]_{ADJ} est analysé comme une réalisation de *gan*¹*jing*⁴ 乾淨 'propre' [AB]_{ADJ} et *man*⁴*teng*¹*teng*¹ 慢騰騰 'lent, traînant' [ABB]_{ADJ} comme celle de *man*⁴*teng*⁰ 慢騰 'lent, traînant' [AB]_{ADJ}, mais, *xing*²*xing*²*se*⁴*se*⁴ 形形色色 'de toutes sortes' [AABB]_{ADJ} et *lü*⁴*you*²*you*² 綠油油 'vert luisant' [ABB]_{ADJ} sont des résultats de la formation lexicale, car, à la base, **xing*²*se*⁴ [AB] et **lü*⁴*you*² [AB] ne sont pas attestés. De ce point de vue, pour déterminer leur statut grammatical, on devrait examiner les adjectifs redupliqués cas par cas.

II. Critère fondé sur l'ajout ou non d'un sens lexical. Selon Zhang (1981) et Liu (1990[2005]), la réalisation syntaxique permet, par définition, de créer un sens supplémentaire d'ordre grammatical, alors que la formation lexicale doit apporter un sens supplémentaire d'ordre lexical. Ainsi, entre *hong*² 紅 'rouge' [A]_{ADJ} et *hong*²*hong*² 紅紅 'rouge' [AA]_{ADJ}, comme il n'y a qu'une «augmentation de degré», sans qu'il y ait de sens lexical ajouté, on doit parler de réalisation syntaxique; mais quand on passe de *hong*² 紅 'rouge' [A]_{ADJ} à *hong*²*yan*⁴*yan*⁴ 紅艷艷 'd'un rouge splendide' [ABB]_{ADJ}, il y a en revanche formation d'un mot différent en raison du sens lexical apporté par [B] *yan*⁴ 艷 'splendide, lumineux'.

III. Critère fondé sur l'ajout ou non du sens grammatical de «quantification». Selon Shi (1996), toute «réduplication syntaxique» est susceptible d'apporter à la forme primitive une valeur supplémentaire de quantification, qui peut être soit une augmentation, telle qu'en crée la reduplication des adjectifs ou des classificateurs, soit une diminution, comme on peut l'observer dans le cas de la répétition des verbes. C'est la raison pour laquelle les «adjectifs à valeur distinctive» (*qu*¹*bie*²*ci*² 區別詞), comme *dan*¹ 單 'impair' ou *zheng*⁴ 正 'principal', qui, par définition, ne se prêtent pas à l'expression de degré, ne saurait être redupliqués. Pour lui, néanmoins, la sous-catégorie d'adjectifs construits par subordination [BA]_{ADJ} semble faire exception: en effet, bien que ces adjectifs soient, eux aussi, incompatibles avec une expression de degré, **hen*³ *bi*³*zhi*² *很筆直 'très droit (comme un pinceau)', on peut parfaitement les redupliquer, ce qui donnera ici *bi*³*zhi*²*bi*³*zhi*² 筆直筆直 'droit (comme un pinceau)' [BABA]_{ADJ}. Toujours

selon Shi, étant donné qu'il n'y a «pas de différence de degré» entre la forme simple [BA] et la forme redoublée [BABA], «ce phénomène peut être considéré comme de la reduplication rhétorique». Guo (2002), également, traite à part les adjectifs [BABA]_{ADJ}, mais pour des raisons différentes (Cf. ci-dessous).

IV. Critère fondé sur l'identité ou non de la fonction syntaxique entre input et output. Selon Guo (2002), on parlera de «fait syntaxique», notion très proche de ce que d'autres auteurs appellent «réalisation syntaxique», si les deux formes, la simple et la redoublée, assument la même fonction syntaxique. Ainsi, il analyse la plupart des adjectifs redoublés comme étant les résultats de la formation lexicale, puisque, par exemple, *gan¹jing⁴* 乾淨 'propre' [AB]_{ADJ} et *gan¹gan¹jing⁴jing⁴* 乾乾淨淨 'propre' [ABAB]_{ADJ}, n'ont pas les mêmes propriétés syntaxiques. Pour lui, deux groupes font exception: le premier rassemble les constructions [AABB]_{ADJ} dans lesquelles [A] et [B] sont des antonymes, comme *da⁴da⁴xiao³xiao³* 大大小小 (grand+ DUP petit+ DUP) 'de tailles différentes (grands et petits)', à cause d'une «reduplication de sens» (*op. cit.*: 42) de chacun des deux morphèmes; le second regroupe les formes [BABA]_{ADJ}, comme *xue³bai²xue³bai²* 雪白雪白 'blanc comme neige', «reduplication syntaxique» de *xue³bai²* 雪白 [AB]_{ADJ}, puisque les deux formes sont des unités lexicales et que leur fonction syntaxique est identique.

V. Critère fondé sur la catégorie des morphèmes en input. Guo (2002) réserve un traitement particulier à certains adjectifs en fonction de la nature grammaticale des morphèmes avant la reduplication. Ainsi, *san¹san¹liang³liang³* 三三兩兩 [AABB]_{ADJ} (trois+ DUP deux+ DUP) 'par petits groupes' et *qian¹qian¹wan⁴wan⁴* 千千萬萬 [AABB]_{NUM} (mille+ DUP dix.mille+ DUP) 'd'un très grand nombre' sont par lui regroupés comme étant générés par la reduplication de numéraux, tandis que *feng¹feng¹huo³huo³* 風風火火 [AABB]_{ADJ} (vent+ DUP feu+ DUP) '(allure) rapide, preste' et *tiao²tiao²kuang⁴kuang⁴* 條條框框 [AABB]_N (article+ DUP règle+ DUP) 'toutes sortes de règles et contraintes' sont analysés comme des reduplication de noms, etc.

Malgré leurs approches différentes, la plupart de ces arguments semblent souffrir d'une même faiblesse: en se plaçant délibérément du côté input, on perd totalement de vue, d'une part, les propriétés fondamentales de la morphologie lexicale et, d'autre part, l'homogénéité fonctionnelle et sémantique d'une catégorie de mots.

5. Restituons aux adjectifs redoublés leur statut lexical

En règle générale et abstraction faite des différentes contraintes liées à leur structure interne et à leur sens lexical, les adjectifs redoublés [DUP.ADJ] peuvent être le modifieur d'un nom (i), et, dans ce cas, un *de* 的 est toujours obligatoire; mais quand

ils assument seuls la fonction de prédicat (ii) ou celle de circonstant préverbal (iii), il peut y avoir, sous certaines conditions, la présence obligatoire ou non d'un *de* (的 ou 地); en outre, ils peuvent aussi être le complément postverbal (iv) introduit par le *de* 得:

- i. [_{SN}[DUP.ADJ] *de* 的 [N]]
- ii. [_P[N] [DUP.ADJ] (*de* 的)]
- iii. [_{SV}[DUP.ADJ] (*de* 地) [V]]
- iv. [_{SV}[V] *de* 得 [DUP.ADJ]]

Nous n'allons pas procéder ici à une description de leurs propriétés syntaxiques communes, ce qui serait un autre sujet d'étude et qui, d'ailleurs, a été déjà abondamment traité. Nous allons seulement proposer quelques arguments pour montrer que, s'il est tout à fait judicieux d'accorder le statut de réalisation syntaxique à la répétition des verbes et au redoublement des classificateurs, il faut en revanche conserver ou restituer leur statut de «mots» aux adjectifs redoublés, tous modèles confondus.

5.1 Homogénéité sémantique des adjectifs redoublés

Dans un long article spécialement consacré aux adjectifs (*xing²rong²ci²* 形容詞),¹⁵ Zhu (1956) les range en deux catégories, l'une regroupant les adjectifs «de forme simple» et l'autre ceux «de forme complexe», la deuxième sous-catégorie étant dérivée de la première. Voici, schématiquement présenté, le principe de ce classement:

Les adjectifs:

I. de forme simple: *da⁴* 大 'grand', *gan¹jing⁴* 乾淨 'propre'

II. de forme complexe

A. les redoublés:

- i. [_{AA}]_{ADJ}, [_{AABB}]_{ADJ}: *da⁴da⁴* 大大 'grand',
gan¹gan¹jing⁴jing⁴ 乾乾淨淨 'propre'
- ii. [_{AliAB}]_{ADJ}: *hu²li⁰hu²tu¹* 糊裡糊塗 'étourdi'

¹⁵ Ici, le terme *xing²rong²ci²* signifie d'office «adjectifs prédicatifs». Cette position est clairement exposée dans Zhu (1982), qui les classe dans la grande catégorie des «mots de nature prédicative» *wei⁴ci²* 謂詞, alors que les lexèmes comme *nan²* 男 'masculin' ou *nü³* 女 'féminin' sont appelés «distinctifs» *qu¹bie²ci²* 區別詞, et font en revanche partie des «mots de nature substantive» *ti³ci²* 體詞.

- B. $[[A_{\text{ADJ}}][BB_{\text{SUFF}}]]_{\text{ADJ}}$: *hei¹hu¹hu¹* 黑乎乎 ‘noir’,
 $[[A_{\text{ADJ}}][XYZ_{\text{SUFF}}]]_{\text{ADJ}}$: *hei¹gu⁰long¹dong¹* 黑咕隆咚 ‘sombre’
- C. $[[B_{\text{MODIF}}][A_{\text{ADJ}}]]_{\text{ADJ}}$: *xue³bai²* 雪白 ‘blanc (comme neige)’,
bing¹liang² 冰涼 ‘glacé’
- D. $[S_{\text{Adj}}[\text{Adv}] + [\text{Adj}]]$: *hen³da⁴* 很大 ‘très grand’,
you⁴gao¹you⁴da⁴ 又高又大 ‘à la fois grand et imposant’

Par la suite, Zhu (1982:73) donnera aux adjectifs de la sous-catégorie II le nom de *zhuang⁴tai⁴xing²rong²ci²* 狀態形容詞, terme que l’on traduit en général par «adjectifs statifs» ou «adjectifs à valeur descriptive». Ceux-ci sont distincts des «qualificatifs» de forme simple (*xing⁴zhi²xing²rong²ci²* 性質形容詞), du fait qu’ils représentent en eux-mêmes un état ou un degré non graduable de la qualité exprimée.

Si l’on écarte les éléments du groupe (D) qui, étant des syntagmes et non des mots, comme le montre à juste titre Paul (2006), n’ont pas leur place dans une classification en tant que catégorie lexicale, les autres groupes rassemblent, à quelques exceptions près (notamment $[[A_{\text{ADJ}}][BCD_{\text{SUFF}}]]_{\text{ADJ}}$ *hei¹gu⁰long¹dong¹* 黑咕隆咚 ‘sombre’), tous les types d’adjectifs construits par reduplication, que celle-ci soit totale ou partielle, y compris même les membres du groupe (C), qui sont aussi parfaitement reduplicables en $[BABA]_{\text{ADJ}}$, *xue³bai²xue³bai²* 雪白雪白 ‘blanc (comme neige)’. Du fait qu’ils sont traités comme une sous-catégorie d’adjectifs, les redupliques doivent être considérés comme appartenant à un ensemble homogène de mots dotés des mêmes caractéristiques inhérentes sur les plans syntaxique et sémantique. Ainsi, il doit exister un principe, susceptible de s’appliquer à tous les éléments, et qui justifie la manière dont on les regroupe.

Or, en chinois tout particulièrement, la seule distribution syntaxique d’un mot, pas plus que sa morphologie, ne permet de définir efficacement sa nature grammaticale. En effet, il ne suffit pas de prouver qu’ils partagent le même environnement syntaxique pour pouvoir affirmer que les prédicats sont de même nature dans *ta¹gao¹* 他高 ‘3SG grand’ = ‘Il est (plus) grand’ et *ta¹qu⁴* 他去 ‘3SG aller’ = ‘Il y va’.

C’est pourquoi, Guo (2002:23-24) propose un critère différent: «Outre son sens lexical, sa morphologie, ses caractéristiques phonologiques et la situation pragmatique, ce qui détermine les propriétés distributionnelles d’un mot, c’est avant tout son sens grammatical, ou plus précisément, sa fonction d’expression».¹⁶ Rappelons que, pour

¹⁶ Guo parle de quatre fonctions: narration, désignation, qualification ou autres rôles secondaires.

Inkelas & Zoll (2005:197) également, une identité sémantique entre les deux éléments en input et un certain sens ajouté en output constituent les conditions nécessaires sans lesquelles on ne peut parler de reduplication morphologique. En accord avec ces points de vue, nous considérons que le véritable trait distinctif qui permet de regrouper tous les adjectifs redupliqués, c'est l'homogénéité de leur valeur sémantique supplémentaire créée par la reduplication.

Cependant, la valeur sémantique n'est pas le sens grammatical. Il arrive que cette valeur sémantique coïncide avec une augmentation (ou une diminution) de degré par rapport à la forme simple, mais elle est plutôt due à des contextes particuliers et/ou au sens lexical de certains mots.¹⁷ En tout cas, cela est foncièrement différent de la «quantification» dont font état plusieurs études (Zhang 2000, Shi 1996, Paris 2007, Chen 2008). En effet, jusqu'ici, les discussions sur ce sujet paraissaient très peu convaincantes, aucun argument ne permettant d'évaluer clairement la différence de degré qui existerait entre, d'une part, *lan² tian¹ bai² yun²* 藍天白雲 'ciel bleu, nuage blanc' ou *gan¹ mian⁴ bao¹* 乾麵包 'pain sec', et, d'autre part, *lan²lan² de tian¹*, *bai²bai² de yun²* 藍藍的天, 白白的雲 ou *gan¹ba¹ba¹ de mian⁴bao¹* 乾巴巴的麵包. En effet, cette différence n'existe pas, pour la simple raison que le locuteur n'a aucune intention de signifier un quelconque degré; il cherche avant tout à faire une image, à décrire un état de la qualité exprimée par l'adjectif, dans le seul but d'agrémenter son discours. Voilà pourquoi les redupliqués constituent l'une des «formes vives» des adjectifs.

Au sujet de ce sens ajouté, rappelons ce passage, malheureusement si souvent négligé, qu'on lit dans Zhu (1956:109): «*da⁴da⁴ de⁰ yan³jing⁰* 大大的眼睛 'grands yeux', *hei¹hei¹ de⁰ pi²fu⁰* 黑黑的皮膚 'peau mate', *gao¹gao¹ de⁰ bi²zi⁰* 高高的鼻子 'nez long', *wan¹wan¹ de⁰ mei²mao⁰* 彎彎的眉毛 'sourcils arqués' ne signifient pas que les qualités de *da⁴* 大 'grand', de *hei¹* 黑 'noir', de *gao¹* 高 'haut' ou de *wan¹* 彎 'courbé' ont atteint leur plus haut degré, mais que leur degré est exactement tel qu'il doit être.¹⁸ [...] Bien que *gao¹ gu¹guai³* 高孤拐 'pommettes saillantes', *duan³ tui³* 短腿 'jambes courtes' ou *huang² lian³* 黃臉 'visage jaunâtre' ne soient pas des images agréables, lorsqu'on utilise les formes redupliquées, ces expressions donnent une impression totalement différente».

En un mot, la forme simple est une dénotation, alors que la forme redupliquée est une CONNOTATION. Et comme toute connotation, c'est une appréciation subjective que le sujet parlant choisit d'ajouter. «Un sens s'ajoute à un autre sens, sans remplacer quoi que ce soit. Il est choisi par le locuteur tout comme l'est un mot: en fonction du sens

¹⁷ Selon Zhu (1956:108), il y a une intensification lorsqu'un redupliqué se trouve en position de circonstant et de complément: *da⁴da⁴ de⁰ qing³ yi¹-ci⁴ ke⁴* 大大地請一次客 (grand+DUP SUB inviter un fois hôte) 'donner une très grande réception'.

¹⁸ *Qia⁴dao⁴ hao³chu⁰* 恰到好處.

exprimé, de l'attitude ou du désir du locuteur, etc. La langue, en tout cas, ne l'impose pas» (Mel'čuk 1993:288). Lorsque le sens lexical le permet, le locuteur y ajoute aussi son propre jugement de valeur en signifiant que la grande taille de cette personne est, pour lui, «exactement telle qu'elle doit être». En ce qui concerne les «Vivid Reduplicates», Chao (1968:209) note aussi les «favorable connotations». Il donne l'exemple suivant, *duan³duanr¹ de⁰ xia⁴bai³, gao¹gaor¹ de⁰ ken⁴* 短短兒的下擺, 高高兒的根 et le traduit lui-même par 'nice and short-length skirt, with nice and high slits', avant d'ajouter qu'il faut voir là «a form of description which one would not use if one did not approve of such dresses». Pour faire comprendre cette valeur connotative, Chao a dû proposer cette traduction un peu forcée, car on sait que la connotation n'est pas réellement traduisible d'une langue à l'autre. Dans (6), à l'instar de Chao, nous avons cherché à rendre cette connotation par un moyen lexical:

- (5) 他個子很高 / 高個子
ta¹ ge⁴zi⁰ hen³ gao¹ / gao¹ ge⁴zi⁰
 3SG taille très grand / grand taille
 'Il est de grande taille.'
- (6) 他個子高高的 / 高高的個子
ta¹ ge⁴zi⁰ gao¹gao¹ de⁰ / gao¹gao¹ de⁰ ge⁴zi⁰
 3SG taille grand+DUP SUB / grand+DUP SUB taille
 'Il est de belle taille.'

La valeur expressive des formes rédupliquées n'est nullement un trait spécifique du chinois. Dans son étude sur le grec, Skoda (1982:37-38) traite le phénomène comme «un universal linguistique» en utilisant selon les contextes les termes de «redoublement expressif» ou «impressif». Car, selon elle, bien que ce dernier terme puisse définir tout mot qui a pour propriété spécifique, par sa structure même, de créer une impression, «il n'en reste pas moins que le redoublement qui fait de ces termes des impressifs est lui-même expressif». En effet, «il répond à une intention du locuteur, qui, usant de ce procédé, exprime ce qu'il ressent en communiquant son impression à son interlocuteur». Labrune (1998:100), à propos des mots dérivés en japonais Yamato, fait aussi remarquer que les valeurs linguistiques des dérivés par reduplication sont «le plus souvent couplées avec une valeur de subjectivité et à une valeur d'expressivité caractéristiques de ce mode de formation lexicale», car «la reduplication est un procédé qui a pour propriété de ne pas référer à du notionnel: il implique que l'on quitte le domaine de la notion pour basculer dans celui de l'occurrence, du particulier» (p.102). Pour Morgenstern & Michaud (2007:123), «l'étude de la reduplication illustre les liens qui unissent le domaine du discours et celui de la langue; autrement dit, le stylistique et le linguistique. Plus

spécifiquement, la reduplication paraît marquer, quelle que soit la partie du discours concernée, un commentaire de l'énonciateur (une prise en charge modale), un retour sur ce qu'il rapporte: événement, propriété, référent... Des faits analogues s'observent dans des langues non apparentées».

De ce point de vue, ce qui distingue fondamentalement les formes simples et les formes redupliquées des adjectifs réside dans les différents niveaux qui séparent la connotation de la dénotation, la subjectivité de l'objectivité, l'occurrence du concept. De par leur fonction d'expression et leurs valeurs sémantiques connotatives, les redupliqués ne peuvent être considérés que comme des unités lexicales à part entière.

5.2 Les deux branches de la morphologie

Que les adjectifs redupliqués constituent un ensemble homogène, c'est là une thèse en faveur de laquelle nous allons maintenant proposer d'autres arguments, mais sous un angle différent.

Comme on l'a vu précédemment, certaines études considèrent que les adjectifs redupliqués dont la base est un mot attesté sont des formes d'une réalisation syntaxique, et que les autres sont des produits issus de la formation lexicale. Pour nous, cette position est le résultat d'une confusion entre les notions de «formation lexicale» et de «structure interne des mots», ou, selon les termes employés par Dong (2004, 2007), entre «morphologie» et «lexique»: on se place du côté input pour examiner les résultats de dérivation en output; on procède à une description du lexique, en l'occurrence, à la définition du statut grammatical de certains de ses éléments, sous l'angle d'un processus qu'est la formation lexicale. Comme le fait remarquer Dong (*op. cit.*: 127), s'il y a eu tant de polémiques sur le statut des «mots» en chinois, c'est parce qu'on n'a pas bien saisi la différence entre ces deux notions.

Il faut rappeler que la morphologie lexicale comporte deux branches: la première étudie la formation lexicale comme un processus de production; la seconde, placée du côté de la réception, examine la structure interne des mots en les traitant comme des produits. C'est d'ailleurs dans ce sens que Mel'čuk (1993:305-308) assigne «deux tâches» à la dérivation, l'une constructive et l'autre descriptive. Si on compare la formation lexicale à une machine à mots qui fonctionne en permanence, les mots attestés ne sont que ses produits à un moment donné, et représentent par définition un état temporaire du lexique, un ensemble limité par rapport à la potentialité infinie de la production discursive. Cette distinction est particulièrement importante lorsqu'on aborde l'étude de la dérivation.

Voici des exemples pour illustrer ces deux «tâches». Dans l'ouvrage intitulé «800 mots du chinois moderne» (Lü 1984), on trouve une présentation détaillée de plus de

trois cents «adjectifs de formes vives» *xing²rong²ci² sheng²dong⁴ xing²shi⁴* 形容詞生動形式; ceux-ci sont répartis selon leurs modèles en quatre listes, que nous schématiserons comme suit:

- (i) [A]_{ADJ} > [AA]_{ADJ}: *jin³jin³ (de⁰)* 緊緊(的) ‘serré’,
hong²hong² (de⁰) 紅紅(的) ‘rouge’
- (ii) [A]_{ADJ} > [ABB]_{ADJ} ou [ACB]_{ADJ} ou [XYZ]_{ADJ}: *yuan²liu¹liu¹* 圓溜溜 ‘rond’,
yuan²de⁰liur¹ 圓得溜兒 ‘rond’ ou
yuan²gu⁰long¹dong¹ 圓咕隆冬
‘rond’
- (iii) [AB]_{ADJ} > [AABB]_{ADJ} ou [AliAB]_{ADJ}: *gu³gu³guai⁴guai⁴* 古古怪怪
‘bizarre’,
gu³li⁰gu³guai⁴ 古里古怪 ‘bizarre’
- (iv) [BA]_{ADJ} > [BABA]_{ADJ}: *xue³bai²xue³bai²* 雪白雪白
‘blanc (comme neige)’

À quelques détails près, cette répartition recoupe le classement de Zhu (1956, 1982), dans la mesure où tous les types d’adjectifs dérivés y sont présentés. Néanmoins, l’approche est différente: cherchant plus ou moins l’exhaustivité, les auteurs proposent quatre listes fermées, établies à partir des seuls mots existants. Bien que ce soit certainement l’une des présentations les plus complètes sur ce sujet, on remarque l’absence de tout un type: en effet, quoique formés selon le même modèle que ceux présentés au groupe III, de nombreux éléments ne figurent pas dans la liste pour la simple raison que leur base n’est pas attestée en tant que dissyllabe dans le lexique. Citons, par exemple, *xing²xing²se⁴se⁴* 形形色色 ‘de toutes sortes’.

Certes, aucune liste ou aucun dictionnaire ne peut prétendre à l’exhaustivité quand il s’agit des mots, qui appartiennent, par définition, à des listes ouvertes. C’est la raison pour laquelle Mel’čuk (1993:312) propose d’inclure dans le dictionnaire uniquement ce qu’il appelle des «dérivés au sens faible», à savoir, ceux qui ne présentent pas de «trace de dérivation» aux yeux du locuteur. Quant aux «dérivés au sens fort», qui sont, en revanche, construits de façon régulière à partir d’un radical attesté, qui sont tout à fait constructibles selon des règles et ne manifestent aucune «phraséologisation», et il n’est pas nécessaire de les intégrer dans le dictionnaire.

C’est grosso modo l’approche adoptée par le «Dictionnaire du chinois moderne» *Xian⁴dai⁴ han⁴yu³ ci²dian³* 現代漢語詞典, si on en juge par les relevés que Zhou (2004: 264-282) a effectués sur l’édition de 1998. En effet, concernant le modèle [AABB]_{ADJ}, le

dictionnaire offre une entrée à tous les «dérivés au sens faible», du type *xing²xing²se⁴se⁴* 形形色色 ‘de toutes sortes’, et seulement à ceux-là, alors qu’il la refuse à tous les dérivés au sens fort comme *gan¹gan¹jing⁴jing⁴* 乾乾淨淨 ‘propre’ ou *hong²hong²* 紅紅 ‘rouge’.¹⁹

Reste qu’on ne saurait confondre un principe adopté par un dictionnaire et la réalité du lexique. Comme le signale Dong (2004:127), il n’y a aucune raison de bannir du lexique des mots construits selon les modèles morphologiques productifs et réguliers, sous prétexte que cela le gonflerait de manière insupportable. En effet, il n’y a pas seulement deux branches de la morphologie, il y a aussi deux notions du lexique.

5.3 Les deux lexiques

Les deux branches de la morphologie lexicale correspondent à deux notions du lexique. En effet, le lexique réel n’est que la partie observable du lexique potentiel. Selon Aronoff (1976:18), «there are two lexicons: a list of morphemes and their meanings which, together with the morphology, defines the class of possible words of a language; and a word lexicon. The actual words are a subset of the possible.» En effet, l’analyse de la structure interne des mots ne porte que sur les seuls éléments réellement attestés dans les dictionnaires ou observables dans les corpus à l’état statique, mais, dès lors qu’on aborde la formation lexicale, on doit prendre en compte les règles imposées par la langue et les possibilités qu’elle offre, qui ont permis et permettront de produire des mots possibles, dont certains existeront probablement un jour.

Traitant des régularités dérivationnelles du français, Corbin (1987:42) va plus loin: selon elle, «la distinction entre mots ‘potentiels’ et mots ‘réels’ n’a pas de pertinence linguistique. Les mots peuvent être attestés ou non», car «leur (non-) attestation n’est qu’un accident sans rapport avec leur ‘existence’ linguistique. Tous les mots construits réguliers sont ‘réels’, même si tous ne sont pas attestés».

Les études de Tang (1982) et de Hua (2002) sont justement placées dans cette optique. Par des méthodes différentes, mais en se fondant sur une analyse des adjectifs redupliqués attestés à partir de leur sens lexical, de leur nature grammaticale, de leur structure interne et de leur registre, ces deux auteurs cherchent à dégager les régularités selon lesquelles certains types d’adjectifs se prêtent plus facilement que d’autres à la reduplication. Selon eux, la catégorie et la structure interne d’un mot peuvent être considérées comme des contraintes grammaticales, alors que le jugement de valeur et le registre sont du domaine des contraintes pragmatiques. Ainsi, un mot ne possèdera la

¹⁹ Néanmoins, on y remarque certains traitements arbitraires et aléatoires. Par exemple, on y retrouve *da⁴da⁴* 大大 ‘grand’ et *zhong⁴zhong⁴* 重重 ‘lourd’, mais pas *xiao³xiao³* 小小 ‘petit’ ni *qing¹qing¹* 輕輕 ‘léger’.

potentialité d’être rédupliqué que lorsqu’il remplira à la fois les conditions grammaticales et pragmatiques. Surtout, ils signalent tous deux que les adjectifs rédupliqués relèvent clairement du registre courant, voire oral, ce qui explique leur emploi quasi nul dans les textes scientifiques, politiques ou administratifs.

Cette dernière caractéristique permet de comprendre pourquoi la reduplication, en tant que procédé, est susceptible d’être pleinement exploitée en littérature. Chez certains écrivains, elle est même poussée à un niveau élevé et présente une remarquable potentialité de création idiosyncrasique.

D’après Cai & Shi (2007), on constate, dans les romans de Yan Lianke 閻連科,²⁰ un emploi surabondant des formes rédupliquées trisyllabiques et quadrisyllabiques appartenant à diverses catégories lexicales. On relève, en particulier, un choix extrêmement libre tant des modèles de construction que des morphèmes entrant dans la reduplication. Cet auteur a créé ainsi de nombreux adjectifs qui ne sont pas attestés dans le lexique, mais qui sont conformes aux règles syntaxiques et parfaitement compréhensibles pour le lecteur. Les procédés utilisés sont variés, qui consistent, sur des modèles existants, à associer librement deux morphèmes jusqu’alors jamais rapprochés, comme *lie⁴lie⁴yan²yan²* 烈烈炎炎 ‘torride’, ou à donner une nouvelle structure à des mots attestés, comme *xian¹ming²ming²* 鮮明明 ‘clair’ ou *ruan³mian²ruan³mian²* 軟綿軟綿 ‘doux, mou’. Cependant, la particularité la plus notable de Yan est la création d’un modèle inexistant d’adjectifs rédupliqués, [ABCC]_{ADJ}.²¹ On a ainsi *xia²chang²wan¹wan¹* 狹長彎彎 ‘étroit et sinueux’ ou *hong²can⁴lan⁴lan⁴* 紅燦爛爛 ‘rouge et brillant’. Les soixante-douze exemples relevés prouvent qu’il ne s’agit certainement pas d’emplois fortuits ou accidentels. Ces créations idiosyncrasiques peuvent paraître de prime abord fantaisistes, mais, nombre d’entre elles sont fondées sur le plan morphologique, et elles réussissent en tout cas à satisfaire un besoin rhétorique propre à l’auteur. Selon Cai & Shi, «le procédé se révèle parfois comme une nécessité grammaticale, mais le plus souvent, il a pour but d’intensifier l’effet d’un mot, d’un sentiment, d’une scène, d’une ambiance, etc. Il peut aussi s’afficher comme une marque d’inventivité ou témoigner de recherches d’expressivités nouvelles». Signalons que c’est aussi dans ce sens que Singh

²⁰ Yan Lianke, originaire du Henan, est considéré et se considère lui-même comme un «écrivain du terroir». Cai & Shi se sont donc posé la question de savoir si, tout compte fait, ses créations libres n’avaient pas tout simplement subi l’influence de sa langue maternelle. D’après leurs investigations, la réponse s’avère négative: non seulement d’autres natifs du Henan ne reconnaissent pas ces formes comme appartenant à leur dialecte, mais de plus, il semblerait que cette pratique se rencontre également chez quelques autres écrivains d’origines différentes, tels que Xing Zhu 星竹, Sun Chunping 孫春平 ou Qiao Ye 喬葉.

²¹ Les adjectifs ABCC existent, mais ils sont de structure différente [[AB]_N [CC]_{ADJ}]_{ADJ}, comme *xi³qi⁴yang²yang²* 喜氣洋洋 ‘joyeux’, *xiao³xin¹yi⁴yi⁴* 小心翼翼 ‘précautionneux’, *wen¹qing²mo⁴mo⁴* 溫情脈脈 ‘affectueux’, etc.

(2005:263) insiste sur le trait [+natif] de la reduplication. Car, «it is true that words are normally not derived but merely plugged in from the mental lexicon, but it is also true that in moments of crisis, speakers actually fashion them on the basis of rules of morphology they have.»

Cependant, les locuteurs ne sont pas totalement maîtres de faire entrer n'importe quel mot nouveau dans le lexique. Comme le signale Aronoff (1976:43-44), il peut y avoir un effet de «blocking», i.e. «the nonoccurrence of one form due to the simple existence of another». Autrement dit, un nouveau membre du lexique doit, lors de sa création, se tailler une place dans la communauté: s'il existe déjà un membre attesté dénotant exactement le même concept, le dernier arrivé risque d'être bloqué à l'entrée. Ainsi, on peut imaginer qu'un nom ou un verbe auront beaucoup plus de mal à se faire accepter en tant que doublet, alors qu'un adjectif strictement descriptif se prête plus facilement à une nouvelle forme redupliquée. C'est sans doute la raison pour laquelle la reduplication des adjectifs représente un terrain de prédilection pour ceux qui sont à la recherche d'une création lexicale idiosyncrasique.

6. Conclusion

Le chinois contemporain possède bien une morphologie, du simple fait qu'il a des règles de réalisation morphologique en syntaxe et des régularités pour les dérivations en lexique. Nous avons voulu montrer que la reduplication des adjectifs chinois est un procédé morphologique productif, que l'on doit formellement distinguer des phénomènes syntaxiques tels que la répétition des verbes ou le redoublement des classificateurs. Cette distinction est essentielle si l'on veut parvenir à une description plus fidèle de la langue et améliorer par là son enseignement. En effet, un fait syntaxique permet au locuteur d'exprimer un sens grammatical, à condition toutefois qu'il se plie à ses règles et ce, avec une «marge de tolérance» si étroite que toute infraction peut conduire à l'incompréhension. La reduplication lexicale, en revanche, offre au locuteur de bien plus grandes possibilités de créations libres. Les mots ainsi créés, même s'ils peuvent paraître surprenants voire choquants, ne sont jamais grammaticalement impossibles.

Références

- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge: MIT Press.
- Cai, Shumei, and Chunhong Shi. 2007. Yan Lianke zuopin zhong de chongdie xingshi tanxi [Examen et analyse des formes rédupliquées dans les œuvres de Yan Lianke]. *Yuyan Jiaoxue yu Yanjiu* [Language Teaching and Linguistic Studies] 2007.4:1-9.
- Cen, Qixiang. 1957. *Yufa Lilun Jiben Zhishi* [Bases de la théorie grammaticale]. Beijing: Time Press.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chen, Guang. 2008. Dui Xiandai Hanyu xingrongci chongdie biao qingwei chengdu de chongxin shenshi [Un réexamen de la diminution d'intensité exprimée par la réduplication des adjectifs en chinois moderne]. *Yuyan Jiaoxue yu Yanjiu* [Language Teaching and Linguistic Studies] 2008.1:35-41.
- Chen, Guanglei. 1994. *Hanyu Cifalun* [Lexicologie chinoise]. Shanghai: Xuelin Press.
- Corbin, Danielle. 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer.
- Dong, Xiufang. 2004. *Hanyu de Ciku yu Cifa* [Le lexique et la morphologie en chinois]. Beijing: Peking University Press.
- Dong, Xiufang. 2007. Cong cihuihua de jiaodu kan nianheshi dongbu jieyou de xingzhi [La nature des [V+Compl] liés à travers la lexicalisation]. *Yuyan Kexue* [Linguistic Sciences] 2007.1:40-47.
- Ge, Benyi. 2001. *Xiandai Hanyu Cihui*. [Lexicologie du chinois moderne]. Ji'nan: Shandong People's Publishing House.
- Guo, Rui. 2002. *Xiandai Hanyu Cilei Yanjiu* [Etudes sur les classes lexicales en chinois moderne]. Beijing: The Commercial Press.
- Hu, Fu, and Lian Wen. 1954. Ci de fanwei, xingtai, gongneng [Les mots: classement, forme et fonction]. *Zhongguo Yuwen* [Studies of the Chinese Language] 1954.8:3.
- Hua, Yuming. 2002. Hanyu chongdie liju: chongdie de zhiyue yinsu [La motivation de la réduplication en chinois: les contraintes]. *Shaoyang Xueyuan Xuebao* [Journal of Shaoyang University] 2002.S1:54-59.
- Inkelas, Sharon, and Cheryl Zoll. 2005. *Reduplication: Doubling in Morphology*. Cambridge & New York: Cambridge University Press.
- Labrune, Laurence. 1998. De l'iconicité en japonais: suffixes en r+voyelle et réduplication. *Cahiers de linguistique-Asie orientale* 27.1:79-118.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

- Li, Yuming. 1996. Shuangyinjie xingzhi xingrongci de ABAB shi chongdie [Réduplication en ABAB des adjectifs qualificatifs dissyllabiques]. *Hanyu Xuexi* [Chinese Language Learning] 1996.4:24-27.
- Liu, Shuxin. 1990[2005]. *Hanyu Miaoxie Cihuixue* [Lexicologie descriptive du chinois]. Beijing: The Commercial Press.
- Lu, Zongda, and Min Yu. 1954. *Xiandai Hanyu Yufa* [Grammaire du chinois moderne], Vol. 1. Beijing: The Masses Press.
- Lü, Shuxiang. 1954. Guanyu Hanyu cilei de yixie yuanzexing wenti [Quelques questions de principe sur les catégories grammaticales du chinois]. *Zhongguo Yuwen* [Studies of the Chinese Language] 1954.9:6-14; 1954.10:16-22.
- Lü, Shuxiang. (ed.) 1984. *Xiandai Hanyu Babai Ci* [800 mots en chinois moderne]. Beijing: The Commercial Press.
- Mel'čuk, Igor. 1993. *Cours de morphologie générale*, Vol. 1. Montréal: Les Presses de l'Université de Montréal, CNRS Éditions.
- Matthews, Peter H. 1974[1991]. *Morphology*. Cambridge & New York: Cambridge University Press.
- Morgenstern, Aliyah, and Alexis Michaud. 2007. La reduplication: universaux iconiques et valeurs en système. *Faits de langues* 29:117-124.
- Packard, Jerome Lee. 2000. *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge & New York: Cambridge University Press.
- Paris, Marie-Claude. 2007. Un aperçu de la reduplication nominale et verbale en mandarin. *Faits de langues* 29:63-76.
- Paul, Waltraud. 2004. Adjectival reduplication in Mandarin Chinese. Communication présentée au 11th International Morphology Meeting, Université de Vienne, 14-17 février.
- Paul, Waltraud. 2006. Zhu Dexi's two classes of adjectives revisited. *Studies in Chinese Language and Culture*, ed. by Christoph Anderl & Halvor Eifring, 303-315. Oslo: Hermes Academic Publishing.
- Paul, Waltraud. 2010. Adjectives in Mandarin Chinese: the rehabilitation of a much ostracized category. *Adjectives: Formal Analyses in Syntax and Semantics*, ed. by Patricia Cabredo Hoffher & Ora Matushansky, 115-152. Amsterdam & Philadelphia: John Benjamins.
- Paul, Waltraud, and Zhitang Yang-Drocourt. 2009. Morphology that meets the eye: adjectival reduplication in Mandarin Chinese. Communication présentée au 17th Congrès de l'Association internationale de linguistique chinoise, Paris, 2-4 juillet.
- Ren, Xueliang. 1981. *Hanyu Zaocifa* [Formation lexicale en chinois]. Beijing: China Social Sciences Press.
- Skoda, Françoise. 1982. *Le redoublement expressif: un universal linguistique, analyse du procédé en grec ancien et en d'autres langues*. Paris: Selaf.

- Shi, Yuzhi. 1996. Shi lun Hanyu de jufa chongdie [La reduplication syntaxique du chinois]. *Yuyan Yanjiu* [Studies in Language and Linguistics] 1996.2:1-12.
- Singh, Rajendra. 2005. Reduplication in Modern Hindi and the theory of reduplication. *Studies on Reduplication*, ed. by Bernhard Hurch, 263-281. Berlin & New York: Mouton de Gruyter.
- Spencer, Andrew, and Arnold M. Zwicky. (eds.) 1998. *The Handbook of Morphology*. Oxford & Malden: Blackwell.
- Tang, Ting-chi. 1982. Reduplication of adjectives in Chinese. *Shida Xuebao* [The Journal of National Taiwan Normal University] 27:279-294.
- Wang, Li. 1956. Guanyu Hanyu you wu cilei de wenti [Existe-il des catégories grammaticales en chinois?]. *Hanyu de Cilei Wenti* [Sur les catégories grammaticales en chinois], Vol. 2, 46-56. Beijing: Zhonghua Book Company.
- Yang-Drocourt, Zhitang. 2008. Un examen critique des études linguistiques existantes sur la reduplication des adjectifs en chinois contemporain. Manuscript.
- Zhang, Guoxian. 2000. Xiandai Hanyu xingrongci de dianxing tezheng [Les particularités des adjectifs en chinois moderne]. *Zhongguo Yuwen* [Studies of the Chinese Language] 2000.5:447-458.
- Zhang, Shoukang. 1981. *Goucifa he Gouxingfa* [Formation lexicale et morphologie]. Wuhan: Hubei People's Publishing House.
- Zhou, Jian. 2004. *Hanyu Cihui Jiegoulun* [Structures du lexique chinois]. Shanghai: Shanghai Lexicographical Publishing House.
- Zhu, Dexi. 1956. Xiandai Hanyu xingrongci yanjiu [Sur les adjectifs du chinois moderne]. *Yuyan Yanjiu* [Studies in Language and Linguistics] 1956.1:83-111.
- Zhu, Dexi. 1982. *Yufa Jiangyi* [Cours de grammaire]. Beijing: The Commercial Press.
- Zwicky, Arnold M. 1990. Inflectional morphology as a (sub)component of grammar. *Contemporary Morphology*, ed. by Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer & John R. Rennison, 217-236. Berlin & New York: Mouton de Gruyter.

Syntactic Realization or Word Formation?

The Grammatical Status of Adjectives Formed by Reduplication in Modern Chinese

Zhitang Yang-Drocourt

INALCO-CRLAO

The reduplication in Modern Chinese is a process used in both the word formation and the syntax. Our discussion concerns the grammatical status of adjectives formed by reduplication. Unlike others studies, which exclusively examine the input forms and therefore consider one part of the reduplicated adjectives as the syntactic realization, we propose to (re)place them in the word formation. In our view, whether there is or not an attested word in the lexicon, and whether this one is itself an adjective or a word of another category, the reduplication enables to build in output a sub-category of lexical units with specific syntactic and semantic properties, different from their simple form. Our analysis is based on the distinction made in morphology between “flexion” and “derivation”, between “internal structure of the word” and “word-formation”.

Key words: modern Chinese lexicon, adjectives, reduplication, morphology, word formation

說“還是” ——祈願情態的浮現*

方 梅

中國社會科學院

本文討論“還是”從連接選擇項到表達祈願情態的演變。“還是”的演變過程說明，在虛詞意義和功能演變的過程中，不僅相鄰詞彙之間可能發生意義的滲透，交際模式所賦予的功能解釋也會參與其中，從而使特定交際模式下的功能進入該交際模式下的詞彙解釋，進而逐步脫離原有交際模式，成為這個虛詞的功能。

關鍵詞：祈願情態，意義滲透，交際模式，詞彙解釋

1. “還是”表達祈願情態

據《現代漢語八百詞》，“還是”作為副詞有兩個：

- (i) 表示行為、動作狀態保持不變或不因上文所說的情況而改變；仍舊；仍然。
- (ii) 表示經過比較、考慮，用“還是”引出所選的一項。

第一類表示仍舊或仍然的“還是”與“只是、老是、總是”等“副詞+是”經詞彙化構成的副詞具有相似的主觀性意義（參看張誼生 2004，董秀芳 2004）。

這類“還是”可以表達客觀事件的延續性，其中的“還是”全都可以被“依然”或者“依舊”替換。例如：

* 本文是國家社會科學基金課題 (05AYY003) “漢語動態呈現語法研究”的部分研究成果，初稿曾在“中國語言學會第 14 次年會”（2008 年 8 月，溫州大學）和第四屆肯特崗國際漢語語言學圓桌會議（2008 年 9 月，新加坡國立大學）宣讀，感謝江藍生、陳一、邵敬敏、吳福祥、張誼生、洪波、孫朝奮、陶紅印教授提出寶貴建議。

- (1) “昨天，我看見你們了，談得那麼親密，連我和阮琳從你們面前走過也看不見。現在你知道她叫什麼了吧？”
“不知道，我還是沒問。”（王朔小說，下文簡稱“王朔”）
- (2) “明天你還是這個時候來取照片。你想放大，拿回底片您另放，這個不包括在內。”（王朔）
- (3) 據我所知，趙蕾至今還是獨身一人。（王朔）

例(1)當說“我還是沒問”的時候，意味著“我過去沒問，這次也沒有問”。例(2)在說一個未然事件，“還是”表示對先前行為的複製。例(3)說“還是獨身一人”的時候，意味著“獨身一人”是一個延續性狀態。表示未然事件和常態的時候，如(2)、(3)，“還是”可以換成“還”，但是表達已然事件，如(1)，不能用“還”替換。

本文討論的“還是”與《現代漢語八百詞》所說的第二類相關。據《現代漢語八百詞》，這類“還是”的用法是：

- (i) 用在動詞或主語前。如：
我看還是去頤和園吧，十三陵太遠。
還是你來吧，我在家等你。
- (ii) 還是+動詞/小句[+的]好。表示經過比較，這樣較為可取。
想來想去，還是親自去一趟[的]好。
我看還是你來辦理[的]好。

上述兩種用法，(ii)是通過詞彙形式“[的]好”把“有所比較”的意思直接表達出來；而(i)裡頭，“有所比較、有所選擇”的意思完全是通過“還是”來解讀的。我們認為，這類“還是”的核心意義是表達言者意願。由於有“的好”的介入，“還是+動詞/小句[+的]好”表達說話人的認識或建議。用法(i)沒有“的好”，則側重表達說話人的意願。

我們發現，與“仍舊”義“還是”比較，這類表達說話人意願的“還是”除了不能替換作“依然”或者“依舊”之外，具有下述特徵：

第一，語音形式上有區別，“仍舊”義“還是”自身總是重音所在，如(4a)；而表達言者意願的“還是”不能重讀，如(4b)。

- (4) a. 這次#還是我去。
b. 還是#我去。

第二，“仍舊”義的“還是”總是指向施事主語。而這類表達言者意願的“還是”，只能指向說話人，表達說話人的選擇。比如下面的例子，儘管(5a)句主語是第一人稱，(5b)句主語是第二人稱，(5c)句主語沒有出現，“還是”都用來引出說話人的意願，而不隨著主語的改變而改變。例如：

- (5) a. 我還是吃了飯送他走。
b. 你還是吃了飯送他走。
c. 還是吃了飯送他走。

第三，表達言者意願的“還是”不參與命題表達，刪除“還是”以後，命題意義不變。

- (5') a. 我[還是]吃了飯送他走。
b. 你[還是]吃了飯送他走。
c. [還是]吃了飯送他走。

因此，可以說，這個“還是”是表現祈願情態 (optative mode) 的副詞。

下面，我們將討論這種祈願情態的“還是”是如何衍生出來的。

2. 共時分布

在現代共時系統中，與祈願情態直接相關的是“還是”表示言者選擇、言者意願、言者認識的用法。祈願情態的“還是”是選擇問句中連詞“還是”用法的延伸。

2.1 言者的選擇

在現代共時系統中，選擇問句的常見形式是“（是）…還是”。如：“是同意還是不同意”，其中的“還是”繫連兩個備選項：

- (6) 您問這幾位老太太，她們過去也幹過小腳偵緝隊，是現在這樣誰也不管誰好呢？還是像過去那樣互相緊盯著好？（王朔）
- (7) 您想怎麼死啊？是活活燒死還是讓我們把您五花大綁拉到郊外斃嘍？（王朔）
- (8) “他胖了還是瘦了？”
“很瘦！那可也許是頭髮欺的。他的頭髮好像有好幾個月沒理過了！頭髮一長，臉不是就顯著小了嗎？”

“（是）A 還是 B”中的 A、B 兩個選項可以分別作為兩個單句出現，如 (6)；也可以整合為一個句子，如 (7)；整合為一個句子的選擇問句，引導第一選項的“是”可以省略，如 (8)。

選擇問句總是要讓對方根據“還是”前後的選項做出回答。

- (9) “你是願意打八圈，還是喝兩盅兒？”
“打八圈吧，今個晚上我的精神很好！”（老舍《四世同堂》）
- (10) “你是這麼想的還是已經這麼幹了？”
“我已經這麼幹過了，否則我怎麼會這麼自信？”（王朔）
- (11) “你們是逮誰捧誰，還是也挑人，單捧有名的？”
“逮誰捧誰！”（王朔）

正是由於選擇問句要求對方在給出的選擇項中挑選，所以才會有下面這樣的對話：

- (12) “那麼我是變醜，還是變美了？”
“別逼著我誇你。”（王朔）

選擇問句的答句可以直接回答所選擇的項目，如上面 (9)、(10)、(11)。另一種情形是採取“還是+選項”的形式。下面例 (13) 是對答，例 (14) 是自問自答。

- (13) “你看，是高慶奎好，還是馬連良好呢？”
“你看呢？”
“恐怕還是高慶奎好一點！”（老舍《四世同堂》）
- (14) 你們打算坐火車還是乘飛機回去？還是乘飛機吧，飛機快，火車晃哩晃當受罪。（王朔）

接下來的問題是，這種“還是”句使用有哪些限制。

如果單看 (9)、(10)、(11)，似乎回答選擇問句不是必“還是”。不過，仔細追究起來，“還是”用與不用是有區別的。從表達命題意義的角度說，在說未然事件的時候，答句用不用“還是”是可選性的，對比 (9) 和 (9')。但是，如果是已然客觀現實，答句則不能用“還是”，如 (10')。

(9) “你是願意打八圈，還是喝兩盅兒？”

“(還是) 打八圈吧，今個晚上我的精神很好！”

(10) “你是這麼想的還是已經這麼幹了？”

“*我還是已經這麼幹過了，否則我怎麼會這麼自信？”（王朔）

當問句可以理解為對已然現實的陳述、也可以理解為對未然事件的陳述的時候，答句有兩種選擇：1) 答句不用“還是”，如 (11)，句子理解為對現實的陳述。2) 答句用“還是”，理解為言者的選擇。如 (11')：

(11) “你們是逮誰捧誰，還是也挑人，單捧有名的？”

“還是單捧有名的！”（王朔）

據我們的考察，用“還是”的答句只能是非現實情態句 (irrealis)，包括：

- 1) 未然事件敘述句。
- 2) 評論句。
- 3) 建議或祈命句。

而答句不含有“還是”的，則沒有這些限制。

“（是）…還是”用在選擇問句時，繫連等立關係的選項，是連詞。與此形成對照的是，在敘事體中，這種連詞的功能仍然保留了下來，用於連接選擇項的“還是”往往可以換作“或（者）”，引出並列項，並不體現言者的選擇。例如：

(15) “我的房門在夜裡一直是上鎖的，不管是有意還是無意，都不可能進來人。”（王朔）

從上面的分析可以看出，“言者的選擇”這個語義解釋，實際是吸收了答句的表達功能，換句話說，“還是”表示言者選擇這個意義是互動交際模式所賦予的。

當“還是”表示言者選擇的時候，它就從作為連詞時的“行域”表達功能轉為“言域”表達功能，即“我告訴你我的選擇”。這種“言域”表達功能是其後一系列衍生用法的基礎。

2.2 言者的意願

在實際語境中，“還是”所引出的陳述不一定與選擇問句“（是）…還是”問句直接對應，選項可以不出現在選擇問句，而是處在自己或對方前面的言談當中。例如：

- (16) “你也離了得了，回頭再勸肖超英也離了，咱們幾個光棍住在一起多樂和。”[陳述句]

“能不離還是不離，能湊合就湊合。你可不知道離回婚多傷身子骨，雖然咱們都是想得開的人。”（王朔）

上面這一例，“還是”句並不針對選擇問句提供的選擇項，“還是”句所表達的選擇內容是在上文的敘述當中。

更進一步，當“還是”不依賴於選擇問句，也不依賴上文，“還是”也就成為直接表達言者的意願的成分了。例如：

- (17) “我願意養你。我們現在伙食費發給個人了，這樣我每個月就能拿二百來塊錢，夠我們倆花了。我們可以安穩穩地過日子。你不是希望我做個賢妻良母嗎？”

“我看，我們還是算了吧。”

- (18) “我還是走吧。”

“別走，就坐這兒。”（王朔）

例(17)的“還是”句之前對方的談話中並沒有提供備選內容，“還是”句是說話人說出自己的決定。而(18)是首發句，說話人覺得自己不宜久留，是說話人對言談現場中情景的直接反應。

可見，從共時的角度看，“言者意願”這個意義的浮現是以選擇項的隱退為前提的。與不用“還是”相比較，用了“還是”表達“我告訴你我的意願”。

2.3 祈命

當“還是”句所述行為關涉聽者的時候，表達對聽者的建議，構成祈命。例如：

(19) “不在不在，就說我不在。”

“你還是跟他說兩句吧。”（王朔）

(20) 家裡沒人，誰照顧你？還是全好了再出院吧。（王朔）

例中儘管含有語氣詞“吧”，但是刪除“吧”，仍是對聽者的建議：

(19') “不在不在，就說我不在。”

“你還是跟他說兩句。”

(20') 家裡沒人，誰照顧你？還是全好了再出院。

表達言者意願與表達對聽者的建議之間，有時並沒有截然的對立，下面的例子裡的“還是”，可以說是表達言者意願，但同時，因為這個意願是直接針對聽者的，因而也可以理解為一種委婉的祈命。例如：

(21) 算了，你既然管不了就別管了，還是讓他們各自去幹自己的那一攤吧。（王朔）

表達說話人的意願可以用“還是…的好”句。例如：

(22) “她老跟我說，特喜歡你。”

“我用得著她喜歡麼？她還是別喜歡我的好。我又不是熊貓不被喜歡就不珍貴了。”（王朔）

這類“還是…的好”裡的“的好”可以刪除。例如：

(22') “她老跟我說，特喜歡你。”

“我用得著她喜歡麼？她還是別喜歡我。我又不是熊貓不被喜歡就不珍貴了。”

表達“建議”這個解釋是通過下面這樣的推理獲得的，即：我認為選擇這樣做比較好，因而“我建議你這樣做”，或者“你應該這樣做”。從發展角度考察，含有“的好”的格式也的確是較早出現的，下面第三節會詳細討論。

2.4 言者的認識

當一個命題不針對聽話人、也不針對說話人的時候，“還是”表達說話人的認識，構成評論句。例如：

(23) “小地方的舞劇，粗糙得很。”

“我們留在北京的同學還沒一個上舞劇的，還是分回省裡強。”（王朔）

(24) “我們後天就放暑假了。”

“還是當學生幸福，每年還有兩個假期。”（王朔）

這類“還是”句中的“還是”都可以換成“我覺得/我認為/我想”。如果“還是”與“我覺得/我認為/我想”同現，往往是說話人特別要著力論證自己的觀點的時候。例如：

(25) “我絕對相信你的能力。你一定能幹好，老同志了嘛。要多做說服解釋工作，我還可以給你多派幾個來。”

“我想我還是不行，群眾更相信領導。”（王朔）

另外，這類“還是”還可以用在含有情態助動詞的句子中，例如：

(26) “孩子還是應該要一個的，一個家麼。”

“不不，堅決不要。人家說了，有孩子夫妻感情就淡了。”（王朔）

因為“還是”的作用在於表達“我告訴你我的看法”，“還是”與“我想”或“應該”同現實構成了一種近義詞連用，因而才會產生強化作用。

3. 歷時演變

我們發現，表示表達言者意願的副詞“還是”，初見於選擇問句的答句。

3.1 明清漢語選擇問句

傅惠鈞(2001)對《老乞大》、《朴通事》、《金瓶梅》和《兒女英雄傳》的考察，認為那一時期選擇問的常見形式有：

- 1) A，B？
- 2) 是A，是B？
- 3) 還是A，還是B？
- 4) 是A，還是B？

徐正考(1996)以《儒林外史》、《紅樓夢》、《俠女奇緣》、《三俠五義》、《老殘遊記》為主要材料，從描寫入手探討清代漢語選擇疑問句系統。認為清代漢語選擇疑問句系統包含簡式系統與繁式系統兩個子系統。其中繁式系統中的主要格式是“是A，是B”式，其主要形式是“是A，是B”、“還是A，還是B”、“是A，還是B”。“還是”置於前項後項都可以，既可單用於前項或後項，又可與“可是”、“是”等配合使用。不過雖然“還是”既可置於前項又可置於後項，但置於後項可以加強句子兩項之間的呼應連接。如：

- (27) 寶玉因問道：“可是病了，還是輸了呢？”（紅·19）/黛玉道：“跌了燈值錢，是跌了人值錢？”（紅·45）/還是與他鬧是不鬧呢？（三·57）/因問我：“還是本地用？遠地用？”（俠·9）（徐正考1996）

我們對會話材料的考察發現，儘管選擇問句的形式各異，但是作為選擇問句的答句，選用哪種形式具有客觀與主觀之別。

如果答句表達的是客觀事實，則直接回答。例如：

- (28) “你算計是坐船貴是坐車貴？”
“坐車比坐船花的多。”（《語言自適集·散語章》）
- (29) “他是見天在家裡吃飯是外頭吃飯哪？”
“他見天吃完了飯才走哪。”（《燕京婦語》，第二課）

選擇問句的各種形式，答句都會遵守這個規律。再如：

- (30) “明兒是您先去我先去呀？”[是 AB 語氣詞]
“明兒早一點兒起來您先去，我打裡頭回來再去也不晚。”（《燕京婦語》，第十二課）
- (31) “咱們是在通興長買是在謙祥益買呀？”[是 A 是 B 語氣詞]
“在通興長那兒買吧，通興長那兒熟布也好。”（《燕京婦語》，第十四課）
- (32) “您要多少尺啊？”
“作個布衫兒十四尺就夠了。”
“是您作是姑娘作呀？”[是 A 是 B 語氣詞]
“我們姑娘作。”（《燕京婦語》，第十四課）

而當答句表達說話人認識的時候，則連同選擇連詞一起截取說出來，而不是僅僅回答選項。回答“是 A 是 B”出現的時候，答句採用“是 B”，也就是說，表達主觀性內容的時候要選擇有標記的形式。例如：

- (33) “貴國的人向來有住店的有住廟的。”
“是店裡好，是廟裡好呢？”
“依我說，是店裡方便些兒。”（《語言自適集·回答章》）

針對選擇問句的答句，表達認識的時候直接截取繫連選擇項的成分，這是一種常見的用法。只不過，早期的答話裡面總是同時出現表示“比較”義的詞。如上面(33)例子中的“些兒”，或者用“還是…（的）好”，如下：

- (34) “請菜。我可不布，都沒有外人，自取罷。”
“那最好！讓大家敬起來，倒顯得拘泥，還是隨便的好。”
（《語言自適集·回答章》）

- (35) “我心裡也遊移，將來是去好啊，是不去好呢。”
“依我說，您還是去好，……”（《談論新編》¹）

當表達說話人的認識和建議的時候有兩種情況。一是選擇內容存在於前面的談話之中，答句中直接用“還是…的好”表示認識，見上例。二是用“還是”句作答，如(36)。

- (36) “這兒的照像館照的有的好有的不好，可是不好的多。”
“是啊，您知道這城裡頭那兒照的好哇？”
“要說這幾處照像館還是霞公府山本照像館那兒照的好。”（《燕京婦語》，第九課）

不過，在早期材料中，上述用法還不普遍，《語言自邇集》中僅 1 例(34)，《燕京婦語》也僅僅有 1 例(36)。

從歷時材料看，《現代漢語八百詞》所說的“還是”表示“經過比較、考慮，有所選擇”，用“還是”引出所選的一項，這種用法起初是在選擇問句的答句這個特定條件下獲得的。這種用法得以廣泛使用，是“還是”置於選擇問句後項的位置（（是）A 還是 B）逐漸成為強勢用法之後的事情。

3.2 認識義和比較義詞語的隱退

這種表達“經過比較、考慮，有所選擇”，早期形式往往有表示認識的詞語共現，或出現“還是…（的）好/不如”句式，或者“依我說”。與此相應，這個時期選擇問句的形式也是多種多樣，“（是）…還是”並沒有成為強勢用法。

在清代的對話材料中，也能見到少量“還是”表示說話人的意願或認識，而不依賴選擇問句的例子。如果“還是”句的上文沒有選擇問句，表示評價意思的

¹ 《談論新編》是清光緒年間的漢語教科書。“書序”：且言語一科，聖門與德行並重，謂非難能之一端乎，然而從未聞以言授受者也，自五洲互市，聘問往來，則言語之授受起焉，而學語之書亦出焉，《語言自邇集》首傳於世，學語者宗之，未嘗非啓發學者之一助，逮至今日，時事屢見更新，語言亦因之變易，金公卓菴曉英文，嫻辭令，博學多識，於授話一道尤為擅長，去歲文部省聘請東來充外國語學校教習，於夏日餘暇同參謀本部平岩道知君合著《談論新編》百章，窮數旬之久，始告厥成，餘觀覽回環，見其事皆目今要務，閱其辭皆通時語言，較諸《自邇集》全部亦有過而無不及焉，善學者苟能簡練揣摩，觸類旁通，施措於官商之際，則博雅善談之名將不難播於海內也，有志華言者，宜鑄金事之。光緒戊戌（戊）秋八月張廷彥序於江戶喜晴樓。

“還是”句總有表示認識的詞語搭配著用。在晚期的材料中，這種把“我想、我看、我說”或“（的）好”直接說出來的表達方式佔絕大多數。例如：

- (37) 你看著怎麼著好呢，在家裡辦，反正也省不了錢，第一這廚行是最難惹的，都包給他們呢，將來是席面上不好看，要給他們開工呢，那就更淘神啦，一樣的酒席，東西是樣樣都得費的，俗語兒說，廚子不偷，五穀不收，這一項人，是最不好辦，並且還得多用許多的茶房，多請幾位助忙的，沒有一樣兒不是費的，外帶著燈油火燭，家傢伙夥的，處處兒都得經心著眼，實在不容易，據我看，還是在堂子裡辦，頭一宗他的人是現成的，跟二只要把單子開出來打算定多少桌都要幾等席，商量好了價錢，交給他們，就算完了，應用的東西差不多他們都是現成的，不用本家兒操心，那夠多們省事呢？（《談論新編》）
- (38) “那個地方兒離京有多遠兒呢？”
“大概有四十多里地。”
“哪們遠咱們走著去行嗎？”
“您怎麼說這個話呢，那個別方兒，可是騎驢也去的了坐車也去的了。不過我的意思還是單要走著去，一來沿路都有景致，可以一邊兒走著一邊兒逛，……”（《京語會話》²）
- (39) “按例可是應當他們眾保家攤賠的，然而辦得到，辦不到，我可也不敢說一定。”“這麼說起來，雖然有聯環保家，還是不如有押款存在衙門裡妥實。”（《談論新編》）
- (40) 那可不好，萬一要是吃錯了，摳不出來，挖不出來，後悔不就晚啦嗎？還是請個先生瞧瞧好。（《京語會話》）

表示認識和比較的詞語隱而不現，是演變的關鍵一步。這種用例在《京語會話》開始出現。例如：

² 本文採用鮮文大學校中韓翻譯文獻研究所印《京語會話》，據鮮文大學中文系朴在淵〈關於《京語會話》的幾點說明〉介紹，《京語會話》“大約成書於 20 世紀第二個十年中期至 20 年代之間”，辛亥革命以後 1920 年前後寫成的一部中國語會話書。由 121 組對話構成，分情景會話和主題會話兩類。這本書原來在中國，是為中國人而作的，後來傳到韓國。另可參看張美蘭、陳思羽（2006）。

- (41) “你這話說的很有道理，人無遠慮必有近憂，早定出個章程來，無論大事小事都按著一定的章程走，當家的人兒，又省心又不至於落出閒話來，這是過日子極好的一個法子，你就擬出一個章程來。”
“這個章程還是得您定，家有千口，主事一人作兄弟的，不過怕您有一時想不到的，提您一個醒兒，至於大主意，還是得您拿。”（《京語會話》）
- (42) 我不怕您過意，您還是別管，我也不是要管，俗語兒說，清官難斷家務事，誰還有願意管的嗎？（《京語會話》）

這裡，表示說話人認為“應該”如何或“最好”如何，直接用“還是”句。“還是”脫離表示認識或比較的詞語而單獨使用，這類用法在《京語會話》有2例。

到了老舍的作品裡，雖然沿襲了舊有的用法。不帶認識義詞語的用法開始多起來。例如：

- (43) “小夥子，我要是錢富裕的話，真想留下！”
“乾脆就留下吧，瞧著辦得了！”
“說真的，小夥子；倒退三十年，這值三個大元寶；現在的年頭，又搭上兵荒馬亂，我——你還是到別處吃喝吆喝去吧！”（老舍《駱駝祥子》）
- (44) “一成家，黑天白日全不閑著，玩完！瞧瞧我的腰，整的，沒有一點活軟氣！還是別跑緊了，一咬牙就咳嗽，心口窩辣蒿蒿的！甭說了，幹咱們這行兒的就得它媽的打一輩子光棍兒！”（老舍《駱駝祥子》）
- (45) “四大爺！你的話說得不錯，可是那作不到啊！你老人家比我的年紀大，總該知道咱們北平人永遠不會造反！還是心平氣和的想辦法吧！”（老舍《四世同堂》）
- (46) “不要聽孫七的，咱們還是老老實實的過日子，別惹事！反正天下總會有太平了的時候！日本人厲害呀，架不住咱們能忍啊！”（老舍《四世同堂》）

從歷時角度歸納起來，如下表：

	非答句	評論句	建議/祈命句
語言自邇集	(+)	我說；(的)好	----
燕京婦語	(+)	+	----
談論新編	+	+	(的)好
京語會話	+	+	(建議類)
老舍小說	+	+	+
王朔小說	+	+	+

3.3 小結

從歷時角度看，“還是”意義和功能的演變，是從連接選擇項到表達祈願情態的演變，即：

繫連選項 > 言者的選擇 > 1. 言者的意願
> 2. 祈命
> 3. 言者的認識

從歷史上看，在“還是”祈願情態這個演變過程中，第一步，“還是”逐步擺脫對選擇問句的依賴，而將選擇問句的答句的功能——說話人的選擇，吸收到了“還是”的意義當中。第二步，“還是”逐步擺脫對認識義和比較義詞語和格式的依賴，同時，吸收了上述句法環境所提供的意義和功能。“還是”從連接選擇項到表達祈願情態的演變過程說明，在虛詞意義和功能演變的過程中，不僅相鄰詞彙之間可能發生意義的滲透，交際模式所賦予的功能解釋也會參與其中，從而使特定交際模式下的功能進入該交際模式下的詞彙解釋，進而逐步脫離原有交際模式，沉澱為這個虛詞的功能。

從現代共時系統上看，“還是”用作連詞繫連選擇項是關係命題意義的，屬於“行域”。“還是”用作副詞，在答句中表達說話人的選擇、意願。當“還是”依賴於問句的時候，“言域”功能還沒有獨立顯現出來。一旦脫離了“答句”這個環境，具有“我告訴你我的選擇”、“我告訴你我的意願”這種功能的時候，“還是”句用來表達說話人的意願、認識，“言域”的功能就完全是由“還是”承載了。而用作建議或祈命的時候，是說話人希望聽者去做某事，就純粹屬於“言域”的了。

引用文獻

- Dong, Xiufang (董秀芳). 2004. 'Shi' de jinyibu yufahua: you xuci dao ci nei chengfen “是”的進一步語法化：由虛詞到詞內成分 [Further grammaticalization of *shi*: from function word to word-internal element]. *Dangdai Yuyanxue* 當代語言學 [*Contemporary Linguistics*] 2004.1:35-44.
- Fu, Huijun (傅惠鈞). 2001. Zhenxingwen yu jiaxingwen: Ming Qing Hanyu xuanze wenju de gongneng kaocha 真性問與假性問：明清漢語選擇問句的功能考察 [Questions and rhetorical questions: a functional study of alternative questions in Ming and Qing Dynasties' Chinese]. *Yuyan Jiaoxue yu Yanjiu* 語言教學與研究 [*Language Teaching and Linguistic Studies*] 2001.3:60-68.
- 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). 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.
- 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.
- 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.
- Xu, Zhengkao (徐正考). 1996. Qingdai Hanyu xuanze yiwenju xitong 清代漢語選擇疑問句系統 [Chinese sentence structure: the forming systems in Qing Chinese]. *Jilin Daxue Shehui Kexue Xuebao* 吉林大學社會科學學報 [*Jilin University Journal Social Sciences*] 1996.5:70-75.
- Zhang, Meilan (張美蘭), and Siyu Chen (陳思羽). 2006. Qingmo Minchu Beijing kouyu zhong de huati biaoji: yi yibai duo nian qian ji bu yuwai Hanyu jiaocai wei li 清末民初北京口語中的話題標記：以 100 多年前幾部域外漢語教材為例 [Topic markers in Beijing dialect during the period between the end of Qing Dynasty and the beginning of the Republic of China]. *Shijie Hanyu Jiaoxue* 世界漢語教學 [*Chinese Teaching in the World*] 2006.2:63-73.
- Zhang, Yisheng (張誼生). 2004. *Xiandai Hanyu Fuci Tansuo* 現代漢語副詞探索 [*Studies of Adverbs in Modern Chinese*]. Shanghai: Xuelin Press.

On *Haishi* 'or' and the Emergence of the Optative Mood

Mei Fang

Chinese Academy of Social Sciences

The present paper discusses the development of *haishi* 'or' from a disjunctive conjunction to a marker of optative mood. The change in meaning and function of *haishi* is not only explained by the lexical influence of neighboring words, it comes also from the functional explanation that emerges from intercommunication. It is assumed that the function originally associated with certain intercommunication markers gradually participate in the lexical meaning of the new intercommunication marker and explain the function of the marker *haishi* 'or'.

Key words: optative mood, lexical influence, intercommunication, lexical meaning

Reflexive Nominal Compounds*

Ekkehard König

Freie Universität Berlin

Reflexive compounds like *self-assessment* in English or *zì-ài* ‘self-love, self-respect’ in Mandarin Chinese have rarely been analyzed in individual languages and even less so from a comparative perspective. Based on a sample of ten languages, this chapter presents a detailed analysis of the salient properties of such compounds in the languages under comparison: their derivation, their argument structure, their meanings, their productivity as well as their categorial properties. Reflexive compounds are typically combinations of event nominalizations of transitive verbs and intensifiers like *self-* in English or *sam(o)-* in Russian and thus manifest a reflexive meaning and a reduced argument structure. From a semantic perspective at least two types can be distinguished, depending on the use of the underlying intensifiers. It is also shown that languages may differ with respect to the intensifiers used in these compounds (one in English, two in German and Finnish, a loanword in French) in the type of nominalizations used as input (event vs. agentive nominalization) and the categorial output (N, ADJ, V¹) of the relevant compounding processes.

Key words: reflexive, compounds, intensifiers, nominalizations, parameters of variation, argument structure, semantic analysis.

1. Introduction

The term ‘reflexive nominal compounds’ will be used in this paper for derived nouns such as *self-control* in English, *Eigenlob/Selbstlob* ‘self-praise’ in German, for *autodérision* ‘self-derision’ in French, for *zì ài* ‘self-love’ in Mandarin, for *samolët* ‘airplane’ in Russian, for *afto-élenxos* ‘self-control’ in Greek, for *ön-becsülés* ‘self-esteem’ in Hungarian and for *oma-kiitos* ‘self-help’ and *itse-luottamus* ‘self-confidence’

* I would like to thank Claire Moyse-Faurie and Wang Lin for providing me with a rich selection of data from French and Mandarin. Thanks are also due to Thekla Wiebusch for comments on an earlier version of this paper.

¹ The following abbreviations are used: ADJ (adjective), DP (determinal phrase), INSTR (instrumental), INT (intensive), N (noun), NEG (negative), NOM (nominative), REF (reflexive), SG (singular), V (verb).

in Finnish. Apart from some occasional remarks here and there (cf. Safir 1996, König & Vezzosi 2004), such deverbal compounds have almost never been analyzed in any detail for individual languages, let alone from a comparative perspective.² This paper will discuss both the formal and semantic properties shared by such derived lexemes — and thus the reason for subsuming them under one meta-grammatical term — as well as the contrasts exhibited by the relevant derivational processes in the languages selected for comparison.³

Among the properties shared by these derived nouns in the languages under comparison we find the following:

- The first component of these nominals can be derived (both diachronically and synchronically) from intensifiers (Engl. (X-)self; Ger. *selbst*, *eigen*; Russ. *sam*; Mand. *zìjǐ*, etc.) and the compounds can be paraphrased in terms of at least one of the uses normally distinguished for intensifiers (cf. König & Gast 2006). Their second component is a deverbal nominalization or deverbal adjective, with very rare exceptions such as German *Eigenkapital* ‘equity’, *Eigenblut* (*therapie*) ‘one’s own blood, autohaemotherapy’, *Selbstbild* ‘self-image’ and, *Selbstzweck* ‘end in itself’, where a non-derived noun is found.
- More often than not, the first part of these derived nouns provides information about the noun’s argument structure. Typically, the two argument positions inherited from the underlying transitive verb must be bound by the same (generalized) quantifier or, putting it somewhat loosely, they must be filled by the same argument. In other words, the relevant nominals express a reflexive meaning.
- Given that we are dealing with the interface of grammar and the lexicon, it should not come as a surprise that we find a certain amount, but not a total degree of semantic compositionality in this domain. As with most patterns of word-formation certain aspects of meaning are simply a matter of lexicalization. Nor do we always find parallel or analogous formations for certain meanings in different languages. German *Selbstentmündigung* does not translate easily into English or French (‘to declare oneself incapable of managing one’s own affairs’; ‘se mettre soi-même sous tutelle de quelqu’un’), especially in its extended metaphorical use, nor does Fr. *autopersuasion* have a straightforward translation into German or English. The negative connotations of German *Selbstzufriedenheit* are not associated with *autosatisfaction* in French and

² Among the rare exceptions we find inter alia Ludwig (1963) and Mutz (2000, 2004).

³ My sample includes the following, mainly European, languages: English, Finnish, French, German, Greek, Hungarian, Italian, Mandarin and Russian.

autojustification is simply translated as *Rechtfertigung* in German. Moreover, there are false friends like German *Selbstbewusstsein* ‘self-confidence’ and Engl. *selfconsciousness*. It is therefore impossible to give parallel glosses in English for all examples and I will sometimes use a French gloss or a verbal paraphrase.

- So far I have used the vague term ‘nominal’ alongside with ‘reflexive nominal compounds’ in order to sidestep the terminological decision between ‘nominalization’ and ‘compound’. In many cases this decision can be based on clear criteria. In the relevant compounds in German (*Selbstinszenierung* ‘self-fashioning, self-styling’, *Eigenkapital* ‘personal capital’) two free forms are combined in the order typical of endocentric compounds (i.e. modifier + head) in Germanic. Other data and other languages present a slightly different picture, however. In English the first component is no longer a free form and some of the languages under comparison use earlier or reduced rather than complex forms of intensifiers in the relevant derivations (e.g. Mand. *zì-* vs. *zìjǐ*). Moreover, the first component may be an affix borrowed from another language (e.g. Greek), as in Fr. *autodérision*, *autogestion* ‘self-administration’, *autodétermination* ‘self-determination’. A further criterion that can be used in some languages is the capacity of these derived nouns to enter additional derivational processes, as is shown by *Selbstdisziplin+los+ig+keit* ‘lack of self-discipline’ in German or *self-disciplined* in English. What we clearly need to distinguish from *self*-compounds are the analogous nominalizations like *control of oneself* in English or *estime de soi* ‘self-esteem’ in French. Such formations manifest all the hallmarks of constructions, e.g. agreement between the subject and the reflexive marker in argument position, as in *mon estime de moi* ‘my self-esteem’ or *he kept control of himself*.
- In addition to varying degrees of lexicalization, there may be several types and layers of such reflexive compounds in a language. In German, both the adnominal intensifier (*selbst*) and the attributive intensifier (*eigen* ‘own’) may occur in the relevant compounds, typically with a clear division of labor, but sometimes also in combination with the same de-verbal nouns or adjectives, as in *Selbstlob* and *Eigenlob* ‘self-praise’. And we also find loanwords such as *Autobiographie* ‘autobiography’, *Automobil* ‘automobile’, *Autodidakt* ‘self-educated person’. In French there are two types based on (i) the Greek prefix *auto-* (e.g. *auto-satisfaction* ‘being pleased with oneself’, *auto-critique* ‘self-criticism’, *auto-punition* ‘self-inflicted punishment’) and (ii) on the attributive intensifier *propre* (e.g. *amour-propre* ‘self-love’, *sens propre* ‘basic meaning’, *nom-propre* ‘proper name’). Identical meanings can, of course, also be

expressed by nominalizations like *mépris de soi* ‘self-loathing’, *confiance en soi* ‘self-confidence’ and by noun phrases with the adjective *personnel* (*initiative personnelle* ‘one’s own initiative’, Ger. ‘Eigeninitiative’, *consommation personnelle* ‘personal use’). In English and Mandarin, by contrast, there is only one possible reflexive marker for such compounds.

- The forms resulting from these processes of composition are typically nouns and adjectives (*self-sufficient*, *self-centered*, *self-evident*), but we also find verbs (e.g. Fr. *s’auto-féliciter* ‘to congratulate oneself’, *s’auto-satisfaire*, *s’autoévaluer* ‘to evaluate oneself’, *s’autosuffire* ‘to be self-sufficient’, etc.).
- Reanalysis of the relevant forms as compounds may give rise to the creation of new abstract nouns such as (*the*) *self* in English. The noun *self* certainly existed before the 20th century, but it denoted a functional concept and was invariably used with a possessive determiner, as for instance in Shakespeare’s sonnets (*thy sweet self*, *thy royal self*). The new abstract noun *self*, which is used with the definite article as an individual concept, is currently expanding its use at an enormous speed and may assume a wide variety of different interpretations (‘identity’, ‘consciousness’, ‘individual’, ‘soul’, ‘character’, ‘perspective’, etc.). The Romance counterparts of Engl. ‘the self’, i.e. *le soi* in French and *il SE* in Italian are much more restricted in their use to formal, mainly philosophical discourse.
- Some languages (e.g. Vedic, Greek) may also have analogous (productive) reciprocal nominalizations with a meaning such as ‘mutual help, mutual respect’ (Gk. *alilo-voithia* ‘mutual help’, Sanskrit *anyonyārakṣa* ‘mutual protection’, *parasparadveṣa* ‘mutual hatred’ cf. Kulikov 2007a, 2007b).

2. Argument structure

Apart from some rare exceptions like Ger. *Selbstläufer*, the nominalizations functioning as heads of reflexive compounds derive from transitive verbs. As far as their semantic type is concerned, they are typically event nominalizations (e.g. *self-determination*, *self-deception*, *self-destruction*), but agent nominalizations are also possible in some languages (Engl. *self-provider*, *self-insurer*; Ger. *Selbsterzeuger* ‘a person who produces his own food’, *Selbstversorger* ‘somebody who is self-reliant’). Nominalizations inherit their argument structure from their underlying verbs, but in contrast to verbs the realization of the relevant arguments for nouns is optional. If the arguments are realized at all, they are expressed by the genitive case or by prepositions in many European languages. In keeping with the reflexive meaning of these compounds, however, only one argument can be realized, viz. the argument corresponding to the

subject of the underlying verb:

- (1) a. The self-determination of the Korean people
- b. *The self-determination of their destiny by the Korean people
- c. John's self-control (*of his company)
- d. John's control of himself/his weaknesses/his actions

The semantic effect of the fusion of a de-verbal nominalization with the intensifier *self-* in English is obviously an operation on the argument structure of the de-verbal noun: The two variables of the underlying verb are unified to a single one (typically x and $f(x)$) and must therefore be bound by the same generalized quantifier (determiner phrase). Or putting it more loosely, the two variables must be replaced by the “same” or by at least two related arguments. As is shown by example (1d), the specification “the same” also includes two arguments related via metonymy. This effect, which is typical of derivation rather than compounding, can roughly be represented as follows for our example *self-control*:⁴

- (2) a. /control/ [+N,...] $\lambda x \lambda y \lambda e [e \text{ inst } [y \text{ CONTROL } x]]$
- b. /self-control/ [+N,...] $\lambda y \lambda e [e \text{ inst } [y \text{ CONTROL } f(y)]]$

Two things are worth noting about this operation: (a) composition does not normally change the argument structure of the head noun, but it affects the way certain argument positions are to be filled. A fire-insurance is an insurance against (the risks of) fire. In the compounds under discussion the effect on the argument structure of the underlying verb consists in the specification that the two variables of the argument structure are unified, i.e. reduced to one. Syntactically, this amounts to a reduction in valency; (b) as is shown by our example, this unification takes the form of changing the variable for the PATIENT into a function of that used for the AGENT. As will be shown by further examples discussed below, this seems to be the more wide-spread situation rather than the exception.⁵

The process of deriving the reflexive compounds—and ultimately the noun *self*—by processes of condensation and abstraction can roughly be represented by the following steps:

⁴ These representations are based on the framework presented in Bierwisch (2009); *e* is an event variable and *inst.* stands for ‘instantiates’.

⁵ Cf. Reuland (2008).

- (3) a. John assesses himself/his achievements.
- b. John's assessment of himself/his achievements
- c. John's self-assessment
- d. the self

A reflexive compound is a condensed version of a reflexive proposition with the corresponding change of its semantic type. In going from such a compound to the noun *self*, abstraction is made from the content of the underlying verbs and from the arguments.

3. Possible interpretations and their derivation

It is a well known fact by now (cf. Edmondson & Plank 1978, König & Siemund 2000, and most recently König & Gast 2006) that there are at least four different uses of intensifiers. In view of the fact that intensifiers play an important role in the derivation of reflexive compounds it is interesting to see whether all of the different meanings exhibited by those uses can be found in the relevant compounds. The relevant uses of intensifiers can be illustrated by the following examples:

- (4) a. (adnominal)
 The Pope himself will come to the rally. (in contrast to one of his clergy)
- b. (adverbial, exclusive)
 Mrs. Dalloway wanted to buy the flowers herself. (alone, no delegation, no help)
- c. (adverbial, inclusive)
 I cannot give you any money. I am a little short of cash myself. (also)
- d. (attributive)
 John's own death surprised me more than that of his brother. (in contrast to s.o. else's death).
- (5) a. John assesses HIMSELF.
- b. Johann bewertet sich selbst.

In these examples the meaning of the relevant use is indicated by a paraphrase, which gives a rough idea of the relevant meaning without necessarily exhausting it. The common denominator of all these uses is that they express emphasis by contrasting a given value, the semantic value of the expression to which they relate, with relevant alternatives. In (4a) we find the adnominal use of the intensifier which contrasts the referent (R) of the preceding DP with an alternative definable in terms of that referent

(f(R)). Sentences of type (4a) illustrate the relevant type but do not form a suitable basis for reflexive compounds, in contrast to (5a) and its German counterpart (5b). The propositional structures underlying *self*-compounds are reflexive structures with an intensifier modifying the reflexive, as indicated by the combination *sich selbst* in the German example (5b). In English combinations of reflexive pronouns and the homophonous intensifiers are excluded, but the relevant focalization and emphasis is indicated by a focal stress on the reflexive. The effect of this focusing could be spelt out by the paraphrase ‘himself rather than other people’. In languages in which intensifiers are formally distinct from reflexive anaphors, as for example in the languages of the European continent (German, Romance, Slavic), reflexive anaphors can be strengthened by intensifiers (*sich selbst*). The use of intensifiers exhibited by such sentences is therefore also an instance of the adnominal use exhibited by sentence (4a), with the sole difference that it is the object argument rather than the subject the intensifier combines and interacts with. Since German illustrates these parallels much more clearly than does English, I will have to use German examples occasionally to discuss the different types of reflexive compounds in detail.

The second and the third uses are instances of an adverbial use, where the intensifier has the function of an adverbial, even though it interacts in its morphological properties with an argument, typically the subject. The contrast expressed by the exclusive adverbial use of the intensifier in (4b) relates not only to the relevant argument, but also to the way an activity is carried out by an agent (alone or by delegation, with some help, etc.). The overall semantic effect of this use therefore is an emphasis on the autonomous action of an agent. The inclusive adverbial use exemplified by (4c) can roughly (though not fully) be paraphrased by *also* and implies that the referent of the argument interacting with the intensifier and the alternative given in the context are both equally affected by a condition that is relevant for some action or state under discussion. In contrast to the adnominal use and the exclusive adverbial use, this inclusive adverbial use is not found in many languages. The final, attributive, use exemplified by (4d) is expressed by a different form in many European languages (i.e. the originally possessive adjective *own*⁶ in English or *eigen* in German and Dutch), but by exactly the same expression found in the nominal use in a wide variety of other languages (e.g. Turkish *kendi*, Mandarin *zìjǐ*). In some dialects of Dutch this attributive intensifier *eigen* ‘own’ can also be used as a reflexive marker. It should, therefore, not come as a surprise that in a variety of languages, for example in German, Greek and Finnish, this attributive intensifier is also used in reflexive compounds.

⁶ That *own* has lost its original meaning of ‘possession’ in many of its uses is clearly shown by examples like *his own losses*, *his own death*, etc.

If we now return to our question as to which of the uses of intensifiers are also found in reflexive compounds we have to exclude subtype (4c) right from the start. There is nothing in the meaning of *self*-compounds that is even vaguely reminiscent of the inclusive use of intensifiers. Furthermore, we will have to exclude sub-type (4a) of the adnominal use, whereas the subtype (5) is highly relevant. ‘Self-pity’ is a state where somebody pities himself (rather than others). Examples of type (4b), in which the two argument positions of a transitive verb are filled by different arguments are relevant for the derivation of agentive compounds like English *self-provider* or Ger. *Selbstabholer* ‘people who pick up things (mail, etc.) themselves’, which will play a marginal role in our discussion. What we also have to include, however, are cases where an exclusive adverbial intensifier combines with a reflexive construction. Compounds like *self-assessment* imply the responsibility of the agent in an event where agent and patient refer to the same person. Finally, as already mentioned, the attributive intensifier may also show up in reflexive compounds. The following table summarizes the preceding discussion:

Table 1: Uses relevant for the derivation of reflexive compounds

Adnominal:	yes, if the reflexive marker is modified
Attributive:	yes (in German, Finnish, Greek)
Inclusive adverbial (4c):	no
Exclusive adverbial (4b, 7):	yes, especially in reflexive construction

In other words, there are strong semantic reasons for the assumption that both the adnominal use exemplified by (5) and the adverbial use exemplified by (4b) are very relevant for the derivation and analysis of reflexive compounds. The semantic distinction in question can be more easily related to formal differentiations in Russian or German than in English or French. Consider the following two examples, where the addition of an adverb and a negation make a difference in constituent order more transparent:

- (6) a. Dieses Dorf [verwaltet sich (normalerweise)] selbst.
 ‘This village has self-administration.’
 b. Dieses Dorf verwaltet sich nicht selbst.
 ‘This village does not have self-administration.’
- (7) a. Karl kritisiert (oft) [sich selbst].
 ‘Charles often criticizes HIMSELF.’
 b. Karl kritisiert nicht sich selbst.
 ‘Charles is not criticizing HIMSELF.’

The first sentence means that the citizens of the village in question are not only the patient — as one would expect — but also the agent in the business of organizing their affairs. The paraphrases used above (without help, interference, delegation, alone) suggest that it is the exclusive adverbial use that we find in this example. The syntactic evidence of this analysis relates to the fact that the intensifier can be separated from the reflexive marker by the negator *nicht*.

In (7a), by contrast, the remarkable implication is that Charles is not only the agent in activities of criticizing, but also chooses himself as a victim, patient or target. The remarkable fact is the choice of the object referent as a target of criticism and we will therefore assume that the intensifier is in construction with the reflexive pronoun. I note in passing that the distinction between two basic verbal constructions is even clearer in languages where intensifiers are inflected for case, as for instance in Russian. In Russian only the intensifiers in (7) would be marked for accusative case:

- | | | | | |
|-----|--------------------------------|-------------|-----------|------------|
| (8) | Narod | upravljajet | sam | soboi |
| | People | govern-3.SG | INT.NOM | REFL.INSTR |
| | ‘The people govern themselves’ | | | |
| | | | | |
| (9) | Narod | upravljajet | samim | soboi |
| | People | govern-3.SG | INT.INSTR | REFL.INSTR |
| | ‘The people govern themselves’ | | | |

In Russian the relevant distinction can be expressed through case marking. In (8) the use of the intensifier corresponding to the German sentences in (6) is marked for nominative case and has the agentive meaning (‘The people have taken their government into their own hands.’). The instrumental marking for both intensifier and reflexive signals the adnominal use of the intensifier in (9) and implies that the people do not govern anybody but themselves. It goes without saying that it is not easy to find a convincing context for (9).

We can now distinguish two groups of reflexive compounds on the basis of these two uses of the intensifier in the propositions underlying the *self*-compounds: the adverbial exclusive one, as illustrated by (6) and (8), and the adnominal one in construction with the reflexive marker, as illustrated by (7) and (9). The relevant test is whether it is the agent or the patient that is a surprising choice for the event in question. The former choice is based on the adverbial exclusive use, the latter on the adnominal use in combination with reflexive pronouns. These two meanings can be distinguished in the underlying verbal constructions, but are neutralized in the formally condensed nominal compounds. The list given in (10) presents a collection of *self*-compounds with remarkable agents. The relevant information in those cases is the one providing the

source or agent of the event in question:

(10) Adverbial reflexive compounds

<i>Selbstzensur</i>	‘self-censorship’,
<i>Selbstbezeichnung</i>	‘self-accusation’,
<i>Selbstbestimmung</i>	‘self-determination’,
<i>Selbsthilfe</i>	‘self-help’,
<i>Selbstachtung</i>	‘self-respect’,
<i>Selbstentmündigung</i>	‘self-incapacitation’,
<i>Selbstentlebung</i>	‘suicide’,
<i>Selbstverpflichtung</i>	‘voluntary acceptance of an obligation’,
<i>Selbstaufgabe</i>	‘self-abandonment’,
<i>Selbstbedienung</i>	‘self-service’, etc.

Representative examples of the other type, i.e. of adnominal reflexive compounds, which characterize the patient or a non-subject as a surprising choice for the event in question, are the following:

(11) Adnominal reflexive compounds

<i>Selbstgefälligkeit</i>	‘complacency’,
<i>Selbstzufriedenheit</i>	<i>Fr. autosatisfaction</i>
<i>Selbstironie</i>	‘self-irony’, ‘self-mockery’,
<i>Selbstinszenierung</i>	‘self-fashioning’, ‘self-styling’,
<i>Selbstüberschätzung</i>	‘overestimate one’s potential’,
<i>Selbstgespräch</i>	‘talking to oneself’,
<i>Selbstvertrauen</i>	‘self-confidence’,
<i>Selbstbefriedigung</i>	‘masturbation’,
<i>Selbsthass</i>	‘self-hate’,
<i>Selbstbild</i>	‘self-image’,
<i>Selbstreflexionen</i>	‘self-reflexion’,
<i>Selbstverständnis</i>	, etc.

Note that the distinction between the two meanings has no formal correlate in the compounds themselves and is even difficult to draw for the underlying verbal constructions in many of the languages under comparison. The best test we can offer for the reality of this distinction is the question whether the subject or the object is a remarkable case for the relevant reflexive relation. It is, for example, not remarkable that we talk to someone, but to choose ourselves as addressees is remarkable and may even be regarded as pathological. *Selbstgespräch* is therefore an instance of an adnominal

reflexive compound. On the other hand, accusations and help may often be directed towards us, but they typically and normally come from others. It is therefore highly remarkable that the subject referent is both the source and the target of such acts. *Self-accusations* and *self-help* are therefore adverbial exclusive compounds in our terminology.

Note furthermore, that the distinction between these two semantic types of nominal reflexive compounds has a clear parallel in the class of reflexive adjectival compounds:

- (12) Adverbial adjectival compounds
self-loading, self-induced, self-appointed, self-inflicted, self-supporting, self-cleaning, etc.
- (13) Adnominal adjectival compounds
self-sufficient, self-reliant, self-respecting, self-addressed, self-satisfied, self-deprecating, etc.

The contrast is easily visible in the relevant paraphrases of the examples in each category:

- (14) a self-inflicted wound → ‘a wound inflicted by someone on him-/herself’
(adverbial exclusive → remarkable agent)
- (15) a self-addressed envelope → ‘an envelope addressed to oneself’
(adnominal → remarkable target)

4. Constraints on productivity

Word formation, both derivation and composition, is partly systematic and partly riddled with idiosyncrasies. In the following section I will take a look at the productivity of the relevant compounds and at the ecology of competing forms, such as, for example, the compounds with *selbst-* vs. those with *eigen-* in German.

In each of the languages under discussion there are a large number of reflexive compounds, only some of which are listed in dictionaries. Moreover, new formations are possible with one or more strategies available in a language. New formations analogous to *self-control* are possible in English, and German allows new reflexive compounds with both *selbst* and *eigen*. There is thus a considerable degree of productivity in this domain. There are, however, also constraints, most of which can probably be formulated in terms of the verbs underlying the nominalization with which the intensifier combines. The following constraints concern those verbal predicates:

- i. The verbs must be transitive, since they must be capable of expressing a reflexive relationship. But this constraint only applies to those cases which we have labeled adnominal compounds. The adverbial type can also be found with underlying intransitive verbs, as is shown by the German examples *Selbstläufer* ‘sure-fire success’, *Selbstheilungskraft* ‘self-healing power’, *Selbstreinigungskraft* ‘self-purifying power’, which are based on the adverbial use of the intensifier (*Selbstheilung* ‘self-healing’ < something heals by itself).
- ii. The constraint that the arguments of the underlying transitive verb must be subcategorized for the feature [+human] applies only to the adnominal type (*Selbsthass*, *Selbstbetrachtung*, *Selbstanklage*), since we also find *Selbstverwaltung* (< ‘jmd. verwaltet etwas selbst’).
- iii. The non-existence of **Selbstrasur* ‘self-shaving’, **Selbstvorbereitung* ‘self-preparation’, **Selbstscham* ‘self-shame’ alongside the existence of *Selbstgespräch* ‘talking to oneself’, *Selbstmord* ‘suicide’, *Selbstkontrolle* ‘self-control’ suggests that a relevant property of the underlying verbs is their other-directed character, that is they denote activities or states typically directed away from the agent. In other words, the adnominal compounds express remarkable reflexivity (cf. König & Vezzosi 2004). Body care and grooming is normally performed on oneself rather than on others and de-verbal nominalizations such as ‘washing’, ‘shaving’, ‘dressing’, ‘preparation’ or ‘defense’ are normally interpreted in a reflexive sense. Compounds like *Selbstverteidigung* ‘self-defense’ seem to provide a problem for this hypothesis, since ‘defend’ is clearly a non-other directed predicate. Note, however, that the German noun has a special non-compositional meaning.

This brief discussion has shown how difficult it is to formulate input-based constraints which allow us to make clear predictions. There are analogous difficulties for any attempt to formulate output-based constraints. The overall impression is that any formation is possible which can be interpreted analogously to certain well-established prototypes.

5. Distribution of competing forms

A comparative analysis of the relevant expressions shows that the set of reflexive compounds in a language comprises several layers, distinguishable both on the basis of their form and on the basis of their integration into a language. In German three different subsets or types can be distinguished. There are, first of all, a small number of lexemes

which have been borrowed from Latin or Romance languages and employ the Greek prefix *auto-* (16a). Secondly there are two major subsets with either *selbst* or *eigen* as markers of the “reflexive” relationship (cf. (16b) and (16c)).

- (16) a. *Autodidakt* ‘self-educated person’; *Autokratie* ‘autocracy’; *Autobiographie* ‘autobiography’.
b. *Selbstbetrug* ‘self-deception’; *Selbstverschulden* ‘one’s own fault’, *Selbstlob* ‘self-praise’.
c. *Eigenleistung* ‘personal contribution’, *Eigenwert* ‘intrinsic value’, *Eigenmittel* ‘one’s own resources’, *Eigenmächtigkeit* ‘unauthorized behavior’.

Compounds with either *selbst-* or *eigen-* are clearly the majority pattern. As is indicated by our examples, these two reflexive ‘prefixes’ overlap to a certain extent in their distribution, but manifest also a clear complementarity. In order to demonstrate this complementarity we have to exclude all those cases which are no longer compositional for various reasons, one being that the second component is no longer a free form and the complex lexeme can be regarded as an instance of derivation: *selbstlos*, *selbständig*, *Eigenschaft*, *Eigentum*, *Eigenheit*. The remaining cases manifest three tendencies differentiating between the distribution of *selbst* and *eigen*:

- i. In contrast to *selbst*, *eigen* also combines with non-derived nouns (*Eigenblut* ‘one’s own blood’, *Eigenheim* ‘one’s own home’, *Eigeninitiative* ‘initiative of one’s own’, *Eigenleben* ‘independent existence’, *der Eigenname* ‘proper name’, etc. (note, however: *Selbstzweck* ‘end in itself’).
- ii. The adverbial exclusive type is rare among compounds with *eigen*, the only exception being adjectival compounds like *eigenhändig* (< *mit eigenen Händen* ‘with one’s own hands’), *eigenmächtig* (< *aus eigener Macht*) ‘high-handed’, which can themselves be used as exclusive adverbial intensifiers.
- iii. A vague possessive relation (*appenance*) is still apparent in compounds with *eigen-*, and more clearly so in the possible paraphrases with attributive *eigen*: (*sein eigener Name* — *sein Eigenname*; *sein Eigenheim* — *sein eigenes Heim* ‘his own home’; *der Eigenwert* ‘intrinsic value’ — *der eigene Wert*; *Eigenkapital* ‘privately owned capital’; *Eigengeräusch* ‘background noise’; *Eigenwärme* ‘body heat’). In a slightly archaic construction the attributive adjective *eigen* can be used without the usual adjectival agreement: *ein eigen(es) Haus* ‘a house of your own’, *ein eigen(es) Leben* ‘individual existence’. It is

plausible that the compounds with *eigen-* developed from these NPs without inflectional marking as a result of frequent use.

Reflexive compounds are not only found in European languages. In Mandarin, reflexive compounds are based on the intensifier *zì*, a morpheme which is also used as the first component of the reflexive marker *zìjǐ*. Available dictionaries (e.g. the *Chinese-English Dictionary*, compiled by Guanghua Wu, University of Transportation, Shanghai, 1998) convey the impression that the relevant pattern of compounding has a high degree of productivity. Here are a few examples:

- | | | |
|------|-----------------|---------------------------|
| (17) | <i>zì ài</i> | ‘self love, self-respect’ |
| | <i>zì ào</i> | ‘self-conceited’ |
| | <i>zì zhuàn</i> | ‘autobiography’ |
| | <i>zì zhì</i> | ‘self-administration’ |
| | <i>zì zhù</i> | ‘self-help’ |

The target of this process of word-formation may be members of different word classes:

- | | | | |
|------|----|---------------|---------------------------------------|
| (18) | a. | <i>zì zhù</i> | ‘self-help’ (noun) |
| | b. | <i>zì mǎn</i> | ‘self-satisfied, conceited’ (adj.) |
| | c. | <i>zì kuā</i> | ‘boast’ (verb) |
| | d. | <i>zìrán</i> | ‘of course’ (adverb), ‘nature’ (noun) |

Considered in this context, the reflexive marker (and intensifier) *zìjǐ* is itself an instance of this compounding process. The following examples illustrate the syntactic properties of some relevant examples:

- | | | | | | |
|------|--|------|---------------|------|------|
| (19) | Zhāngsān | hěn | zì’ài. | | |
| | Zh. | very | self love | | |
| | ‘Zhangsan has strong self-love.’ | | | | |
| (20) | Zhāngsān | hěn | zì’ào. | | |
| | Zh. | very | self arrogant | | |
| | ‘Zhangsan is very conceited.’ | | | | |
| (21) | Zhāngsān | bù | zhī-dào | zì | bá |
| | Zh. | NEG | know | self | pull |
| | ‘Zhangsan does not know how to extricate himself.’ | | | | |

As many idiomatic expressions reflexive compounds are frequently used in so-called double phrasal structures of the type [zìV₁ zìV₂] or [zìN₁ zìN₂] (cf. Meng Ji 2005). The following examples are cases in point:

- (22) Zhāngsān zì yán zì yǔ.
 Zh. self talk self talk
 ‘Zhangsan is talking to himself.’
- (23) Zhāngsān hěn zì ài zì zhòng.
 Zh. very self love self esteem
 ‘Zhangsan has strong self-love and strong self-esteem.’

6. Contrasts

Reflexive compounds are found in all the languages under comparison. Moreover, there are typically several types of such expressions. German has as many as three types, three compounding strategies (*selbst-*, *eigen-*, *auto-*) in addition to the reflexive nominalization construction (*Unzufriedenheit mit sich selbst* ‘discontent with oneself’). Finnish has two compounding strategies parallel to those with *selbst-* (*itse-*) and *eigen* (*oma-*) in German and with *afto-* and *iðio-* in Modern Greek. In English there is the productive strategy with *self-* and a short list of compounds with the Greek root *auto-*. In Chinese there seems to be only one strategy based on the intensifier *zì*. The different types differ in their productivity. In German compounding based on the Greek intensifier *auto-* ‘self-’ is only possible with Greek roots. The compounds with *selbst-* listed in dictionaries, by contrast, are only a small subset of those currently in use.

Languages may differ in the category of possible targets of such compounding processes: German and English allow nouns and adjectives, French also allows verbs (*s’autoféliciter* ‘congratulate oneself’, *s’auto-organiser* ‘organize oneself’). But there are also differences in the types of admissible nouns: German allows the derivation of agentive nouns (*Selbstversorger* ‘somebody who is self-reliant’, *Selbsterzeuger* ‘somebody who produces his own food’), French does not. Translation from one language into the other must be based on a search through all the different options in a language and requires a thorough familiarity with the target language.

Several types of nominal reflexive compounds can be distinguished on the basis of their meaning. There is first of all the contrast between underlying adnominal uses and adverbial uses of intensifiers in the interpretation of de-verbal event nominalizations, i.e. the contrast between a remarkable object referent and a remarkable agentive subject referent for the reflexive compounds. Moreover, we have agentive compounds in German and in English.

The languages under comparison use various formal means to derive the relevant compounds: intensifiers (German), shorter and older versions of intensifiers (English, Mandarin, Hungarian), uninflected intensifiers (Russian) and prefixes borrowed from Greek (Romance). Despite some overlap there is some kind of complementarity in the distribution and meaning of the relevant formal elements. In those cases where two compounds with different intensifiers are found there is usually a difference in meaning, as is illustrated by minimal pairs in German like *Selbstwert* ‘self-esteem’ vs. *Eigenwert* ‘intrinsic value’ or *Selbständigkeit* ‘(economic) independence’ vs. *Eigenständigkeit* ‘(intellectual) independence’.

Reflexive compounds are highly condensed expressions of very complex meanings and seem to be found primarily in languages with a long tradition of literacy, philosophical writing and psychological reflections. They are not found in all languages. In Arabic, Hebrew, Basque and Turkish the relevant notions are always expressed by nominalizations of the French type *confiance en soi*, which do occur in all the languages under comparison. We can thus formulate the following hierarchy for the encoding of reflexive relations: compounds > nominalizations > sentential constructions. Whether or not a language also has the relevant compounds in addition to the nominalizations depends on the availability of compounding processes in the relevant language. Apart from that, language contact seems to have played an important role. More often than not the relevant pattern seems to have been borrowed from neighboring languages and Greek seems to have played an important role as a source.

References

- Bierwisch, Manfred. 2009. Nominalization: lexical and syntactic aspects. *Quantification, Definiteness, and Nominalization*, ed. by Anastasia Giannakidou & Monika Rathert, 281-319. Oxford & New York: Oxford University Press.
- Edmondson, Jerry A., and Frans Plank. 1978. Great expectations: an intensive self analysis. *Linguistics and Philosophy* 2.3:373-413.
- Ji, Meng. 2007. From etymology to modern phraseology: a corpus-based study of structural variants of Chinese idioms in naturally occurring contexts. *Journal of Chinese Language and Computing* 17.2:67-82.
- König, Ekkehard, and Peter Siemund. 2000. Intensifiers and reflexives: a typological perspective. *Reflexives: Forms and Functions*, ed. by Zygmunt Frajzyngier & Traci S. Curl, 41-74. Amsterdam & Philadelphia: John Benjamins.
- König, Ekkehard, and Volker Gast. 2006. Focused assertion of identity: a typology of intensifiers. *Linguistic Typology* 10.2:223-276.
- König, Ekkehard, and Letizia Vezzosi. 2004. The role of predicate meaning in the development of reflexivity. *What Makes Grammaticalization? A Look from its Fringes and its Components*, ed. by Walter Bisang, Nikolaus Himmelmann & Björn Wiemer, 213-244. Berlin & New York: Mouton de Gruyter.
- Kulikov, Leonid. 2007a. Reciprocal constructions in Vedic. *Reciprocal Constructions*, Vol. 2, ed. by Vladimir P. Nedjalkov, 709-738. Amsterdam & Philadelphia: John Benjamins.
- Kulikov, Leonid. 2007b. The system of reciprocal pronouns in Sanskrit: paths of grammaticalization and typologically rare semantic oppositions. Manuscript.
- Ludwig, Hans-Werner. 1963. *Die Self-Komposita bei Thomas Carlyle, Matthew Arnold and Gerard Manley Hopkins*. Tübingen: Niemeyer.
- Mutz, Katrin. 2000. Polysemie von AUTO- in komplexen Wörtern des Italienischen. Manuscript.
- Mutz, Katrin. 2004. Reflexivierung und Wortbildung im Französischen. *Semantische Rollen*, ed. by Rolf Kailuweit & Martin Hummel, 355-374. Tübingen: Gunter Narr Verlag.
- Reuland, Eric. 2008. Anaphoric dependencies: How are they encoded? Towards a derivational-based typology. *Reciprocals and Reflexives: Typological and Theoretical Explorations*, ed. by Ekkehard König & Volker Gast, 499-556. Berlin & New York: Mouton de Gruyter.
- Safir, Ken. 1996. Semantic atoms of anaphora. *Natural Language and Linguistic Theory* 14.3:545-589.

Case, Tense and Clauses

Yen-Hui Audrey Li (李豔惠)

University of Southern California

The theory of Case has played an important role in deriving categorial selection (subcategorization) from the semantic properties of lexical items. Some of such analyses have been built on the claim that noun phrases require Case but non-gerundival clauses are not in Case-assigned positions. In Chinese, the issue is whether different types of clauses are distinguished and whether the notion of Case is relevant in Chinese, particularly to (different types of) clauses. This paper shows that the notion of abstract Case indeed is relevant in Chinese but clauses are not in Case-assigned positions. When a non-Case marked clause appears in a Case position, it is actually contained in a complex noun phrase and Case is assigned to the containing NP instead. This analysis of clauses captures many interesting semantic and syntactic properties of clauses in various contexts.

Key words: Case, clause, tense, c-selection, s-selection, conjunction

1. Introduction

Chinese does not have much inflectional morphology. It is not surprising that the literature frequently debated on whether Chinese makes some of the similar types of distinctions as those demonstrated by inflectional morphemes in other languages. For instance, the issue has been raised whether Chinese has number distinctions (plurality; see, among others, Iljic 1994, 2001, 2005, Li 1999). Much debate has also generated arguments for and against the distinction between finite (tensed) clauses and non-finite clauses in Chinese (see, for instance, Huang 1982[1998], Li 1985, 1990, C. Tang 1990, T. Tang 2000, Hu, Pan & Xu 2001, among many others). Recent works have considered the role of semantic and syntactic tense anchoring in “tensed” clauses (Lin 2003a, 2003b, 2006, Sybesma 2007, Tsai 2008).

Focusing on the question of whether different types of clauses need to be distinguished, we note that the issue is not only interesting in itself but also important in helping us understand the behavior of clauses in Chinese and the more general studies concerning order and constituency in Universal Grammar. Claims have been made in the literature that certain constructions allow only specific kinds of clauses. Take for instance Kayne’s anti-symmetric approach to word order and hierarchical structures in

Universal Grammar. In such an approach, relative constructions in all languages have the same base structure and head-final relative clauses — the modifying relative preceding the modified phrase — are, roughly speaking, derived by moving the relative clause across the modified phrase to a higher (left) position. Due to some well-formedness conditions on the movement process, head-final relative clauses are possible only when such relatives are non-tensed (for details of the analysis, see Kayne 1994). That is, a tensed relative clause should not precede the modified phrase. In this regard, being able to define clearly the nature of clauses in Chinese and determine their distribution would be a good testing ground for the said approach to a universal relative construction.

The distinction of different types of clauses also bears on the characterization of Case theory and the Visibility condition on thematic role assignment (Stowell 1981, Plann 1986, Li 1985, 1990, Lin 2003a, 2003b, 2006, Sybesma 2007, Tsai 1995, 2008). Empirically, whether or not clauses need Case has been controversial.¹ Pesetsky (1982) distinguishes between NPs and clauses categorially: only the former need Case. Accordingly, the subcategorization properties of heads can be determined by the s(ematic)-selection properties of a head, coupled with the Case assigning ability of the head. C(ategorical)-selection is derived from s-selection interacting with Case. Stowell (1981) argues that some clauses are inherently visible for theta-assignment, such as infinitives and the tensed complements of heads expressing psychological states. Other tensed clauses are assigned Case but cannot stay in Case positions. Such Case properties affect the distribution of clauses and related grammatical processes. For instance, demonstrating the role of Case on the distribution of clauses is the generalization that non-gerundive clauses do not occur in the object positions of Ps.

Li (1985, 1990) was the first systematic attempt to define the role of Case in a language like modern Chinese, which does not have any overt Case markings. Regarding clauses, the poverty of morphological inflections in Chinese makes it quite challenging to determine if different types of clausal structures can be clearly distinguished in this language. The study needs to rely on indirect evidence. Appropriate interpretations of the relevant data are critical. The works by Huang (1982[1998]), Li (1985, 1990), C. Tang (1990), T. Tang (2000), Tsai (1995), Hu, Pan & Xu (2001), among many others, represent the continuous serious efforts to solve the challenges. As theories evolve and more empirical generalizations are discovered, better tools become available to explore the relevant issues in greater depth. This work builds on the more recent observations and analyses regarding tense anchoring, conjunction and ellipsis structures in Chinese (Aoun & Li 2003, Huang 2006, Li 2005, 2007, 2008, Sybesma 2007, Tsai 1995, 2008, Zhang 2009) and re-examine the Case status and the structures of the clausal constituents

¹ Following a widely adopted convention, we use the capitalized “Case” to refer to the notion of abstract Case in Case theory.

in postverbal and preverbal positions. It brings new empirical generalizations from the constructions involving null objects, conjunction, topicalization of clauses, and extraction from within complement clauses to properly define the nature of clauses in clearly Case-marked positions in Chinese. The study will help resolve the conflict between the claim that clauses in Chinese are always Case-marked when they are assigned thematic roles (Li 1985, 1990, Tsai 1995) and the modified claim that some of such clauses are not assigned Case (Li 2008). It will be shown that clauses in Case positions behave as if they are nominal expressions.² This follows if Case is assigned to NPs, rather than clauses, as in Pesetsky (1982). Chinese is not much different from English in regard to the Case requirement on clauses and NPs (following Pesetsky's analysis of English). The conclusion has significant consequences on how Case should be characterized in the grammar.

Empirically, this work will focus on the clauses in the object positions of verbs and prepositions, leaving other possibilities to a separate work because of the limited space. I will first review the issues regarding Case and clauses in §1 and discuss the distribution of “tensed” clauses in §2. Section 3 brings a different perspective to the issues regarding the categorial status of “tensed clauses” — conjunction. An NP structure will be proposed for clauses in clearly Case-marked positions (§4). It will be concluded that NPs and clauses do not behave alike regarding Case theory (§5).

2. Case and clauses

Case theory was a major tool in the government and binding theory to capture the generalizations regarding order and constituency in natural languages (Chomsky 1980, 1981, 1986). For instance, NPs must be assigned Case — the Case filter (Chomsky 1981).

(1) the Case filter

*_[NP e]

where NP has phonetic content but not Case

The function of (1) is to ensure that noun phrases appear in the appropriate positions structurally, such as the object position of verbs and prepositions and the subject position

² Nominal expressions in argument positions will be labeled as NPs in this work, with reference to the classic terms such as Complex NP constraint, the Case filter applying to NPs. The distinction between NPs and DPs is not relevant in this work.

As will be clear in the text, the question of whether Chinese distinguishes tensed clauses from infinitives would not be resolved in this paper because both do not occur in Case-marked positions.

of tensed clauses in English — right next to the Case assigners, V, P and Tense.³ The Case filter (and the more general Case theory) was further connected to the Theta theory: Case must be present when a theta-role is assigned — the Visibility Condition on theta-role assignment. Every theta-role must be assigned to an argument and every argument must be assigned a theta-role.⁴ The requirement of Case is therefore reduced to the need of theta-roles to be properly assigned to arguments and arguments properly receiving theta-roles. Nonetheless, the NP Case filter and the Visibility condition have different empirical coverage. The former applies to NPs; whereas the latter is relevant to all the complements assigned thematic-roles, including clauses. To determine if the Case filter can indeed be reduced to the Visibility condition, it is important to determine if clauses are subject to the Case filter. Stowell (1981) extensively discusses this issue and argues that different types of clauses should be distinguished according to their Case properties: (i) tensed clauses need Case but cannot appear in Case position, (ii) infinitival clauses (having a tense feature but not \pm past, i.e., not having lexical content) are intrinsically Case-marked, just like PPs, (iii) certain predicates allow their tensed clausal complements to be assigned thematic roles without Case, and (iv) gerundival clauses, not specified for Tense, are assigned Case. In other words, only gerundival clauses in (iv) occur in the positions where Case is assigned. Infinitives are not relevant because they are Case-marked intrinsically. Tensed clauses are most interesting because of the following condition proposed by Stowell (Stowell 1981:146):

(2) The Case-Resistance Principle (CRP)⁵

Case may not be assigned to a category bearing a Case-assigning feature.

Tense is a Case assigner (assigning Case to the subject of the clause), making tensed clauses unable to stay in Case positions, according to the CRP. When assigned Case,

³ Not all Case markers are overt. For instance, Larson (1985) suggests that bare NP adverbs of time, location like *tomorrow*, *now*, *here*, *someplace warm and sunny* and a limited set of bare-NP adverbs of manner are inherently Case marked.

⁴ A common assumption is that only subcategorized complements are assigned thematic roles. However, a more inclusive view has also been proposed, such as the following condition on adverbial theta-role assignment (Larson 1985:606):

(i) *Adverbial θ -Role Assignment*

Assign an adverbial θ -Role to α , where α is any phrase.

If this is adopted, the Visibility condition does not exempt adverbial NPs from the Case filter.

⁵ Li (1985, 1990) suggests that the Chinese data favor the Thematic Resistance Principle as in Koopman (1984) over Stowell's CRP, because thematic and Case positions are not identical in this language.

clauses⁶ need to be moved away from the Case-marked position via extraposition or topicalization. This accounts for why the sentences in (3a-c), are not acceptable, where tensed clauses appear in Case-marked positions, in contrast to (3d-e), where the tensed clause has been vacuously moved away from the subject and object position:

- (3) a. *He blamed it [on [that Bill was too strict]].
 b. *I consider [[that John came home] to be fortunate].
 c. *John wondered [how [that Bill arrived late] to upset Mary].
 d. [That he is here] is important.
 e. I know [that he is the best].

In contrast to English clauses, Li (1985, 1990) notes that clauses in Chinese behave like nominal phrases and occur in Case-marked positions. Tsai (1995) further explored the similarity between clauses and nominal phrases with respect to their sensitivity to the Case requirement. The following examples are from Tsai (1995:282-285),⁷ illustrating the relevance of Case to clauses as well as to NPs.

- (4) wo [* (dui) [Akiu weishenme bu lai]] hen guanxin.
 I about Akiu why not come very care
 'I care about why Akiu will not come.'
- (5) wo [* (dui) [Akiu bu lai]] hen zaiyi.
 I about Akiu not come very mind
 'I do mind Akiu will not come.'
- (6) wo [* (dui) zhe-jian shi] hen guanxin.
 I about this-CL matter very care
 'I care about this matter.'
- (7) wo [* (dui) zhe-jian shi] hen zaiyi.
 I about this-CL matter very mind
 'I do mind this matter.'

⁶ We will simply use "clauses" as abbreviation for "tensed clauses" in the rest of this work, because tensed clauses present most interesting complications regarding Case requirements: they cannot be in case-assigned positions in English.

⁷ Tsai's (1995) translation includes the intensifier *do*: *I do care about...* The *hen* here need not be interpreted as a real intensifier, as the deletion of *hen* makes the sentence unacceptable (see, for instance, Li & Thompson 1981 for *hen* without its intensifier interpretation).

These examples show that a complement clause and a complement NP in the preverbal position equally require a Case-marker *dui* (also see Paul 2002:699).⁸

Postverbally, either a clause or an NP complement is assigned Case by the verb; therefore, the Case-marking *dui* does not appear:

- (8) wo hen guanxin [(*)*dui*] [Akiu weishenme bu lai] / zhe-jian shi].
 I very care about Akiu why not come / this-CL matter
 'I care about why Akiu will not come/this matter.'
- (9) wo hen zaiyi [(*)*dui*] [Akiu bu lai] / zhe-jian shi].
 I very mind about Akiu not come / this-CL matter
 'I do mind Akiu will not come/this matter.'

Other prepositions behave like *dui*:

- (10) a. cong [Akiu jinlai zheli] dao [ta likai],
 from Akiu enter here to he leave
 Lisi yi-ju hua dou mei shuo.
 Lisi one-CL word all have-not speak
 'From the moment Akiu entered here to the moment he left, Lisi did not say a word.'

⁸ Chinese allows SOV word order without a Case marker preceding the O, especially when the preposed object has a special discourse function (such as contrast, focus):

- (i) wo ji bu chi.
 I chicken not eat
 'I don't eat chicken.'
- (ii) ta zhe-jian shi zhidao le.
 he this-CL matter know *le*
 'He knows about this matter.'

Unexpectedly, a clause generally is not quite natural in such an object position:

- (iii) ??ta ni bu neng lai zhidao le.
 he you not can come know *le*
 'He knows that you cannot come.'

Dui is not possible when the verb is *zhidao* and some others (which might easily take *ba*):

- (iv) *ta dui ni bu neng lai zhidao le.
 he to you not can come know *le*
- (v) ta ba/**dui* haizi da le ji ci.
 he child hit *le* several times

This suggests that *dui* is not simply a Case marker for a preverbal object.

- b. cong [Akiu shenmeshihou qichuang] dao [ta zai nali chifan],
 from Akiu when get-up to he at where eat
 Lisi dou dating-de yiqingerchu.
 Lisi all investigate thorough
 ‘From the question of when Akiu wakes up to the question of where he eats, Lisi made a thorough investigation.’

The following examples, with sentential subjects inside relative clauses, illustrate the possibility of a clause staying in subject positions.

- (11) a. [henduo [[Akiu neng huo-zhe hui-lai]
 many Akiu can live-Dur back-home
 shi tamen_i jingya] de ren_i] dou mei lai.
 make them surprised *de* people all have-not come
 ‘[Many people to whom it is surprising [that Akiu can come back alive]] did not come.’
- b. [henduo [[Akiu neng-bu-neng huo-zhe hui-lai]
 many Akiu can-not-can live-DUR back-home
 gen tamen_i wuguan] de ren_i] dou mei lai.
 to them irrelevant *de* people all have-not come
 ‘[Many people to whom it is irrelevant [whether Akiu can come back alive or not]] did not come.’

The similarity in the positions where clauses and NPs are possible suggests that clauses are assigned Case and the Case filter can be appropriately reduced to the Visibility condition on theta-assignment.

Unfortunately, the data are more complicated. The parallelism in distribution between NPs and clauses fails in the following instances, where a clausal complement is acceptable postverbally but not an NP complement (Tsai 1995:301-302, ex.(51)-(52)).

- (12) *wo hen haoqi [zhe-jian shi de qiyin].
 I very curious this-CL matter *de* cause
 ‘I am curious about the cause of this matter.’
- (13) wo hen haoqi [Akiu weishenme bu lai].
 I very curious Akiu why not come
 ‘I am curious why Akiu will not come.’

Tsai suggests that the Visibility condition is obeyed consistently, i.e., both clausal and NP complements should be assigned Case in order to receive theta-roles. The difference in the above examples is simply that *haoqi* selects a clause, not an NP.⁹ In other words, *haoqi* specifies a categorial selection requirement (c-selection): the complement following *haoqi* must be a clause (or a PP, see note 9). C-selection is arbitrary: the c-selection requirement for each lexical item must be listed. This contrasts with Pesetsky's (1982) proposal that c-selection should follow from the semantic selection (s-selection) properties of lexical heads and the abilities of specific heads to assign Case. According to Pesetsky, the following contrasts show that the verb *ask* in English assigns Case to its complement, not *wonder*, even though they both require a question complement:

- (14) a. John asked the question.
 b. John asked what the time was.
- (15) a. *John wondered the question.
 b. John wondered what the time was.

These pairs of sentences demonstrate that not all verbs in English assign Case and accept NPs as their complements. Clauses do not need Case; therefore, they can be complements of the verbs unable to assign Case.

If, according to Tsai, Chinese does not derive complement selection from s-selection interacting with Case and must rely on c-selection, we seem to be missing a generalization: it generally is impossible to find the type of verbs opposite of the Chinese *haoqi* 'curious', labeled as *HAOQI*, which s-selects a proposition but only c-selects an NP. That is, we do not seem to have such a hypothetical *HAOQI* that can appear in (a), but not in (b).¹⁰

- (16) a. wo hen HAOQI [zhe-jian shi / zhe-ge wenti].
 I very curious this-CL matter / this-CL question
 'I am curious about this matter.'

⁹ Tsai (1995) notes that *haoqi* can also select a PP, although he does not discuss further why PPs appear preverbally (see Li 1985, 1990 for the split between the Case directionality requirement and the head parameter, which would need to be recast in different terms in the current framework):

(i) wo [*(dui) [Akiu weishenme bu lai] / na-jian shi hen haoqi.
 I about Akiu why not come that-CL matter very curious
 'I am curious why Akiu will not come.'

¹⁰ If a verb s-selects an individual, only an NP is a possible complement.

- b. *wo hen HAOQI [Akiu weishenme bu lai].
 I very curious Akiu why not come
 ‘I am curious why Akiu will not come.’

In this respect, it is interesting to point out that in English, even though argument NPs always require Case, it is not the same with complement clauses. In discussing the Case requirement of clauses in English, Stowell observes that some heads license their clausal complements without Case assignment.

- (17) a. Mary is happy that Charles is leaving home.
 b. Kevin is certain that the tent is in the car.
 c. Neil is afraid that the computer will break down.
 cf.
 (18) a. Kevin is [certain of Ray’s genius].
 b. Neil is [afraid of Constable O’Malley].
 (19) a. *Kevin is certain Ray’s genius.
 b. *Neil is afraid Constable O’Malley.

The English translations for (12)-(13) also show that the clausal complement in the English sentence corresponding to (13) is not assigned Case. Stowell (1981:204) suggests that “these psychological-state-denoting adjectives have a special property that excludes them from the general requirement that theta-roles can only be assigned to A’-chains headed by PRO or Case ... the adjective phrases [in these cases] instantiate a special case of theta-role assignment, which is limited to relations of awareness or recognition of the propositional content of a complement clause.” It is proposed that theta-roles can be assigned to clauses when the head has a lexical feature [+R]. Nonetheless, the postulation of [+R] has the same effect as saying that clauses are not assigned Case in these instances.¹¹

With respect to Chinese, the following generalization has been put forward in Li (2005, 2007) on empty categories: in Chinese, when a verb allows a nominal object, it also accepts a null object; when a verb only allows a clausal object, it does not accept a null object.

¹¹ Stowell notes that these clauses cannot be topicalized, leaving variables, which require Case.

- (i) a. *[That Charles is leaving], I believe that Mary is [happy ____].
 b. *[That the computer will break down], I know that Neil is [afraid ____].

- (20) a. If a verb is subcategorized for a nominal object, the object can be empty.
 b. If a verb is subcategorized for a clausal object only, the object cannot be empty.

Verbs allowing a nominal object and a null object:

- (21) a. wo tingdao-le na-jian shi.
 I hear-*le* that-CL matter
 'I heard that matter.'
 b. wo tingdao ta de-le da jiang le; ta ye tingdao-le.
 I heard he get-*le* big prize *le* he also hear-*le*
 'I heard that he got a big prize; he also heard.'

Verbs disallowing a nominal object and a null object. *Zheme(yang)* 'so' must appear:

- (22) a. *wo renwei/yiwei na-jian shi.
 I think/think that-CL matter
 'I thought/thought that matter.'
 b. wo renwei/yiwei ta hen congming; tamen ye *(zheme(yang))
 I think/think he very smart they also so
 renwei/yiwei.
 think/think
 'I thought that he was smart; they also thought.'
 (23) a. *wo cai na-jian shi.
 I guess that-CL matter
 'I guess that matter.'
 b. wo cai ta hen congming; tamen ye *(zheme(yang)) cai.
 I guess he very smart they also so guess
 'I guess that he is smart; they also guess.'
 (24) a. *wo dasuan na-jian shi.
 I plan that-CL matter
 '*I planned that matter.'

- b. wo dasuan mingtian qu; tamen ye *(zheme(yang)) dasuan.¹²
 I plan tomorrow go they also so plan
 ‘I planned to go tomorrow; they also planned.’

This correlation can be subsumed under the following condition.

- (25) The Visibility Condition on empty categories:

Empty categories in argument positions should be assigned Case or be in a chain containing Case.

This condition means that a null object is possible only if Case is assigned to the position and it may capture the different possibilities of a null object between English and Chinese. As is well-known, Chinese, not English, allows its object to be null:

- (26) a. John saw him. *Mary saw, too.
 b. I like him. *She doesn’t like.
 cf.
 (27) a. John kanjian-le ta; Mary ye kanjian-le.
 John see-*le* him Mary also see-*le*
 ‘John saw him; Mary also saw.’
 b. wo xihuan ta; ta bu xihuan.
 I like him he not like
 ‘I like him; he doesn’t like.’

This contrast can be captured by an adapted inverse Case filter in Bošković (1997: 134-142).

- (28) English, not Chinese, requires Case to be realized on a lexical item.

The obligatoriness of overtly realizing Case features and the requirement on null arguments to be Case marked conspire to rule out any null objects in English. In this language, if a Case feature is available, it must be overtly realized on a lexical item; if such a feature is not available, a null argument is not licensed. These considerations also capture the fact that the object CPs in the following instances cannot be “deleted” (cf. Lobeck 1995, Merchant 2001 for the impossibility of CP deletion in English). The

¹² If *dasuan* is only subcategorized for an infinitival clause, it would not affect the discussion on the Case status of clauses in this work, as infinitivals are not assigned Case, as in Stowell (1981).

relevant object positions are not assigned Case.

- (29) a. *Mary was afraid that the idea wouldn't work and Bill was [_{AP} [happy [_{CP} e]]].
b. *I suppose that he will come and they suppose [_{CP} e], too.

The facts presented so far reveal two conflicting generalizations: those in (4)-(11) show that clauses are like NPs and are Case-marked in Chinese. In contrast, other facts, such as those related to the generalization in (20), suggest that clauses in Chinese, as in English (14)-(19), are not like NPs and are not Case-marked. How can this conflict be resolved? Three logical options suggest themselves:

- (30) a. Clauses must always be assigned Case in Chinese but not in English. (20) should not be accommodated by Case.
b. Different types of clauses must be recognized in Chinese. That is, we need to recognize finer peripheral structures for clauses (Cinque 1999, 2002, Rizzi 1997, 2004). A clause may have some or all of the following projections at the left periphery: Force Phrase, Topic Phrase, Operator Phrase, etc. Case is required with certain projections but not the others.
c. Clauses in Chinese are not in Case positions, just as in English. The occurrence of clauses in Case-marked positions is only apparent.

In the following discussions, let us, for the sake of convenience, refer to the complement positions of the verbs in (22)-(24) as non-Case positions and those in (21), together with the objects of prepositions and subjects, as Case positions. It will be shown that the option in (30c) is more adequate than the other two, contra the observations and analyses in Li (1985, 1990) and Tsai (1995).

3. Different types of clauses?

As noted, the lack of inflectional morphology in Chinese makes it more difficult to identify distinct categories. For instance, the tensed clausal complements in (31a-b) and (33a-c) can all be translated as clauses in Chinese ((32), (34) respectively), even though only the verbs in (31) and (32) can assign Case to the complements.

- (31) a. I know [that he does his work]/it.
b. I like his/him doing this work/it.

- (32) a. wo zhidao [ta zuo zhe gongzuo] / zhe-jian shi.
I know he do this work / this-CL matter
'I know that he does this work/this matter.'
- b. wo xihuan [ta zuo zhe gongzuo] / zhe-jian shi.
I like he do this work / this-CL matter
'I like his/him doing this work/this matter.'
- (33) a. He is curious (about) who is doing this work.¹³
b. He is happy that he is doing this work.
c. He prefers for him to do this work.
- (34) a. ta hen haoqi shei zuo zhe gongzuo.
he very curious who do this work
'He is curious who is doing this work.'
- b. ta hen gaoxing ta zuo zhe gongzuo.
he very happy he do this work
'He is happy that he is doing this work.'
- c. ta ningyuan ta zuo zhe gongzuo.
he prefer he do this work
'He prefers for him to do this work.'

¹³ An important difference between tensed clauses headed by *wh*-words and other tensed clauses is that the former seem to optionally appear in Case-marked positions, but the latter generally do not (except some limited cases such as *except that*: "It would be funny *except that it's not*."):

- (i) a. It's the question of whether God Himself operates outside His own law.
b. The question of whether Abrams' STAR TREK is truly canon has been answered.
- (ii) a. The question whether Obama is black enough should end.
b. On the question whether temperature determines the distribution of marine species of animals in depth.

Chinese also allows questions in both contexts:

- (iii) a. ni renwei ta weishenme mei lai/ta lai-bu-lai?
you think he why not come/he come-not-come
'Why do you think he didn't come/Do you think he will come or not?'
- b. wo hen haoqi shei mei lai.
I very curious who not come
'I am curious who did not come.'
- (iv) a. wo [dui [ta yuan-bu-yuanyi lai]] hen haoqi.
I to he willing-not-willing come very curious
'I am curious about whether he is willing to come.'
- b. wo juede [shei yinggai lai] bu shi zhongyao de wenti.
I feel who should come not be important *de* question
'I feel who should come is not an important question.'

Even without overt morphological markings, the question can still be asked: does Chinese distinguish different types of clauses in the relevant contexts? Might different types of clauses be responsible for the seemingly contradictory patterns: some clauses appear in Case-marked positions and some others do not? The data for the clauses in Case and non-Case positions seem to suggest that both allow the same type(s) of clauses. They can be *wh*-questions, as demonstrated earlier. In addition, topic and focus (indicated by the focus marker *shi* ‘be’ or *daodi* ‘on earth’) elements are also allowed in both contexts.

- (35) a. wo zhidao na-jian shi.
I know that-CL matter
‘I know that matter.’
b. wo dui na-jian shi hen haoqi.
I to that-CL matter very curious
‘I am curious about that matter.’
c. *wo yiwei na-jian shi.
I thought that-CL matter
- (36) a. wo zhidao **shi** ta na-le na-ben shu.
I know be he take-*le* that-CL book
‘I know that *he* took that book.’
b. wo dui ta **daodi** bu yuanyi zuo **shenme** hai mei gao qingchu.
I to he to-end not willing do what still not make clear
‘I am still not clear what on earth he is not willing to do.’
c. wo yiwei **shi** ta na-le na-ben shu.
I thought be he take-*le* that-CL book
‘I thought that *he* took that book.’
- (37) a. wo zhidao na-ben shu, ta na-le.
I know that-CL book he take-*le*
‘I know that book, he took.’
b. wo dui na-jian shi shei yuanyi zuo hen haoqi.
I to that-CL matter who willing do very curious
‘I am curious about, that matter, who will do (it).’
c. wo yiwei na-ben shu ta na-le.
I thought that-CL book he take-*le*
‘I thought that book, he took.’

This suggests that the typical left-peripheral elements such as question operators, topic and focus elements are allowed in the complement positions of Ps and verbs assigning Case and the non-Case-assigning verbs. What is left is the option of tense: could it be that Chinese does distinguish tensed clauses from non-tensed ones and that verbs like *renwei* ‘think’, *cai* ‘guess’ require tensed clausal complements but Case-marked positions are occupied by non-tensed clauses?

The answer to this question is dependent on the availability of a syntactic representation of tense in Chinese. Tsang (1981), Huang (1982[1998]), Li (1985, 1990), among others, argue that Chinese distinguishes infinitival clauses from tensed clauses and modal-like words such as *hui* can serve as a tense marker. In contrast, Hu, Pan & Xu (2001) argue that Chinese does not make such a distinction. Lately, Lin (2003a, 2003b, 2006), Sybesma (2007) and Tsai (2008) revisited the issue of whether Chinese has a tense projection syntactically. In the following paragraphs, we show that even if we follow the claim by Sybesma and Tsai that Chinese does express tense syntactically and accept all their arguments for the claim, such tensed clauses seem to still appear in the typical Case-marked positions.

Let us illustrate this point with the most recent work, Tsai (2008). According to Tsai, there is some “incompleteness” effect observed in Chinese for sentences like the following:

- (38) a. %Akiu pao-zhe.
 Akiu run-DUR
 b. %Akiu kan-zhe dianshi.
 Akiu watch-DUR TV
- (39) %Akiu na-le shu.
 Akiu take-PRF book
 ‘Akiu took books.’

Tsai claims that these cases sound incomplete because of their failure in tense anchoring, i.e., to guarantee a proper temporal reference of a given sentence through syntactic measures. He analyzes tense anchoring as a process of spelling out an underlying event argument by a variety of morpho-syntactic means. This process may involve event coordination, event subordination, event modification, event quantification, or verb raising to *v/T*.

- (40) a. Akiu yizhi pao-zhe.
 Akiu continuously run-DUR
 ‘Akiu is running continuously.’

- b. Akiu yibian kan-zhe dianshi, yibian
 Akiu on.the.one.hand watch-DUR TV on.the.other
 xie-zhe baogao.
 write-DUR report
 ‘Akiu is watching TV and writing the report at the same time.’
- c. Akiu na-le san-ben shu.
 Akiu take-PRF three-CL book
 ‘Akiu took three books.’
- d. Akiu yinggai/mei na shu.
 Akiu should/have.not take book
 ‘Akiu should take/have not taken books.’
- e. Akiu na-le shu jiu pao.
 Akiu take-PRF book then run
 ‘Akiu ran away immediately after taking the book.’
- f. Akiu na-le shu le.
 Akiu take-PRF book PRT
 ‘(As for now,) Akiu has taken the book.’

Regardless of the mechanisms to encode the notion of tense, what is pertinent to our discussion is that even if we follow Tsai’s arguments and agree that Chinese expresses tense syntactically (tense anchoring), these “tensed clauses” seem to comfortably appear in typical Case positions, including the object of Case-assigning verbs and the object of prepositions:

- (41) a. wo zhidao [Akiu yizhi pao-zhe].
 I know Akiu continuously run-DUR
 ‘I know that Akiu is running continuously.’
- b. wo zhidao [Akiu na-le shu jiu pao].
 I know Akiu take-PRF book then run away
 ‘I know that Akiu ran away immediately after taking the book.’
- c. wo zhidao [Akiu name kuai jiu na-le shu le].
 I know Akiu that fast then take-PRF book PRT
 ‘I know that Akiu has taken the book that fast.’
- (42) a. wo [dui [Akiu yizhi pao-zhe] hen bu gaoxing.
 I to Akiu continuously run-DUR very not happy
 ‘I am not happy that Akiu is running continuously.’

- b. wo [dui [Akiu yibian kan-zhe dianshi, yibian
I to Akiu on.the.one.hand watch-DUR TV on.the.other
xie-zhe baogao]] hen bu gaoxing.
write-DUR report very not happy
'I am not happy that Akiu is watching TV and writing the report at the
same time.'
- c. wo [dui [Akiu na-le san-ben shu]] hen bu gaoxing.
I to Akiu take-PRF three-CL book very not happy
'I am not happy that Akiu took three books.'
- d. wo [dui [Akiu yinggai/mei na shu]] hen bu gaoxing.
I to Akiu should/have.not take book very not happy
'I am not happy that Akiu should take/have not taken books.'
- e. wo [dui [Akiu na-le shu jiu pao]] hen bu gaoxing.
I to Akiu take-PRF book then run very not happy
'I am not happy that Akiu ran away immediately after taking the book.'
- f. wo [dui [Akiu name kuai jiu na-le shu le]] hen bu gaoxing.
I to Akiu that fast then take-PRF book PRT very not happy
'I am not happy that Akiu has taken the book that fast.'

The range of possibilities shown above is similar to that for the patterns with verbs not allowing NP complements, such as *renwei/yiwei/cai* 'think/guess'.

- (43) a. wo yiwei [Akiu yizhi pao-zhe].
I think Akiu continuously run-DUR
'I thought that Akiu was running continuously.'
- b. wo yiwei [Akiu yibian kan-zhe dianshi,
I think Akiu on.the.one.hand watch-DUR TV
yibian xie-zhe baogao].
on.the.other write-DUR report
'I thought that Akiu was watching TV and writing the report at the same
time.'
- c. wo yiwei [Akiu na-le san-ben shu].
I think Akiu take-PRF three-CL book
'I thought that Akiu took three books.'
- d. wo yiwei [Akiu yinggai/mei na shu].
I think Akiu should/have.not take book
'I thought that Akiu should take/had not taken books.'

- e. wo yiwei [Akiu na-le shu jiu pao].
 I think Akiu take-PRF book then run
 'I thought that Akiu ran away immediately after taking the book.'
- f. wo yiwei [Akiu name kuai jiu na-le shu le].
 I think Akiu that fast then take-PRF book PRT
 'I thought that Akiu had taken the book that fast.'

The lack of contrast in acceptability between (41)-(42) and (43) shows that, if indeed there is tense anchoring, distinguishing tensed clauses from non-tensed ones, the types of clauses grouped under the tensed ones can appear in those positions allowing NPs (Case-marked positions), as well as the positions not allowing NPs.

A cautionary remark should be made concerning cases like (41), those with verbs allowing both a postverbal nominal and clausal complement. Although we recast this pattern in terms of Case marking — the verbs in this pattern can assign Case, it should be noted that the complement clause might not always be in Case positions, as such a clause might be extraposed (Stowell 1981). Therefore, we will focus on the patterns in (42) and (43).

Even though (42) and (43) show that the same types of clauses can appear in clearly Case marked positions (prepositional object) and non-Case marked positions (those disallowing NPs), the two patterns do not share the entire range of possibilities. For instance, the focus marker *shi* is quite natural in the clausal complements of the verbs disallowing NP complements (44); whereas the sentences in (45) show that the focus marker *shi* is much less acceptable in clearly Case-marked positions:

- (44) wo yiwei [Akiu shi yizhi pao-zhe].
 I think Akiu be continuously run-DUR
 'I thought that Akiu indeed was running continuously.'
- (45) a. wo [dui [Akiu (*shi) yizhi pao-zhe]] hen bu gaoxing.
 I to Akiu be continuously run-DUR very not happy
 'I am not happy that Akiu indeed was running continuously.'
- b. wo [ba [Akiu (*shi) yizhi pao-zhe]] dangzuo shi hen
 I ba Akiu be continuously run-DUR regard be very
 zhongyao de shi.
 important *de* matter
 'I took it as important that Akiu indeed was running continuously.'
- c. wo [bei [Akiu (*shi) yizhi pao-zhe]] fansi le.
 I bei Akiu be continuously run-DUR annoyed
 'I was annoyed by Akiu's indeed running continuously.'

The unacceptability of (45) is interesting. Have we finally found a clue to distinguishing the type of clauses that does occur in Case marked positions and the other type that does not? The answer to this question will be shown to be negative. Indeed, the next section will show that the fact from conjunction argues for a nominal structure for the clauses in the object position of prepositions. Together with the fact regarding the overt co-occurrence of nouns (phrases) with clauses, it will be proposed that clauses in prepositional object positions are actually nominal expressions. Accordingly, there is no compelling reason to recognize the occurrence of clauses in prepositional object positions; there is no support for the claim that clauses themselves are assigned Case, governed by the Case filter.

4. Surprising conjunction facts

As mentioned, the relative paucity of inflectional morphology in Chinese makes it challenging to distinguish categories. Indirect evidence is necessary to help with the task. The choice of conjunction words provides such help (see Aoun & Li 2003, Huang 2006, Li 2008 and Zhang 2009). Pertinent to this work is the fact that the conjunction words *erqie* ‘and’ and *he/gen* ‘and’ are used to conjoin different categories of phrases: *he* and *gen* conjoin nominal phrases and *erqie*, non-nominal constituents such as clauses. The distinction is illustrated below.

- (46) Zhangsan *he/gen*/**erqie* Lisi *dou* *hen* *congming*.
 Zhangsan and Lisi all very smart
 ‘Zhangsan and Lisi are both smart.’
- (47) Zhangsan *hen* *congming* *erqie*/**he*/**gen* Lisi *ye* *hen* *congming*.
 Zhangsan very smart and Lisi also very smart
 ‘Zhangsan is smart and Lisi is also smart.’
- (48) *wo renwei/yiwei/cai* Zhangsan *hen* *congming* *erqie*/**he*/**gen* Lisi
 I think/thought/guess Zhangsan very smart and Lisi
ye *hen* *congming*.
 also very smart
 ‘I think/thought/guess Zhangsan is/was smart and Lisi is/was also smart.’

Importantly, not all clauses require *erqie* as the conjunction word. The nominal conjunction words, *hen* and *gen*, are possible in such contexts as the object of some verbs, the object of P, and the subject of a sentence.¹⁴

¹⁴ *Erqie*, not *he/gen*, seems to require the occurrence *ye* ‘also’ in the second conjunct.

- (49) wo xiang-zhidao Zhangsan zuole shenme he/gen/erqie Lisi (ye)
 I want-know Zhangsan did what and Lisi also
 zuole shenme.
 did what
 ‘I want to know what Zhangsan did and what Lisi (also) did.’
- (50) a. Zhangsan neng-bu-neng lai he/gen Lisi neng-bu-neng lai
 Zhangsan can-not-can come and Lisi can-not-come come
 dou bu shi wenti.¹⁵
 all not be question
 ‘Whether Zhangsan can come and whether Lisi can come are not
 questions.’
- b. Zhangsan de da jiang he/erqie Lisi ye de da jiang shi hen
 Zhangsan get big prize and Lisi also get big prize be very
 zhongyao de.
 important *de*
 ‘That Zhangsan gets a big prize and Lisi also gets a big prize are very
 important.’
- (51) wo dui Zhangsan keyi lai erqie/he/gen Lisi ye keyi lai mei yijian.
 I to Zhangsan can come and Lisi also can come not opinion
 ‘I have no opinion on (the fact) that Zhangsan can come and Lisi can come too.’

¹⁵ If *erqie* conjoins clauses, *dou* is not possible.

- (i) Zhangsan bu lai erqie Lisi ye bu lai (*dou) shi wenti.
 Zhangsan not come and Lisi also not come all be problem
 ‘That Zhangsan can come and Lisi can also come is a problem.’

Erqie conjoins two CPs to make one CP. The entire CP can still be a clause followed by a singular noun:

- (ii) wo dui Zhangsan bu lai erqie Lisi ye bu lai zhe-ge wenti
 I to Zhangsan not come and Lisi also not come this-CL problem
 hen danxin.
 very worried
 ‘I am worried about the problem that Zhangsan cannot come and Lisi cannot come
 either.’

Indeed, a noun phrase expressing plural entities seems impossible:

- (iii) *wo dui Zhangsan bu lai erqie Lisi ye bu lai zhe-liang-ge
 I to Zhangsan not come and Lisi also not come this-two-CL
 wenti hen danxin.
 problem very worried
 ‘I am worried about the two problems that Zhangsan cannot come and Lisi cannot
 come either.’

- (52) wo ba Zhangsan keyi lai he/gen Lisi ye keyi lai dangzuo shi
 I *ba* Zhangsan can come and Lisi also can come regard be
 hen zhongyao de shi.
 very important *de* matter
 ‘I take the fact as important that Zhangsan can come and Lisi can come too.’
- (53) wo bei Zhangsan keyi lai he/gen Lisi ye keyi lai xiadao le.
 I *bei* Zhangsan can come and Lisi also can come shocked
 ‘I was shocked by the fact that Zhangsan can come and Lisi can come too.’
- (54) [cong [[Zhangsan jinlai] he/gen [Lisi jinlai]] dao [xianzai]],
 from Zhangsan enter and Lisi enter to now
 wo dou mei shuo hua.
 I all have-not speak word
 ‘From the time Zhangsan entered and Lisi entered till now, I did not say a word.’

Why is it that the nominal conjunction is possible with (49)-(54) but not (46)-(48)? The translation of (54) provides a clue: it requires the use of nominal expressions like *time*, *moment*. The obligatory use of time expressions in the translation for (54) makes sense because the object for the preposition *cong* ‘from’ and *dao* ‘to’ should not be a proposition. Rather, the relevant objects should express temporal points. (54) is synonymous with the one below, which contains nominal temporal expressions (even though the repetition of *na shihou* ‘that time’ sounds somewhat redundant):

- (55) [cong [[Zhangsan jinlai] (na shihou) he/gen [Lisi jinlai] na shihou]
 from Zhangsan enter that time and Lisi enter that time
 dao [xianzai]], wo dou mei shuo hua.
 to now I all have-not speak word
 ‘From the time Zhangsan entered and (the time) Lisi entered till now, I did not say a word.’

Indeed, the instances possible with *hen/gen* as the conjunction word can all have the addition of a noun phrase: [Clause + Noun Phrase].

- (56) wo xiang-zhidao Zhangsan zuole shenme he/gen Lisi (ye) zuole
 I want-know Zhangsan did what and Lisi also did
 shenme zhe liang-jian shi.
 what these two-CL matter
 ‘I want to know the two questions of what Zhangsan did and what Lisi (also) did.’

- (57) Zhangsan neng-bu-neng lai he/gen Lisi neng-bu-neng lai
Zhangsan can-not-can come and Lisi can-not-can come
zhe liang-ge wenti dou bu zhongyao.¹⁶
these two-CL questions all not important
'The questions of whether Zhangsan can come and whether Lisi can come
are not important.'
- (58) wo dui Zhangsan keyi lai he/gen Lisi ye keyi lai
I to Zhangsan can come and Lisi also can come
zhe liang-jian shi mei yijian.
these two-CL matter not opinion
'I have no opinion on the two matters that Zhangsan can come and Lisi can
also come.'
- (59) wo ba Zhangsan keyi lai he/gen Lisi ye keyi lai zhe liang-jian
I ba Zhangsan can come and Lisi also can come these two-CL
shi dangzuo shi hen zhongyao de shi.
matter regard be very important *de* matter
'I take the two matters as important that Zhangsan can come and Lisi can
also come.'
- (60) wo bei Zhangsan keyi lai he/gen Lisi ye keyi lai zhe
I bei Zhangsan can come and Lisi also can come these
liang-jian shi xiadao le.
two-CL matter shocked
'I was shocked by the two matters that Zhangsan can come and Lisi can
come too.'

In contrast, those clauses not allowing *he/gen* as conjunction words do not accept an accompanying noun phrase:

¹⁶ The two occurrences of *wenti* 'question' in the following example sound redundant:

- (i) Zhangsan neng-bu-neng lai he/gen Lisi neng-bu-neng lai zhexie wenti
Zhangsan can-not-can come and Lisi can-not-come come these questions
dou bu shi wenti.
all not be question
'The questions of whether Zhangsan can come and whether Lisi can come are not
questions.'

- (61) a. wo renwei/yiwei/cai Zhangsan keyi lai erqie Lisi ye keyi lai
 I think/thought/guess Zhangsan can come and Lisi also can come
 (*zhe liang-jian shi).
 the two-CL matter
 ‘I think/thought/guess (*the (two) matter(s)) that Zhangsan can come
 and Lisi can also come.’
- cf. b. wo zhidao Zhangsan keyi lai he/gen Lisi ye keyi lai
 I know Zhangsan can come and Lisi also can come
 (zhe liang-jian shi).
 the two-CL matter
 ‘I know (the two matters) that Zhangsan can come and Lisi can also
 come.’

These contrasts suggest that what appears as a clause in a Case position might actually be a more complex nominal structure containing a clause and a covert noun (phrase), equivalent of the overt expression ‘(the) question’, ‘(the) matter’, ‘the time’, etc. The relation between the clause and the noun (phrase) might be an apposition or a noun complement structure. We will not pursue in this work the precise characterization of the nature of the covert noun (phrase) and its relation with the clause but simply refer to the nominal clausal structure as a complex NP structure. What is important is that recognizing a clause in Case positions as having a more complex structure provides a better understanding of the facts discussed so far in this work and other related phenomena.

5. Complex NP structures

Adopting a complex NP structure for the clauses in Case positions accommodates the facts described so far. First, because what are conjoined are noun phrases, it is expected that the nominal conjunction word *he* or *gen* is used. Two entities are conjoined; therefore, the distributive marker *dou* is possible, as in (50a). Note that the use of *erqie* indicates that the conjunction creates one CP, which can be accompanied by a noun phrase expressing single, not plural entities. The occurrence of *dou* is impossible (see note 15):

- (62) Zhangsan bu-neng lai erqie Lisi ye bu-neng lai (*dou)
 Zhangsan not-can come and Lisi also not-come come all
 bu shi wenti.
 not be question
 ‘Whether Zhangsan can come and whether Lisi can come are not problems.’

cf.

- (63) Zhangsan bu-neng lai he/gen Lisi bu-neng lai dou bu shi wenti.
 Zhangsan not-can come and Lisi not-come come all not be question
 ‘Whether Zhangsan can come and whether Lisi can come are not problems.’

In addition, it is expected that the focus marker *shi*, glossed as ‘be’, does not appear within a complex NP, illustrated below:

- (64) a. [[Akiu (**shi*) yizhi pao-zhe] zhe-jian shi]
 Akiu be continuously run-DUR this-CL matter
 ‘the matter that Akiu is indeed running continuously’
 b. [[Akiu (**shi*) neng-bu-neng lai] zhe wenti]
 Akiu be can-not-can come this question
 ‘the question whether Akiu indeed can come’

The distribution of the focus marker *shi* is sensitive to island conditions, which is expected if the focus marker *shi* needs to be raised to the matrix clause at LF, which cannot cross island boundaries (complex NP constraint in this case — no extraction crossing the boundary of a complex NP, one of the island constraints as in Ross 1967, Chomsky 1981).

The island effects created by the complex NP structure are manifested not only in where the focus marker *shi* can appear but also in the constructions involving overt extraction.¹⁷ Consider the patterns containing a clause assigned Case by *dui* ‘to’, *ba* and *bei*. Extraction of the subject from within such a clause is not possible as in (65), in contrast to (66), which allows the embedded subject to be topicalized:

- (65) a. Lisi, wo [dui [**(ta)* bu hui lai]] hen bu gaoxing.
 Lisi I to he not will come very not happy
 ‘Lisi, I am not happy that (he) will not come.’
 b. Lisi, wo [ba [**(ta)* bu hui lai]] dangzuo shi hen zhongyao de shi.
 Lisi I *ba* he not will come regard be very important *de* matter
 ‘Lisi, I took it as important that (he) was not coming.’
 c. Lisi, wo [bei [**(ta)* yizhi ku-zhe]] fansi le.
 Lisi I *bei* he continuously cry-DUR annoyed
 ‘Lisi, I was annoyed by (his) crying continuously.’

¹⁷ It is not possible to extract from within an appositive clause, either, even though the term ‘complex NP constraint’ generally is not used to accommodate apposition cases.

- (66) Lisi, wo renwei/yiwei/cai (ta) bu hui lai.
 Lisi I think/thought/guess he not will come
 ‘Lisi, I think/thought/guess that (he) would not come.’

In addition, only argument question words can appear inside a clause in the relevant Case positions and be interpreted as having scope outside the clause. Adjunct question words are not possible.

- (67) ta dui shei hui bei qing lai yanjiang hen gaoxing ne?
 he to who will *bei* invite come speak very happy Q
 ‘Who(x) is he happy (x) will speak?’
- (68) a. *ta dui ni weishenme yao yanjiang hen gaoxing ne?
 he to you why will speak very happy Q
 ‘Why(x) is he happy he will speak (x)?’
 cf. b. ta dui ni weishenme yao yanjiang hen haoqi ma?
 he to you why will speak very curious Q
 ‘Is he curious about why you will speak?’
- (69) a. *ta dui ni yao-bu-yao yanjiang hen gaoxing ne?
 he to you will-not-will speak very happy Q
 cf. b. ta dui ni yao-bu-yao yanjiang hen haoqi ma?
 he to you will-not-will speak very curious Q
 ‘Is he curious about whether you will speak?’
- (70) ta ba shei yao yanjiang kande hen yanzhong ne?
 he to who will speak regard very serious Q
 ‘Who(x) does he take seriously (x) will speak?’
- (71) a. *ta ba ni weishenme yao yanjiang kande hen yanzhong ne?
 he to you why will speak regard very serious Q
 ‘Why(x) does he take seriously he will speak (x)?’
 cf. b. ta dui ni weishenme yao yanjiang kande hen yanzhong ma?
 he to you why will speak regard very serious Q
 ‘Does he take seriously about why you will speak?’
- (72) a. *ta ba ni yao-bu-yao yanjiang kande hen yanzhong ne?
 he to you will-not-will speak regard very serious Q
 cf. b. ta ba ni yao-bu-yao yanjiang kande hen yanzhong ma?
 he to you will-not-will speak regard very serious Q
 ‘Does he take seriously whether you will speak?’

These patterns demonstrate that the clauses following *dui/ba/bei* behave like islands, which seems to support the proposal that these clauses are not what they appear to be. The structures are more complicated: there is a covert noun (phrase) — complex NP structures. Unfortunately, resorting to the complex NP constraint is not the only possibility. The unacceptable patterns discussed above involve extraction from a constituent on the left branch of the tree structures. Even though the nature of the left-branch condition is not clear (e.g., see Kennedy & Merchant 2000 for the claim that the left-branch condition is a PF phenomenon and see the variations regarding the relevance of left-branch condition in different types of languages such as Bošković 2005, Corver 1990, 1992, among many others), it is still a possible factor. Therefore, we can only claim that the facts regarding extraction are compatible with a complex NP structure but do not exclusively argue for it.¹⁸ Still, there is some evidence from postverbal clausal complements that supports the more complex structure. Recall that some verbs can assign Case to their clausal complements and the nominal conjunction word *he* or *gen* can conjoin such clausal complements, as in (61b), repeated below:¹⁹

- (61) b. wo zhidao Zhangsan keyi lai he/gen Lisi ye keyi lai
 I know Zhangsan can come and Lisi also can come
 (zhe liang-jian shi).
 the two-CL matter
 ‘I know (the two matters) that Zhangsan can come and Lisi can also come.’

The complex NP analysis would predict that the *shi* focus is not possible in this pattern, nor an adjunct *wh*-question or A-not-A (alternative) question in the clausal complement but with scope outside the clause. These predictions are born out:

¹⁸ Nonetheless, it is relevant to point out that the constituents following *dui*, *ba* and *bei* are all arguments (thematically-marked). Tsai (1995) notes that the *dui* phrase is selected by the verb or adjective. The object of *ba/bei* are both arguments among the layers of structures constituting the predicates of sentences (for the details of possible analyses on *ba/bei*, see Li 2006 on *ba*, Huang, Li & Li 2009 on *ba* and *bei*). Chinese may be considered as a head-final language (Li 1985, 1990). It would be important to investigate the nature of the Left Branch Condition in head-final languages.

¹⁹ The postverbal position can be ambiguous: Case marked or non-Case marked. A verb can optionally assign Case in Chinese, as suggested in Li (1985, 1990). An extraposition option might also be entertained.

- (73) *wo zhidao Zhangsan shi keyi lai he/gen Lisi ye shi keyi lai
 I know Zhangsan be can come and Lisi also be can come
 (zhe liang-jian shi).
 the two-CL matter
 ‘I know (the two matters) that Zhangsan indeed can come and Lisi indeed can also come.’
- (74) a. *ni xiangxin Zhangsan weishemne keyi lai he/gen Lisi
 you believe Zhangsan why can come and Lisi
 ye weishenme keyi lai (zhe liang-jian shi) ne?
 also why can come the two-CL matter Q
 ‘Why(x) you believe Zhangsan can come(x) and Lisi can also come(x)?’
- b. *ni xiangxin Zhangsan ke-bu-keyi lai he/gen Lisi
 you believe Zhangsan may-not-may come and Lisi
 ke-bu-keyi lai (zhe liang-jian shi) ne?
 may-not-may come the two-CL matter Q
 ‘You believe Zhangsan may or may not come and Lisi may or may not come?’

6. Conclusion

Case-marked positions allow NPs and NPs are assigned Case. The question is whether clauses in Case positions are indeed what they are: are they really assigned Case by the relevant Case assigners? We argued that such instances of clauses in Case-marked positions actually are deceptive. The tools available from the study of conjunction and null objects helped to make the discovery. The relevant clauses have more complex structures. They are complex nominals and Case is assigned to the nominal. The clause itself is not assigned Case, which is why clauses are possible in the positions where Case is not assigned and NPs are not possible. This amounts to saying that it is still possible to consider C-selection as derivative of s-selection interacting with Case, as proposed by Pesetsky.²⁰ The Case filter applies to NPs, not to clauses.²¹ To complete the paradigm, we should point out that, just like their English counterparts, non-Case marked

²⁰ Our work focuses on the role played by Case and clauses. However, the general issue of whether c-selection can be completely derived from a very well articulated theory of s-selection (such as the work on lexical semantics by Levin & Rappaport Hovav 1996, 2005 and many others) is still debatable. Further discussion on this issue is beyond the scope of this paper.

²¹ Li (1985, 1990) argues that non-argument NPs such as duration phrases also receive Case. Some languages overtly case mark such NPs, such as Korean.

clauses in Chinese cannot undergo A'-movement, leaving variables in need of Case. Recall that the complement clauses of *happy/afraid* in the English sentences (75a-b) are not assigned Case, as illustrated by the unacceptability of topicalization in (76a-b), in contrast to the possibility of topicalizing the clausal complement when Case is available as in (78a-b):²²

- (75) a. I believe that Mary is happy that Charles is leaving.
 b. I know that Neil is afraid that the computer will break down.
- (76) a. *[That Charles is leaving], I believe that Mary is happy ____.
 b. *[That the computer will break down] I know that Neil is afraid ____.
- (77) a. [That Charles is leaving], I believe that Mary knows ____.
 b. [That the computer will break down] I know that Neil understands ____.

The same contrast is found in Chinese:

- (78) a. ta shuo [tamen renwei/cai ta hui lai].
 he say they think/guess he will come
 ‘He said they thought/guessed that he would come.’
 b. *ta shuo, [ta hui lai, tamen renwei/cai].²³
 he say he will come they think/guess
- (79) a. ta shuo [tamen zhidao/bu xiangxin ta hui lai].
 he say they know/not believe he will come
 ‘He said they knew/did not believe that he would come.’

²² We should make a cautionary note about the possibility of inferences from the discourse when a complement is not present syntactically. For instance, even though the example in (i) seemingly has the reading according to which John’s curiosity is about his not being selected. However, if a reflexive is used, which favors a syntactic position in the complement position, the sentence becomes much less acceptable:

- (i) ni weishenme mei bei xuanshang, wo zhidao ni hen haoqi.
 you why not be.elected I know you very curious
 ‘Why you were not elected, (I) know that you are curious.’
 (ii) *ziji weishenme mei bei xuanshang, wo zhidao ni hen haoqi.
 self why not be.elected I know you very curious
 ‘Why self was not elected, I know you are curious.’

²³ A more complex sentence is created here to avoid the possibility that ‘they/think/guess’ is used as an afterthought remark.

- b. ta shuo, [ta hui lai, tamen zhidao/bu xiangxin].
 he say he will come they know/not believe
 'He said, that he would come, they knew/did not believe.'

These generalizations suggest that among the three options in (30a-c), only (30c) is adequate empirically — Case is relevant to NPs, not clauses, and the occurrence of clauses in Case-marked positions is only apparent.

References

- Aoun, Joseph, and Y.-H. Audrey Li. 2003. *Essays on the Representational and Derivational Nature of Grammar: The Diversity of Wh-Constructions*. Cambridge: MIT Press.
- Bošković, Željko. 1997. *The Syntax of Nonfinite Complementation: An Economy Approach*. Cambridge: MIT Press.
- Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59.1:1-45.
- Chomsky, Noam. 1980. On binding. *Linguistic Inquiry* 11.1:1-46.
- Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford & New York: Oxford University Press.
- Cinque, Guglielmo. 2002. *Functional Structure in DP and IP*. Oxford & New York: Oxford University Press.
- Corver, Norbert. 1990. *The Syntax of Left Branch Extractions*. Tilburg: Tilburg University dissertation.
- Corver, Norbert. 1992. On deriving certain left branch extraction asymmetries: a case study in parametric syntax. *Proceedings of the 22nd Annual Meeting of the North East Linguistic Society (NELS 22)*, 67-84. Amherst: GLSA.
- Hu, Jianhua, Haihua Pan, and Liejiong Xu. 2001. Is there a finite vs. nonfinite distinction in Chinese? *Linguistics* 39.6:1117-1148.
- Huang, C.-T. James. 1982[1998]. *Logical Relations in Chinese and the Theory of Grammar*. MIT dissertation, published by New York: Garland.
- Huang, C.-T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge & New York: Cambridge University Press.

- Huang, Shizhe. 2006. Property theory, adjectives, and modification in Chinese. *Journal of East Asian Linguistics* 15.4:343-369.
- Iljic, Robert. 1994. Quantification in Mandarin Chinese: two markers of plurality. *Linguistics* 32.1:91-116.
- Iljic, Robert. 2001. The origin of the suffix *-men* 們 in Chinese. *Bulletin of the School of Oriental and African Studies* 64.1:74-97.
- Iljic, Robert. 2005. Personal collective in Chinese. *Bulletin of the School of Oriental and African Studies* 68.1:77-102.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.
- Kennedy, Christopher, and Jason Merchant. 2000. Attributive comparative deletion. *Natural Language and Linguistic Theory* 18.1:89-146.
- Koopman, Hilda Judith. 1984. *The Syntax of Verbs*. Dordrecht: Foris.
- Larson, Richard K. 1985. Bare-NP adverbs. *Linguistic Inquiry* 16.4:595-621.
- Larson, Richard K. 2009. Chinese as a reverse Ezafe language. *Yuyanxue Luncong [Essays on Linguistics]* 39:30-85. Beijing: The Commercial Press.
- Levin, Beth, and Malka Rappaport Hovav. 1996. Lexical semantics and syntactic structure. *The Handbook of Contemporary Semantic Theory*, ed. by Shalom Lappin, 487-507. Oxford: Blackwell.
- Levin, Beth, and Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge & New York: Cambridge University Press.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- 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: Kluwer.
- Li, Y.-H. Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8.1:75-99.
- Li, Y.-H. Audrey. 2005. Shenglue yu chengfen queshe [Ellipsis and missing objects]. *Yuyan Kexue [Linguistic Sciences]* 2005.2:3-19.
- Li, Y.-H. Audrey. 2006. Chinese *ba*. *The Blackwell Companion to Syntax*, Vol. 1, ed. by Martin Everaert & Henk van Riemsdijk, 374-468. Malden: Blackwell.
- Li, Y.-H. Audrey. 2007. Kongyulei lilun he Hanyu kongyulei de bianzhi yu zhicheng yanjiu [The theory of empty categories and the identification of empty categories in Chinese]. *Yuyan Kexue [Linguistic Sciences]* 2007.2:37-47.
- Li, Y.-H. Audrey. 2008. Case, 20 years later. *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, Vol. 1, 41-68. Columbus: The Ohio State University.

- Lin, Jo-wang. 2003a. Selectional restrictions of tenses and temporal reference of Chinese bare sentences. *Lingua* 113.3:271-302.
- Lin, Jo-wang. 2003b. Temporal reference in Mandarin Chinese. *Journal of East Asian Linguistics* 12.3:259-311.
- Lin, Jo-wang. 2006. Time in a language without tense: the case of Chinese. *Journal of Semantics* 23.1:1-53.
- Lin, T.-H. Jonah. 2001. *Light Verb Syntax and the Theory of Phrase Structure*. Irvine: University of California dissertation.
- Lobeck, Anne C. 1995. *Ellipsis: Functional Heads, Licensing, and Identification*. Oxford & New York: Oxford University Press.
- Merchant, Jason. 2001. *The Syntax of Silence*. Oxford: Oxford University Press.
- Paul, Waltraud. 2002. Sentence-internal topics in Mandarin Chinese: the case of object preposing. *Language and Linguistics* 3.4:695-714.
- Pesetsky, David. 1982. *Paths and Categories*. Cambridge: MIT dissertation.
- Plann, Susan. 1986. On case-marking clauses in Spanish: evidence against the Case Resistance Principle. *Linguistic Inquiry* 17.2:336-345.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. *Elements of Grammar: Handbook in Generative Syntax*, ed. by Liliane M. V. Haegeman, 281-337. Dordrecht: Kluwer.
- Rizzi, Luigi. 2004. *The Structure of CP and IP*. Oxford & New York: Oxford University Press.
- Ross, John R. 1967. *Constraints on Variables in Syntax*. Cambridge: MIT dissertation.
- Stowell, Timothy Angus. 1981. *Origins of Phrase Structure*. Cambridge: MIT dissertation.
- Sybesma, Rint. 2007. Whether we tense-agree overtly or not. *Linguistic Inquiry* 38.3: 580-587.
- Tang, Chih-Chen Jane. 1990. *Chinese Phrase Structure and the Extended X'-Theory*. Ithaca: Cornell University dissertation.
- Tang, Ting-chi. 2000. Finite and nonfinite clauses in Chinese. *Language and Linguistics* 1.1:191-214.
- Tsai, Wei-Tien Dylan. 1995. Visibility, complement selection and the case requirement of CP. *Journal of East Asian Linguistics* 4.4:281-312.
- Tsai, Wei-Tien Dylan. 2008. Tense anchoring in Chinese. *Lingua* 118.5:675-686.
- Tsang, Chui-Lim. 1981. *A Semantic Study of Modal Auxiliary Verbs in Chinese*. Stanford: Stanford University dissertation.
- Zhang, Niina Ning. 2009. *Coordination in Syntax*. Cambridge & New York: Cambridge University Press.

漢語句法的雙層式分析法： 功能與形式的會聚

曹逢甫

國立清華大學

主語與主題角色的認定一直是漢語語法學界爭論不休的問題。本文主張以雙層式分析法區別大句和小句兩個層次，主語與主題的功用才能清楚地顯現。文中論證漢語基本的篇章單位是主題鏈，主題鏈等同於英語的句子，由一個主題與一句或數句與此主題相關的簡句組成。主題可以藉由主題鏈輕易辨識出來，而主語則必須是小句中的 DP/NP，與句子的謂語構成某種選擇關係。

釐清主語與主題的角色後，本文接著檢視漢語主題鏈中的子句結構來看漢語的句式。在第五節，我們討論如何界定漢語書面語的表層句以及主題鏈與 FSP 的關係。最後並將我們對主題鏈的定義與中國詠物或詠地的古詩做連結，說明這些古詩其實就是以物或地為主題的超大主題鏈。

關鍵詞：主語，主題，主題鏈，大句層次，小句層次

1. 引言

筆者在二十幾年前曾經給英語系大學四年級的學生做過四段中英文標句號的小實驗，結果出人意料地發現，這一班的學生雖然都以中文為母語、英文為外語，但他們在中文方面的表現離原作者甚遠，而且他們之間彼此的差異也很大。為了方便引述我把這四段文章以及學生表現的統計表呈現於下：

- (1) 曾經是歷史上最光輝的拳王阿里 近年來勝利後 總是說要退休 但總未退休 結果敗在初出道的史賓柯斯手下 本來可以光榮退休的 卻不道落成這樣一個下場
- (2) 而你不會相信那些傳說的 正像我也很難相信它一樣 因為我們全是有知識 有理性 有獨立思考能力的現代人 我們聽取了這一類的傳說立

即會指陳它的荒謬 它的無稽來 但任何荒謬無稽的傳言都曾被一代一代裡更多的人們相信過 不然它不會像風一樣的衍傳 更不會傳到我們的耳中了

- (3) At Dallas airport nearly 5000 people were waiting. The President in a dark-blue suit stepped from his plane smiling happily. He and Jackie were met by a committee that gave her a bouquet of red roses. Their car was ready to leave but Kennedy had to shake hands with some voters. Jackie, her roses cradled in her left arm, also touched the outstretched hands. After a few minutes she started to walk away, but noticing that her husband was still at it, smiled fondly and returned.
- (4) It seldom rains. The geography books credit this portion of Utah nine to ten inches of precipitation. Actual rainfall and snowfall vary widely from year to year. There are a few perennial springs hidden in secret places known only to the deer and coyotes, to myself and a few friends, but the water does not flow far before vanishing into the air and under the ground. Even the rain when it comes does not always fall to the ground, but can be seen evaporating halfway down — curtain of blue rain dangling out of reach, torture by tantalizing.
- (5) 學生在句逗號使用測試的表現

	中文		英文	
	(1)	(2)	(3)	(4)
原文句數	1	2	6	5
學生句數的平均數	2.53	3.82	5.49	4.94
範圍	1-4	2-5	4-6	3-6
與作者句號使用完全符合者人數	1	1	8	6

對於這個讓人困惑的問題，本人一直認為，問題的關鍵在於漢語的句子與詞組（或稱短語）的界限並不很清楚，而二者不容易區分的原因又是因為漢語的主語沒有英語明顯，而主語不明顯除了與漢語本身缺乏語法標誌有關外，還有個更根本的原因，那就是漢語的主語與主題經常混雜在一起，時而合而為一，時而分別出現。如就兩者之間的語義關係而言，則變化更多；它們時而密切，時而疏遠，有時更看不出有任何關係。因此，二者角色的認定一直是漢語語法學界爭論不休的課題。這一點，我們從漢語學界幾位代表作者對漢語主語和主題的討論就可看出。

2. 前人有關主語與主題的分析

我們曾經仔細檢視下列諸家對於如何界定漢語主語與主題的討論，發現王力先生 (1956) 雖然在主語的界定上還算有相當的說服力，¹ 但因為他完全忽視主題的問題，因此像 (6) 這樣的句子，其中 C₂ 小句中的「肚子」在句中所扮演的角色，王先生就無法交代清楚：

- (6) 小明 一整天沒吃飯，肚子好餓。
 C₁ C₂

趙元任先生 (1968) 承襲了早期美國結構學派的句法分析法，把所有的組成成分都再細分為兩個直接組成成分，因此，漢語跟英語一樣，S 底下分為主語與謂語。但稍微做過英漢句子比較研究的人都知道，漢語的句子與英語的句子是很不同的。趙先生當然也很清楚，因此，他接著說漢語句法上的主語與謂語在語義上應解釋為「主題」與「評述」。趙先生所做的調整雖然能解釋像 (7) 這樣的句子，但卻硬生生地把 (8) 中的「三隻大黃狗」認定為賓語：

- (7) 前面有三個人。
 (8) 前面來了三隻大黃狗。

另外，趙先生在處理所謂「雙主格」句子如 (9) 的三個小句子時，在句法上把除了第一個名詞組之外的成分分析為謂語，這在句法分析上雖然不是絕無僅有，但也是相當罕見的。

- (9) 那個人 頭腦簡單，話都說不清楚，什麼事都辦不成。
 C₁ C₂ C₃

在語義上，他認為整個句子的第一個名詞組與第二個名詞組之間有某種語義關係，但這種語義關係有時緊密，有時鬆散：像「那個人」與 (9C₁) 和 (9C₂) 的兩個名詞組「頭腦」、「話」都是領屬關係，這種關係就叫緊密；但在 (9C₃) 中兩者之間就說不出有什麼明確關係，因此，我們可以說在趙先生的分析框架裡，它們之間的共現關係既不能用語法也不能用語義來加以限定。

¹ 王力 (1956) 是這樣界定「主語」的：主語是句子的組成部分，它通常是由名詞、代名詞或具有名詞用途的詞（有時加上附加語）來表現的；它指稱事物，謂詞所指稱的行為（包括主動、被動）、性質或屬性是屬於這一事物的。

在 (10C₁) 與 (10C₂) 中，第一個名詞組與第二個名詞組的關係固然可以理解成不可分割的領屬關係，但 (10C₃) 的「房子」是身外物，明顯屬於可分割物，但卻同在以「張大姐」為第一個名詞組的句子中出現，因此，湯先生的限制肯定是過嚴了。

3. 本文的分析法

(一) 主語與主題的確在句中有句法、語義與指稱上的不同特徵，而且我們可以根據這些特徵對二者加以區辨。(11) 跟 (12) 就是根據筆者 (2004) 的研究所提出的主語與主題的區辨特徵。

- 主語總不帶介詞；
- 就位置而言，主語是動詞左邊的第一個有生名詞組；或者就是緊靠動詞前的名詞組；
- 主語總與句子的主要動詞構成某種選擇關係；
- 主語傾向於是特指的 (specific)；

- e. 主語在下面的同指名詞組代名詞化或刪除過程中有重要作用：真反身化、祈使化以及同等 NP 刪除。

(12) 主題的區辨特徵

- a. 主題總佔據主題鏈第一個子句的 S 首位的位置；
- b. 主題可以由停頓助詞「啊」、「呀」、「呢」、「麼」、「吧」與句子的其餘部分隔開；
- c. 主題總是有定的；
- d. 主題是一個語篇概念 (discourse notion)，它可以而且經常將其語意管轄範圍擴展到多個子句；
- e. 主題在主題鏈中控制所有的代名詞化或同指名詞組刪除；
- f. 主題，除非同時也是主語，否則不參與真反身化、同等 NP 刪除以及祈使化過程。

(二) 爲了避免傳統的局限於小句的討論並落實 (11c) 與 (12a)、(12d) 的論點，本文認爲，在從事漢語語法分析時我們應該分兩個層次來進行，也就是分別從大句 (= 主題鏈) 及小句的層次來探討，因爲只有分這兩個層次來探討，主題與主語的功用及扮演的角色才能清楚地顯現。這基本上就是本文所主張的雙層分析法。

以下我們就選幾個句子來顯示如何進行雙層式的分析：

- (13) 這個英文句子 我不懂，_____ 拿去請教三叔公，
 Topic C₁ C₂
 _____ 他也不懂。
 C₃

- (14) 這個英文句子，真難。
 Topic & Subject C₁

根據雙層式分析法，(13) 總共有三個小句，但因爲三個小句的主題都是「這個英文句子」，因此它就可以被抽離出來成爲大句的共同主題，有了共同主題及三個評述子句，這整個大句就形成了一個主題鏈，但其中的三個小句卻分別有其主語：第一個小句爲「我」，第二小句也是「我」，但這裡承前省略。第三小句則是「他」。再來看第 (14) 句。根據我們的分析，「這個英文句子」是主題兼主語。有人看到這裡或許要問：既然第 (14) 句只有一個小句，那我們還要認定它爲「主

題」並稱它為「主題鏈」嗎？答案是肯定的。這樣做一方面是爲了理論的普遍性，也就是在「主題鏈」這個集合裡我們可以設定它可以有 1...N 個相同的主題；另一方面，我們得指出它的確含有主題的特質。這一點只要我們把 (13) 和 (14) 合併就可以看得很清楚：

- (15) 這個英文句子 真難， 我不懂， 拿去請教三叔公，
 Topic C₁ C₂ C₃
 他也不懂。
 C₄

(15) 的主題鏈是以「這個英文句子」爲共同主題的，但它在每個小句中所扮演的語法角色卻不一樣。在第一小句 (C₁) 中它也是主語，第二小句 (C₂) 及第三小句 (C₃) 中它分別是動詞「懂」與「拿」的賓語，而主語都是「我」。在第四小句 (C₄) 中，它也是賓語，而主語是「他」。

下面我們另外列了兩個例子。各位不妨利用我們剛才所談的雙層分析法看看能不能對它們的主題與主語以及句子的大框架進行分析。²

- (16) 這棵樹，花小，葉子大，很難看，我沒買。
 (17) 台北市天文台台長蔡章獻，他有個計畫，他要把滿天星斗搬到天文台的屋裡，讓大家看得更容易更仔細，對天象更瞭解。這個計畫，如果經費有了著落，很快就可以實現了。

我們在閱讀古代典籍時，正如預期地，也發現了許許多多長主題鏈的例子。由於篇幅所限，先以《史記·李將軍列傳》中一個很有名的句子爲例：

² 我們的分析如下：

- (16) 這棵樹，花小，葉子大，很難看，我沒買。
 Topic C₁ C₂ C₃ C₄
 (17) 台北市天文台台長蔡章獻，他有個計畫，他要把滿天星斗搬到天文台的屋裡，
 Topic₁ C₁ C₂
 讓大家看得更容易更仔細，對天象更瞭解。這個計畫，如果經費有了著落，
 C₃ Topic₂ C₁
 很快就可以實現了。
 C₂

- (18) 廣 出獵，見草中石，以為虎而射之，中石，沒鏃，視之石也。
 Topic C₁ C₂ C₃ C₄ C₅ C₆
因復更射之，終不能復入石矣。
 C₇ C₈

由 (18) 的例子，我們可以清楚地看出主題鏈的妙用，如果運用得宜，它可以無限延伸，用來陳述一個事件或一則小故事。從這個觀點來看，特別值得一提的是 C₄ 和 C₅ 中的動詞「中」與「沒」，如果照一般的用法，這兩個動詞的主語應該是「箭」和「箭尖」，但如果以「箭」或「箭尖」為主語，整個主題鏈就會因此而中斷，因此為了不讓主題鏈中斷，「中」和「沒」都發展了使動的用法，也就是「中」和「沒」都應理解為「致使」動詞，即「廣」射中石頭，和使箭尖沒入石中的意思。

(19) 是《史記》中摘出的另一個例子，(20) 來自《春秋左傳》，而 (21)-(23) 則來自唐宋古文。

- (19) 景公召穰苴，與語兵事，大說之，以為將軍。（《史記·司馬穰苴列傳》）
 (20) 晉侯作二軍，公將上軍，太子申生將下軍，趙夙御戎，畢萬為右，以滅耿、滅霍、滅魏。還，為太子城曲沃，賜趙夙耿，賜畢萬魏，以為大夫。（《春秋左傳·閔公·傳元年》）
 (21) 吾年未四十，而（吾）視茫茫，而（吾）髮蒼蒼，而（吾）齒牙動搖；（吾）念諸父與諸兄，皆康強而早世，（吾）（念）如吾之衰者，其能久存乎？（韓愈〈祭十二郎文〉）
 (22) 臣伏見天后時，有同州下邳人徐元慶者，父爽，為縣尉趙師韞所殺，卒能手刃父讎，束身歸罪。（柳宗元〈駁復讎議〉）
 (23) 予獨愛蓮之出淤泥而不染，濯清漣而不妖，中通外直，不蔓不枝，香遠益清，亭亭淨植，可遠觀而不可褻玩焉。（周敦頤〈愛蓮說〉）

此外，我們又分別在說漢語的學生學英語的中介語以及洋涇浜英語中找到了不少疑似主題鏈的句子，現各舉一例：

- (24) Taiwan is island, lack of resources and capital, has an amount of labours.
 (Chinese-English interlanguage)³

³ (24) 是筆者某年在台灣舉辦的留學考試作文卷中發現的例子。

- (25) Taipan can sen wun piece coolly come my shop looksee.
'Your Honor can send a porter to come to my shop to have a look?' (*Canton English*, 1857)

4. 從主題鏈看漢語的句式

如果各位可以同意漢語的句子基本上是個主題鏈的話，現在我們可以回頭來看看在整個主題鏈的大框架底下各個小句是以何種方式呈現並且組織起來的。我們發現漢語最基本的句式是動詞前只有一個名詞組的簡句。這個名詞組既作主語又作主題，如 (26) 所示：

- (26) a. 她很好看。
b. 他來了。
c. 他打了球。
d. 他叫張三。

簡句當然還可以根據動詞的不同分成不同的句型。由於篇幅的關係，這裡不擬進一步說明，有興趣的讀者請參見曹逢甫 (1990, 2004)。稍微複雜一點的結構是動詞前面有兩個或三個名詞組。我們根據它們的先後順序亦即它們各自轄域 (scope) 的大小，分別稱之為大主題、次主題、三主題等。句子除了主題數的多寡不同，還可能因為所含小句句式的不同而又有變化。以有兩個主題的句子而言，就有雙主格句 (27a)、賓語提前句 (27b)、把字句 (27c)、動詞重複句 (27d)、副詞作次主題句 (27e) 以及介詞組作主題句 (27f)。

- (27) a. 他耳朵很靈。
b. 他數學習題做完了。
c. 他把錢看得很重。
d. 他打籃球打得很好。
e. 他昨天覺得很舒服。
f. 他為了你吃了很多苦。

這些句子中的次主題在適當的主題下都可變為大主題，如 (28) 的對應句所示：

- (28) a. 耳朵他很靈。
 b. 數學習題他做完了。
 c. 錢他把它看得很重。
 d. 打籃球他打得很好。
 e. 昨天，他覺得很舒服。
 f. 為了你，他吃了很多苦。

介紹完簡句我們要進入複句的討論。複句根據主題的不同可以分成三種：(一)普通主題鏈，(二)引介主題鏈，(三)兼語主題鏈。在此，我們只能做簡單的說明並各舉一二例。

(一)普通主題鏈：普通主題串是由一個或一個以上的簡句組成，並由一個共有的主題引領的句式；再根據中間是否有聯結成分又可分為有標主題鏈和無標主題鏈：

- (29) a. 他雖然很用功，(他)卻沒有考及格。(有標主題鏈)
 b. 銀行九點鐘開門，(銀行)開始營業。(無標主題鏈)

(二)引介主題鏈：「引介主題鏈」主題為未知；因其為未知，故需要一個前行分句加以引介。最常見的引介詞是存在動詞，如「有」、「是」以及文言文的「無」、隱現動詞如「出現」、「消失」和姿態動詞如「站著」、「掛著」、「飄來」等。

- (30) a. 他有一個妹妹，(這個妹妹)是中文系畢業的，(這個妹妹)卻很喜歡洋歌。
 b. 牆上掛著一幅畫，(這一幅畫)是齊白石的，(這一幅畫)很值錢。

(三)兼語主題鏈：這種句式可以同時「兼顧」引介的主題和被引介的主題。

- (31) a. 他們種那種菜給張三吃。
 b. 他們種那種菜_____吃。
 c. 他們種那種菜給自己吃。

最後，因為主題鏈等於句子(S)，因此，一定也可以以名詞子句的方式出現在更大的句子中，(32)就是很好的例子：

- (32) a. 星期六銀行要到十點鐘才開門，開始營業是很平常的。
b. 我希望你明天早點回家，幫我燒幾道菜。

主題鏈就像英語的 S 或 CP 一樣當然也能出現在漢語的關係子句（也稱形容詞子句）和副詞子句，如 (33) 和 (34) 所示：

- (33) 那本[s 印刷精美，內容卻平淡無奇的]s 小說，銷路果然奇差無比。
(34) 為了[s 那家銀行能在三月一日開門開始營業]s 她把所有的田產全賣了。

但有一點需要特別注意的：漢語的關係子句出現的位置是在中心語名詞的左方，因此通常都不會太長，因為太長了會增加短程記憶很大的負擔，也因為這個緣故，主題鏈出現在關係子句的情形就更少了，這一點和英語的情形是正好相反的。

5. 餘論

在這一節裡我們將繼續討論與雙層式分析法有關的三個議題：首先，我們要回來討論在一個單獨小句沒有大句有關確定主題的資訊時如何來進行分析的問題。其次我們稍稍回應我們一開始提出的有關句號和逗號的使用問題並提出二個具體的建議。最後一小節擬延伸觸角進入主題選擇與段落連貫的問題。

5.1 主題與主語在獨立小句的分析問題

許多受過傳統語言學由小單位擴大到大單位分析的語言學者看過了前面的分析以後也許會問到說：漢語固然有不少大大小小的主題鏈，但漢語也有不少單獨出現的獨立小句，那麼在那種情形下主語和主題又如何決定？這是一個值得深入討論的問題，目前，老實說，我們並沒有很好的答案，但至少我們可以提供幾個思考的方向。先舉個實際的例子來看：

- (35) a. 風浪海上那一天很大。
b. 海上那一天風浪很大。
c. 那一天海上風浪很大。

根據前頭所提之處理原則，謂語「很大」與「風浪」有語意選擇關係，因此可以確定後者為主語。而「海上」與「那一天」一為地點一為時間之修飾語，而且在 (35a) 它們也分別出現在時間副詞與方位副詞經常出現的位置，其為狀語成分似乎也沒有問題。因此 (35a) 句應有如下的句法分析也是可以確定的：

主語＋方位狀語＋時間狀語＋謂語

這是在小句層次的分析。至於大句的層次我們前頭也講過：在沒有別的競爭者的情形之下，主語也是當然的主題，所以就大句的層次分析「風浪」是主題，就小句層次而言「風浪」是主語。

再來看 (35b)。在小句層次，「風浪」自然是主語，而「海上」和「那一天」則分別是狀語。在大句層次則情況與 (35a) 有了變化，「海上」與「那一天」分別是因為主題的關係被提到主語外面，「海上」在前「那一天」在後，所以「海上」為大主題而「那一天」為次主題。(35c) 的情形基本上跟 (35b) 相似，只是在該句中「那一天」的轄域大於「海上」，因此「那一天」為大主題，「海上」為次主題。⁴

有人或許要問，有沒有其他的證據來支持我們的分析？在筆者的書中（曹逢甫 1990, 2004）曾提到主題的功用有二：即鏈接 (chaining) 功能與對照 (contrastive) 功能，分別由 (36) 與 (37) 所示：

- (36) 張大偉_i 個子很高， _i 手掌很大， _i 速度也不慢， _i 合適打籃球。
- (37) 錢永祥_i 個子_k很高，左逢源_j _k卻很矮。

運用這個原理，只要我們把 (35b) 與 (35c) 中的大主題與次主題延伸成主題鏈，並且檢視一下它們有沒有上述兩項功能，就可以看出它們是不是主題了。

- (35) b'. 海上_i 那一天_j 風浪很大， _i 第二天_k 卻風平浪靜。
- c'. 那一天_i 海上_j 風浪很大，昨天_k _j 卻風平浪靜。

在 (35b') 中大主題「海上」有鏈接功能，次主題「那一天」與「昨天」有對照功

⁴ 如果採用 GB 理論結構樹來表示，那麼大主題 (Tp1) 與次主題 (Tp2) 就可以用 Chomsky-adjunction 的方式移到 CP 的上面，大主題在最上面，次主題其次依序排下來。

能。在 (35c') 中大主題「那一天」與「昨天」有對照功能而次主題「海上」有鏈接功能。由此可見 (35b') 與 (35c') 中前後兩小句主語前的兩個成分都是主題。

5.2 漢語中句號與逗號使用的問題

我們先前的小測試清楚地顯示出說漢語的學生在句號的使用上會有許多問題。其實這種情形先前對教授漢語有經驗的教師已多有報導（尹銘 1978，陳紀滢 1978，楊遠 1983），楊遠 (1983) 更給它取個名字叫「一逗到底」，意思是學生在寫文章時常一直使用逗號直到一段文字的末端非使用句號時才使用句號。可惜這些作者都只注意到這個現象並未能提出任何具體可行的改進建議。

在經過先前的討論之後我們可以有信心地提出一個大原則就是：在每個大句的後頭加句號，在每個小句的後頭加逗號，除非該小句是個獨立小句。分別如 (38)、(39) 與 (40) 所示：

- (38) 這棵樹，花小，葉子大，很難看。⁵（帶 3 個小句的大句）
 Topic C1 C2 C3
- (39) 這個中文句子，真難，我看了三遍，也看不懂。（帶 3 個小句的大句）
 Topic C1 C2 C3
- (40) 天氣終於放晴了。（獨立小句）

這一條大原則在語言理論上來講是絕對可以站得住腳的，但在實際運用時可能需要加以限制，這是因為任何一個大句在理論上講可能包含的小句數目是沒有上限的。就以先前所舉的 (38) 與 (39) 來說，雖然它們都只有 3 個小句，但如果要表達的意思有需要的話，我們還可以在第 3 小句之後再加其他的小句如 (38') 與 (39') 所示：

- (38') 這棵樹，花小，葉子大，很難看，我沒買。
 Topic C1 C2 C3 C4
- (39') 這個中文句子，真難，我看了三遍，也看不懂，拿去請教老師，
 Topic C1 C2 C3 C4
 他也看不懂。
 C5

⁵ 主題，尤其是長一點的主題，其後也常有逗號。

這種帶有許多小句的大句在文言文文中是常有的，如先前所引的 (18) 句所示：

- (18) 廣 出獵，見草中石，以為虎而射之，中石，沒鏃，視之石也。
 Topic C₁ C₂ C₃ C₄ C₅ C₆
 因復更射之，終不能復入石矣。
 C₇ C₈

這在文言文裡是常有的。因為文言文多單音節詞而且句式較簡省，因此即使多到八個小句讀起來還不會太冗長，也不容易造成理解上的困難。但如果用現代白話來說，可能句子會變成兩、三倍長，這就容易造成理解上的困難，因此可能要把 (18) 句拆成三、四句為佳，如 (18') 所示：

- (18') 李廣，出去打獵，看見草中的大石，以為是老虎，就拿起弓箭來射。
 他不但打中了石頭還連箭尖都沒入石頭中。他走向前去檢視才發現原來射中的是一塊大石。他倒退到原來的地點重新舉箭來射，但再也不能射入石頭中了。

回到先前的討論，我們發現在現代漢語裡要寫出像 (18) 句一樣帶有八個小句的長句肯定會帶來理解上的困難，這是因為一個閱讀者在理解一個句子時會先把已處理部分的資訊儲存在短程記憶 (short-term memory) (參閱 Clark & Clark 1977: Ch. 4, Shulman 1970, 1972)。句子如果太長了，就可能超出了記憶的容量而導致理解上的困難。可惜在漢語句子理解過程的了解目前還太少，因此我們也沒辦法準確的掌握到底句子要多長才不會把我們的短程記憶擠爆。目前只能先抓一個數目，研究者的直覺是四個小句，有了這個約數我們可以把它運用到先前測驗的第 (1) 句，如 (1') 所示：

- (1') 曾經是歷史上最光輝的拳王阿里，近年來勝利後，總是說要退休，但總未退休，結果敗在初出道的史賓柯斯手下。他本來可以光榮退休的，卻不道落成這樣一個下場。⁶
 Topic C₁ C₂
 C₃ C₄ Topic C₁
 C₂

⁶ 當我們把原來 (1) 的六個小句斷成二句時，在第二句主題的地方最好有個與第一句同指的代名詞。

這樣子的斷句方式比起原先的方式感覺上好多了，同時也更接近受測十八個學生的平均數 2.53。但這條規則顯然還有缺陷，這一點和漢語中連詞的使用有關。漢語中的子句連詞往往是可有可無的（見王力 1955，曹逢甫 1979, 1990），當它不出現時，就不會影響我們原先所設定的以四小句為最大量的斷句原則，如 (41a) 所示。

- (41) a. 李小虎，真不錯，速度快，又肯下功夫練習，一定能取得好成績。
 Topic C₁ C₂ C₃ C₄

但如果加入一個像「如果」這樣的小句連詞，那麼因為語意連貫的需要就只能在第三句斷句或第五句斷句，分別如下列 (41b)、(41c) 二句所示。如果在第四句後斷句就會如 (41d) 句所示有語意不連貫的問題。

- b. 李小虎，真不錯，速度快，又肯下功夫練習。如果他能有個好教
 Topic C₁ C₂ C₃ C₄
練，一定能取得好成績。
 C₅

- c. 李小虎，真不錯，速度快，又肯下功夫練習，如果他能有個好教
 Topic C₁ C₂ C₃ C₄
練，一定能取得好成績。
 C₅

- ??d. 李小虎，真不錯，速度快，又肯下功夫練習，如果他能有個好
 Topic C₁ C₂ C₃ C₄
教練。一定能取得好成績。
 C₅

如果把上面的因素納入考量，那麼我們的第二條規則就得稍做調整成為：漢語的主題鏈應以 4±1 為理想的斷句單位。

5.3 主題鏈與 FSP

Functional Sentence Perspective (FSP) 是功能語法學派的學者在談到句子語法結構與篇章連貫性的關係時常常引用的原則（詳見 Daneš 1974，Firbas 1974，

Kuno 1978)。Kopple (1982) 爲了要用實驗的方法證實 FSP 在英文的篇章連貫性起著一定的作用，故意把 (42a) 的英語段落改寫成 (42b)，並要學生閱讀後表示他們對這兩段文章閱讀容易度與篇章連貫性的看法。受測的學生明顯地顯示 (42a) 在篇章連貫性與閱讀容易度上的優勢。

- (42) a. Currently *the Marathon* is the best waxless ski for recreational cross-country skiing. *Its weight* is a mere two pounds. Yet *its two-inch width* allows the skier to break a trail through even the heaviest snow. *Its most unique characteristic* is the fishscale design for its bottom. *The Marathon* is almost as effective as most waxable skis. In fact, *it* is even better than some waxable skis when the snow is very wet. *The Marathon* can be used with most conventional bindings. However, *it* works best with the Suomi double-lock. Finally, *the Marathon* is available in six different colors.
- b. Currently *the best waxless ski for recreational cross-country skiing* is the Marathon. *A mere two pounds* is its weight. Yet *the skier* can break a trail through even the heaviest snow with its two-inch width. *The fishscale design for its bottom* is its most unique characteristic. *Most waxable skis* are only slightly more effective than the Marathon. In fact, *some waxable skis* are not as good as it when the snow is very wet. *Most conventional bindings* can be used with the Marathon. However, *the Suomi double-lock* works best with it. Finally, *six different colors* are available for the Marathon.

因爲英語相較於漢語而言，是一個主語明顯的語言，現在只要我們比較一下 (42a) 跟 (42b)，我們很容易就可以看出 (42a) 中所有句子的主語不是 *the Marathon* 就是與 *the Marathon* 同指的 *it*，或是與它有關的屬性如 *its width*（它的寬度），*its characteristic*（它的特性），因此要把這些句子用一個主題鏈把它們串起來是一件很容易的事。但在 (42b) 的段落裡，就缺乏這些特性，因此想用主題鏈的方式把它們全都連結在一起幾乎是不可能的事。各位在念完這兩段翻譯後一定也有相同的感覺。

- (43) a. 目前馬拉松是娛樂越野滑雪的最佳無蠟雪橇，重量僅兩磅，但是它兩英吋的寬度讓滑雪者即使在最大的雪中都能滑行。它最大的特色在於底部的魚鱗設計。馬拉松幾乎跟可上蠟的雪橇一樣好用。事實

上，當下溼雪時，它甚至於比可上蠟雪橇還好用。馬拉松可以和大部分的傳統綁帶合用但它和所米雙重鎖同用效果最佳。最後，馬拉松共有六種顏色可供挑選。

- b. 目前娛樂越野滑雪的最好的無蠟雪橇是馬拉松。兩磅是它的重量，但是滑雪者穿上了這種寬僅兩吋的雪橇卻可以在最大的雪中滑行。底部魚鱗式的設計是它最獨特的特色。大部分可上蠟的雪橇都只比它稍微好用。事實上，有些上蠟雪橇在溼雪中還不如它好用。大部分的傳統綁帶都能跟馬拉松合用，但是所米雙重鎖與它合用效果最佳。最後，還有六種顏色可供馬拉松使用者挑選。

其實跟 (42a) 與 (43b) 有同樣篇章安排的，在漢語的分類中還有許多的詠某物或詠某地的詩，今舉劉長卿〈登餘干古縣城〉為例：

(44) 孤城上與白雲齊
萬古荒涼楚水西
官舍已空秋草沒
女牆猶在夜烏啼
平沙渺渺迷人遠
落日亭亭向客低
飛鳥不知陵谷變
朝來暮去弋陽溪

這一首詩共有八行，其中第一、二行與七、八行是二行形成一個主題鏈，其主題分別是孤城，即餘干城，與飛鳥，即餘干城的飛鳥。其餘的三、四、五、六行則各自形成一個主題鏈並分別以官舍、女牆、平沙、落日為主題，而它們在語意上當然都和餘干城有部分與全體的關係，也都有以部分代表全體的作用。從這個角度來看這一首詩，我們幾乎可以說它和 (43a) 一樣都是一個超大主題鏈。又從 (42a) 與 (43a) 的比較，我們很清楚地可以看出英語與漢語在篇章組織方面其實有共同的原則，只是或許它在英文中的使用沒有漢語頻繁，因此沒有語法化成一種句法構式。這一不同也正反映了漢英在類型學上的一個重要不同：漢語比英語要來得更加篇章取向 (discourse-oriented)。

6. 結論

從先前的討論，我們可以得到下面四個結論：

(一) 主語和主題在漢語的語法分析中是可以區分的，而且是有必要的區分。主語主要的作用在於小句，而主題的作用在於大句，只要我們認清它們個別的範域與句法功能，它們中間的區分是明顯易見的，這也是我們為什麼一再強調漢語的語法分析需要分成大句與小句兩個層次。至於獨立小句的分析則因為大句與小句在此合而為一，所以主語與主題的區分需要別的輔助標準。

(二) 我們對大句和小句所做的區分，對於漢語書面語句逗號的使用有相當大的幫助。我們可以簡單地立個規則，大句的後頭加句號，大句的小句與小句之中以逗號分開。只是因應大句中子句的數目可以無限延伸的特性，我們提議漢語的大句的長度以 4 ± 1 的小句為最合適。

(三) 從漢英語言比較的角度而言，英文的主題與主語是趨向於合而為一的，又因為英語有比較顯著的構詞學上的標記，所以主語的認定問題比較小。反觀漢語則因為主語和主題在許多情形下是分開的，又缺乏明顯的標記，因為主題與主語的認定問題較多。但從另一個角度來觀察，漢語因為主題較為明顯，因此在檢驗一段文章是否語意連貫方面較為容易，只要檢驗段內各大句的主題是否語意連貫就行了。而英語的主語因為身兼二職，即有些主語是言談語意的需要而產生的，有些是文法的需要而產生的，因此要檢驗一個段落是否連貫就得先排除那些因文法需要而產生的主語，如 (45) 句中之 'it'。

(45) It is believed that college students generally work harder than high school students.

(四) 最後我們必須一提的是，有許多學者雖然認為漢語的分析一定得帶進主題和主語的概念，但卻認為主語才是句法的概念，而主題是篇章語用的概念，因此不宜在句法學討論。這種說法似是而非，主題的概念無疑地是來自篇章語用的考量，但從語法化的角度來看，這個篇章語用的概念也會因為經常被使用而發展成為語法概念，所以主題是否應納入語法的討論端視它在語言的語法中所扮演的角色而定，是不能先驗地加以排除的。

引用文獻

- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chen, Chi-ying (陳紀滢). 1978. Duo duanju, shao jingtan 多斷句，少驚歎 [More punctuation, less exclamation]. *Central Daily* Sept. 18, page 10.
- Clark, Herbert H., and Eve V. Clark. 1977. *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Daneš, František. 1974. Functional sentence perspective and the organization of the text. *Papers on Functional Sentence Perspective*, ed. by František Daneš, 106-128. The Hague: Mouton.
- Firbas, Jan. 1974. Some aspects of the Czechoslovak approach to problems of functional sentence perspective. *Papers on Functional Sentence Perspective*, ed. by František Daneš, 11-37. The Hague: Mouton.
- Kopple, William J. Vande. 1982. Functional sentence perspective, composition, and reading. *College Composition and Communication* 33.1:50-63.
- Kuno, Susumu. 1972. Functional sentence perspective: a case study from Japanese and English. *Linguistic Inquiry* 3.3:269-320.
- Kuno, Susumu. 1978. Generative discourse analysis in America. *Current Trends in TextLinguistics*, ed. by Wolfgang U. Dressler, 275-294. Berlin & New York: Walter de Gruyter.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Shaw, Wilkinson J. 1897. Canton English. *New Review* 16:548-555.
- Shulman, Harvey G. 1970. Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior* 9.5:499-508.
- Shulman, Harvey G. 1972. Semantic confusion errors in short-term memory. *Journal of Verbal Learning and Verbal Behavior* 11.2:221-227.
- Tang, Ting-chi (湯廷池). 1972. *A Case Grammar of Spoken Chinese*. Taipei: Haiguo Book Company.
- Tang, Ting-chi (湯廷池). 1988. *Hanyu Cifa Jufa Lunji 漢語詞法句法論集* [Studies on Chinese Morphology and Syntax]. Taipei: Student Book.
- Tsao, Feng-fu (曹逢甫). 1979. *A Functional Study of Topic in Chinese: The First-Step toward Discourse Analysis*. Taipei: Student Book.
- Tsao, Feng-fu (曹逢甫). 1990. *Sentence and Clause Structure in Chinese: A Functional Perspective*. Taipei: Student Book.

- Tsao, Feng-fu (曹逢甫). 2004. *Hanyu de Juzi yu Ziju Jiegou* 漢語的句子與子句結構 [*Sentence and Clause Structure in Chinese*], translated by Jing Wang. Beijing: Beijing Language and Culture University Press.
- Wang, Li (王力). 1955. *Zhongguo Yufa Lilun* 中國語法理論 [*Principles of Chinese Grammar*]. Beijing: Zhonghua Book Company.
- Wang, Li (王力). 1956. Zhuyu de dingyi ji qi zai Hanyu zhong de yingyong 主語的定義及其在漢語中的應用 [Definition of subject and its application in Chinese]. *Yuwen Xuexi* 語文學習 [*Chinese Language Learning*] 1956.1:21-25. Beijing: People's Education Press.
- Yang, Yuan (楊遠). (ed.) 1982. *Biaodian Fuhao Yanjiu* 標點符號研究 [*Studies on Punctuation*]. Hong Kong: Tianjian Press.
- Yin, Ming (尹銘). 1978. Biaodian fuhao yu Guoyuwen jiaoxue 標點符號與國語文教學 [Punctuation marks and Chinese teaching]. *Central Daily* Oct. 8, page 10.

A Double-tier Approach to Chinese Sentential Structure: Convergence of Form and Function

Feng-fu Tsao

National Tsing Hua University

The problem of whether both subject and topic should be posited in Chinese grammar has long been debated. This paper is an attempt to solve the problem by proposing a two-tiered model consisting of a sentential tier and a clausal tier. It is argued and demonstrated that the basic discourse unit in Chinese, equivalent to a surface sentence in English, is a topic-chain, a stretch of discourse comprising from one to many clauses and headed by a common topic, which is what all the clauses are about. It is also argued that topic can be identified as soon as a topic chain is established. Subject, on the other hand, should always be determined within a clause as a DP/NP that bears some selectional relation with the predicate.

With subject and topic thus established, we then go into the clausal patterns that Chinese sentence has by examining the clausal structures that can occur in a topic chain. In the fifth section, I take up the problems of how to determine a surface sentence in Chinese written discourse and the relationship of topic chain and Functional Sentence Perspective. The last part of the section is devoted to the demonstration of the relation between our definition of a topic chain and a special type of Chinese classical poetry called locale poetry whose form is in effect a super topic chain with the locale as its topic.

Key words: subject, topic, topic chain, sentential tier, clausal tier

Verbal Reduplication and Verbal Classifiers in Chinese*

Marie-Claude Paris (白梅麗)

*Université Paris Diderot – UMR 7110
CNRS, Laboratoire de linguistique Formelle*

In this paper, after presenting previous analyses of verbal classifiers in Chinese — henceforth abbreviated Clv — we describe their syntactic and semantic characteristics. We compare the sequence V Q Clv to two reduplicated verb forms: (i) VV and (ii) V-*yi*-V. We show that verb reduplication cannot be equated with V Q Clv and that Clv are not DPs, but MPs. We posit a distinction between weak and strong Clv, according to their quantificational and aspectual properties. In conclusion a parallelism between a Chinese V Q Clv sequence and a light verb followed by a deverbal noun in English is established.

Key words: (a)telicity, [+control] predicates, (delimitative) aspect, light verb + deverbal noun, verbal reduplication, verbal classifiers

1. Introduction

Though the lexical category ‘verbal classifier’¹ is not commonly found in general

* I am very happy to dedicate this paper to my dear friend and colleague Beiluobei (貝羅貝). We struck up a friendship 40 years ago. I hope this ‘bond’ will last another 40 years. I would never have been able to carry out research on this topic and then, finally, get down to writing it without the help of many colleagues and institutions both at home and, most of all, abroad.

In chronological order, I first thank Thekla Wiebusch, who included me in her European Commission-COST project, N° A31 “Stability and Adaptation of Classification Systems in a Cross Cultural Perspective”. Thekla has given me precious diachronic information on verbal classifiers, which, for lack of time, I have not been able to include here. I presented a first version of this paper in the panel she organized in Cracow, during the 10th International Cognitive Linguistics Conference (July 2007). Following this, I had to temporarily discontinue my work on this topic, but thanks to two institutions in Taiwan, I was able to resume my research in 2009. During a three-week stay as a visiting scholar at the National Taiwan Normal University (March) and a two-week stay as a guest of the Taiwanese Ministry of Education at the National Sun Yat-sen University (October), I was able to check data and hypotheses with colleagues. I would like to warmly thank Li Duanduan (University of British Columbia/NTNU), Tsai Mei-chih and Shyu Shu-ing (National Sun Yat-sen University) for sharing their time and intuition

linguistics books, this word class is nevertheless of utmost importance when describing East Asian languages in general, and Chinese in particular.² In this paper, I would like to give a quantificational/aspectual definition of what a verbal classifier stands for in Chinese. After presenting previous work on this topic, I will compare verbal classifier constructions with other constructions, and in particular with two forms of verbal reduplication: VV and V-*yi*-V. I will show that there are, at least, two main classes of verbal classifiers,³ which I label ‘weak verbal classifiers’ and ‘strong verbal classifiers’. Their use depends on the aspectual properties of the predicate they are in construction with. I extend my results to English light verb constructions where the light verb is followed by a deverbal noun.

2. Previous works on the topic

As far as contemporary⁴ Chinese is concerned, the most detailed presentations of verbal classifiers I know of are from Chao (1968) for a semantic classification and Paris (1981) for a syntactic description.

2.1 Chao (1968)

Whereas almost no Chinese grammar book lists the term ‘verbal classifier’ in their indexes,⁵ Chao’s grammar (1968) contains the most complete list of such items and

with me. Thanks to travel grants from Paris 7 – Denis Diderot University and from the EACL (European Association of Chinese Linguistics), I presented two other versions of this paper at NACCL-21 (Bryant University, USA, June 2009) and at EACL-VI (Poznan, Poland, August 2009).

¹ Though Aikhenvald (2000) included verbal classifiers in her work, her description does not apply very readily to the Chinese language. In this paper I will abbreviate the expression ‘verbal classifier’ as Clv. The other abbreviations used are: Cln: Nominal Classifier; Q: Quantifier; and, Sfx: Suffix.

² For Chinese verbal classifiers, see Zhang (2002); for various East and Southeast Asian languages, see Gerner (2007, 2009); and, for Vietnamese, see Nguyen (1995:89-106).

³ In Arabic and in Chinese, Fassi-Fehri & Vinet (2008) posit a four-way distinction to account for the classification of both nouns and verbs, based on their mereological properties. “... achievements are equated with individuals, accomplishments with groups, activities with kinds and states with masses” (*op. cit.*: 56).

⁴ Some diachronic work has been done on Mandarin Clv, see, for instance, Hung (2000).

⁵ The category Clv does not appear in the indexes of the following reference grammars: Cheung et al. (1994), Huang, Li & Li (2009), and Ross & Ma (2006). I have found this label in Abbiati (1998) and Xu (1996) only.

labels them ‘measures for verbs’, abbreviated as *Mv*. In this paragraph I will use Chao’s abbreviation, i.e. *Mv*, but I will use *Clv* in the remainder of the paper. Chao’s semantic definition reads as follows: “a measure for verbs of action expresses the number of times an action takes place.” (*ibid.*: 615). From this definition, one gathers that an *Mv* is equivalent to a frequency marker. Chao lists 40 *Mvs*, which he distributes into three main (semantic) categories. These *Mvs* are also labelled ‘cognate objects’ (*ibid.*: 312, 615-620). These classes express:

1. the action of the verb (18 *Mvs*),
2. the body part which performs the action (7 *Mvs*) and
3. the instrument with which the action is performed (15 *Mvs*).

The first category is the biggest one. It includes the following items (among others): the verb itself, as in *kan yi kan* ‘to take a look’ or as in *kankan* ‘to have a look’. Other frequent *Mvs* are: *hui* ‘time’; *ci* ‘time’; *bian* ‘time’; *tang* ‘trip/time’; *xia* ‘stroke/time’; *fan* ‘a series of/time’, as in *jiaoxun yi fan* ‘to give a good scold’; *quan* ‘a round’ as in *rao yi quan* ‘to go for a round’; *jiao* ‘a sleep’ as in *shui ji jiao* ‘sleep on several occasions’; and *guan* ‘barrier’ as in *guo wu guan* ‘go through five passes’.

The second category comprises seven items. *Shou* ‘hand’, *jiao* ‘foot’, *bazhang* ‘palm’ as in *da-le ta liang bazhang* ‘slapped him twice’; *quan*₁ ‘(a single encounter of the) fist’ as in *ta zou-le wo yi quan* ‘he gave me a blow’ and *quan*₂ ‘(a hit of the) fist’; and *kou* ‘mouth’, *yan* ‘eye’.⁶

The most common *Mvs* which belong to the third category are: *dao* ‘knife’; *jianzi* ‘scissors’; *futou* ‘hatchet’ as in *kan futou* ‘chop slash’; *bianzi* ‘whip’, as in *da bianzi* ‘to whip’; *zhen* ‘needle’ as in *zha zhen* ‘to stick a(n acupuncture) needle’; *jian* ‘arrow’, as in *she jian* ‘to shoot an arrow’; and *qiang* ‘pistol’, as in *fang qiang* ‘to shoot’.

It is somewhat difficult to understand what Chao’s underlying classification criterion or criteria are.

First, many examples given in all three classes do not correspond to the semantic definition proposed. *Guan* ‘barrier’ in category one, *quan* ‘fist’ in category two or *zhen* ‘needle’ in category three, for example, do not correspond to the function and the meaning of a frequency marker, that is ‘time’.

⁶ For this category, it is interesting to note that Shao (2007) shows how in contemporary spoken Northern Mandarin the use of the verbal classifier *yi ba* is expanding, losing its original meaning of marking an action involving a hand, as in *la yi ba* ‘to give a hand’, ‘to help’. Thus, for example, according to Shao (*ibid.*: 16, 19), *yi ba* is found in collocations such as *zai nuli yi ba* ‘be more diligent’ or in *shuang yi ba* ‘to have some fun’. Some speakers of Taiwan Mandarin do not accept the latter examples, but would accept *shuang yi shuang* or *shuang yi xia* ‘to have some fun’.

Second, as an Mv is not defined syntactically, one gets the impression that whereas examples like *hui* ‘time’, *ci* ‘time’, *bian* ‘time’, *tang* ‘time’, and *xia* ‘time’, in category one are genuine classifiers, other terms such as, for instance, *bazhang* ‘palm’, *zhen* ‘needle’ or *qiang* ‘pistol’ can also be nouns. In the sequence *yi*⁷ + *xia* (or *bian*, *tang*,...), *xia* (or *bian*, *tang*,...) is a genuine classifier: it cannot be preceded by other classifiers nor can it function as an argument. On the contrary, *zhen* ‘needle’ can be preceded by a classifier and can be an argument. Compare *xia* ‘time’ and *zhen* ‘suture’/‘needle’ in the three pairs (1)-(6). Examples (1)-(2) look alike on the surface, but (3)-(4) and (5)-(6) show that *zhen* ‘needle’, contrary to *xia* ‘time’, evidences nominal properties. Whereas *yi gen zhen* ‘a needle’ is the object of the verb *huan* ‘change’ in (6), *yi xia* ‘a bit’/‘a while’ is not the object *feng* ‘to sew’ in (5).

- | | |
|--|--|
| (1) <i>yi xia</i>
one time
‘once’/‘one time’ | (2) <i>yi zhen</i>
one stitch/needle
‘a stitch’/‘a suture’ |
| (3) * <i>yi ge xia</i>
one Cln time | (4) <i>yi gen zhen</i>
one Cln needle
‘a needle’ |
| (5) <i>feng yi xia</i>
sew one time
‘to sew a bit’/‘a short while’ | (6) <i>huan yi gen zhen</i>
change one Cln needle
‘to change needle’ |

In sum, Chao not only acknowledges the existence of a lexical category in Chinese, which he labels Mv, but he also presents a useful list of examples. The drawback of his categorization is that it mingles items having different syntactic properties — ‘real’ classifiers and classifiers which are also nouns. Moreover, his definition of Mvs, which is basically semantic, is too narrow. Mvs may indeed indicate frequency, but their semantic role is much wider than that. In the next section I will present some syntactic characteristics of Clvs.

2.2 Paris (1981)

In Paris (1981)⁸ I studied the different distributions of nominal and verbal classifiers in postverbal positions. I used nine tests⁹ to show that while the sequence Q Cln N does

⁷ Here *yi* ‘one’ stands for the whole paradigm of numerals *er* ‘two’, *san* ‘three’, etc.

⁸ The following is only a brief summary of the results I obtained in 1981.

⁹ The nine tests include demonstrative and adjectival insertion, passive *bei* NP and active *ba* NP

form a constituent — a nominal one — the sequence Q Clv N does not. Q(uantifier) subsumes the list of all possible numerals.

Thus, though (7) and (8) share the same surface structure, they are constructed very differently: the postverbal constituents, namely *yi ben shu* and *yi xia shu* respectively, do not allow the same bracketing. Compare (9) and (10). In (9) the sequence *yi ben shu* is a nominal constituent, in (10) *yi xia shu* is not a nominal constituent.

(7) Ta kan-le yi ben shu.
he read-Sfx one Cln book
'He read a book.'

(8) Ta kan-le yi xia shu.
he read-Sfx one Clv book
'He did a bit of reading.'

(9) Ta kan-le_{NP}[(*yi ben*) shu]_{NP}

(10) Ta kan-le *NP[(*yi xia*) (shu)]*NP

In terms of surface constituents, (10) should be analyzed as (11).

(11) [ta]_{Pred P}[_{VP}[_V[[kan-le]_V]_{Q Clv}[*yi xia*]_{Q Clv}]_{VP NP}[shu]_{NP}]_{Pred P}.

The sequence (*yi xia*) (*shu*) is composed of two autonomous sub-constituents. *Yi xia* is part of the VP and *shu* of the PredP. The relationship between V and NP is equivalent to that between a verb and its object.

Four more tests can be added to the nine constituency tests I used previously to ascertain that Q Cln N and Q Clv N do not share the same structure. These tests are:

- (i) question formation by means of *duoshao* 'how much'/'how many?',
- (ii) movement of the quantifier phrase (QP) past the object noun,
- (iii) modification by the approximative marker *zuoyou* 'around' and
- (iv) sentence coordination.

constructions, interrogation by means of *shenme*, topicalization and apposition, the possibility of being the head of a relative clause and the possibility of belonging to a pseudo-cleft sentence. I do not share Sybesma's analysis that "... frequentative expressions form one constituent with the Object NP" (Sybesma 1999:129).

2.2.1 Question formation by means of *duoshao* ‘how much’, ‘how many’?

A [Q Cln] constituent can stand as an answer to a question containing the interrogative (nominal) quantifier *duoshao* ‘how much?’/‘how many?’, as (12) and (13) show. But the [Q Clv] sequence *yi yan* ‘a glance’ in (15) cannot constitute an answer to *duoshao* in (14).

- (12) Ta mai-le *duoshao*?
 he buy-Sfx how many
 ‘How many [books] did he buy?’
- (13) Ta mai-le *yi ben*.
 he buy-Sfx one Cln
 ‘He bought one [book].’
- (14) *Ta kan-le *duoshao*?
 he look-Sfx how.much
- (15) Ta kan-le *yi yan*.
 he look-Sfx one eye/Clv
 ‘He cast a glance.’

2.2.2 The position of [Q Clv] with respect to the postverbal noun

While a nominal QP cannot occupy both the pre- and post-nominal object positions, a verbal QP can. From (16), (17) cannot be formed, but from (18), (19) can.

- (16) Ta xiang mai *yi ben* shu.
 he want buy one Cln book
 ‘He would like to buy a book.’
- (17) *Ta xiang mai shu *yi ben*.
 he want buy book one Cln
- (18) Ta xiang hui *yi tang* jia.
 he want return one Clv home
 ‘He would like to go home once.’
- (19) Ta xiang hui jia *yi tang*.
 he want return home one Clv
 ‘He would like to go home once.’

2.2.3 Modification by means of *zuoyou* ‘about’

While a nominal QP can be followed by *zuoyou* ‘around’ within an NP, as shown in (20), a verbal QP cannot, as seen in (21).¹⁰

- (20) Ta kan-le san shi *chang* *zuoyou* de Jingju.
 he look-Sfx three ten Cln approximately *de* Peking opera
 ‘He saw around thirty performances of Peking opera.’

- (21) *Ta qu-le san shi *ci* *zuoyou* de Beijing.
 he go-Sfx three ten Clv approximately *de* Peking

More crucial to the idea that Q Clv forms a constituent with the verb that precedes it (and not with the noun that follows it) is the test of sentence coordination.

2.2.4 Sentence coordination

Example (22) below is not acceptable because the second Q Clv has been elided in sentence-final position. The fact that *yi xia* ‘a bit’ is compulsory after the verb *tian* ‘to lick’, as can be seen in (23), shows that *yi xia* is part of the VP and that it carries relevant information. If it did not, it would be deleted. (23) is ambiguous between a strict and a sloppy reading of the anaphoric possessive pronoun *tade* ‘his’.

- (22) *Zhangsan tian-le yi xia ta de bingbang, Lisi ye tian-le Ø.
 Zhangsan lick-Sfx one Clv he *de* ice bar Lisi also lick-Sfx Ø
- (23) Zhangsan tian-le yi xia ta de bingbang, Lisi ye tian-le
 Zhangsan lick-Sfx one Clv he *de* ice bar Lisi also lick-Sfx
yi xia.
 one Clv
 a. ‘Zhangsan licked his ice bar a bit and Lisi also licked Zhangsan’s ice bar.’
 b. ‘Zhangsan licked his ice bar a bit and Lisi also licked his (own) ice bar a bit.’

In sum, while in the “V [Q Cln N]” sequence, Q Cln and the following N entertain a relationship of dependence and of quantification, in the “V [Q Clv N]” sequence, it is Q Clv and the preceding V which entertain a relationship of dependence and of quantification. Clv and N do not entertain a relationship of dependence. By ‘dependence’

¹⁰ See Paris & Vinet (2010). The intended meaning of (21) is: ‘He went to Peking approximately 30 times’.

I mean selectional restrictions. Seen in this light, nominal and verbal classifiers share important common points. The bracketing of the constituents as in (11) above:

$$\text{Pred P}[_{\text{VP}}[\text{V-le}] \text{ Q Clv}[_{\text{Q Clv}}[\text{yi Clv}] \text{ Q Clv}] \text{ VP NP}[_{\text{NP}}[\text{N}]] \text{ Pred P}$$

is meant to account for the dependency between V and its following Q Clv phrase.

In this section, I have presented Chao's classification of 'measures for verbs', as markers of frequency. Then I have proposed the results of my own syntactic analysis of the sequence Q Clv N. I have laid emphasis on the syntactic/semantic dependency between the Q Clv constituent and its preceding V.

In the following sections, I will compare constructions containing a reduplicated verb with constructions containing a verb followed by Q Clv. In Chao's and Li & Thompson's books, verbal reduplication patterns such as VV and V-yi-V are said to be (semantically) equivalent to a V Q Clv construction. I will try to evaluate in what respect this proposition is well grounded.

3. The comparison between the reduplication patterns VV and V-yi-V, on the one hand and V yi Clv, on the other

Verb reduplication in Mandarin can take (at least) two forms: either the repetition of the verb, or the repetition of the verb preceded by the quantifier *yi* 'one', giving the following two surface patterns: VV and V-yi-V. Such forms are labelled a 'delimitative aspect'.¹¹ Both Chao (1968) and Li & Thompson (1981) draw parallels between these two types of verb reduplications, equating the former with the latter.¹² Moreover, both forms of verb reduplication are also equated to the sequence V Q Clv, i.e. verb followed by a quantifier + a verbal classifier. In other words, the three patterns VV, V-yi-V and V Q Clv are considered to be equivalent. For Li & Thompson (*ibid.*: 233), "the *yi* plus the reduplicated syllable functions like a quantity adverbial", as in *san ci* 'three times'.

If it is certainly true that these three forms do share some aspectual and pragmatic properties, I will show, however, that none of them is strictly equivalent to the other two. I will first present the similarities between these three patterns and then stress their dissimilarities.

¹¹ This terminology is commonly used to describe this aspect not only in Sinitic languages. For uses referring to Mandarin and Cantonese, see — among others — Xiao & McEnery (2004) and Matthews & Yip (1994), respectively. For references regarding this terminology as used in Slavic linguistics, see Filip (2000).

¹² "This reduplication may optionally involve the morpheme *yi* between the verb and the reduplicated syllable", Li & Thompson (1981:232).

3.1 VV, V-*yi*-V and V *yi* Clv: identity of behaviour

In this section I will first study the similar co-occurrence restrictions among the three verbal forms VV, V-*yi*-V and V Q Clv regarding:

- (i) durative markers — such as the adverb *zhengzai* and the final particle *ne*, and
- (ii) the degree adverbial *shaowei* ‘somewhat’.

Then I will look at the degree to which the behaviour of these three patterns is or is not identical.

3.1.1 Similarities: aspectual and pragmatic properties

The durative adverb *zhengzai* ‘precisely’ cannot occur with VV, with V-*yi*-V nor with V *yi* Clv, as (24)-(26) illustrate.

- (24) *Ta *zhengzai* *kan-kan* shu.
 he precisely.at look-look book
- (25) *Ta *zhengzai* *kan-yi-kan* shu.
 he precisely.at look-one-look book
- (26) *Ta *zhengzai* *kan* *yi* *xia* shu.
 he precisely at look one while book

The final particle *ne*, which, among other things, indicates continuity, is also incompatible with VV, with V-*yi*-V and with V *yi* Clv, as (27)-(29) illustrate.

- (27) *Tamen *ting-ting* yinyue *ne*.
 they listen-listen music *ne*
- (28) *Tamen *ting-yi-ting* yinyue *ne*.
 they listen-one-listen music *ne*
- (29) *Tamen *ting* *yi* *xia* yinyue *ne*.
 they listen one while music *ne*

In an imperative sentence, the degree adverbial *shao(wei)* ‘a little’, which attenuates the illocutionary force of the order, is compatible with the three forms VV, V-*yi*-V and V Q Clv, as shown in (30)-(32).

- (30) Ni shaowei dengdeng!
 you a bit wait-wait
 ‘Please wait a bit.’
- (31) Ni shaowei deng-yi-deng!
 you a bit wait-one-wait
 ‘Please wait a bit!’
- (32) Ni shaowei deng yi xia!
 you a bit wait one while
 ‘Please wait a bit.’

But note that all Clvs are *not* interchangeable. Thus, while the Clv *xia* ‘a while’ is acceptable in (32), it cannot be replaced by the Clv *ci* ‘time’, contrary to Li & Thompson’s claim. Thus, example (33), where *ci* ‘time’ is a substitute for *xia* ‘time’ in (32), is not well formed.

- (33) *Ni shaowei deng yi ci!
 you a bit wait one time

In the following I will present some more dissimilarities between the three forms VV, V-*yi*-V and V *yi* Clv.

3.1.2 VV, V-*yi*-V and V *yi* Clv: non-identity of behaviour

What (32) and (33) above show is that *xia* cannot be used in the same context as *ci*. Hence the general definition which Chao provides for Clvs does not hold (see §2.1 above). If it were true that “a measure for verbs of action expresses the number of times an action takes place”, then both (32) and (33) should be equally acceptable. From the pair of examples (32) and (33), we can posit the existence of two main types of verbal classifiers:

- (i) those such as *ci* ‘time’ which *count* the number of occurrences an event takes place, and
- (ii) those such as *hui* ‘a while’, or *zhenzi* ‘a moment’, which *measure* a (short) interval during which an event takes place.

The verbal classifier *xia* ‘time’/‘a while’ is ambiguous: it can either mark the frequency of an event, and thus be substituted by *ci* ‘time’ or it can be substituted by *hui*

‘a while’, when it indicates an interval. Example (34) below is ambiguous: it has two readings (34a) and (34b). In (34a), the predication is punctual/telic, hence *xia* indicates a frequency; in (34b), the predication is durative, hence *xia* indicates an interval. Thus there exists a *mutual dependency* between the *aspectual value* of the *predicate* and the Q Clv it is in construction with.

- (34) Ta tui-le wo yi xia.
 he push-Sfx I one time/while
 a. ‘He pushed me once.’
 b. ‘He pushed me a while.’

Now that I have established two (semantic) classes of Clvs — the quantitative and the measure Clvs — and their relationships with the predicates that precede them, I will use well formed V Q Clv sequences and compare them to VV and V-*yi*-V constructions. I will use both syntactic and morphological arguments to oppose the two groups.

From the grammatical Q *yi* Clv examples (35) and (37), one cannot derive a corresponding VV form, nor a V-*yi*-V form, as shown in (36) and (38), respectively.

- | | |
|---|--|
| <p>(35) Ta bing-le yi zhenzi.
 he be.sick-Sfx one while
 ‘He was sick for a while.’</p> | <p>(36) *Ta bing-bing
 he be.sick-be.sick</p> |
| <p>(37) Tui teng-guo yi ci.
 leg hurt-Sfx one Clv
 ‘[My/his...] leg hurt once.’</p> | <p>(38) *Tui teng-yi-teng
 leg hurt-one-hurt</p> |

Examples (35) and (37) contain [–control] predicates *bing* ‘to be sick’ and *teng* ‘to hurt’. As is well known, such types of predicates can hardly be repeated in Mandarin.¹³ This is why (36) and (38) are ungrammatical. Hence the set of verbs which allow reduplication — be it in the form of the VV or the V-*yi*-V pattern — is smaller than the set of verbs which allows a Q Clv. V Q Clv do not co-occur with markers of ‘total’ duration as *zhengzai* or *ne* — see (26) and (29) — but they do with markers indicating ‘partial’ or short duration, see (34b).

Secondly, the V Q Clv sequence is (morphologically) different from both the VV and the V-*yi*-V patterns in that Q Clv is a free form. On the contrary, the second V in VV and the *yi*-V in V-*yi*-V are not autonomous: they are bound forms. Whereas (V) (Q Clv) is composed of two constituents, both (VV) and (V-*yi*-V) constitute only one

¹³ See, among others, Paris (2007a).

single constituent: they are X° . Examples (40) and (42), which are derived from (39) and (41), are ill formed; the direct object *Xiao Li* has moved between the two syllables of *song-song* in (40) and between *deng* and *yi deng* in (42).

- (39) Ni qu song-song Xiao Li.
 you go send.off-send.off Xiao Li
 'Please, send Xiao Li off.'
- (40) *Ni qu song Xiao Li -song
 you go send.off Xiao Li -send.off
- (41) Ni deng-yi-deng Xiao Li!
 you wait-one-wait Xiao Li
 'Wait for Little Li a bit!'
- (42) *Ni deng Xiao Li -yi-deng
 you wait Xiao Li -one-wait

Contrary to the pairs (39)-(40) and (41)-(42), (43)-(44) is well formed. Both the surface orders (Q Clv) (Object) and (Object) (Q Clv) are permitted.

- (43) Ni deng yi xia Zhangsan!
 you wait one while Zhangsan
 'Wait for Zhangsan for a while!'
- (44) Ni deng Zhangsan yi xia!
 you wait Zhangsan one while
 'Wait for Zhangsan for a while!'

Thirdly, in the so-called 'tentative construction', i.e. a construction where a reduplicated verb is followed by the verb *kan* 'to look', the VV and the V-yi-V forms are not equivalent at all. While the VV pattern is acceptable, as in (45), V-yi-V is not, as in (46).

- (45) Ni pao-pao kan!
 you run-run look
 'Let's see how you run!'
- (46) *Ni pao-yi-pao kan!
 you run-one-run look

Fourthly, the dissimilarity between the V-yi-V pattern, as in (49) on the one hand, and the V+Q Clv sequence, as in (47) and (48) on the other hand, manifests itself clearly in the *lian...ye/dou* (negative) constructions. Both *yi ci* in (47) and *yi xia* in (48) indicate a

minimum quantity — be it one time [+count] or one interval [–count]/[+measure] — that the event of reading requires to take place. This is why such quantifying expressions can enter the (negative) *lian...ye/dou* ‘even’ construction, as in (50) and (51). But as *yi kan* in (52) renders the (negative) *lian...ye/dou* construction ungrammatical, one can only conclude that [V] *yi* Clv and [V] *yi* V do not carry the same semantic function. If they did, (52) should be acceptable, but it is not.

- (47) Kan *yi ci*!
look one time
‘to look [at something] once’
- (48) Kan *yi xia*!
look one moment
‘to look [at something] a bit’
- (49) Kan *yi kan*!
look one look
‘to cast a glance at’
- (50) Ta *lian yi ci dou bu xiang kan*.
he even one time all Neg want look
a. ‘He does not even want to look once.’
b. ‘He does not even want to look at all.’
- (51) Ta *lian yi xia dou bu xiang kan*.
he even one time/while all Neg want look
a. ‘He does not even want to look once.’
b. ‘He does not even want to look a bit.’
- (52) *Ta *lian yi kan dou bu xiang kan*
he even one look all Neg want look

Finally, note in passing, that if VV and V-*yi*-V share many syntactic and semantic characteristics, they are not identical phonologically. In a VV form, the second form is toneless, but in a V-*yi*-V form, the second V keeps its tone.

In sum, in this section, by means of five tests, I have shown that the two verbal reduplicated patterns — VV and V-*yi*-V — do not behave alike syntactically, semantically or phonologically. I have opposed the reduplicated forms, V-*yi*-V, to the V Q Clv pattern. The (syntactic) distribution of V Q Clv is wider than that of V-*yi*-V, because it is less constrained semantically.

V Q Clv is acceptable with [–control] predicates, but VV and V-*yi*-V are not.

In conclusion, VV cannot be equated to V-*yi*-V, which itself cannot be equated to V Q Clv, contrary to what has very often been claimed in the literature.

In the following, I will dwell a little more on the opposition between two classes of Clv, which I mentioned above in §3.1.2.

4. The two classes of Clvs: weak and strong

In this section, I will try to justify why it is necessary to posit that there are (at least) two classes of verbal classifiers in Mandarin. Then I will very briefly present some constructions in English and in French that function similarly to Chinese Clvs.

4.1 Weak and strong Clvs in Chinese

The syntactic properties of Q Clv that I have described above all indicate that Q Clv are not DPs:¹⁴ they are measure phrases (MPs) — just like *yi sui* ‘one year’ or *yi mi* ‘one meter’ in (53) and (54).

- (53) Ta lao-le *yi* *sui*.
 he be.old-Sfx one year
 ‘He aged one year.’

- (54) Ta pao-le *yi* *mi* lu.
 he run-Sfx one meter road
 ‘He ran one meter.’

Q Clv expressions are not referential. They do not introduce a discourse referent, nor can they be the head of a relative clause. They are incorporated¹⁵ into the verb. Two types of Clvs can be distinguished as follows: those which allow Q to be the whole range of numbers (*yi* ‘one’, *liang* ‘two’, *san* ‘three’, etc.), and those which only allow *yi* ‘a’, a marker of indefiniteness. I will call the first ones “strong Q Clvs” and the second ones “weak Q Clvs”.

The distinction between these two classes is grounded on syntactic evidence. Strong Q Clvs can have scope, as seen is the (negative) *lian...ye/dou* construction above in

¹⁴ Note that whereas some Chinese linguists, like Ding (1961) label them as *zhun binyu* ‘quasi objects’, some others, like Liu et al. (1983) label them as *shuliang buyu* ‘quantificational complements’. For me, verbal classifiers are predicational, hence they are not objects (*binyu*), but complements (*buyu*).

¹⁵ I use the term ‘incorporated’ in the sense of Farkas & De Swart (2003).

examples (50) and (51). They indicate the number of times an event takes place. In a word, they are markers of frequency. As opposed to strong Q Clvs, “weak Q Clvs” do not have a broad scope, they are scopally inert and simply indicate an occurrence of an event; see examples (61) and (64)–(65).

Examples (55) and (56) contain strong Clvs. *Yi* ‘one’ can be substituted by other numbers and with the indefinite *ji* ‘some’ (‘1 to 9’), see (57) and (58). Such strong Clvs can enter a (negative) *lian...ye/dou* ‘even’ construction, see (59) and (60).

- (55) Ta kan-le ni yi yan.
he look-Sfx you one eye
‘He cast a glance at you.’
- (56) Ta qu-le yi tang.
he go-Sfx one Clv
‘He went once.’
- (57) Ta kan-le ni ji yan.
he look-Sfx you several eye
‘He cast a glance at you several times.’
- (58) Ta qu-le ji tang.
he go-Sfx several Clv
‘He went several times.’
- (59) Ta lian yi yan ye mei kan ni.
He even one eye also Neg look you
a. ‘He did not even cast a glance at you.’
b. ‘He did not look at you at all.’
- (60) Ta lian yi tang ye mei qu.
he even one Clv also Neg go
a. ‘He did not even go once.’
b. ‘He did not go at all.’

The weak Q Clv *yi tiao* ‘a startle’, contained in (61) below, allows neither the substitution of *yi* with *san* ‘three’, for example — as shown in (62) — nor the movement of *yi tiao* in preverbal position in a *lian...ye/dou* ‘even’ construction, as in (63). Example (63) shows that *yi tiao* has a narrow scope.

- (61) Ta xia-le wo yi tiao.
he frighten-Sfx I one Clv
‘He frightened me.’

(62) *Ta xia-le wo san tiao.
he frighten-Sfx I three Clv

(63) *Ta lian yi tiao dou mei xia wo.
he even one Clv all Neg frighten I

Other examples of weak Clvs can be found in (64) and (65). In this pair of examples, *yi* ‘a’ cannot be substituted by *liang* ‘two’.

(64) Ta ma-le Zhangsan yi/*liang dun.
he curse-Sfx Zhangsan one/*two a while
‘He gave Zhangsan a good scolding.’

(65) Ta la-le wo yi/*liang ba.
he pull-Sfx I one/*two handle
a. ‘He gave me a hand.’
b. ‘He helped me.’

The opposition between weak and strong Clvs that I have proposed for Chinese is not *ad hoc*. It can also be used to characterize some parallel phenomena in Indo-European languages, such as English and French.¹⁶

4.2 V a N in English and V yi Clv in Mandarin

In my view, the counterparts to Chinese Clv constructions can be found in English, in predicative constructions composed of a light (causative) verb — such as *to have*, *to take*, *to give* — followed by a nominalized verb. In (66) the light verb is *have*, in (67) *take* and in (68) *give*.

(66) To *have* a chat, a taste, a try, a cry, a fall, a go...

(67) To *take* a peek at, a glance at, a swim, a seat...

(68) To *give* a gasp, a kiss, a read, a fright, a hand...

The semantic distinction between verbs like *to chat*, *to swim* and *to gasp*, etc. and their light verbs counterparts *to have a chat*, *to take a swim* and *to give a gasp*, etc. is *aspectual*. Whereas the former pattern indicates unbounded/homogeneous predications,

¹⁶ I will not treat French here as I did in my oral presentations at Krakow and Poznan. The noun *coup* in French functions as an equivalent to a verbal classifier in Mandarin. For an illuminating analysis, see Gross (1984).

i.e. mass predication, the latter indicates bounded/punctual, i.e. countable predication. Hence, the function of the *a* N in an English <light V *a* N> construction is similar to that of *yi* Clv in a <V *yi* Clv> construction in Mandarin. The function of Mandarin *yi* Clv is to bound, to limit the predication that V indicates. Such a limitation can be twofold. It can either create a sub-interval where the predication is homogeneous, as in (5) above, or it can introduce an end point, as in (15) above. Examples (5) and (15) are repeated here as (69) and (70) for the sake of convenience.

- | | |
|--|--|
| (69) feng <i>yi</i> <i>xia</i>
= sew one time | (70) Ta kan-le <i>yi</i> <i>yan</i> .
= he look-Sfx one Clv |
| (5) 'to sew a bit'/'a short while' | (15) 'He cast a glance.' |

That Q Clv does bound the predication is proven by the fact that durative aspectual markers such as the preverbal *zai* 'be...ing' or the verbal suffix *-zhe* '-ing' are incompatible with Q Clv, as shown in (71) and (72). The elision of the Q Clv *yi xia* would render (71) grammatical. The substitution of *-zhe* with the telic marker *-le* would render (72) well formed.

- | |
|--|
| (71) *Wo <i>zai</i> deng <i>yi</i> <i>xia</i> huoche.
I be...ing wait a while train |
| (72) *Ta kan- <i>zhe</i> wo <i>yi</i> <i>yan</i> .
he look-Sfx I one eye |

Only if the function of V *yi* Clv is understood in aspectual terms can we understand why the Q Clv sequence *yi xia* is ambiguous, see (73) (same as (34) above). When the predication is atelic, *yi xia* bounds it by restricting it to a sub-interval; *yi xia*, then, translates as 'a bit', 'a short while' and can be substituted by *yi hui(r)* 'a short while', as in (73) and (74).

When the predication is telic, hence it is seen as a whole or restricted to a point, *yi xia* translates as 'once'/'one time', that is to say it is a frequency marker. Only in the latter case are *xia* and *ci* 'time' synonymous. See (73) and (75).

- | |
|--|
| (73) Ta tui-le wo <i>yi</i> <i>xia</i> .
= He push-Sfx I one time/while |
| (34) a. 'He pushed me once.' |
| b. 'He pushed me a while.' |

- (74) Ta tui-le wo yi *huir/xia*.
He push-Sfx I one while
'He pushed me a while.'
- (75) Ta tui-le wo yi *ci/xia*.
He push-Sfx I one time
'He pushed me once.'

The syntactic and semantic parallelism between the two patterns V *yi* Q Clv in Mandarin and V *a* N in English can be drawn further. In the same way as *yi* 'a'/'one' can be interpreted either as an indefinite article 'a' or as a cardinal number 'one', *a* in English can also be understood in two different ways.

When *a* reads as the indefinite article, it cannot be substituted by *one*, *two* etc. All the examples in (76) are ill formed.

- (76) *to give two frights, *to give two reads, *to take two glances at, *to have two struggles with, *to have two goes

When *a* is equivalent to the cardinal number 'one', substitution with other numbers is allowed, as in (77). As is well known, in stage-level predications, quantifying on the direct object can be interpreted as quantifying on the number of events, see Krifka (1989, 1992). Hence *two* in example (77) below is interpreted as the frequency marker *twice* or as *on two occasions*, as shown in (78). Saying that two events described as 'someone making a phone call' took place is equivalent to saying that 'someone made a phone call twice'.

- (77) to make/give *two* phone calls, to take/have *two* drinks, to take *two* breaks...
(78) to call *twice*, to have a drink/to take a break *on two occasions*...

If Mandarin *yi* or English *a* are not interpreted as marking numbers, it is then impossible to say that *yi* Q Clv or *a* N are equivalent to frequency markers. When the number of times the event took place must be indicated, this is done by having recourse to a frequency marker; compare (79) and (80).

- (79) *to have *two* reads, to have *three* struggles with someone etc.
(80) to have *a* read *twice*, to have *a* struggle with someone *three* times etc.

5. Concluding remarks

In this paper I have tried to highlight the tight aspectual relationship which exists between the V and its following Q Clv in a V Q Clv construction in Chinese.¹⁷ I have shown that, contrary to the standard assumptions:

- (i) V Q Clv and VV/V-*yi*-V reduplication patterns do *not* function alike, neither syntactically nor semantically, and
- (ii) it is not accurate to characterize a Clv *only* as an indicator of *frequency*.

Because verb reduplication — whether in the form of a VV or V-*yi*-V pattern — only limits a predication and does not make it countable, one can predict that it cannot be equivalent to a V Q Clv construction. This is all the more true given that, in a V Q Clv sequence, Q can stand for *liang* ‘two’, *san* ‘three’, etc. but in V-*yi*-V reduplicated pattern, *yi* can never be substituted with *liang* ‘two’, *san* ‘three’, or any other number.

A Clv is a measure phrase: it limits and bounds a stage-level predication it is in construction with in two different ways. With what I have called ‘weak verbal classifiers’, a Clv simply bounds the predication. *Yi*, which precedes Clv, is understood as an indefinite marker — it cannot be substituted by *liang* ‘two’, etc. and *yi* Clv is scopeless. With strong verbal classifiers, a Clv bounds the predication, renders it telic, and thus makes it countable. *Yi* is understood as a number and can be substituted by *liang* ‘two’, *san* ‘three’, etc.; *yi* Clv can have a broad scope in a *lian...ye/dou* ‘even’ construction.

Drawing a distinction between the two classes of verbal classifiers in terms of the roles they play in the quantificational domain is, I hope, more fruitful than listing verbal classifiers according to their lexical properties. Otherwise, for example, how could we explain that even though both *zui* ‘mouth’ and *jiao* ‘foot’ are body parts, they do not function alike? From (81) one cannot predict that (82) cannot be derived, whereas (84) can be derived from (83).

(81) cha *yi* zui
 insert one mouth
 ‘to interrupt’

(82) *cha *ji* zui
 insert several mouth

(83) ti *yi* jiao
 kick one foot
 ‘to give a kick’

(84) ti *ji* jiao
 kick several foot
 ‘to give several kicks’

¹⁷ See also Jaxontov (1997), Xiao & McEnery (2004), and Fassi-Fehri & Vinet (2008).

In short, even though the label ‘verbal classifier’ does not belong to the set of meta-linguistic terms used to describe languages in the Indo-European tradition, this label could be used to describe some constructions in English (or in French). By drawing a comparison between *yi Clv* in Mandarin and *a N* in English, where N is a verbal noun, I have tried to show that these constructions share similar properties. These similarities probably explain why the equivalents of ‘a hand’, which are italicized in the three examples (85)-(87) — *a hand* in English, *yi ba* in Mandarin and *un coup de main* in French — could adequately be labelled verbal classifiers.

(85) to lend (X) *a hand*

(86) la (X) *yi ba*
 ‘to lend (X) a hand’/‘to help (X)’

(87) donner *un coup de main* (à X)
 give a Clv of hand (to X)
 ‘to help (X)’

References

- Abbiati, Magda. 1998. *Grammatica di cinese moderno*. Venezia: Libreria editrice Cafoscarina.
- Aikhenvald, Alexandra Y. 2000. *Classifiers: A Typology of Noun Categorization Devices*. Oxford & New York: Oxford University Press.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Cheung, Hung-nin Samuel et al. 1994. *A Practical Chinese Grammar*. Hong Kong: The Chinese University Press.
- Ding, Shengshu. 1961. *Xiandai Hanyu Yufa Jianghua* [Lectures on Modern Chinese Grammar]. Beijing: The Commercial Press.
- Fassi-Fehri, Abdelkader, and Marie-Thérèse Vinet. 2008. Verbal and nominal classes in Arabic and Chinese. *Recherches Linguistiques de Vincennes* 37:55-83.
- Farkas, Donka F., and Henriëtte de Swart. 2003. *The Semantics of Incorporation: From Argument Structure to Discourse Transparency*. Stanford: CSLI.
- Filip, Hana. 2000. The Quantization Puzzle. *Events as Grammatical Objects: The Converging Perspectives of Lexical Semantics and Syntax, Logical Semantics and Syntax*, ed. by Carol L. Tenny & James Pustejovsky, 3-60. Stanford: CSLI.
- Gerner, Matthias. 2007. Verb classifiers in East and Southeast Asia. Paper presented at the 10th International Cognitive Linguistics Conference (ICLC 10), July 15-20, 2007. Poland: Krakow University.
- Gerner, Matthias. 2009. Instruments as verb classifiers in Kam (Dong). *Linguistics* 47.3: 697-742.
- Gross, Gaston. 1984. Etude syntaxique de deux emplois du mot *coup*. *Linguisticæ Investigationes* 8.1:37-61.
- Huang, C.-T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge & New York: Cambridge University Press.
- Hung, I-Fang. 2000. *Dunhuang Tulufan Wenshu zhong zhi Liangci Yanjiu* [A Study of Classifiers in Tun-Huang and Turfan Documents]. Taipei: Wenjin Press.
- Jaxontov, Sergej Je. (=Yakhontov, Sergey E.) 1997. The problem of describing plurality of actions in Chinese. *Typology of Iterative Constructions*, ed. by Viktor S. Xrakovskij, 441-452. München: LINCOM Europa.
- Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. *Semantics and Contextual Expressions*, ed. by Renate Bartsch, Johan van Benthem & Peter van Emde Boas, 75-115. Dordrecht: Foris.
- Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. *Lexical Matters*, ed. by Ivan A. Sag & Anna Szabolcsi, 29-53. Stanford: CSLI.

- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Reference Grammar*. Berkeley: University of California Press.
- Liu, Yuehua, Wenyu Pan, and Wei Gu. 1983. *Shiyong Xiandai Hanyu Yufa* [*The Modern Pragmatic Chinese Grammar*]. Beijing: Foreign Language Teaching and Research Press.
- Matthews, Stephen, and Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. London & New York: Routledge.
- Nguyen, Phu-Phong. 1995. *Questions de linguistique vietnamienne*. Paris: Presses de l'Ecole Française d'Extrême Orient.
- Paris, Marie-Claude. 1981. *Problèmes de syntaxe et de sémantique en linguistique chinoise*. Institut des Hautes Etudes Chinoises. Paris: Collège de France.
- Paris, Marie-Claude. 2007a. Un aperçu de la reduplication nominale et verbale en mandarin. *Faits de langue* 29:65-76.
- Paris, Marie-Claude. 2007b. On event quantification. Paper presented at the 10th International Cognitive Linguistics Conference (ICLC 10), July 15-20, 2007. Poland: Krakow University.
- Paris, Marie-Claude, and Marie-Thérèse Vinet. 2010. Approximative *zuǒyòu* 'around, about' in Chinese. *Language and Linguistics* 11.4:767-801.
- Ross, Claudia, and Jing-heng Sheng Ma. 2006. *Modern Mandarin Chinese Grammar*. London & New York: Routledge.
- Shao, Jingmin. 1996. Dongliangci de yuyi fenxi ji qi yu dongci de xuanze guanxi [A semantic analysis of verbal classifiers and their selectional restrictions with verbs]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 1996.2:100-109.
- Shao, Jingmin. 2007. Shuo 'V yi ba' zhong V de fanhua yu 'yi ba' de cihuihua [Generalization of V and lexicalization of *yi ba* in the "V + *yi ba*" construction]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 2007.1:14-19.
- Sybesma, Rint. 1999. *The Mandarin VP*. Dordrecht: Kluwer.
- Xiao, Richard, and Tony McEnery. 2004. *Aspect in Mandarin Chinese: A Corpus-based Study*. Amsterdam & Philadelphia: John Benjamins.
- Xu, Dan. 1996. *Initiation à la syntaxe chinoise*. Paris: L'Asiathèque.
- Zhang, Niina Ning. 2002. Counting and classifying eventualities in Chinese. Mimeo USC. <http://www.usc.edu/dept/LAS/ealc/chinling/articles/Classification>.

Decoding the Multifunctional Usages of *gen* in Chinese^{*}

Stephen Yun-hsiang Chuang (莊雲翔)¹ and Elizabeth Zeitoun (齊莉莎)²

*Kinmen Senior High School*¹

*Academia Sinica*²

Due to the grammaticalization processes it has undergone, Chinese *gen* ‘with, and’ now displays multi-functional usages. Originally a noun meaning ‘heel’, it has come to function as a verb meaning ‘to follow’. It can also serve as a preposition marking comitativity, reciprocity, unidirectionality, and as a conjunction entailing either a distributive or a collective reading. Syntactic tests have been proposed since the early 1970s to disambiguate the grammatical status of *gen* as a preposition and as a conjunction. While incorporating findings of earlier contributions, the present paper shows that each function and meaning of *gen* is influenced in large part by verbal semantics, i.e., verbal semantics and the syntactic position of modifying elements interact closely and give rise to the different interpretations and functions of *gen*. At one end of the scale, *gen* functions only as a conjunction; at the other end, it can be analyzed as a verb, as a preposition and as a conjunction; between these two ends, it can be treated as a preposition and a conjunction.

Key words: grammaticalization, conjunction, preposition, coordination, verbal semantics, comitativity

1. Introduction

Stassen (2000) and Haspelmath (2004) point out that the world’s languages vary in terms of their strategy to mark coordination and comitativity, i.e., while some languages exhibit a distinction between coordination and comitativity, others do not. As further

^{*} The present paper represents an abridged and revised version of Stephen Y. Chuang’s MA thesis (Chuang 2007), written under the co-supervision of Elizabeth Zeitoun and Lillian M. Huang. This research was financially supported by a project from the National Science Council (NSC 95-2411-H-158-004 — A Typological Study of Coordination and Comitativity in French, Mandarin and Some Formosan Languages) conducted in collaboration with the French CNRS, co-directed by Lillian M. Huang and Elizabeth Zeitoun in Taiwan and by Marie-Claude Paris in France. We are grateful to Marie-Claude Paris, Lillian M. Huang, Lily I. Hsu, Stacy F. Teng, Shu-chuan Tseng, Jackson T.-S. Sun, Cheng-sheng Liu, One-Soon Her and two anonymous reviewers for their comments and suggestions at different stages of this study.

shown in Paris (2005), English and French, for instance, display a morpho-syntactic contrast between coordination, which is marked respectively by *and* and *et* as in (1a-b),¹ and comitativity, which is manifested by *with* and *avec* as in (2a-b).

- (1) a. John and Mary read some novels.
- b. Jean et Marie lisent des romans.
- (2) a. John read some novels with Mary.
- b. Jean lit des romans avec Marie.

Chinese, on the other hand, makes use of a single marker, i.e. *gen*, to encode coordination and comitativity. As illustrated in (3),² *gen* is ambiguous: it can be perceived as a conjunction ‘and’ or as a preposition ‘with’ marking comitativity.

- (3) 張三 跟 瑪莉 讀 了 一些 小說。
- Zhangsan *gen* Mali du le yixie xiaoshuo.
- Zhangsan *GEN* Mali read PFV some novel
- i. ‘Zhangsan and Mali read some novels.’
- ii. ‘Zhangsan read some novels with Mali.’

Gen can also function as a verb, as exemplified in (4).³

- (4) 我 跟 著 他。
- Wo *gen* zhe ta.
- 1SG *GEN* DUR 3SG
- ‘I am following him.’

¹ Unless mentioned otherwise, all the examples provided in this paper were collected or constructed by the first author. We have also consulted *Academia Sinica Balanced Corpus of Modern Chinese* (<http://www.sinica.edu.tw/SinicaCorpus/>) and Dr. Tseng Shu-chuan’s *Archives and Linguistic Representations of Spoken Taiwan Mandarin* (<http://mmc.sinica.edu.tw/>).

² The abbreviations used in this paper include: ASP: Aspect, CL: Classifier, COLL: Collective, COM: Comitative, CONJ: Conjunction, COP: Copula, DIST: Distributive, DUR: Durative, EXP: Experiential, G: Goal-oriented, INTR: Intransitive, Non-unidir: Non-unidirectional, NP: Noun Phrase, PFV: Perfective, PREP: Preposition, REC: Reciprocal, S: Source-oriented, TR: Transitive, V: Verb, VO: Verb Object compound, UNIDIR: Unidirectional, 1SG: First Person Singular, and 3SG: Third Person Singular.

³ It has been demonstrated that *gen* was originally a noun meaning ‘heel’ (Furukawa 2000) and that through metaphorical change, it came to be used as a verb meaning ‘to follow’ before further undergoing two grammaticalization processes, the first from a verb to a preposition meaning ‘with’ and the second from a preposition to a conjunction meaning ‘and’ (Liu & Peyraube 1994).

The multifunctional usages of *gen* have aroused the interest of many scholars (*cf.* Teng 1970, Tang 1990, Furukawa 2000, and Paris 2005, 2007a among others). It has been shown that in subject position, *gen* can function as a preposition or a conjunction; in object position, it can only be treated as a conjunction. When *gen* functions as a preposition, it can serve to mark comitativity (e.g., X *gen* Y *chuqu* ‘X went out with Y’), reciprocity (e.g., X *gen* Y *chaojia* ‘X and Y quarreled’) and unidirectionality (e.g., X *gen* Y *jieshu* ‘X borrowed a book from Y’). When *gen* functions as a conjunction, it serves to conjoin two NPs to form a larger syntactic unit. Two different readings need to be distinguished. When *gen* co-occurs with *dou* ‘all’, it entails a distributive reading. When *gen* co-occurs with *yiqi* ‘together’, it has a collective reading. In ambiguous examples, a number of tests (e.g., insertion of a temporal adverb, negation, clefting, anaphora and relativization) can be used to draw a distinction between *gen* as a preposition and *gen* as a conjunction (entailing either a distributive or a collective reading, see Paris 2005 for details).

In this paper, we adopt an approach which integrates the findings briefly summarized above but highlights the interface between syntax and semantics by arguing that verbal semantics needs also to be taken into consideration to account for the different functions/meanings of *gen*, i.e., the behavior of *gen* is influenced in large part by verbal semantics. As an illustration, consider (5a-b), where the interpretation of *gen* is different though it occurs in the same construction, i.e. *wo gen Lisi*. In (5a), the verb *xihuan* ‘to like’ is a stative verb; *gen* serves to conjoin two NPs and is treated as a conjunction. In (5b), the verb *jie* ‘to borrow’ is a source-oriented verb; *gen* is more likely to be analyzed as a preposition which introduces the *source* of the denoted event. These findings can be further corroborated and substantiated by the syntactic position of temporal adverbs such as *zuotian* ‘yesterday’ and the co-occurrence of *dou* ‘all’ and *yiqi* ‘together’, as demonstrated in earlier studies, *cf.* (5c-d).

- (5) a. 我 跟 李四 喜歡 打 籃球。
 Wo *gen* Lisi xihuan da lanqiu.
 1SG *GEN* Lisi like hit basketball
 ‘Lisi and I like to play basketball.’
- b. 我 跟 李四 借 了 三 百 元。
 Wo *gen* Lisi jie le san bai yuan.
 1SG *GEN* Lisi borrow PFV three hundred dollar
 ‘I borrowed three hundred dollars from Lisi.’
- c. 我 跟 李四 都 喜歡 打 籃球。
 Wo *gen* Lisi dou xihuan da lanqiu.
 1SG *GEN* Lisi all like hit basketball
 ‘Both Lisi and I like to play basketball.’

- d. 我 以前 跟 李四 借 了 三 百 元。
 Wo yiqian *gen* Lisi jie le san bai yuan.
 1SG before *GEN* Lisi borrow PFV three hundred dollar
 ‘I’ve previously borrowed three hundred dollars from Lisi.’

The purpose of this paper is to show that a scale can be set up, whereby verbal semantics and the syntactic position of modifying elements interact closely and give rise to the different interpretations and functions of *gen*. At one end of the scale, *gen* functions only as a conjunction; at the other end, it can be analyzed as a verb, as a preposition and as a conjunction; between these two ends, it can be treated as a preposition and a conjunction. We will only consider sentences of the type *X gen Y Verb* where ambiguity is more likely to arise and exclude sentences like *X Verb gen Y Verb* (e.g., 我喜歡跟李四打籃球 *Wo xihuan gen Lisi da lanqiu*, ‘I like to play basketball with Lisi’), where *gen* can only be analyzed as a preposition.

In §2, we propose a verbal taxonomy that accounts for the proper interpretation of *gen*, whose behavior is largely determined by the verb type it co-occurs with (§3). Section 4 summarizes our findings.

2. Verbal taxonomy

Verbs are classified along the following parameters:

- (i) Verbal semantics: verbs are classified in terms of (a) stativity, (b) reciprocity, and (c) motion.
- (ii) Verb subcategorization: verbs are distinguished according to whether they take an object, a clausal complement or they form a VO compound.
- (iii) Aspectual variations: verbs are also differentiated in terms of whether they are marked for aspect.

We briefly explain these three parameters in the following subsections.

2.1 Verbal semantics

This verbal taxonomy basically takes a binary-distinction approach⁴ toward verbal semantics and posits that verbs can be classified into:

⁴ Whenever irrelevant for the present discussion, however, binary features will not be used.

- (i) stative vs. non-stative verbs
- (ii) reciprocal vs. non-reciprocal verbs
- (iii) motion vs. non-motion verbs

2.1.1 Stative vs. non-stative verbs

As has been shown in previous classification frameworks (e.g., Vendler 1967, Chao 1968, Teng 1977, and Her 1991 among others), verbs can be basically divided into stative and non-stative verbs. Stative verbs usually refer to mental states or physical characteristics and do not involve any action. They include verbs like *xihuan* ‘to like’, *haipa* ‘to be afraid’, *xiwang* ‘to hope’, *yongyou* ‘to possess’, *renwei* ‘to think’, *gao* ‘to be tall’, *qiong* ‘to be poor’, *kaixin* ‘to be happy’, *hao* ‘to be good’, *man* ‘to be slow’, etc. Stative verbs also include copular predicates which mainly consist of nouns such as *xuesheng* ‘student’, and *laoshi* ‘teacher’ introduced by the copula *shi* ‘to be’, as in *shi xuesheng* ‘to be a student’, and *shi laoshi* ‘to be a teacher’.

Non-stative verbs, however, constitute a complex category, which encompasses different types of verbs, e.g., motion verbs such as *lai* ‘to come’, and *qu* ‘to go’, non-motion verbs such as *tiaowu* ‘to dance’, *he* ‘to drink’, and *xie* ‘to write’, unidirectional verbs, such as *xue* ‘to learn’, *jie* ‘to borrow’, and *jiang* ‘to speak’.

2.1.2 Reciprocal vs. non-reciprocal verbs

Following Lichtenberk (1994), reciprocal verbs are defined as requiring two (or more) participants involved in the same action. These two participants may either act one upon the other, e.g., *chaojia* ‘to quarrel’, and *xiang’ai* ‘to love each other’ or take part in the same action together, e.g., *taolun* ‘to discuss’, and *hezuo* ‘to collaborate’. Verbs like *taolun* and *hezuo* are termed “collective” verbs in this study.⁵

As mentioned in Liu (2000:123), in Mandarin Chinese, the common way of marking reciprocity is to add an adverbial (e.g., *xianghu* or *huxiang* ‘mutually’) in front of the verb. Thus, verbs that are originally not reciprocal can be made reciprocal through the addition of *hu* or *huxiang* ‘mutually’, e.g., *taoyan* ‘to dislike’ vs. *huxiang taoyan* ‘to dislike each other’, *ma* ‘to scold’ vs. *huma* ‘to scold each other’. Reciprocal verbs can be stative in nature, as in *huxiang xihuan* ‘to like each other’ or non-stative, as in *huxiang chaoxiao* ‘to laugh at each other’.⁶

⁵ These verbs are termed “collective” verbs, based on Frajzyngier (2000) and are treated as a subtype of reciprocal verbs. We find that they also behave like reciprocal verbs with respect to the interpretation of *gen*.

⁶ We will show below that *stative* and *non-stative reciprocal* verbs behave the same with respect to the interpretation of *gen*.

A certain type of stative reciprocal verb is headed by the copula *shi* taking a nominal predicate such as *fuqi* ‘couple’, *tongxue* ‘classmate’ and *pengyou* ‘friend’, as in *shi fuqi* ‘to be husband and wife’. These nominal predicates serve to establish a “converse”⁷ relationship between the two referents conjoined by *gen* (see Lichtenberk 2000).

2.1.3 Motion/Non-motion verbs

Motion verbs involve spatial movement from one place to another while non-motion verbs do not feature any movement. Motion verbs include verbs like *qu* ‘to go’, *lai* ‘to come’, and *dao* ‘to reach’. Non-motion verbs include verbs like *chi* ‘to eat’, *shui* ‘to sleep’, *zuo* ‘to sit’, *shuo* ‘to say’ and *jie* ‘to borrow’.

Certain non-motion verbs like *shuo* ‘to say’ and *jie* ‘to borrow’ behave differently with respect to the interpretations of *gen* as opposed to other non-motion verbs like *zuo* ‘to sit’ and *shui* ‘to sleep’. When *gen* co-occurs with verbs like *shuo* ‘to say’ and *jie* ‘to borrow’, it can mark a unidirectional relationship between the two referents. Accordingly, these verbs are labeled “unidirectional” verbs here. Unidirectional verbs can be further divided into two types of verbs:

- (i) Source-oriented verbs: *jie* ‘to borrow’, *xue* ‘to learn’, *lesuo* ‘to extort’, *mai* ‘to buy’, and *yao* ‘to ask for’.
- (ii) Goal-oriented verbs: *shuo* ‘to say’, *jieshi* ‘to explain’, *xuanbu* ‘to announce’, *jieshao* ‘to introduce’, *tiqi* ‘to mention’ and *baogao* ‘to report’.

2.2 Verb subcategorization

In terms of subcategorization frame, verbs can be basically defined as taking an object or not, i.e. transitive or intransitive.

Stative verbs which are intransitive encompass verbs like *gao* ‘to be tall’, *xiao* ‘to be small/little’, *hao* ‘to be good’, *huai* ‘to be bad’, *kuai* ‘to be fast’, and *man* ‘to be slow’. Stative verbs which are transitive include verbs like *xihuan* ‘to like’, *ai* ‘to love’, *taoyan* ‘to dislike’, *haipa* ‘to fear’, and *xiangxin* ‘to believe’.

Non-stative verbs which are intransitive include verbs like *tiao* ‘to jump’, *youyong* ‘to swim’, and *gongzuo* ‘to work’. Non-stative verbs which are transitive include verbs like *da* ‘to beat’, *chi* ‘to eat’, *jie* ‘to borrow’, *shuo* ‘to say’, and *taolun* ‘to discuss’. Non-stative transitive verbs can also form a VO compound, e.g., *jiehun* ‘to get married’, *dajia* ‘to fight’, *qianshou* ‘to hold hands’, *jiewen* ‘to kiss each other’, and *jianmian* ‘to meet’.

⁷ When two referents are in a converse relationship, “A is related in a certain way to B, and B is conversely related to A” (Lichtenberk 2000:34).

2.3 Aspectual variations

Aspectual encoding in reciprocal verbs can play a role with regard to the function of *gen*. In (6a), *gen* in the reciprocal event *jiehun* ‘to get married’ serves as a conjunction entailing a collective reading. Therefore, the insertion of *dou* ‘all’ yields the ungrammaticality of sentence as in (6b).

- (6) a. 張三 跟 瑪莉 結婚。
 Zhangsan *gen* Mali jiehun.
 Zhangsan *GEN* Mali get-married
 ‘Zhangsan and Mali got married.’
- b. *張三 跟 瑪莉 都 結婚。
 Zhangsan *gen* Mali dou jiehun.
 Zhangsan *GEN* Mali all get-married

However, when an aspect marker is added to the compound, the reciprocal verbal compound does not necessarily require a subject with plural reference, as shown in (7a-b). In (7b) — as opposed to (6b) — the insertion of *dou* is possible and yields a distributive meaning.

- (7) a. 張三 跟 瑪莉 昨天 結 了 婚。
 Zhangsan *gen* Mali zuotian jie le hun.
 Zhangsan *GEN* Mali yesterday tie PFV marriage
 ‘Zhangsan and Mali got married yesterday.’
- b. 張三 跟 瑪莉 昨天 都 結 了 婚。
 Zhangsan *gen* Mali zuotian dou jie le hun.
 Zhangsan *GEN* Mali yesterday all tie PFV marriage
 ‘Both Zhangsan and Mali got married yesterday (separately).’

2.4 Summary of the verbal taxonomy

Integrating the above factors (i.e. verbal classification, verbal subcategorization and aspectual variations), the verbal taxonomy can be schematized as follows:

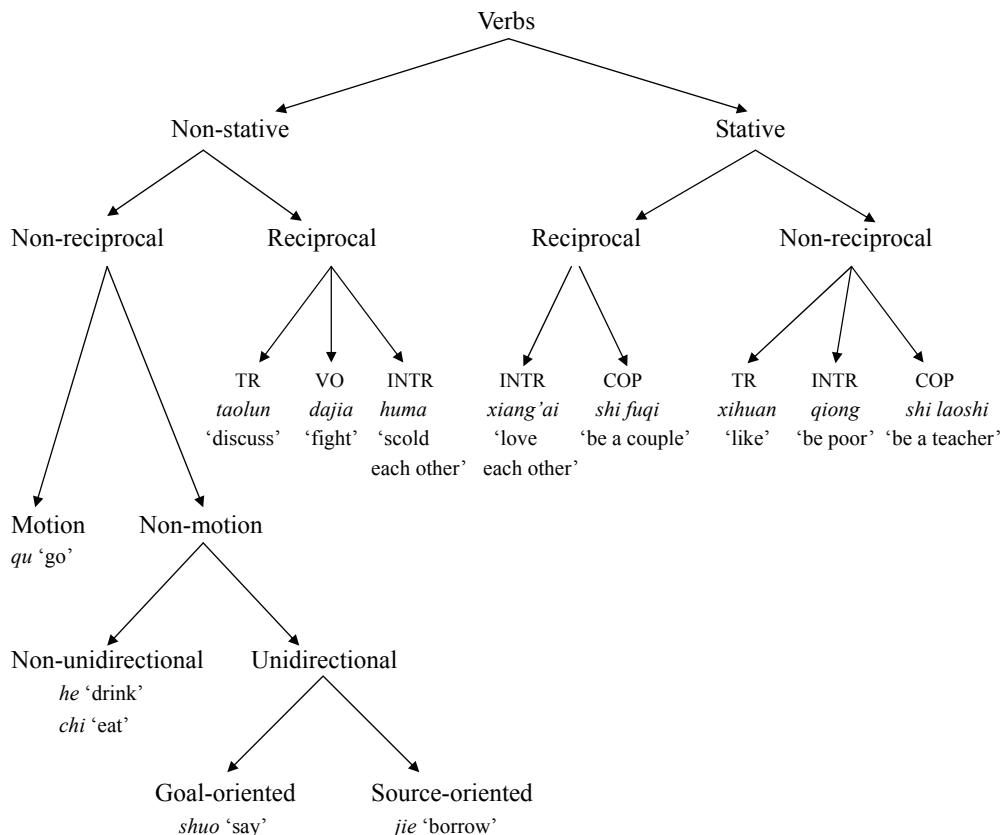


Figure 1: Verbal taxonomy to be used to account for the multi-functions of *gen*

3. Interaction between different types of verbs and the behavior of *gen*

As shown in §2.1, verbs are divided into two major categories, i.e., stative and non-stative verbs, each of which can be subdivided into reciprocal and non-reciprocal verbs. In the following sections, we show how these different types of verbs interact closely with the interpretations of *gen*. In §3.1, we examine *gen* in co-occurrence with stative verbs. In §3.2, we focus on *gen* in co-occurrence with non-stative verbs.

3.1 *Gen* co-occurring with stative verbs

3.1.1 *Gen* in co-occurrence with stative non-reciprocal verbs

When *gen* occurs with stative non-reciprocal verbs, as in (8a-c), its behavior is quite unitary: it can only function as a conjunction.

- (8) a. 張三 跟 李四 是 老師。
 Zhangsan *gen* Lisi shi laoshi.
 Zhangsan *GEN* Lisi COP teacher
 ‘Zhangsan and Lisi are teachers.’
- b. 張三 跟 李四 很 窮。
 Zhangsan *gen* Lisi hen qiong.
 Zhangsan *GEN* Lisi very poor
 ‘Zhangsan and Lisi are very poor.’
- c. 張三 跟 李四 喜歡 打 籃球。
 Zhangsan *gen* Lisi xihuan da lanqiu.
 Zhangsan *GEN* Lisi like hit basketball
 ‘Zhangsan and Lisi like to play basketball.’

In these three examples, the insertion of an adverb such as *yiqian* ‘before’ between *Zhangsan* and *Lisi* gives rise to the ungrammaticality of the sentence in question. Compare the grammaticality of (9a-c) and (9a'-c'):

- (9) a. *張三 以前 跟 李四 是 老師。
 Zhangsan yiqian *gen* Lisi shi laoshi.
 Zhangsan before *GEN* Lisi COP teacher
- a'. 張三 跟 李四 以前 是 老師。
 Zhangsan *gen* Lisi yiqian shi laoshi.
 Zhangsan *GEN* Lisi before COP teacher
 ‘Zhangsan and Lisi were teachers before.’
- b. *張三 以前 跟 李四 很 窮。
 Zhangsan yiqian *gen* Lisi hen qiong.
 Zhangsan before *GEN* Lisi very poor
- b'. 張三 跟 李四 以前 很 窮。
 Zhangsan *gen* Lisi yiqian hen qiong.
 Zhangsan *GEN* Lisi before very poor
 ‘Zhangsan and Lisi were very poor before.’
- c. *張三 以前 跟 李四 喜歡 打 籃球。
 Zhangsan yiqian *gen* Lisi xihuan da lanqiu.
 Zhangsan before *GEN* Lisi like hit basketball
- c'. 張三 跟 李四 以前 喜歡 打 籃球。
 Zhangsan *gen* Lisi yiqian xihuan da lanqiu.
 Zhangsan *GEN* Lisi before like hit basketball
 ‘Zhangsan and Lisi liked to play basketball before.’

As pointed out above, as a conjunction, *gen* can have two different readings: a distributive or a collective reading. Since only *dou* ‘all’ can be inserted (see (10a-c)), *gen* can only be given a distributive reading.

- (10) a. 張三 跟 李四 以前 都 是 老師。
 Zhangsan *gen* Lisi yiqian dou shi laoshi.
 Zhangsan *GEN* Lisi before all COP teacher
 ‘Both Zhangsan and Lisi were teachers before.’
- a'. *張三 跟 李四 以前 一起 是 老師。
 Zhangsan *gen* Lisi yiqian yiqi shi laoshi.
 Zhangsan *GEN* Lisi before together COP teacher
- b. 張三 跟 李四 以前 都 很 窮。
 Zhangsan *gen* Lisi yiqian dou hen qiong.
 Zhangsan *GEN* Lisi before all very poor
 ‘Both Zhangsan and Lisi were very poor before.’
- b'. *張三 跟 李四 以前 一起 很 窮。
 Zhangsan *gen* Lisi yiqian yiqi hen qiong.
 Zhangsan *GEN* Lisi before together very poor
- c. 張三 跟 李四 以前 都 喜歡 打 籃球。
 Zhangsan *gen* Lisi yiqian dou xihuan da lanqiu.
 Zhangsan *GEN* Lisi before all like hit basketball
 ‘Both Zhangsan and Lisi liked to play basketball before.’
- c'. *張三 跟 李四 以前 一起 喜歡 打 籃球。
 Zhangsan *gen* Lisi yiqian yiqi xihuan da lanqiu.
 Zhangsan *GEN* Lisi before together like hit basketball

It is important to note, however, that when *gen* co-occurs with certain stative transitive verbs like *yongyou* ‘to own’ and *xuyao* ‘to need’, the conjunction *gen* can be given two different interpretations (distributive or collective), as shown in (11a-a') and (11b-b'):

- (11) a. 張三 跟 瑪莉 都 擁有 五棟 房子。
 Zhangsan *gen* Mali dou yongyou wudong fangzi.
 Zhangsan *GEN* Mali all own five-CL house
 ‘Both Zhangsan and Mali own five houses.’

- a'. 張三 跟 瑪莉 一共 擁有 五棟 房子。
 Zhangsan *gen* Mali yigong yongyou wudong fangzi.
 Zhangsan *GEN* Mali altogether own five-CL house
 'Zhangsan and Mali own five houses in total.'
- b. 張三 跟 瑪莉 都 需要 兩瓶 水。
 Zhangsan *gen* Mali dou xuyao liangping shui.
 Zhangsan *GEN* Mali all need two-CL water
 'Both Zhangsan and Mali need two bottles of water.'
- b'. 張三 跟 瑪莉 一共 需要 兩瓶 水。
 Zhangsan *gen* Mali yigong xuyao liangping shui.
 Zhangsan *GEN* Mali altogether need two-CL water
 'Zhangsan and Mali need two bottles of water in total.'

On the other hand, when no additional markers such as *dou* 'all' and *yigong* 'altogether' occur, *gen* is more likely to be perceived as entailing a collective reading. Consider the following sentences in (12a-b).

- (12) a. 張三 跟 瑪莉 擁有 五棟 房子。
 Zhangsan *gen* Mali yongyou wudong fangzi.
 Zhangsan *GEN* Mali own five-CL house
 'Zhangsan and Mali own five houses (in total).'
- b. 張三 跟 瑪莉 需要 兩瓶 水。
 Zhangsan *gen* Mali xuyao liangping shui.
 Zhangsan *GEN* Mali need two-CL water
 'Zhangsan and Mali need two bottles of water (in total).'

3.1.2 *Gen* in co-occurrence with stative reciprocal verbs

In co-occurrence with stative reciprocal verbs, the behavior of *gen* becomes complex.

When *gen* co-occurs with nominal predicates such as *shi fuqi* 'to be husband and wife', and *shi tongxue* 'to be classmates', as in (13a), it can be analyzed as a conjunction, as in (13b), or as a preposition, as in (13c).

- (13) a. 張三 跟 瑪莉 是 夫妻。
 Zhangsan *gen* Mali shi fuqi.
 Zhangsan *GEN* Mali COP couple
 'Zhangsan and Mali are husband and wife.'

- b. 張三 跟 瑪莉 以前 是 夫妻。
 Zhangsan *gen* Mali yiqian shi fuqi.
 Zhangsan *GEN* Mali before COP couple
 ‘Zhangsan and Mali were husband and wife before.’
- c. 張三 以前 跟 瑪莉 是 夫妻。
 Zhangsan yiqian *gen* Mali shi fuqi.
 Zhangsan before *GEN* Mali COP couple
 ‘Zhangsan was Mali’s husband before.’

As a conjunction, *gen* can only be given a collective reading; the addition of *dou* ‘all’ causes the ungrammaticality of the sentence in question, as shown in (13d).

- (13) d. *張三 跟 瑪莉 (以前) 都 是 夫妻。
 Zhangsan *gen* Mali (yiqian) dou shi fuqi.
 Zhangsan *GEN* Mali (before) all COP couple

When *gen* co-occurs with stative reciprocal verbs formed by the addition of the reciprocal markers *huxiang*, *xianghu* or the reduced forms *xiang*, *hu* ‘mutual’ to a stative verb, as in (14a), it can only function as a conjunction (collective reading). The insertion of the temporal adverb *yiqian* ‘before’ in (14b-c) illustrates this point. Furthermore, the insertion of *dou*, as in (14d), leads to the ungrammaticality of the sentence in question.

- (14) a. 張三 跟 瑪莉 互相 喜歡。
 Zhangsan *gen* Mali huxiang xihuan.
 Zhangsan *GEN* Mali mutual like
 ‘Zhangsan and Mali like each other.’
- b. 張三 跟 瑪莉 以前 互相 喜歡。
 Zhangsan *gen* Mali yiqian huxiang xihuan.
 Zhangsan *GEN* Mali before mutual like
 ‘Zhangsan and Mali liked each other before.’
- c. *張三 以前 跟 瑪莉 互相 喜歡。
 Zhangsan yiqian *gen* Mali huxiang xihuan.
 Zhangsan before *GEN* Mali mutual like
- d. *張三 跟 瑪莉 以前 都 互相 喜歡。
 Zhangsan *gen* Mali yiqian dou huxiang xihuan.
 Zhangsan *GEN* Mali before all mutual like

Table 1 summarizes the above discussion.

Table 1: Functions of *gen* co-occurring with stative verbs

Stative verbs			Functions of <i>gen</i>			Co-occurrence constraints	
			Conj		Prep	<i>dou</i> 'all'	<i>yiqi</i> 'together'
			Distributive reading	Collective reading			
Non-Reciprocal	Copula		+	–	–	+	–
	Intransitive		+	–	–	+	–
	Tr	Others	+	–	–	+	–
		<i>you</i> 'to have'	+	+	–	+	– <i>yigong</i> 'altogether'
Reciprocal	Copula		–	+	+(Rec)	–	–
	Intr <i>huxiang</i> 'mutual'		–	+	–	–	–

3.2 *Gen* co-occurring with non-stative verbs

In this section, we deal with all the types of non-stative verbs included in the above verbal semantic framework. Section 3.2.1 is devoted to *gen* in co-occurrence with non-stative reciprocal verbs and §3.2.2 focuses on *gen* in co-occurrence with the non-stative non-reciprocal verbs.

3.2.1 *Gen* co-occurring with non-stative reciprocal verbs

One type of non-stative reciprocal verb takes the overt markers *huxiang*, *xianghu* or the reduced forms *xiang*, *hu* 'mutual' as shown in (15a). When *gen* co-occurs with such reciprocal verbs, it can only function as a conjunction (collective reading) as in (15b-c), but never as a preposition, see (15d).

- (15) a. 張三 跟 李四 互 罵。
 Zhangsan *gen* Lisi hu ma.
 Zhangsan *GEN* Lisi mutual scold
 'Zhangsan and Lisi scold each other.'
- b. 張三 跟 李四 常常 互 罵。
 Zhangsan *gen* Lisi changchang hu ma.
 Zhangsan *GEN* Lisi often mutual scold
 'Zhangsan and Lisi often scold each other.'

- c. *張三 跟 李四 都 常常 互 罵。
 Zhangsan *gen* Lisi dou changchang hu ma.
 Zhangsan *GEN* Lisi all often mutual scold
- d. *張三 常常 跟 李四 互 罵。
 Zhangsan changchang *gen* Lisi hu ma.
 Zhangsan often *GEN* Lisi mutual scold

Non-stative reciprocal verbs which take the form of VO compound as in (16a) or which are collective verbs as in (17a) behave alike. In co-occurrence with such reciprocal verbs, *gen* can function either as a conjunction (collective reading) — compare the grammaticality of (16b-c) and (17b-c) — or as a preposition, *cf.* (16d) and (17d).

- (16) a. 張三 跟 瑪莉 吵架。
 Zhangsan *gen* Mali chaojia.
 Zhangsan *GEN* Mali quarrel
 ‘Zhangsan and Mali quarrel.’
- b. 張三 跟 瑪莉 常常 吵架。
 Zhangsan *gen* Mali changchang chaojia.
 Zhangsan *GEN* Mali often quarrel
 ‘Zhangsan and Mali often quarrel.’
- c. *張三 跟 瑪莉 都 常常 吵架。
 Zhangsan *gen* Mali dou changchang chaojia.
 Zhangsan *GEN* Mali all often quarrel
- d. 張三 常常 跟 瑪莉 吵架。
 Zhangsan changchang *gen* Mali chaojia.
 Zhangsan often *GEN* Mali quarrel
 ‘Zhangsan often quarrels with Mali.’
- (17) a. 張三 跟 瑪莉 討論 這件 事。
 Zhangsan *gen* Mali taolun zhejian shi.
 Zhangsan *GEN* Mali discuss this-CL matter
 ‘Zhangsan and Mali discuss this matter.’
- b. 張三 跟 瑪莉 常常 討論 這件 事。
 Zhangsan *gen* Mali changchang taolun zhejian shi.
 Zhangsan *GEN* Mali often discuss this-CL matter
 ‘Zhangsan and Mali often discuss this matter.’
- c. *張三 跟 瑪莉 都 常常 討論 這件 事。
 Zhangsan *gen* Mali dou changchang taolun zhejian shi.
 Zhangsan *GEN* Mali all often discuss this-CL matter

- d. 張三 常常 跟 瑪莉 討論 這件事。
 Zhangsan changchang *gen* Mali taolun zhejian shi.
 Zhangsan often GEN Mali discuss this-CL matter
 ‘Zhangsan often discusses this matter with Mali.’

It is interesting to note that when an aspect marker is added to VO compounds or collective verbs, the conjunction *gen* can also entail a distributive reading. Compare (18a-b) and (19a-b).

- (18) a. *張三 跟 瑪莉 都 吵架。
 Zhangsan *gen* Mali dou chaojia.
 Zhangsan GEN Mali all quarrel
 b. 張三 跟 瑪莉 都 吵 了 架。
 Zhangsan *gen* Mali dou chao le jia.
 Zhangsan GEN Mali all disturb PFV quarrel
 ‘Both Zhangsan and Mali have quarreled (with someone).’
 (19) a. *張三 跟 瑪莉 都 討論 這件事。
 Zhangsan *gen* Mali dou taolun zhejian shi.
 Zhangsan GEN Mali all discuss this-CL matter
 b. 張三 跟 瑪莉 都 討論 過 這件事。
 Zhangsan *gen* Mali] dou taolun guo zhejian shi.
 Zhangsan GEN Mali all discuss EXP this-CL matter
 ‘Both Zhangsan and Mali have discussed this matter (with someone).’

When aspect is encoded into the VO compound *chaojia* ‘to quarrel’ or the collective verb *taolun* ‘to discuss’, the two referents conjoined by *gen* no longer share a reciprocal relationship: the denoted reciprocal event becomes distributive and refers to each referent separately (another participant in this reciprocal event is left unspecified).

3.2.2 *Gen* co-occurring with non-stative non-reciprocal verbs

When *gen* co-occurs with non-stative non-reciprocal verbs, the interpretation of *gen* is varied. Among non-stative non-reciprocal verbs, we note that there is a binary distinction between motion (*cf.* §3.2.2.1) and non-motion verbs, and among non-motion verbs, a distinction exists between non-unidirectional verbs (*cf.* §3.2.2.2) and unidirectional verbs (*cf.* §3.2.2.3).

3.2.2.1 *Gen* co-occurring with motion verbs

When *gen* co-occurs with motion verbs like *qu* ‘to go’, the meaning of *gen* is likely to be ambiguous between a preposition marking comitativity and a verb meaning ‘to follow’. Sentence (20a) entails two readings, paraphrased in (20b) and (20c):

- (20) a. 張三 昨天 跟 李四 去 台北。
 Zhangsan zuotian *gen* Lisi qu taibei.
 Zhangsan yesterday *GEN* Lisi go Taipei
 b. 張三 昨天 跟 著 李四 去 台北。
 Zhangsan zuotian *gen* zhe Lisi qu taibei.
 Zhangsan yesterday follow DUR Lisi go Taipei
 ‘Zhangsan followed Lisi to Taipei yesterday.’
 c. 張三 昨天 跟 李四 一起 去 台北。
 Zhangsan zuotian *gen* Lisi yiqi qu taibei.
 Zhangsan yesterday *GEN* Lisi together go Taipei
 ‘Zhangsan went to Taipei with Lisi yesterday.’

Besides functioning ambiguously as a verb and as a preposition, *gen* is also likely to be analyzed as a conjunction (entailing either a collective or distributive reading). As an illustration, consider (21a-b):

- (21) a. 張三 跟 李四 昨天 一起 去 台北。
 Zhangsan *gen* Lisi zuotian yiqi qu taibei.
 Zhangsan *GEN* Lisi yesterday together go Taipei
 ‘Zhangsan and Lisi went to Taipei together yesterday.’
 b. 張三 跟 李四 昨天 都 去 台北。
 Zhangsan *gen* Lisi zuotian dou qu taibei.
 Zhangsan *GEN* Lisi yesterday all go Taipei
 ‘Both Zhangsan and Lisi went to Taipei yesterday.’

In other words, when *gen* co-occurs with motion verbs such as *qu* ‘to go’, *lai* ‘to come’ and *paodao* ‘to run to’, it can be taken to embody three functions: verb, preposition, conjunction (entailing two different readings, collective or distributive).

3.2.2.2 *Gen* co-occurring with non-motion/non-unidirectional verbs

When *gen* co-occurs with non-motion verbs, like *chi* ‘to eat’, it is unambiguously

interpreted as a preposition marking comitativity, *cf.* (22a). The addition of *yiqi* in (22b) renders this meaning even clearer.

- (22) a. 我 昨天 跟 他 吃 晚餐。
 Wo zuotian *gen* ta chi wancan.
 1SG yesterday *GEN* 3SG eat dinner
 ‘I had dinner with him yesterday.’
- b. 我 昨天 跟 他 一起 吃 晚餐。
 Wo zuotian *gen* ta yiqi chi wancan.
 1SG yesterday *GEN* 3SG together eat dinner
 ‘I had dinner with him yesterday.’

Besides functioning as a preposition, when *gen* co-occurs with these non-motion verbs, it can also be treated as a conjunction; the collective/distributive readings depend on the occurrence of *yiqi* ‘together’ or *dou* ‘all’, as shown in (23a-b).

- (23) a. 我 跟 他 昨天 一起 吃 晚餐。
 Wo *gen* ta zuotian yiqi chi wancan.
 1SG *GEN* 3SG yesterday together eat dinner
 ‘He and I had dinner together yesterday.’
- b. 我 跟 他 昨天 都 吃 了 晚餐。
 Wo *gen* ta zuotian dou chi le wancan.
 1SG *GEN* 3SG yesterday all eat PFV dinner
 ‘Both he and I had dinner yesterday.’

3.2.2.3 *Gen* co-occurring with unidirectional verbs

When the preposition *gen* co-occurs with certain non-motion verbs, it can serve to indicate a one-way relationship between the two referents (Teng 1970), as in (24a-b).

- (24) a. 張三 昨天 跟 李四 借 了 三本 書。
 Zhangsan zuotian *gen* Lisi jie le sanben shu.
 Zhangsan yesterday *GEN* Lisi borrow PFV three-CL book
 ‘Zhangsan borrowed three books from Lisi yesterday.’
- b. 張三 昨天 跟 李四 說 了 一些 話。
 Zhangsan zuotian *gen* Lisi shuo le yixie hua.
 Zhangsan yesterday *GEN* Lisi say PFV some word
 ‘Zhangsan said some words to Lisi yesterday.’

In the above two examples, *gen* serves to introduce the source and the goal of the denoted event. When *gen* co-occurs with source-oriented verbs like *jie* ‘to borrow’, *xue* ‘to learn’, and *mai* ‘to buy’, as in (24a), it introduces another referent which is the *source* of the denoted event. When *gen* co-occurs with goal-oriented verbs like *shuo* ‘to say’, *jieshi* ‘to explain’, and *xuanbu* ‘to announce’, as in (24b), it introduces another referent which is the *goal* of the denoted event.

Note, however, that if *yiqi* is added, *gen* serves as a comitative preposition instead. Compare (24) and (25):

- (25) a. 張三 昨天 跟 李四 一起 借 了 書。
 Zhangsan zuotian *gen* Lisi yiqi jie le shu.
 Zhangsan yesterday *GEN* Lisi together borrow PFV book
 ‘Zhangsan, together with Lisi, borrowed books yesterday.’
 b. 張三 昨天 跟 李四 一起 說 了 一些 話。
 Zhangsan zuotian *gen* Lisi yiqi shuo le yixie hua.
 Zhangsan yesterday *GEN* Lisi together say PFV some word
 ‘Zhangsan, together with Lisi, said some words yesterday.’

Besides, it is possible for *gen* in these cases to function also as a conjunction entailing either a collective reading as in (26a-b) or a distributive reading as in (26c-d).

- (26) a. 張三 跟 李四 昨天 一起 借 了 書。
 Zhangsan *gen* Lisi zuotian yiqi jie le shu.
 Zhangsan *GEN* Lisi yesterday together borrow PFV book
 ‘Zhangsan and Lisi borrowed books together yesterday.’
 b. 張三 跟 李四 昨天 一起 說 了 一些 話。
 Zhangsan *gen* Lisi zuotian yiqi shuo le yixie hua.
 Zhangsan *GEN* Lisi yesterday together say PFV some word
 ‘Zhangsan and Lisi said some words together yesterday.’
 c. 張三 跟 李四 昨天 都 借 了 書。
 Zhangsan *gen* Lisi zuotian dou jie le shu.
 Zhangsan *GEN* Lisi yesterday all borrow PFV book
 ‘Both Zhangsan and Lisi borrowed books yesterday.’
 d. 張三 跟 李四 昨天 都 說 了 一些 話。
 Zhangsan *gen* Lisi zuotian dou shuo le yixie hua.
 Zhangsan *GEN* Lisi yesterday all say PFV some word
 ‘Both Zhangsan and Lisi said some words yesterday.’

In other words, when *gen* co-occurs with unidirectional verbs, it can function as a preposition marking comitativity, a preposition marking unidirectionality or a conjunction (entailing a distributive or collective reading).

3.2.3 Summary

We have shown that when *gen* co-occurs with non-stative verbs, the meanings of *gen* are varied. When *gen* co-occurs with non-stative reciprocal verbs, the interpretations of *gen* are similar to those of *gen* in co-occurrence with stative reciprocal verbs. On the other hand, when *gen* co-occurs with non-stative non-reciprocal verbs, *gen* basically stands to mean one of two possibilities: it can function as a preposition or as a conjunction (two readings: collective or distributive). Furthermore, when it co-occurs with motion verbs, it is likely to be perceived as a verb and when it co-occurs with unidirectional verbs, it can develop another function, i.e. marking a unidirectional relationship. Table 2 provides a summary of this section.

Table 2: Functions of *gen* co-occurring with non-stative verbs

Non-stative verbs			Functions of <i>gen</i>						Co-occurrence restrictions	
			Conj		Prep			V	<i>dou</i>	<i>yiqi</i>
			Dist	Coll	Rec	Com	Unidir			
Reciprocal	Intr	<i>huxiang</i> 'mutual'	–	+	–	–	–	–	–	–
	VO	–ASP	–	+	+	–	–	–	–	–
	Compound	+ASP	+	+	+	–	–	–	+	+
	Transitive	–ASP	–	+	+	–	–	–	–	–
	(Collective)	+ASP	+	+	+	–	–	–	+	+
Non-reciprocal	Motion		+	+	–	+	–	+	+	+
	Non-motion	Non-unidir	+	+	–	+	–	–	+	+
		Unidir	S	+	–	+	+	–	+	+
			G	+	–	+	+	–	+	+

3.3 A binary-distinction analysis at two levels

The behavior of *gen* is subject to the following two parameters, stativity and reciprocity:

(i) When *gen* co-occurs with *stative non-reciprocal* verbs like *qiong* 'to be poor' and *xihuan* 'to like', *gen* can only serve as a conjunction (only one reading: distributive). In co-occurrence with some *measure* or *possessive* verbs (stative/non-reciprocal) such as

you ‘to have, to own’, *gen* can function as a conjunction entailing either a distributive or a collective reading.

(ii) When *gen* co-occurs with *reciprocal* verbs, whether *stative* or *non-stative*, *gen* can usually function either as a preposition or as a conjunction (collective reading), although reciprocal verbs formed by *huxiang* or *xianghu* ‘mutual’ make *gen* function only as a conjunction (only one reading: collective). In co-occurrence with *VO compound* or *collective* verbs marked for aspect, *gen* can also function as a conjunction entailing a distributive reading.

(iii) When *gen* co-occurs with *non-stative non-reciprocal* verbs, the meanings of *gen* are varied. *Gen* can basically serve as a preposition or as a conjunction (two readings possible: distributive or collective). When *gen* co-occurs with *motion* verbs, it is said to retain some verbal traits and can be perceived as a verb. When *gen* co-occurs with certain *non-motion/unidirectional verbs* like *shuo* ‘to say’ and *jie* ‘to borrow’, *gen* as a preposition not only can function as a comitative marker but can also serve to indicate a unidirectional relationship between the two referents.

The two levels of analyses can be schematically represented by Figure 3. Table 3, which combines Table 1 and Table 2, presents in a more comprehensive manner how *gen* behaves in co-occurrence with stative and non-stative verbs.

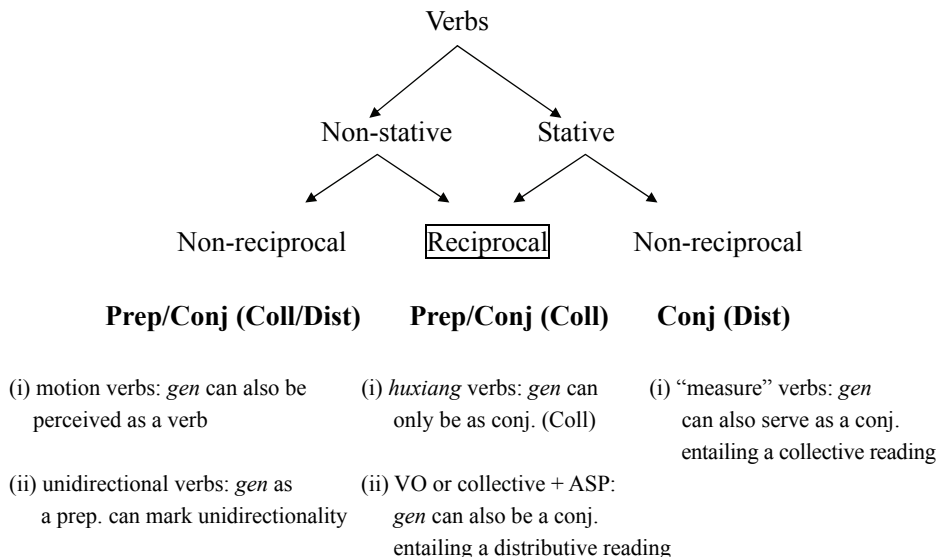


Figure 2: Interpretations of *gen* under a two-level analysis

Table 3: Functions of *gen* co-occurring with stative and non-stative verbs

Types of verbs			Functions of <i>gen</i>						Co-occurrence restrictions		
			Conj		Prep			V	<i>dou</i>	<i>yiqi</i>	
			Dist	Coll	Rec	Com	Unidir				
Stative verbs											
Non-reciprocal	Copula		+	–	–	–	–	–	+	–	
	Intransitive		+	–	–	–	–	–	+	–	
	Tr	Others	+	–	–	–	–	–	+	–	
		<i>you</i> ‘to have’	+	+	–	–	–	–	+	–	
Reciprocal	Copula		–	+	+	–	–	–	–	–	
	Intr	<i>huxiang</i> ‘mutual’	–	+	–	–	–	–	–	–	
Non-stative verbs											
Reciprocal	Intr	<i>huxiang</i> ‘mutual’	–	+	–	–	–	–	–	–	
	VO Compound	–ASP	–	+	+	–	–	–	–	–	
		+ASP	+	+	+	–	–	–	+	+	
	Transitive (Collective)	–ASP	–	+	+	–	–	–	–	–	
		+ASP	+	+	+	–	–	–	+	+	
Non-reciprocal	Motion		+	+	–	+	–	+	+	+	
	Non-motion	Non-unidir	+	+	–	+	–	–	+	+	
		Unidir	S	+	+	–	+	+	–	+	+
			G	+	+	–	+	+	–	+	+

4. Conclusion

It has been shown in the past that:

(i) Chinese uses *gen* as a way to encode comitativity and coordination. This leads to the ambiguity of *gen* in some sentences.

(ii) It is possible to distinguish the ambiguous functions and meanings of *gen* by means of syntactic and semantic tests (insertion of a temporal adverb, negation, clefting, anaphora, and relativization).

(iii) As a preposition, *gen* can also serve to mark unidirectionality and reciprocity in addition to comitativity.

With an attempt to incorporate these earlier contributions into the present paper, we bring forth a model to account for the different behaviors of *gen*. By way of explicit illustration, we take recourse in two figures. Figure 3 features a graded scale, which

includes seven positions from (a) to (g) whereby *gen* co-occurs with different types of verbs and shows different behaviors. Figure 4 integrates this graded scale into our proposed verbal taxonomy and indicates the position at the bottom of each verb type.

VERB/PREP/CONJ				PREP/CONJ		CONJ	
	(a) <i>qu</i> 'go'	(b) <i>chi</i> 'eat'	(c) <i>jie</i> 'borrow'	(d) <i>dajia</i> 'fight'	(e) <i>xiang'ai</i> 'love each other'	(f) <i>you</i> 'have'	(g) <i>qiong</i> 'be poor'
Verb type	Non-stative			Non-stative/Stative		Stative	
Function	Non-reciprocal			Reciprocal		Non-reciprocal	
Verb	+	–	–	–	–	–	–
Prep (Com)	+	+	+	–	–	–	–
Prep (Unidir)	–	–	+	–	–	–	–
Prep (Rec)	–	–	–	+	–	–	–
Conj (Coll)	+	+	+	+	+	+	–
Conj (Dist)	+	+	+	–	–	+	+

Figure 3: Behavior of *gen*

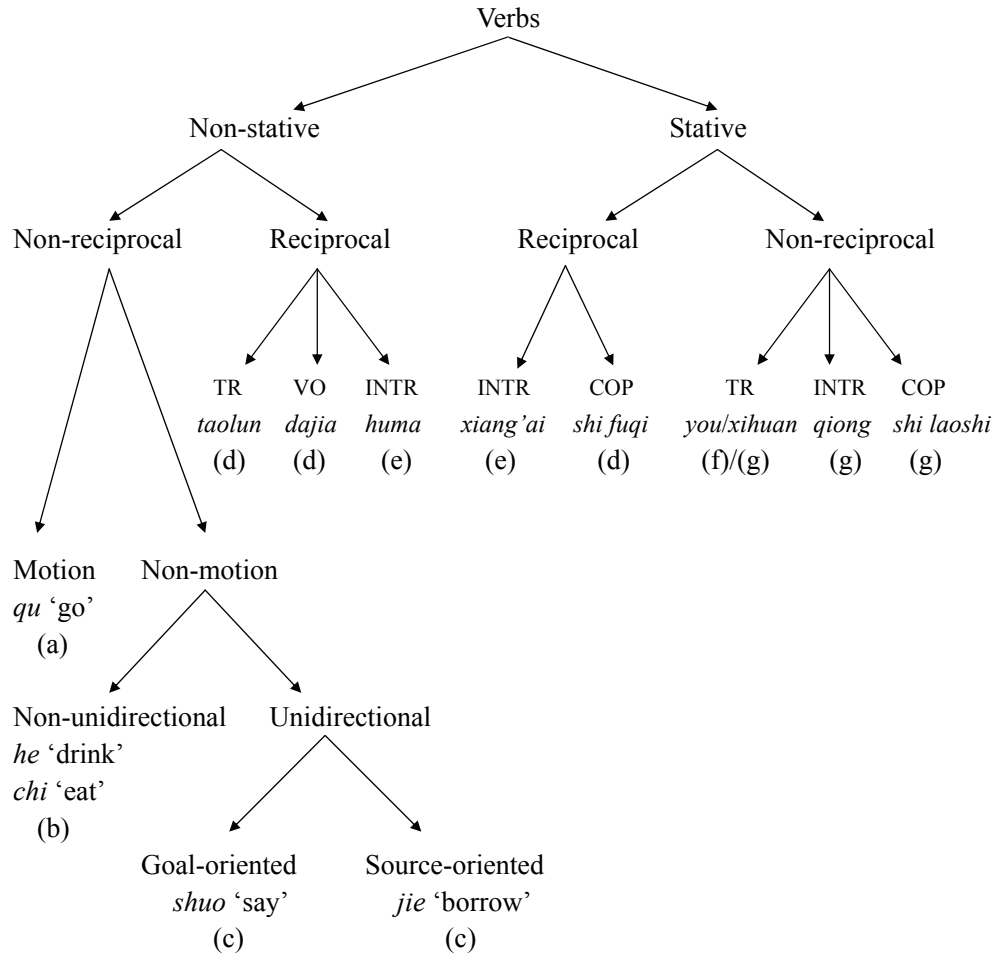


Figure 4: Behaviour of *gen* within our verbal taxonomy

References

- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chuang, Stephen Yun-hsiang. 2007. *A Semantic and Syntactic Study of gen in Taiwan Mandarin*. Taipei: National Taiwan Normal University MA thesis.
- Frajzyngier, Zygmunt. 2000. Coding of the reciprocal function: two solutions. *Reciprocals: Form and Function*, ed. by Zygmunt Frajzyngier & Traci S. Curl, 179-194. Amsterdam & Philadelphia: John Benjamins.
- Furukawa, Yutaka. 2000. 'Gen' zi de yuyi zhixiang ji qi renzhi jieshi [The semantic orientation and cognitive interpretation of *gen*]. *Yuyan Jiaoxue yu Yanjiu* [Language Teaching and Linguistic Studies] 2000.3:37-44.
- Haspelmath, Martin. 2004. Coordinating constructions: an overview. *Coordinating Constructions*, ed. by Martin Haspelmath, 3-39. Amsterdam & Philadelphia: John Benjamins.
- Her, One-Soon. 1991. *Grammatical Functions and Verb Subcategorization in Mandarin Chinese*. Taipei: Crane.
- Lakoff, George, and Stanley Peters. 1969. Phrasal conjunction and symmetric predicates. *Modern Studies in English*, ed. by David A. Reibel & Sanford A. Schane, 113-142. Englewood Cliffs: Prentice-Hall.
- Lichtenberk, Frantisek. 1994. Reflexives and reciprocals. *The Encyclopedia of Language and Linguistics*, ed. by R. E. Asher & J. M. Y. Simpson, 3504-3509. Oxford & New York: Pergamon Press.
- Lichtenberk, Frantisek. 2000. Reciprocals without reflexives. *Reciprocals: Form and Function*, ed. by Zygmunt Frajzyngier & Traci S. Curl, 31-62. Amsterdam & Philadelphia: John Benjamins.
- Liu, Jian, and Alain Peyraube. 1994. History of some coordinative conjunctions in Chinese. *Journal of Chinese Linguistics* 22.2:179-201.
- Liu, Mei-chun. 2000. Reciprocal marking with deictic verbs *come* and *go* in Mandarin. *Reciprocals: Form and Function*, ed. by Zygmunt Frajzyngier & Traci S. Curl, 123-131. Amsterdam & Philadelphia: John Benjamins.
- Paris, Marie-Claude. 2004. On parts of speech in contemporary Chinese (and elsewhere). Paper presented at the Syntactic Categories and their Interpretation in Chinese: A Workshop on the Syntax and Semantic of Chinese. Budapest: Research Institute of Linguistics, Hungarian Academy of Science.
- Paris, Marie-Claude. 2005. Mandarin *gen* and French *et/avec*: another look at distributivity and collectivity. Paper presented at the 3rd International Conference in Contrastive Semantics and Pragmatics, September 16-18, 2005. Shanghai: Shanghai International Studies University.

- Paris, Marie-Claude. 2007a. L'interaction entre *gen* et la négation en chinois contemporain. *CLALIX* 20:147-166.
- Paris, Marie-Claude. 2007b. Nouveau regard sur "Le problème linguistique des prépositions et la solution chinoise". *Combat pour les langues du monde, Hommage à Claude Hagège, sous la responsabilité de M. M. J. Fernandez-Vest*. Collection Grammaire et Cognition 4 & 5: 401-410. Paris: L'Harmattan.
- Paris, Marie-Claude. 2008. On parts of speech in Chinese: *gen*. *The Linguistic Review* 25.3-4:347-366.
- Paris, Marie-Claude. 2010. Mandarin *gen* and French *et/avec*: another look at distributivity and collectivity. *Contrasting Meaning in Languages of the East and West*, ed. by Dingfang Shu & Ken Turner, 517-530. Oxford & New York: Peter Lang.
- Stassen, Leon. 2000. AND-languages and WITH-languages. *Linguistic Typology* 4.1:1-54.
- Tang, Ting-chi. 1990. *Studies in Chinese Syntax*. Taipei: Student Book.
- Teng, Shou-hsin. 1970. Comitative versus phrasal conjunction. *Papers in Linguistics* 2.2:315-358.
- Teng, Shou-hsin. 1977. *A Semantic Study of Transitivity Relations in Chinese*. Taipei: Student Book.
- Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca: Cornell University Press.

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

Liejiong Xu (徐烈炯)
City University of Hong Kong

The A-not-A question is a special form of Chinese interrogative sentences. Some grammarians have claimed that it is a special type of wh-questions. This paper shows that this claim is unconvincing because the arguments based on island constraints and intervention effects are not well supported by evidence. It concludes that it is more natural to classify A-not-A questions as alternative questions.

Key words: A-not-A question, alternative question, wh-question, island constraint, intervention effect, focus

1. Typology of questions

I classify questions pragmatically into three types. The person that asks the question may present a proposition and ask for confirmation or disconfirmation. I call this type of question “confirmatory question”. The person may present more than one option and ask for a choice. I call this type of question “alternative question”. The person may ask for a choice among options not explicitly mentioned, but the domain is understood or demarcated to a certain extent. This is what is commonly called “constituent question” or “wh-question”.

The above classification is pragmatic. Each type may be syntactically realized in various forms across languages and even in one language. Each form may have a special term in the grammar of a language. For instance, the so-called yes-no question in English is a kind of confirmatory question and the particle question in Chinese ending in *ma* is also a kind of confirmatory question. They have different forms and different names.

Now let us consider a special kind of Chinese interrogative form known as “A-not-A question” or “V-not-V question” since Chao (1968). It is formed by duplicating the predicate (typically a verb or an adjective) or part of it with the negation morpheme between the two copies, e.g. *lai-bu-lai?* ‘come-not-come’; *leng-bu-leng?* ‘cold-not-

* I would like to thank D. Terence Langendoen for his comments on this paper.

cold'.¹ Its form suggests that it belongs to the category of alternative question. However, it has been claimed by some linguists that it is more similar to wh-questions since the two share two important properties, viz. island constraints and intervention effects. I will address the two topics respectively in the next two sections.

Before going any further, let us clarify a minor argument in the literature in favor of lumping the A-not-A question and the wh-question together into one class. It is so argued in Hsieh (2004), etc. on the basis of the fact that neither of them can be answered by *shi de* 'yes' and that either of them may co-occur with emphatic adverbs like *daodi* 'at all', *jiujing* 'ever'. But these facts only show that the A-not-A question is not a confirmatory question. Only confirmatory questions can be answered by yes or no. One cannot use yes or no to indicate a choice between two options presented without any inclination toward one or the other. Only confirmation questions exclude the appearance of adverbs like *daodi*, *jiujing*. One does not use an adverb that expresses high uncertainty when one has formed one's opinion and merely asks for confirmation. Since these tests only single out the confirmatory question but do not distinguish the alternative question from the wh-question, they fail to prove that the A-not-A question is a kind of wh-question.

2. Island constraints

It was first observed in Huang (1982) that the A-not-A question, unlike the alternative question, cannot appear in island constructions. On the basis of this observation he hypothesized that the A-not-A question contains an abstract question morpheme (Q-operator), which, just like a wh-operator, undergoes LF-movement to the scope position of CP. Such a movement is subject to the familiar island conditions such as the Complex Noun Phrase Constraint and the Sentential Subject Constraint. (Examples will be provided shortly.)

But not all wh-expressions in Chinese are sensitive to these constraints. All nominal wh-expressions are not, so a more precise characterization is that:

- (1) Questions with adjunct wh-expressions and A-not-A questions are subject to the island constraints.

It was pointed out in Xu (1990), etc. that even this generalization is too strong. In fact, some adjunct wh-expressions behave like nominal wh-expressions. Although wording

¹ Following the convention in grammar, I put hyphens to call attention to A-not-A forms like *lai-bu-lai*. There is no hyphen in writing Chinese characters.

like (1) continues to appear in the recent literature, what it actually means is that

- (2) Questions with *weishenme* ‘why’ and A-not-A questions are subject to the island constraints.

If island constraints are diagnostic criteria in the classification of questions, then (2) disconfirms rather than confirms that the A-not-A question should be grouped together with the wh-questions because most wh-questions are insensitive to them.

Since the A-not-A question is conceptually a kind of alternative question, the next step is to see if other alternative questions are also subject to island constraints. If some of them, not necessarily all of them, do, then island constraints should not be used as criteria for the classification of questions at all.

To ascertain whether the A-not-A question and the alternative question differ systematically it is necessary to test sentences in which each of the two forms appears in island configurations. In so doing we must be aware of the fact that grammaticality or acceptability is complicated by other factors as well.

Let us consider simple alternative questions and A-not-A questions first. The following three forms have no difference in acceptability.

- (3) a. Ta renshi ni haishi bu renshi ni?
 he know you or not know you
 ‘Does he know you or doesn’t know you?’
 b. Ta renshi haishi bu renshi ni?
 he know or not know you
 c. Ta renshi bu renshi ni?
 he know not know you

Hagstrom (2006) uses each of the above as a relative clause and reports that in the two alternative questions with *haishi* ‘or’, (4a) and (4b), are equally good in contrast to the starred A-not-A question (4c).²

- (4) a. Ni xihuan [renshi ni haishi bu renshi ni] de ren?
 you like know you or not know you MM³ person
 ‘Do you like people who know you or people who don’t know you?’

² (4a) is his example (30a), (4b) is his (30c) and (4c) is his (32).

³ The Chinese relative clause precedes its head and is closed by the particle *de*, which, in fact, is a modifier marker (glossed MM), used to mark modifiers of any syntactic category.

- b. Ni xihuan [renshi ni haishi bu renshi] de ren?
you like know you or not know MM person
'Do you like people who know you or don't?'
- c. *Ni xihuan [renshi-bu-renshi ni] de ren?
you like know-not-know you MM person
'Do you know people who know you or don't know you?'

In my intuitive judgments, not only (4c) is less acceptable, but (4b) also requires more time and effort to process and understand than (4a). I suspect that the English translation of (4b) is less acceptable compared with that of (4a). My suspicion is confirmed by intuitions of some native speakers of English. The elliptical forms of both languages may be less acceptable for the same reason. Whatever it is, it cannot be fully accounted for by singling out the A-not-A form.

Alternative questions can appear in various forms as the result of repeated application of conjunction reduction or anaphoric ellipsis. The more reduction or ellipsis takes place, the more time it takes for mental processing and understanding. The following forms show a decreasing degree of acceptability.

- (5) a. Ni xihuan renshi ni de ren haishi xihuan bu renshi ni de ren?
- b. Ni xihuan renshi ni de ren haishi bu renshi ni de ren?
- c. Ni xihuan renshi ni de haishi bu renshi ni de ren?
- d. Ni xihuan renshi ni haishi bu renshi ni de ren?
- e. Ni xihuan renshi haishi bu renshi ni de ren?
- f. Ni xihuan [renshi bu renshi ni] de ren?

Although the A-not-A question (5f) may not be syntactically derived in the same way as the other forms, it is obviously even less complete than any of those above it and therefore not surprisingly less acceptable.

The so-called complex NP in Chinese is in the form of a relative clause followed by a head noun. An alternative question anchored in such a construction is potentially open to two interpretations. Let us call them the N-alternative reading and the V-alternative reading. Take for example *ni maile haishi mei mai de shu* 'you-bought-or-not-bought-MM-book'. The N-alternative reading asks for a choice between two entities. One is a book which you bought and the other is a book which you didn't buy. The V-alternative reading asks for a choice between the two descriptions of the same entity. In such a case there is only one book involved. The question is whether it should be described as a book which you bought or as a book which you didn't buy. The latter interpretation leads to unacceptability. So what is crucial for acceptability is whether the expression under consideration can be interpreted as N-alternative.

As has been shown, ellipsis affects acceptability. But it is not the only factor. In the above examples in (4) and (5) the head of the relative clause *ren* ‘person/people’ represents a generic class of people. If the head refers to a specific individual, acceptability decreases in most of the forms.⁴ Compare (4a) with (6) below.

- (6) *Ni xihuan nage renshi ni haishi bu renshi ni de ren?
 you like that know you or not know you MM person
 ‘*Do you like the person who knows you or who doesn’t know you?’

If only one specific person is involved, it is impossible to obtain an N-alternative reading. The English counterpart of (6) is equally impossible.

Chinese generally does not use morphological forms to distinguish the singular from the plural or the generic from the specific; nevertheless, such an underlying difference does exist and it affects the acceptability of A-not-A questions in different configurations. Compare the following pairs.

- (7) a. *Xin bu xin Shangdi de ren zai yanshuo?
 Believe-not-believe God MM person is speaking
 b. Ni ziji shuyu xin bu xin Shangdi de ren?
 yourself belong-to believe-not-believe God MM people
 ‘Do you yourself belong to people who do or don’t believe in God?’
- (8) a. *Jia bu jia tang de kafei zai zhuozi shang?
 Add-not-add sugar MM coffee PREP table on
 b. Ni pingshi he jia bu jia tang de kafei?
 you normally drink add-not-add sugar MM coffee
 ‘Do you normally drink coffee with or without sugar?’

Why are the b-forms not as bad as the a-forms? In the a-forms, the A-not-A question is part of the subject. It is well known that the subject tends to serve as the topic and therefore tends to be interpreted as specific. Some examples cited in the literature and intended to illustrate island effects are unacceptable on this account. In the b-forms, the A-not-A question is contained in the predicate or in the object and is therefore non-specific and more acceptable.

⁴ Only in (5a) and (5b) are there two coordinated head nouns. Hence it is possible to have an N-alternative reading. They are alternative questions. There cannot be two head nouns in an A-not-A question.

The specific-nonspecific contrast exists in alternative questions as well as in A-not-A questions.

- (9) a. *Xin haishi bu xin Shangdi de ren zai yanshuo?
believe or not believe God MM person is speaking
b. Ni ziji shuyu xin haishi bu xin Shangdi de ren?
yourself belong-to believe or not believe God MM people
'Do you yourself belong to people who do or don't believe in God?'
- (10) a. *Jia haishi bu jia tang de kafei zai zhuozi shang?
add or not add sugar MM coffee PREP table on
b. Ni pingshi he jia haishi bu jia tang de kafei?
you normally drink add or not add sugar MM coffee
'Do you normally drink coffee with or without sugar?'

So while we do see ungrammatical sentences cited in the literature with the A-not-A form appearing in island configurations, it needs more solid evidence to prove that those sentences are free of other complications and that A-not-A questions and alternative questions differ systematically with regard to the Complex NP Constraint.

Let us move on to sentential subjects. Some A-not-A questions contained in a sentential subject, too, are not absolutely unacceptable. Their difference from the corresponding alternative questions is minimal.

- (11) a. [Jiu li jia bu jia bingkuai] hao?
wine in add not add ice.cubes better
'Which is better, adding ice cubes in wine or not?'
- b. [Jiu li jia haishi bu jia bingkuai] hao?
wine in add or not add ice.cubes better

To the best of my knowledge, such examples have never been documented in the literature. There may be other reasons why some A-not-A questions are less acceptable than others. It needs further research. If the common belief that all A-not-A questions obey the island constraints and no alternative questions do is unreliable, one is forced to further revise (2) as (12).

- (12) Some wh-questions, some A-not-A questions and some alternative questions are subject to the island constraints.

3. Intervention effects

The combination of a wh-word with a quantificational or focusing element leads to ungrammaticality in certain configurations. This is known as intervention effects. This issue has received attention for more than ten years since Beck (1996).

Soh (2005) cites examples to show intervention effects in Chinese.

- (13) *Ta ye qu-bu-qu?
 he also go not go
 ‘Is or isn’t he also going?’
- (14) *Ta ye weishenme ma ta?
 he also why scold he
 ‘Why did he also scold him?’

From such examples she makes the following generalization:

- (15) Questions with wh-adjuncts and A-not-A questions are sensitive to intervention effects.

Why are sentences like (13) and (14) ungrammatical? Soh assumes that the Q-operator in the A-not-A question and the wh-word *weishenme* ‘why’ must undergo covert feature movement to C and Spec C respectively at LF and that the movement is blocked because the focus element *ye* intervenes between its original position and landing position. In note 3 of Soh (2005:145) she tries to justify her treatment of the Q-operator in an A-not-A question as an adverbial wh-expression. It is because it patterns like *weishenme* in its inability to escape an island. But it is not clear why a construction that violates island constraints should show intervention effects as well.

In the following sections we consider why focus elements and quantifiers give rise to intervention effects and whether alternative questions as well as A-not-A questions are sensitive to such effects. Finally, we briefly observe whether wh-nominals as well as wh-adjuncts are sensitive to intervention effects.

3.1 Focus elements

The focus elements, e.g. *ye* ‘also’, *zhi* ‘only’, cited in Soh (2005) are taken from a category of adverbs that have been found to affect the grammaticality or acceptability of A-not-A questions in previous research. This phenomenon has been well documented in the literature, e.g. Li & Thompson (1979), Tang (1981, 1984), Ernst (1994), Zhang (1996),

Wu (1997a, 1997b) and is revisited more recently in Law (2006). The researchers all agree that the presence of some adverbs, but not others, restrict the formation of A-not-A questions, but they use different names to call them and propose different accounts. Zhang (1996) provides a list of 14 types of adverbs, which he calls “presupposition triggers”. With the presence of such an adverb, the verb is interpreted as presupposed. He states the constraint as follows: V of A-not-A questions cannot be presupposed. Law (2006) calls those adverbs “predicate-related adverbs” and observes that a sentence that contains a predicate-related adverb entails one without it. For example, *ta ye qu* ‘he is also going’ entails *ta qu* ‘he is going’. He prefers to use the logical term entailment, but it is not different from Zhang’s presupposition in effect. The distinction between his predicate-related adverbs and predicate unrelated adverbs is semantic, even though his solution appears to be syntactic, which hypothesizes that a predicate-related adverb is a potential A-bar binder that prevents the A-not-A operator from moving across it at LF. This account is similar to the one proposed in Soh (2005), in which the blocker is called an intervener.

If presupposition or entailment, whatever it is called, is the reason for the unacceptability of A-not-A questions like (13), then one naturally expects the alternative questions parallel to them to be similarly restricted. The expectation is borne out. As we can see, (16) is indeed as unacceptable as (13).

- (16) *Ta ye qu haishi bu qu?
 he also go or not go
 ‘Is or isn’t he also going?’

Wu (1997a, 1997b) provides an alternative explanation for the unacceptability of sentences like (13). It is a model-theoretic account, but it also involves the notion of presupposition. Questions can be viewed as partitions in the light of Groenendijk & Stokhof (1984) and Higginbotham (1993). A question, expressed by an interrogative form, is a partition of the possible states of nature into cells. An A-not-A question makes a bipartition into two cells C_1 and C_2 , one representing the affirmative and the other the negative. The two cells must be mutually exclusive in the sense that they do not overlap. The partition must be exhaustive in the sense that there cannot be a state that is neither C_1 nor C_2 . These are the semantic conditions an A-not-A question must satisfy. In the case of (13), the partition is between C_1 (the person is going with someone else or going like someone else) and C_2 (the person is not going with someone else or not going like someone else). Thus C_1 presupposes someone else is going but C_2 presupposes someone else is not going. It is obviously a contradictory presupposition. If Wu’s account is correct, it would be strange that it should not apply to alternative questions like (16) as well.

Let us compare another minimal pair.

- (17) *Ta you qu-bu-qu?
 he again go not go
 ‘Is or isn’t he going again?’
- (18) *Ta you qu haishi bu qu?
 he again go or not go

Both the A-not-A question and the alternative question are asking for a choice between C_1 (the person is going again) and C_2 (the same person is again not going). The former presupposes that he went last time and the latter presupposes that he didn’t go last time. Once again there is obviously a contradiction in presupposition. Thus the partition view of questions rules out the alternative question as well as the A-not-A question.

Now we insert the focus element *you* ‘again’ into another alternative question in a different syntactic construction. Likewise, it is not acceptable.

- (19) *Ta you he kafei haishi cha?
 he again drink coffee or tea
 ‘Did he drink coffee or tea again?’

This alternative question is asking for a choice between C_1 (he drank coffee again) and C_2 (he drank tea again). The former presupposes that he had drunk coffee last time and the latter presupposes that he had drunk tea last time. Once again, the alternative question involves contradictory presuppositions. But this time there is no A-not-A question parallel to it. If one follows Soh (2005)’s proposal, it would be most natural to assume that there is a Q-operator in the alternative question and that intervention effects obtain where the focusing adverb blocks its covert feature movement to C. But she only points out that A-not-A questions are comparable to wh-questions with *weishenme*, but does not mention that they share the relevant property with alternative questions.

3.2 Quantificational expression

Another kind of intervener is quantifiers. In Soh (2005)’s analysis the movement of the A-not-A operator will be blocked if a quantificational expression like *meiyou ren* ‘nobody’ appears along its path of movement.

- (20) *Meiyou ren hui-bu-hui qu Meiguo?
 no one will not will go US
 ‘Will or won’t no one go to the US?’

Again, it is not mentioned that quantificational expressions produce similar intervention effects in other alternative questions. But evidently they do as well.

- (21) *Meiyou ren hui qu haishi bu hui qu Meiguo?
 no one will go or not will go US
 ‘Will or won’t no one go to the US?’

- (22) *Meiyou ren hui qu Meiguo haishi Yingguo?
 no one will go US or UK
 ‘Will no one go to the US or UK?’

The partition account can deal with these sentences. The questions in (20) and (21) ask for a choice between C_1 (nobody will go to the US) and C_2 (nobody won’t go to the US). But this is not an exhaustive partition as it leaves the possible state that some will go and others won’t, thus violating the semantic condition of exhaustivity on the partition view of questions. The same condition is imposed on the A-not-A question and on the alternative question. The questions in (22) ask for a choice between C_1 (nobody will go to the US) and C_2 (nobody won’t go to the UK). This, too, is not an exhaustive partition as it leaves the possible state that some will go to the US and others will go to the UK. If a referential expression replaces the quantifier in the above examples, the sentences become acceptable since now the partition is exhaustive.

One may observe that the unreduced forms of alternative questions are more acceptable. Compare (23) with (21) and (24) with (22).

- (23) Meiyou ren hui qu Meiguo haishi meiyou ren bu hui qu Meiguo?
 no one will go US or no one not will go US
 ‘Will no one go to the US or will no one not go to the US?’

- (24) Meiyou ren hui qu Meiguo haishi meiyou ren hui qu Yingguo?
 no one will go US or no one will go UK
 ‘Will no one go to the US or will no one go to the UK?’

Does this show that the partition account leaks and that, after all, alternative questions are different from A-not-A questions? No. What is the context in which questions like (23) and (24) are appropriately used? They are typically used in the contexts of echo questions. For instance, A hears B say nobody would like to go to the US but in A’s opinion UK is the last place one may choose to go to. Then A may use the unreduced form (24) to ask B for confirmation, but he will not use the reduced form (22) in such a case. B’s reply to question (23) is probably (25a) but not (25b).

- (25) a. Meiyou ren hui qu Meiguo.
 no one will go US
 ‘No one will go to the US.’
 b. Meiguo.
 US

3.3 Focus crossing

Hagstrom (2006) explicitly denies that intervention effects appear in alternative questions with the explicit presence of the conjunction *haishi*. He cites an example that never appears elsewhere in the literature, which is numbered below as (26) in contrast to the A-not-A question in (27).⁵

- (26) Ni zhi xihuan yige ren haishi bu zhi xihuan yige ren?
 you only like one person or not only like one person
 ‘Do you only like one person or not only like one person?’
 (27) *Ni zhi xi bu xihuan yige ren?
 you only like-not-like one person
 ‘Do you only like one person?’

This appears to be new evidence in support of his claim that the A-not-A question differs from the alternative questions. Although we native speakers of Chinese agree with him about the grammaticality judgment of (26), we find on closer inspection that this sentence does not form a minimal pair with (27). A significant comparison should be made between (27) and (28).

- (28) *Ni zhi xihuan yige ren haishi zhi bu xihuan yige ren?
 you only like one person or only not like one person
 ‘Do you only like one person or only not like one person?’

In (28) the focus element *zhi* precedes and takes scope over the negation marker *bu* just as it does in (27), but in (26) their order is reversed. A-not-A questions (and their corresponding alternative questions) ask for a choice between V and not-V, in this case for a choice between *xihuan* ‘like’ and *bu-xihuan* ‘not like’, not one between *zhi* ‘only’ and *bu zhi* ‘not only’. This correction is crucial because now the alternative question with *haishi* turns out to be as unacceptable as the corresponding A-not-A question without the presence of the conjunction. Whereas (26) supports his claim about the difference

⁵ (26) is his example (61) and (27) is his example (52).

between the two question forms, (28) is a counterexample to it.

Why is (26) acceptable and why is (28) unacceptable? The contrast lies in the latter part of the conjoined VP, repeated below for easy comparison.

- (29) a. *bu zhi xihuan yige ren*
 not only like one person
 b. **zhi bu xihuan yige ren*
 only not like one person

Both of (29a) and (29b) contain two adverbs, *bu* ‘not’ and *zhi* ‘only’. Each of them is a focus-sensitive operator to be associated with an expression interpreted as focus. Focus-sensitive operators in Chinese are relatively flexible in word order. The general rules are (i) that it must be placed in a position that c-commands the expression targeted for focalization and (ii) that the former must be placed as close to the latter as possible. In (29a) the focus of *bu* is *zhi xihuan yige ren* ‘only like one person’ or simply *zhi* ‘only’ and the focus of *zhi* is most likely to be *yige ren* ‘one person’. In (29b) the focus of *zhi* is also most likely to be *yige ren* but then the association between the operator and the focus is intervened by another operator *bu*. Where there are two operators and two foci, their association chains should be in serial order but do not cross each other. This accounts for the contrast between (26) and (28) and accounts for the unacceptability of (27) and (28). This issue is not relevant to the difference between the A-not-A question and the alternative question.

3.4 Other wh-questions

Soh (2005) provides examples to show that *weishenme* ‘why’ is sensitive to intervention effects, but claims that all wh-adjuncts are, including the A-not-A operator. In this last section we briefly consider questions with wh-nominals. Compare the following sentences:

- (30) a. **Ta zhi shenme shihou mai-le diannao?*
 he only what time bought computer
 ‘Only when did he buy a computer?’
 b. **Ta zhi weishenme mai-le diannao?*
 he only why bought computer
 ‘Only why did he buy a computer?’

Soh takes expressions like *shenme shihou* ‘what time’, *nali* ‘what place’ as wh-nominals. With the presence of the focus element *zhi* ‘only’, neither of the above sentences is

good. Some speakers may find (30a) better than (30b). It is when they take it as an echo question. We know that a question asking for reasons cannot be an echo question (Hu 2002).

Now let us try wh-questions containing a quantifier as a potential intervener.

- (31) a. Shenme jijie meiyou ren qu gongyuan?
 what season no one go park
 ‘In what seasons no one goes to park?’
 b. *Meiyou ren shenme jijie qu gongyuan?
 no one what season go park

The contrast between (31a) and (31b) is parallel to Soh’s examples cited below:

- (32) a. Weishenme meiyou ren cizhi?
 why no one resign
 ‘Why did nobody resign?’
 b. *Meiyou ren weishenme cizhi
 no one why resign

If (32b) is ungrammatical because *weishenme* undergoes covert feature movement to Spec C at LF and thus the movement is blocked by the intervening quantifier, then the same thing must happen with wh-nominals like *shenme jijie* ‘what season’ in (31b). So there is no difference between wh-adjuncts and wh-nominals in this respect.

Therefore the generalization stated in (15) should be revised as

- (33) Wh-questions, alternative questions and A-not-A questions are all sensitive to intervention effects.

With regard to intervention effects, it is unnecessary to split wh-questions into nominal wh-questions and adverbial wh-questions, it is unnecessary to take A-not-A questions out of alternative questions and it is unnecessary to join A-not-A questions with adverbial wh-questions.

Now we return to the observation made at the beginning of this article. The A-not-A question is similar to alternative questions in form and in meaning. However, for a quarter of a century some linguists have claimed that it is a special type of wh-question. I have shown here that the claim is unconvincing because the arguments based on island constraints and intervention effects are not well supported by evidence. The conclusion is that it is more natural to classify A-not-A questions as alternative questions.

References

- Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4.1:1-56.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Ernst, Thomas. 1994. Conditions on Chinese A-not-A questions. *Journal of East Asian Linguistics* 3.3:241-264.
- Groenendijk, Jeroen, and Martin Stokhof. 1984. On the semantics of questions and pragmatics of answers. *Varieties of Formal Semantics*, ed. by Fred Landman & Frank Velman, 143-170. Dordrecht: Foris.
- Hagstrom, Paul. 2006. A-not-A questions. *The Blackwell Companion to Syntax*, Vol. 1, ed. by Martin Everaert & Henk van Riemsdijk, Ch.7: 173-213. Malden: Blackwell.
- Higginbotham, James. 1993. Interrogatives. *The View From Building 20*, ed. by Kenneth Hale & Samuel Jay Keyser, 195-228. Cambridge: MIT Press.
- Hsieh, Miao-Ling. 2004. On the licensing of A-not-A forms in Chinese and the DP hypothesis. *Concentric: Studies in Linguistics* 30.2:68-92.
- Hu, Jianhua. 2002. *Prominence and Locality in Grammar*. Hong Kong: City University of Hong Kong dissertation.
- Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge: MIT dissertation.
- Law, Paul. 2006. Adverbs in A-not-A questions in Mandarin Chinese. *Journal of East Asian Linguistics* 15.2:97-136.
- Li, Charles N., and Sandra A. Thompson. 1979. The pragmatics of two types of yes-no questions in Mandarin and its universal implications. *Chicago Linguistic Society (CLS)* 15:197-206. Chicago: Chicago Linguistic Society.
- Lü, Shuxiang. 1985. Yiwen, fouding, keneng [Question, negation, possibility]. *Zhongguo Yuwen [Studies of the Chinese Language]* 1985.4:241-250.
- Soh, Hooi Ling. 2005. Wh-in-situ in Mandarin Chinese. *Linguistic Inquiry* 36.1:143-155.
- Tang, Ting-chi. 1981. Guoyu yiwenju de yanjiu [A study of interrogative sentences in Mandarin]. *Journal of Taiwan Normal University* 26:219-277.
- Tang, Ting-chi. 1984. Guoyu yiwenju xulun [A further study of interrogative sentences in Mandarin]. *Journal of Taiwan Normal University* 29:381-437.
- Wu, Jianxin. 1997a. A model-theoretic approach to A-not-A questions. *U. Penn Working Papers in Linguistics* 4.2:273-289.
- Wu, Jianxin. 1997b. More on A-not-A questions: a model-theoretic approach. *Proceedings of the 16th West Coast Conference on Formal Linguistics (WCCFL 16)*, ed. by Emily Curtis, James Lyle & Gabriel Webster, 463-477. Stanford: CSLI.

- Xu, Liejiong. 1990. Remarks on LF-movement in Chinese questions. *Linguistics* 28.2: 355-382.
- Zhang, Zhengsheng. 1996. Focus, presupposition and the formation of A-not-A questions in Chinese. *Journal of Chinese Linguistics* 25.2:227-257.

詞義演變三例

蔣紹愚
北京大學

本文討論了三種不同類型的詞義演變及其演變的原因：(一)“快”：愉快——快速。“慢”：怠慢——緩慢。演變的原因是人們覺得“愉快——快速”和“怠慢——緩慢”之間有關係。(二)“念”：心念——口誦。演變的原因是“念”有一個“心念口誦”的過渡階段。(三)“要”：想要——如果。演變的原因是“語境吸收”。

關鍵詞：快，慢，念，要，詞義演變

1. 引言

本文討論三種不同類型的詞義演變，著重分析其演變的原因。

2. “快”、“慢”的詞義演變

“快”在上古是“愉快”的意思，屬於情感、情緒範疇。後來變成“快速”的意思，屬於空間位移或時間推移的速度範疇。這在汪維輝(2000)《東漢—隋常用詞演變研究》中已經有了很好的說明，此處不贅。本文要討論的是：“快”的詞義為什麼能從“愉快”變為“快速”？這兩個意義之間是否有聯繫？

和“快”的詞義演變相關的是“慢”。“慢”原來的意思是“傲慢”、“怠慢”，也屬於情感、情緒範疇。後來變成“緩慢”，也屬於空間位移或時間推移的速度範疇。我們也要問：“慢”為什麼能從“傲慢”、“怠慢”變成“緩慢”？這兩個意義之間是否有聯繫？

2.1 “快”的詞義演變

從詞義分析的角度看，“快”的“愉快”義是人的主觀感覺，“快”的“快

速”是事物的客觀屬性，但在某種情況下，這兩者是有一定聯繫的。有的時候，“快”的意義既不是完全主觀的“愉快”，也不是完全客觀的“快速”，而是介於這兩者之間的“暢快/痛快”，“暢快/痛快”既可以是人的主觀感覺，又可以引起人們這種主觀感覺的事物的客觀性狀。而這種客觀性狀的一個主要因素就是“快速”。古漢語中一些作修飾語的“快”，如果譯成現代漢語，最準確的應該是“暢快/痛快”。請看下面幾個詞語。

2.1.1 “快戰”

- (1) 今日固決死，願為諸君快戰。（《史記·項羽本紀》）
- (2) 進不得快戰以徼功，退不得溫飽以全命。（《後漢書·皇甫規傳》）

這裡的“快”都是“暢快/痛快”的意思。但什麼樣的戰鬥能使人感到暢快/痛快呢？項羽說：“為諸君潰圍、斬將、刈旗。”皇甫規說：“願假臣兩營二郡屯列坐食之兵五千，出其不意，與護羌校尉趙沖共相首尾……”都是突擊性的速決戰。

後代用“快戰”一詞也都是此義，如：

- (3) 曷飛等見萇兵少，悉眾來攻。萇固壘示弱，潛遣其子崇率精騎數百，出不意乘敵後，敵擾逆擊，遂潰，斬曷飛及將士萬餘。……萇曰：“吾自束髮以來，未嘗快戰如此，以千六百人破三萬眾，國勢由此克舉。”（《晉略·國傳6》）
- (4) （虞允文）乃復書……願快戰以決雌雄。亮得書大怒……召諸將，約以三日濟江，否則盡殺之。（《宋史紀事本末·卷18》）
- (5) 力在快戰不願守，直令疾鬪而血鏖。（王令《寄王正叔》）

可見“快戰”都是突擊性的速決戰，而且常常是出其不意的，不會持久。這種戰鬥令人感到暢快/痛快。

2.1.2 “快雨”

- (6) 向來快雨，想君佳，方得此雨為佳，深為欣嘉。（王羲之《雜帖》）

“快雨”是使人感到暢快/痛快的雨。但什麼雨能使人感到暢快/痛快呢？

- (7) 過清河倪太守時，天旱，倪問輅雨期。輅曰：今夕當雨。是日暘燥，晝無形似，府丞及令在坐咸謂不然。到鼓一中，星月皆沒，風雲並起，竟成快雨。注：黃昏之後，雷聲動天。到鼓一中，星月皆沒，風雲並興，玄氣四合，大雨河傾。（《三國志·魏志·管輅傳》）
- (8) 快雨霽亭午，晴曦作春妍。（葛勝仲《晚歲卜居……》）

可見，“快雨”是下得快也停得快的暴雨。

2.1.3 “快雪”

- (9) 快雪時晴，佳。（王羲之《雜帖》）

“快雪”是使人感到暢快/痛快的雪。什麼雪能使人感到暢快/痛快？

“快雪”只找到了宋人的例句。這些後來的例句當然不能看作詞義演變的橋樑，但可以用來說明“暢快/痛快”和“快速”這兩個詞義的人們心理上的聯繫。

- (10) 去年秋冬交，白晝雷虺虺。隆冬忽暄燠，青紅開百卉。雨澤不時下，快雪傾復霽。（李呂《乙巳四月比屋多疹痘瘡痼間作五月盡猶未已病者延綿在寢治療不獲其法往往失生理殊可憫痛乃作民病書事一首以紀之》）
- (11) 夜光吟屋炯無眠，曉色梅簷重粲然。芒端殘臘洗晴雪，花底好春回隔年。（陳傑《臘月中旬立春快雪過晴》）

這些說明“快雪”是下得快也晴得快的雪。

正因為在某些場合“愉快”和“快速”之間有一定的聯繫，所以，在詞義演變過程中，有的“快”究竟是“愉快”義還是“快速”義，有時不易確定。還有的“快”應是“暢快”義，但看起來很像“快速”義。汪維輝(2000)各舉了四例，今轉引一例：

- (12) 配名天地厚陰陽，從石入金快翱翔。（《老子化胡經玄歌·老君十六變詞》）

本文再補充一例：

- (13) 隱柔桑之稠葉兮，快啁號以遁暑。（曹植《蟬賦》）

從當時“快”的常用意義來看，“快啁號”應是說蟬愉快地啁號，但理解成快速地啁號也可以，因為，在事實上，蟬愉快地啁號就是快速地啁號。這些例子，正可用作“快”在語言運用中發生重新分析，逐漸導致詞義演變的例證。

2.2 “慢”的詞義演變

“慢”的詞義演變過程還沒有見到詳細的分析，所以要多說幾句。

“慢”在上古是“怠慢、傲慢”義。例略。也有表“緩慢”義的用法，但例子很少：

- (14) 叔馬慢忌。毛傳：慢，遲。（《詩經·鄭風·大叔于田》）
(15) 譚諧慢易、繁文簡節之音作，而民康樂。（《禮記·樂記》）慢，疏也。（《正義》）慢猶緩也，對忽之言。（衛湜《禮記集說》引輔廣）

在中古時期，“慢”的“緩慢”義也很少見。如：

- (16) 俗間有戲婦之法，於稠眾之中，親屬之前，問以醜言，責以慢對，其為鄙黷，不可忍論。（《抱朴子·疾謬》）
(17) 典略曰：自此紹貢御希慢，私使主簿耿苞密白曰：赤德衰盡，袁為黃胤，宜順天意。（《三國志·魏志·袁紹傳》）

“慢”的“緩慢”義比較常用是在唐代。但唐代散文中也不多見，《全唐文》中僅見2例：

- (18) 元和九年……不克如其鄉。大懼緩慢茲久，哭命其子某：……（《全唐文·卷590》柳宗元《續榮澤尉崔君墓誌》）

- (19) 夫琴之甚病有七：……調弄節奏，或慢或急，任已去古。七也。
(《全唐文·卷 818》薛易簡《琴訣》)

《六祖壇經》和《入唐求法巡禮行記》中無。
《敦煌變文集新書》中僅 2 例：

- (20) 風慢香煙滿殿飛，人人盡有祝堯詞。(《長興四年中興殿應聖節講經文》)
(21) 飲酒勾巡一兩盃，徐徐慢怕管絃催。(《難陀出家變文》)

《祖堂集》中僅 1 例：

- (22) 雖慢而不遲，雖急而不燥[躁]。(《祖堂集·卷 1》)

但在唐詩中相當多，下面僅舉 5 例：

- (23) 人間流更慢，魚戲波難定。(劉長卿《湘中紀行十首·花石潭》)
(24) 臺上看山徐舉酒，潭中見月慢回舟。(劉禹錫《和思黯憶南莊見示》)
(25) 緊搏紅袖欲支頤，慢解絳囊初破結。(元稹《山枇杷》)
(26) 緩歌慢舞凝絲竹，盡日君王看不足。(白居易《長恨歌》)
(27) 輕攏慢撚抹復挑，初為霓裳後六么。(白居易《琵琶行》)

那麼，“怠慢”義和“緩慢”義是否有聯繫呢？有的。“怠”是心態，“緩”是行事，“怠”的心態會造成“緩”的行事。請看下面的例句：

- (28) 沈於樂者洽於憂，厚於味者薄於行，慢於朝者緩於政。(《管子·中匡》)

又如：

- (29) 故天子諸侯無靡費之用，士大夫無流淫之行，百吏官人無怠慢之事。
(《荀子·君道》)

“怠慢”是主觀心態，“事”是客觀行事。用怠慢的主觀心態來對待客觀行事，就表現為行事之“緩”。這從下面一段話可以看到，群吏“怠”的結果是“吏不治”，庶人“怠”的結果是“農事緩”：

- (30) 有強執有命說議曰：“壽夭貧富，安危治亂，固有天命，不可損益。窮達賞罰幸否有極，人之知力，不能為焉。”群吏信之，則怠於分職；庶人信之，則怠於從事。吏不治則亂，農事緩則貧。（《墨子·非儒》）

所以，“緩怠”有時可以一起用：

- (31) 人心之變：有餘則驕，驕則緩怠。夫驕者，驕諸侯。驕諸侯者，諸侯失於外。緩怠者，民亂於內。諸侯失於外，民亂於內，天道也。此危亡之時也。若夫地雖大，而不並兼，不攘奪；人雖眾，不緩怠，不傲下；國雖富，不侈泰，不縱欲；兵雖彊，不輕侮諸侯，動眾用兵，必為天下政理：此正天下之本，而霸王之主也。（《管子·重令》）

在《論語》中有一段話：

- (32) 子曰：不教而殺謂之虐；不戒視成謂之暴；慢令致期謂之賊。（《論語·堯曰》）

什麼叫“慢令”？劉寶楠《論語正義》解釋說：

“慢令者，《新序·雜事篇》：‘緩令急誅，暴也。’緩令即慢令。《說文》訓‘慢’為‘惰’，凡怠惰，則致緩也。”

按：“緩令急誅”在秦漢文獻中出現不止一次。如：

- (33) 樂武子曰：“居上位而不恤其下，驕也；緩令急誅，暴也；取人之善言而棄其身，盜也。”文公曰：“善。”還載老古，與俱歸。（《新序·雜事》）
- (34) 孔子曰：“夫以不肖伐賢，是謂奪也；以賢伐不肖，是謂伐也；緩其令，急其誅，是謂暴也；取人善以自為己，是謂盜也。”（《說苑·政理》）

- (35) 孔子曰：“汝未之詳也。夫以賢代賢，是謂之奪；以不肖代賢，是謂之伐；緩令急誅，是謂之暴；取善自與，謂之盜。”（《孔子家語·辯政》）

而在《春秋繁露》中，作“慢令急誅”：

- (36) 聽謁受賂，阿黨不平，慢令急誅，誅殺無罪。（《春秋繁露·五行相勝》）

“慢令”和“緩令”意思差別不大，只不過“慢”是從心態講的，“緩”是從行事方式講的；而二者有密切的關係：正如劉寶楠所說，“凡怠惰，則致緩也。”所以，當“慢（怠慢）”用來修飾一個行為動詞時，它本身就隱含了“緩慢”的意義。這就是“慢令”可以說成“緩令”的原因，也是“慢”的詞義從“怠慢”演變為“緩慢”的理據。

實詞的詞義演變，從 A 義演變為 B 義，一般來說，A 和 B 是有意義聯繫的。只不過這種聯繫有的明顯，有的隱晦。對於後一種，要多花一些力氣，找出其間的聯繫。

3. “念”的詞義演變

在現代漢語中，“念”有三個意義：(一) 思念。(二) 朗讀。(三) 學習。第二義和第三義的聯繫比較清楚。第一義和第二義是什麼關係？這個問題要從“念”的歷史演變來回答。

3.1 “念”在中古佛典中的用法

“念”在上古是“思念”之義，例多不舉。在中古佛典中，“念”仍是“思念”之義，如《正法華經·受記品第九》：“心念此已。”梵文原文是“cintayitvā anuvicintya”，“cintayitvā (√cint)”是“思維、考慮”之意。

魏晉南北朝的佛典中的“念佛”、“念經”，並不是後來所說的口念“阿彌陀佛”、口念佛經的意思，而是心中思念佛、思念經。如：

- (37) 當習六思念。六思念者，念佛、念經、念僧、念施、念戒、念天。
（《弘明集·卷13》）
- (38) 以深心念佛，修持淨戒故。（《妙法蓮華經·方便品》）

對應的梵文是“maranto (√sm?)”，表示“記得、想起”的意思。

那些佛典中所說的“念言”，也不是在口中說，而是在心中想（“念言”的“言”如同“想道”的“道”，並不表示言語行爲，而只是提示想的內容）。如：

- (39) 昔有一人。其婦端正，唯其鼻醜。其人出外，見他婦女，面貌端正，其鼻甚好。便作念言：“我今寧可截取其鼻，著我婦面上不亦好乎？”即截他婦鼻，持來歸家，急喚其婦：“汝速出來與汝好鼻。”
（《百喻經·28》）

還有多處作“心自念言”，如：

- (40) 梵志欣豫。心自念言。（《生經·卷1》）

更清楚地說明“念言”是心理活動而不是言語行爲。

- (41) 於是賢者阿難自念言。（《正法華經·授阿難羅云決品第九》）

梵文原文是“cintayāmāsa (√cint)”是“思維、考慮”之意。

直到唐代義淨譯的《破僧事》中，“念”絕大多數還是表示心理活動，只有2例是言語行爲，這可能是受了唐代口語的影響。如：

- (42) 作是念已，即說頌曰：“未得財時起貪愛，求不得時生苦惱。設得財物貪不息，故知財利招無利。念此頌已，便即睡著。”（《破僧事·卷9》）
- (43) 彼仙弟子遙見二十客仙來至，即報仙師曰：“有諸客仙今來至此。”時獨角仙口念：“善來。喚令入室。”（《破僧事·卷12》）

3.2 “念”表示言語行爲（起源於唐代）

到唐代，“念”仍然表示“思，想”，如：

- (44) 爐上香雲天斷絕，心中憶念法花經。（《敦煌變文集新書·妙法蓮華經講經文（一）》）
- (45) 惠遠於旃檀和尚處，常念正法，每睹直（真）經。（《敦煌變文集新書·廬山遠公話》）
- (46) 不念名聞衣食，不貪一切功德利益。（《祖堂集·卷14》）

《祖堂集》中的“念言”，多數仍是表思想活動。如：

- (47) 道吾念言：“他與藥山有因緣矣。”便卻共他去藥山。（《祖堂集·卷16》）

但“念”更多的是表示言語行爲。如《六祖壇經》中的一段話：

- (48) 能大師言：“善知識，淨心念摩訶般若波羅蜜法。大師不語，自淨心神。良久乃言：……摩訶般若波羅蜜者，西國梵語，唐言大智慧到彼岸。此法須行，不在口念；口念不行，如幻如化。”

第一個“淨心念摩訶般若波羅蜜法”究竟是心念還是口念很難斷定。但下面說“口念”，就肯定是言語行爲了。

在《入唐求法巡禮行記》中有如下例句，說明“念”是口念：

- (49) 讀齋文了，唱念“釋迦牟尼佛”，大眾同音稱佛名畢。（《入唐求法巡禮行記·卷1》）
- (50) 大眾同音念“尺迦牟尼佛、彌勒尊佛，大聖文殊師利菩薩、一萬菩薩、一切菩薩摩訶[薩]。”如次學詞同念。念佛了，打槌隨意，大眾散去。（《入唐求法巡禮行記·卷2》）

在《敦煌變文集新書》和《祖堂集》中可以看得更清楚：

- (51) 能者高聲念阿彌陀佛。（《敦煌變文集新書·佛說阿彌陀經押座文》）
- (52) 白莊曰：“念經即是閑事，我等各自帶煞，不欲得聞念經之聲。”遠公曰：“既不許念經，不要高聲，默念得之已否？”（《敦煌變文集新書·廬山遠公話》）
- (53) 惠[慧]能得錢，卻出門前，忽聞道誠念《金剛經》。（《祖堂集·卷2》）
- (54) 宗和尚喝云：“什摩念經，恰似作曲唱歌相似，得與摩不解念經。”（《祖堂集·卷18》）

在《酉陽雜俎·卷7》中有這樣的例句：

- (55) 江陵開元寺般若院僧法正日持金剛經三七遍。長慶初得病卒。至冥司，見若王者問：“師生平作何功德？”答曰：“常念金剛經。”乃揖上殿，令登繡坐，念經七遍。侍衛悉合掌階下，拷掠論對皆停息。而聽念畢後，遣一吏引還。王下階送云：“上人更得三十年在人間，勿廢讀誦。”
- (56) 石首縣有沙彌道蔭，常持念金剛經。寶曆初，因他出夜歸，中路忽遇虎，吼擲而前。沙彌知不免，乃閉目而坐，但默念經，心期救護。虎遂伏草守之。
- (57) 至歲除日，（劉氏）請僧受八關，沐浴易衣，獨處一室，趺坐，高聲念經。
- (58) 其人惟念金剛經，其聲哀切，晝夜不息。

在唐詩中也有：

- (59) 庵前古折碑，夜靜念經時。月皎海霞散，露濃山草垂。鬼聞拋故塚，禽聽離寒枝。想得天花墜，馨香拂白眉。（周朴《贈念經僧》）
- (60) 誰知閑退跡，門逕入寒汀。靜倚雲僧杖，孤看野燒星。墨霑吟石黑，苔染釣船青。相對惟溪寺，初宵聞念經。（齊己《和鄭谷郎中幽棲之什》）

除了佛教信徒“念佛”、“念經”外，文人也“念經（念儒家經典）”和“念書”。如：

- (61) 建隆四年詔：尚書禮部所補太廟齋郎，自今每歲齋郎以十五人為額，其蔭補人並須年貌合格，試念書精熟。……乾德五年，虞部郎中趙元拱、國子監奏（丞）延緒坐試齋郎念經不實，皆責授。（章如愚《群書考索·卷17》）

3.3 “念”從“心想”到“口念”的演變途徑

說到這裡，我們需要討論一個問題：“念”從“心想”演變為“口念”，這兩者之間有沒有聯繫？

要回答這個問題，首先要弄清楚：唐代“口念”的“念”究竟是什麼意思？和現代漢語中“念報”的“念”是否一樣？

請看下面一段話：

- (62) 周世宗廢併寺院，有詔約束云：“男年十五以上，念得經文一百紙，或讀得五百紙，女年十三以上，念得經文七文七十紙，或讀得三百紙者，經本府陳狀乞剃頭，委錄事參軍本判官試驗。”……念經、讀經之異，疑為背誦與對本云。”（《容齋三筆·卷9》）

從周世宗的詔書看，“念經”和“讀經”是有區別的。區別是什麼？宋代的洪邁已不十分清楚，只是說“疑為背誦與對本云”。但從下面的資料可以看到，洪邁的說法還是有根據的。

- (63) 新羅誦經儀式（大唐喚作‘念經’）：打鐘定眾了。下座一僧起打槌，唱“一切恭敬敬禮常住三寶”。次一僧作梵，“如來妙色身”等兩行偈。音韻共唐一般。作梵之會，一人擎香盆曆行眾座之前。急行行便休。大眾同音誦“摩訶般若”題數十遍也。有一師陳申誦經來由了。大眾同音誦經。或時行經本，或時不行經本。念經了。導師獨唱“歸依佛，歸依法，歸依僧”。次稱佛菩薩號。導師唱云“南無十二大願”。大眾云“藥師琉璃光佛”。導師云“南無藥師也”。大眾同音云“琉璃光佛”。導師云“南無大慈悲也”，大眾同音云“觀世音菩薩”。餘皆如是。禮佛了，導師獨結願回向。回向稍長。回向之後，導師云“發心”，大眾同音亦云“發心”。次導師唱發願，已

竟，頂禮三寶。次施主擎施物坐。導師與咒願。便散去。（《入唐求法巡禮行記·卷2》）

- (64) 應補齋郎，並須引驗正身，以防偽濫。……限念書十卷，試可則補。（《舊五代史·卷39》）
- (65) 童子每當就試，止在念書背經，則雖似精詳，對卷則不能讀誦。（《舊五代史·卷148》）
- (66) 長興元年敕：童子準往例委諸道表薦，不得解送。每年所放不得過十人，仍所念書並須是正經，不得以諸子書虛成卷數。（《文獻通考卷35》）

《入唐求法巡禮行記》說“同音誦經。或時行經本，或時不行經本。念經了……”，可見“念經”是出聲“誦”的；而且有時不用經本，則肯定是背誦。《舊五代史》和《文獻通考》的記載說明“念書”是一種考試方式，上引《群書考索》例也是說“念經”、“念書”是用來考童子的，肯定不會是對本宣讀。《舊五代史·卷148》說，有的童子念書精熟，但“對卷則不能讀誦”，很清楚地表明“念書”和“讀誦”是不一樣的。背熟了，即使不懂文義甚或不識字也能流利地背誦，但“讀誦”文章，則不識字或不會斷句，就讀不下去。

憑記憶口誦自己所作的詩文也叫“念”。如：

- (67) 開元中有李幼奇者，以藝干柳芳，念百韻詩。芳便暗記，題之於壁，謂幼奇曰：“此吾之詩也。”幼奇大驚。徐曰：“相戲耳！此君所念詩也。”因謂幼奇，更念他新著文章，一遍皆能記。（《唐語林·卷3》）
- (68) 忽遇上幸維所，浩然錯愕伏床下。維不敢隱，因之奏聞。上欣然曰：“朕素聞其人。”因得詔見。上曰：“卿將得詩來耶？”浩然奏曰：“臣偶不齋所業。”上即命吟。浩然奉詔拜舞，念詩曰：“北闕休上書，南山歸臥廬。不才明主棄，多病故人疎。”（《唐摭言·卷11》）

直到宋代，“念”還有用作“背誦”義的。如：

- (69) 孝宗小年極鈍。高宗一日出對廷臣云：“夜來不得睡。”或問：“何故？”云：“看小兒子讀書，凡二三百遍，更念不得，甚以為憂。”

某人進云：“帝王之學，只要知興亡治亂，初不在記誦。”（《朱子語類·卷127》）

顯然，“看小兒子讀書，凡二三百遍，更念不得”，其意思是說讀了二三百遍還不會背。

既然“念”是背誦，那就需要先在心裡“念（想）”熟，然後在口中“念（讀）”出來。唐代有的語料很好地反映了這兩者的關係：

- (70) 文殊師利言：“我於海中唯常宣說妙法華經。”智積問文殊師利言：“此經甚深微妙，諸經中寶，世所稀有。頗有眾生勤加精進修行此經，速得佛不？”文殊師利言：“有娑竭羅龍王女，年始八歲，智慧利根，善知眾生，諸根行業。得陀羅尼，諸佛所說，甚深祕藏，悉能受持。深入禪定，了達諸法。於剎那頃，發菩提心。得不退轉，辯才無礙。慈念眾生，猶如赤子。功德具足，心念口演。微妙廣大，慈悲仁讓。志意和雅，能至菩提。”（《妙法蓮華經·卷4》）
- (71) 念念念兮入惡易，念念念兮入善難。念經念佛能一般，愛河竭處生波瀾。言公少年真法器，白晝不出夜不睡。心心緣經口緣字，一室寥寥燈照地。（齊己《贈念法華經僧》）

這樣，“念”的意義就從“心念”演變為“心念口演（背誦）”。

不過，佛經數量眾多，有些佛經篇幅也很長，“念經”要全部背誦也是不可能的。所以，“念經”有時也指誦讀佛經。上引《入唐求法巡禮行記·卷2》說“或時行經本”，這就是照著經本誦讀。又如：

- (72) 蓮經七軸六萬九千字，日日夜夜終複始。乍吟乍諷何悠揚，風篁古松含秋霜。但恐天龍夜叉乾闥衆，聒塞虛空耳皆聾。我聞念經功德緣，舌根可算金剛堅。（齊己《贈持法華經僧》）

“諷（背誦）”和“吟（吟詠，誦讀）”都是“念”。

3.4 “念”表示“學習”義

“念”既然有“背誦”、“誦讀”義，就很容易發展為“學習”義。此義在

敦煌變文中就有，如：

- (73) 舜即歸來書堂裡，先念論語孝經，後讀毛詩禮記。（《敦煌變文集新書·舜子變》）
- (74) 女即使聞周氏教，兒還教念百家詩。（《敦煌變文集新書·左街僧錄大師壓座文》）

這兩處“念”，理解為“誦讀”或者“背誦”也可以，但理解為“學習”更好。

《太平御覽·卷 276》有一條記載：

- (75) 《會稽典錄》曰：張立之為人剛毅，志意慷慨。太祖嘗抑之曰：爾不念詩書慕聖道而好乘汗馬擊劒，此一夫之用何足貴也。

《會稽典錄》為晉虞預所作，如果《御覽》所引可靠，則早在晉代“念”就有“學習”義了。不過，這似乎不符合“念”詞義演變的規律，所以，這條材料應當存疑。

宋代以後，“念”表示“學習”義的逐漸增多。下面略舉數例：

- (76) 先生問壽昌：“近日教浩讀甚書？”壽昌對以方伯謨教他午前即理理論語，仍聽講，曉些義理；午後即念些蘇文之類，庶學作時文。（《朱子語類·卷 118》）
- (77) 他念了兩年書，而今跟在班裡記帳。（《儒林外史·25 回》）
- (78) 近來，蒙少爺的教訓，才曉得書是該念的。（《儒林外史·31 回》）
- (79) 你們成日家跟他上學，他到底念了些什麼書！倒念了些流言混話在肚子裡，學了些精緻的淘氣。（《紅樓夢·9 回》）
- (80) 昨兒聽見太太說，蘭哥兒念書真好，他打學房裡回來，還各自念書作文章，天天晚上弄到四更多天才睡。（《紅樓夢·92 回》）
- (81) 寶玉認真念了幾天書，巴不得玩這一天。（《紅樓夢·92 回》）
- (82) 十三歲上就把《四書》、《五經》念完，開筆作文章、作詩、都粗粗的通順。（《兒女英雄傳·1 回》）
- (83) 你們爺兒們的書也不知都念到那兒去了！（《兒女英雄傳·40 回》）

更值得注意的是《紅樓夢·9 回》中的一句話：

(84) 倒念了些流言混話在肚子裡，學了些精緻的淘氣。

這裡的“念”顯然就不能理解為“讀”，而只能理解為“學”了。

4. “要”的詞義演變

“要”在現代漢語中有假設連詞的用法。這個用法是怎樣演變來的？

4.1 “要”作為假設連詞最早的例句（明代）

據現有的資料，“要”作為假設連詞，最早見於明代。如《金瓶梅》中有以下例句：

- (85) 金蓮道：“琴童兒是他家人，放壺他屋裡，想必耍瞞昧這把壺的意思。要叫我，使小廝如今叫將那奴才來，老實打著，問他個下落。”（31回）
- (86) 楊姑娘道：“還是姐姐看的出來，要著老身就信了。”（40回）
- (87) 伯爵道：“休說五兩的話。要我手段，五兩銀子要不了你的，我只消一言，替你每巧一巧兒，就在裡頭了。”（45回）
- (88) 我和他緊隔著壁兒，要與他一般見識起來，倒了不成！（51回）

在《醒世姻緣傳》中更多，下面選5例：

- (89) 你要有些差池，我只好跑到你頭裡罷了！（2回）
- (90) 我明日治一根菜兒，家裡也沒去處，就在前頭廟裡請二位約長吃三鐘。要肯光降，我就好預備。（34回）
- (91) 小冬子要不早娶了巧妮子去，只怕賣了妹子嫖了也是不可知的！（52回）
- (92) 狄希陳在外一邊掙，一邊說道：“二位哥體量我，到家就來。要扯了謊，就是個禽獸畜生！”（66回）
- (93) 孫氏道：“大閨女二十五歲哩。要閨女不嫌，可就好。”（72回）

到清代，假設連詞“要”就用得更多。例子就不舉了。

4.2 假設連詞“要”的產生及其與助動詞“要”的關係

假設連詞“要”是怎樣產生的？和“要”原有的意義有沒有關係？

我認爲是從助動詞“要”（義同“欲”）發展來的。先看《朱子語類》的例句：

- (94) 易是變易，陰陽無一日不變，無一時不變。莊子分明說“易以道陰陽”。要看易，須當恁地看，事物都是那陰陽做出來。（卷 62）
- (95) 問：“用方知，不用則不知。”曰：“……恐伊川那時自因問答去，今不可曉。要附在‘至之道可以前知’解中，只攪得鶻突，沒理會。”（卷 97）
- (96) 小人不可與君子同處於朝。昔曾布當建中靖國初，專欲涵養許多小人，漸漸被他得志，一時諸君子皆為其所陷。要之，要出來做時，小人若未可卒去，亦須與分明開說是非善惡，使彼依自家話時，卻以事付之。若分明說是非，不依自家話時，自家只得去了。（卷 130）
- (97) 本朝韓魏公為相。或謂公之德業無愧古人，但文章有所不逮。公曰：“某為相，歐陽永叔為翰林學士，天下之文章，莫大於是！”自今觀之，要說他自不識，安能知歐陽永叔，也得。（卷 135）

上述例句中的“要”，應該是助動詞，可以用“欲”來替換。但如果用“若”來替換，句子意思也基本上沒有變。這告訴我們，“要”的助動詞用法和假設連詞用法之間是有聯繫的。

但僅僅憑語感是不夠的，還需要進一步分析。助動詞“要”和假設連詞“要”究竟有什麼聯繫，有什麼區別呢？在什麼條件下，“要”可以從助動詞發展為假設連詞？

在現代漢語中，假設連詞“要”和助動詞“要”的區別是很明顯的。如下面兩個句子：

- (98) a. 我要走了，明天再來。
b. 我要走了，就不來了。

(98a) 中的“要”只能換成“想/打算”，不能換成“如果”；(98b) 中的“要”只能換成“如果”，不能換成“想/打算”。(98a) 中的“要”是助動詞，表示主

語“我”的意願，“走”是想施行而還沒有施行的主觀意圖；“我要走了”可以是一個獨立的句子，也可以再接續一個小句。(98b)中的“要”是假設連詞，表示假設某種情況的出現，“走了”是假想中的一種已經出現的客觀情況；“我要走了”是條件複句中的前一個小句，後面必須接續另一個小句，表示當這種假設實現時，會出現什麼情況。

不過，助動詞“要”和假設連詞“要”後面的“VP”有一個共同點：不論是想施行而還沒有施行的主觀意圖，還是假想中的一種已經出現的客觀情況，都是在現實中沒有出現的，這為“要”從助動詞演變為假設連詞提供了條件（見下）。

從歷史上看，漢語中的“要”用作助動詞在前，用作假設連詞在後。前者可以換成“欲”，例多不舉。後者可以換成“若”，上述《金瓶梅》、《醒世姻緣傳》中的例句都是這一類。兩者區分也和上述(98a)、(98b)兩類一樣，是很清楚的。

但是助動詞“要”也可以出現在條件複句的前一小句中。如《朱子語類》中的例句：

(99) 又曰：“性最難說，要說同亦得，要說異亦得。”（卷4）

(100) 今之官司合用印處，緣兵火散失，多用舊印。要去朝廷請印，又須要錢，所以官司且只苟簡過了。（卷106）

這種“要”，粗看起來，似乎也能翻譯成“如果”，整句可以譯成“如果說同也行，如果說異也行”，“如果去朝廷請印，就又得要錢”。但仔細分析，就會看到，“如果”所表達的假設之意是整個句式帶來的，不是“要”所表達的。漢語表示假設關係，很多不用假設連詞，自古至今都是如此。而“要”在這裡表示的仍是“想/打算”之意，這從後面的“VP”可以看出來：“去朝廷請印”、“說同”、“說異”都是想施行而還沒有施行的主觀意圖，而不是假想中的一種已經出現的客觀情況。所以，準確的翻譯應是：“如果想去朝廷請印，就又得要錢”，“如果想說同也行，如果想說異也行”。這說明助動詞“要”和假設連詞“要”是有區別的，但也說明，這種區別不是不可逾越的。

如果再進一步，助動詞“要”不但出現在條件複句的前一小句中，而且後面的“VP”究竟是想施行而還沒有施行的主觀意圖，還是假想中的一種已經出現的客觀情況也不容易分清，這時，助動詞“要”和假設連詞“要”的區別就更模糊了。造成這種不易分清的原因在於：（一）上面說過，不論是想施行而還沒有施行

的主觀意圖，還是假想中的一種已經出現的客觀情況，都是在現實中沒有出現的；(二) 漢語中動詞的體貌沒有很清楚的標記，所以，在某些情況下，一些“VP”究竟是將然而未然，還是假設的已然，是不易分清的。上引《朱子語類》中的四個例句，“要”後面的“VP”都是如此，所以，“要”換成“欲”和“若”都可以。

但《朱子語類》中還有一例：

- (101) 問：“‘全體大用，無時不發見於日用之間。’如何是體？如何是用？”曰：“體與用不相離。且如身是體，要起行去，便是用。”
(卷 16)

此例中“要”後面的“起行去”，很難說是某人或人們的意圖，而只能說是假設中的客觀情況，所以，這裡的“要”不能換成“欲”，只能換成“若”，應該看作假設連詞了。但在《朱子語類》中僅此一例，而且，據筆者不完全的檢索，在宋代以至元代都沒有發現這樣的例句，為慎重起見，我們不說假設連詞“要”在宋代已經出現，而說假設連詞“要”出現在明代。

但《朱子語類》中這些兩可的例句給我們提供了“要”從助動詞“要”到假設連詞“要”的演變途徑，這是沒有問題的。概括起來說，助動詞“要”首先要出現在條件複句的前一小句中，這種句法環境就賦予這個小句表示假設的語義。其次，“要”後面的“VP”究竟是想施行而還沒有施行的主觀意圖，還是假想中的一種已經出現的客觀情況不易分清，因此，“要”表示意願的語義就減弱。這樣，當說話者說出“要出來做時”（卷 130）這個句子時，聽話者就可能有兩種理解；一種(A)認為是“欲出來做時”之意，一種(B)認為是“若出來做時”之意，而且這兩種理解差別不太大：(A)中“欲”表示的是意願，但整個小句是假設；(B)中“若”表示的是假設，但“出來做”是施事的意願。所以兩者都是意願和假設兼表的。這樣的言語行為反覆出現，就會使語言的使用者形成一種印象：“要”在表示意願的同時，還可以表示假設。

再下一步的發展，就是人們把“要”作為假設連詞，和“若”一樣的使用。這時，“要”就不局限於上述句法、語義條件了，像上述《金瓶梅》、《醒世姻緣傳》中的例句，“要”的句法、語義條件都和助動詞“要”很不一樣：“要”可以出現在主語前面，可以出現在助動詞“肯”前面，後面的“VP”可以是違反施事意願的行為（扯了謊）。這些例句中的“要”就不再理解為“欲”和“若”兩可的了，而是只能理解為“若”。這種“要”就超出了重新分析的階段，發展為成熟的假設連詞了。

4.3 假設連詞的來源

Elizabeth C. Traugott 有一篇文章〈Conditional markers〉專門討論連詞的來源。文章綜合了對多種語言的連詞的考察，把連詞的來源概括為五類：

- (一) 由表情態的詞演變而來。
- (二) 由繫詞結構演變而來。
- (三) 由疑問式演變而來。
- (四) 由表示某物是已知的或給予的詞演變而來。
- (五) 由原本是表時間的詞演變而來。

關於第一類，文章特別提到了認識情態詞和祈願情態詞 (epistemic modality, (the modality of possibility and doubt), and optative modality (the modality of wishing))，舉了五種語言的例子，包括英語的 *suppose* 和漢語的“要”。

這告訴我們，由情態詞演變為連詞，是在多種語言中存在的，漢語表意願的助動詞“要”演變為連詞，是符合這種較普遍的演變規律的。

〈Conditional markers〉還對這種演變的動因做了解釋。文章說：條件句表述的是一種假設的情景，是一種可能性（包括真實的可能、想像的可能和違反事實的可能），這會使人們用一個認識情態詞作為條件分句的標記；一個想像的情景常常是人們所希望的，這會使人們用一個祈願情態詞作為條件分句的標記。

〈Conditional markers〉一文的這一層意思，本文在上面已經說過：表意願的助動詞“要”之所以能演變為假設連詞，很重要的一個原因是因為“不論是想施行而還沒有施行的主觀意圖，還是假想中的一種已經出現的客觀情況，都是在現實中沒有出現的”。但是，僅僅指出這一點，還不足以充分說明“要”是怎樣從表意願的助動詞演變為假設連詞的。因為，儘管表意願的助動詞和假設連詞有表示假設情景的共同性，但畢竟兩者還有區別，表意願的助動詞“欲”和假設連詞“若”不能等同。所以，“要”的語義從表示“欲”演變為表示“若”，還必須消除“欲”的意義，而獲得“若”的意義，這個過程還必須加以解釋。這一點，本文在上面已經說過：“要”的演變應當具備兩個條件：“助動詞‘要’首先要出現在條件複句的前一小句中，這種句法環境就賦予這個小句表示假設的語義。其次，‘要’後面的‘VP’究竟是想施行而還沒有施行的主觀意圖，還是假想中的一種已經出現的客觀情況不易分清，因此，‘要’表示意願的語義就減弱。”第二個條件使“要”原來的意義（表示意願）弱化，這就是通常所說的

“semantic bleaching (語義漂白)”；第一個條件使“要”具有了新的意義(表示假設)，而這種意義原來是語境所有的，現在，通過語用推理 (pragmatic inference)，成為“要”這個詞所有了。這就是通常說的“absorption of context (語境吸收)”。(“absorption of context”是 Bybee, Perkins & Pagliuca 的 *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World* 首先提出來的，其詳細論述在此不贅。)也就是說，表意願的助動詞“要”之所以能演變為假設連詞，除了兩者之間有表示假設情景的共同點之外，還必須經過“語義漂白”和“語境吸收”。這是“要”的完整的演變途徑。

4.4 複合的假設連詞“若要”

在討論假設連詞“要”的產生與來源時，還應說到複合的假設連詞“若要”。“若要”最初不是一個詞，“若”和“要”不在同一層次上。如：

- (102) 若要和市和雇，先依時價付錢。(《唐大詔令集·唐中宗即位赦文》)
- (103) 若要聽道安講者，每人納絹一疋，……要聽道安講者，每人納錢一百貫文。(《敦煌變文集新書·廬山遠公話》)

顯然，這些例句的切分應是：“若/要……”，“若”是假設連詞，“要”是表意願的助動詞。在唐宋時期的語料中，“若要”都不是一個詞。如：

- (104) 又問：“此四句若要連續看，如何？”(《朱子語類·卷34》)

根據現代的語感，這個“若要”似乎應該是一個詞。但聯繫下文“才要連續，便是說文字，不是要著實做工夫。若著實做工夫，便一句自是一句”，“若”和“要”還是分開的。又：

- (105) 若要欲得眼親逢，學取經文便合同。(《敦煌變文集新書·解座文匯劄》)

這個例句比較特殊。按常規，“要欲”不可能是一個詞，也不可能是“要”和“欲得”連用，那麼，就應該是“若要/欲得”。但如果說“若要”在敦煌變文

中就已經是一個詞，而從唐宋到元代，再也找不出另一個例子，這又違反語言演變的規律。因此，只能讀作“若/要/欲得/眼親逢”，但這是一個有語病的句子。

複合的假設連詞“若要”見於明代：

(106) 若要死貪無厭足，家中金鑰教誰收？（《金瓶梅·12回》）

(107) 你問他要了那三兩買藥的銀子來。你若要不將來，我坐你的工價。
（《醒世姻緣傳·67回》）

明代已經有假設連詞“要”，假設連詞“要”和假設連詞“若”並列使用，並且凝固為一個詞。這是合乎詞語演變的規律的。

“若要”的形成過程和“若使”類似。“若使”本來也是“若/使……”，“使”是“使之”的意思。如：

(108) 若使遭遇時君，更閱賢知，為所稱善，則必度越諸子矣。（《漢書·揚雄傳》）

後來，“若使”是兩個假設連詞“若”和“使”的並用，而且逐步凝固為一個複合的假設連詞。如：

(109) 若使陛下出不世之詔，效臣錐刀之用，使得西屬大將軍，當一校之隊，若東屬大司馬，統偏舟之任，必乘危蹈險，騁舟奮驪，突刃觸鋒，為士卒先。（《三國志·魏志·曹植傳》）

(110) 潁川李曇，德行純備，著於人聽。若使擢登三事，協亮天工，必能翼宣盛美，增光日月矣。（《後漢書·徐穉傳》）

5. 小結

本文討論的詞義演變的三個案例，兩個是實詞，一個是虛詞。實詞的詞義演變，舊義和新義之間都有一定的聯繫，但聯繫的情況兩個案例各不相同：

“快”、“慢”是舊義和新義所表達的概念在人們的心理上有一定的聯繫：使人們心理上感到暢快的事物常常是快速的，心理上的怠慢往往會導致行動的緩慢。

“念”是舊義和新義所反映的行為在實際上有某種關係：背誦必須是“心念+口

言”。虛詞的詞義演變和實詞不同，往往和句法環境有關，就本文所討論的案例來說，助動詞“要”（表意願）如果出現在假設複句的條件小句中，就可能使人感到“要”有表假設的意義，如果後接的“VP”究竟是想施行而還沒有施行的主觀意圖，還是假想中的一種已經出現的客觀情況不易分清，“要”表示意願的語義就會減弱。在反覆進行的語言交際過程中，具備上述條件的“要”就會產生“欲/若”的歧解，再進一步發展，就由舊義“表意願”外產生一個新義“表假設”。這三個案例可以簡單地表列如下：

〈表 1〉“快、慢、念、要”演變的階段

	舊義	中間階段	新義
快、慢	A	$A \leftrightarrow B$	B
念	A	$A + B$	B
要	A	(B) A-	B

還有一點是要說明的：詞義演變的“個性”比較強。像本文所討論的三個案例，其演變途徑只適用於這三個案例本身，而不能推演到其他同類詞語。“快”可以通過“愉快→暢快→快速”的途徑演變，但“樂、說（悅）”就不能。“念”可以通過“心想→背誦→誦讀→學習”的途徑演變，但“思、想”就不能。“要”可以通過“意願→假設”的途徑演變，但“欲、願”就不能。當然，也會有幾個詞沿著同一途徑演變的，實詞如“族、眾、庶”都由“眾多”演變為“一般”，虛詞如“教、給、讓”都由“使役”演變為“被動”。語言研究似乎對這些“成批”演變的更感興趣，確實，從這裡可以概括語言演變的規律。但對那些“零散”的案例也有研究的必要，因為從個別案例出發，多加積累，就能更清楚地看到語言演變的多樣性，而且通過豐富多樣的案例，也可以從“個性”中發現“共性”，同樣可以概括出語言演變的規律。而且，經過深入的研究，也許能找出為什麼某個詞能發生演變，而其他同類的詞不能發生演變的原因。具體語言現象的考察、分析、研究永遠是語言研究的基礎，只有在堅實的基礎上，才能有理論的昇華。

引用文獻

- Bybee, Joan L., Revere D. Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- 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.
- Ma, Beijia (馬貝加). 2002. *Jindai Hanyu Jieci* 近代漢語介詞 [*Preposition in Early Modern Chinese*]. Beijing: Zhonghua Book Company.
- Traugott, Elizabeth Closs. 1985. Conditional markers. *Iconicity in Syntax*, ed. by John Haiman, 289-307. Amsterdam & Philadelphia: John Benjamins.
- Wang, Weihui (汪維輝). 2000. *Donghan-Sui Changyongci Yanbian Yanjiu* 東漢—隋常用詞演變研究 [*A Study on Common Words in East Han to Sui Dynasties*]. Nanjing: Nanjing University Press.

The Semantic Evolution of *kuài/màn*, *niàn*, and *yào*

Shaoyu Jiang

Peking University

This paper discusses three types of shift in lexical meaning and the reasons of these developments. 1. *kuài*: joyful-fast; *màn*: idle-slow. The reason for such a lexical shift resides in the fact that people think that there is relation between these two concepts. 2. *niàn*: think-chant. The reason for such a development is that the meaning of *niàn* in the transitional stage is “chant while thinking”. 3. *yào*: want-if. The reason for such a development is “absorption of context”.

Key words: *kuài*, *màn*, *niàn*, *yào*, development of lexical meaning

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
香港教育學院
北京大學

漢語歷史研究除了要揭示漢語在不同共時層面所表現出的各種變化，更要揭示導致變化的原因和總結變化的規律。然而，這並非易事。儘管古代文獻浩如煙海，卻因為用基本脫離口語的文言文寫成，其中很少或根本沒有記錄下變化的過程環節；生活在變化發生幾百年甚至上千年之後的我們，對當時的語言基本沒有自省的能力，無法瞭解這些變化究竟是如何發生的，以及它們的性質。

基於社會語言學「用現在解釋過去」的研究方法，作者嘗試觀察漢語在當今所發生的變異，發現這些變異的過程和原因，並與漢語在古代出現的種種變化進行比較。本文即是這種努力的初步成果。文章的結論是，使用者對語言固有成分在理解和使用上發生「言語錯誤」，是古今漢語詞彙演變的共同原因之一。而這些錯誤的發生，多與漢字有關。

關鍵詞：漢語歷史研究，社會語言學，語言演變，言語錯誤，漢語詞彙，漢字

1. 引言

做語言歷史研究，除了要揭示語言在不同共時層面所表現出的各種變化，更要揭示導致變化的原因和總結變化的規律——找出變化的原因，可以更好地瞭解歷史，這個歷史不僅是語言的，還可以是文化的、民族的；總結變化的規律，可

* 貝羅貝教授多年致力於將現代語言學的理論介紹和運用到漢語歷史研究當中來，貢獻良多。本文是筆者對此的一個響應。文章構思多年，有關的內容曾在四川大學和北京大學的課堂上講過，亦在學術會議上作過報告。初稿寫成後曾在香港教育學院「方潤華講座」(2009)演講，亦在國際中國語言學學會第18屆年會(IACL-18, Boston, 2010)交流。在這些過程中，幸獲指教，受益良多。尤其是匿名審稿人和同門汪維輝教授對語料的使用提出了很有價值的訂正意見，在此一併表示感謝。本文的研究得到香港教育學院研究啟動項目(RGB55/2008-2009)的資助。

以預知未來。就漢語史研究而言，過去的工作以挖掘語言事實為主。在繼續這樣的工作的同時用更多的精力對已知的語言事實之來龍去脈進行解釋，相信這將是今後研究的重點所在。

然而，這並不是一件輕鬆的工作。儘管漢語有世界最多的古代書面文獻保留至今，但是這些文獻絕大多數是由我們稱之為「文言」或者「半文言」（古白話文）¹ 的書面語寫成的。除了經常提到的「言、文脫節」外，它們還含有一定數量的跨時空的成分（方言和仿古）。這對於十分強調系統性的語言研究來說，是一個大的障礙。而進一步的困難是，一方面，在文言文獻中，研究者觀察到的，主要是已經完成變化的語言 (language) 現象。這些變化是何時發生的？是如何發生的？變化的過程是什麼？機制是什麼？文言文能夠提供的有價值的信息相當有限；另一方面，不論是文言文還是白話文，其中必然會含有一定數量的個人言語創新，它們不被或者尚未被社會接受（即大多數時人認為不可說的成分）。這都給今後以揭示變化原因和總結變化規律為重點的漢語歷史研究造成了難以想像的障礙。

舉例來說，《莊子·外篇·知北遊》有一則故事，說的是東郭子問莊子所謂「道」在什麼地方。莊子回答說：無所不在。東郭子要莊子說具體一些。莊子就從螻蟻說起，直到稗草、磚瓦、大小便溺等，都是道所在的地方。東郭子見莊子越說越低下，就不再回應了。莊子說：要滿足你的要求把道的真相說清楚，就像市場上的仲介商用腳碰豬腿來估量它的肥瘦一樣，「每下愈況」。² 莊子的「每下愈況」後來成為一個成語。它本來的意思是說，越是踏在豬的下部（即腳脛上），就越能看出它是否真肥（因為腳脛是最難肥的部位），比喻越從低微的事物上推求，就越能看清事物的真相。然而，至遲到了宋代，這個成語被用來形容情況愈來愈糟。³

在漢語的歷史演變過程當中，類似的變異不知出現過多少，過去的研究者對此多有揭示。但是，面對語言學研究更關心的問題，即這些變異為什麼會發生，

¹ 古代漢語書面語中的白話文，最初是作為文言文與口語的混合體出現的，實質上是文言文的變體。

² 《莊子·知北遊》：「東郭子問於莊子曰：『所謂道惡乎在？』莊子曰：『無所不在。』東郭子曰：『期而後可。』莊子曰：『在螻蟻。』曰：『何其下耶？』曰：『在稗稊。』曰：『何其愈下耶？』曰：『在瓦甓。』曰：『何其愈甚耶？』曰：『在屎溺。』東郭子不應。莊子曰：『夫子之問也，固不及質。正獲之問於監市履豨也，每下愈況。』唐陸德明《經典釋文》卷 27：「李云：正，亭卒也；獲，其名也。監市，市魁也。豨，大豕也。履，踐也。夫市魁履豕，履其股腳豨難肥處，故知豕肥耳。問道亦況下賤則知道也。」

³ 本文使用的「言語錯誤」的例子，大部分為學術界人所共知者，而非作者的發見。這些例子散見各處，多為輾轉引用，始作者難以確考，故不再羅列其原始出處。

或者說是什麼引起了這些變異的發生，我們要麼毫不關心，要麼束手無策。原因之一，就是儘管文獻浩如煙海，卻很少或根本沒有記錄下變化的過程環節；生活在變異發生幾百年甚至上千年之後的我們，對當時的語言基本沒有內省 (intuition) 的能力，難以感知這些變異的性質。

眾所周知，社會語言學研究語言變異現象，它觀察語言在現時的變化，並通過這種觀察，發現語言演變的原因和規律。社會語言學研究語言共時的變異，根本目的還是要找出語言歷史演變的原因和規律。基於「一致原則」(uniformitarian principle)，即「當代引起變異的語言動力與過去歷史上起作用的動力相似」(Crystal 2000:374)，社會語言學家提出了「用現在解釋過去」(the use of the present to explain the past) 這一重要的研究方法 (Weinreich et al. 1968, Labov 2001)。簡要地說，就是把對當代語言的共時描寫跟對該語言的歷史演變研究有機地結合起來，從當代共時語料中觀察語言正在發生的變化，用來推知和解釋歷史上的變化。

上世紀七十年代以來，一些歷史語言學家運用社會語言學的方法，在其他一些語言的歷史研究中取得了非常重要的進展，給漢語歷史研究有益的啓示。如果做漢語歷史研究的人，也能注意觀察漢語活的言語活動，從中發現事實和規律，用來幫助古代漢語的研究，同樣會對古代漢語的研究產生積極的推動力。其實反之亦然：如果做漢語（現代）共時研究的人，關注一下漢語在歷史上的變化，或可從中得到啓示，來預測共時變異的可能結果。

2. 觀察：現時活的言語活動

讓我們先來觀察現代日常的漢語言語活動。有以下幾種情形可能，或者正在導致，或者已經產生新「說法」。

2.1 認錯字、讀錯音

日常生活中的「認錯字」現象可以分爲兩類。一類指因字「形近」致誤，即致誤者將形體相近而意義迥然不同的兩個字相混；一類指因字「同形」致誤，即致誤者不瞭解「同形詞」現象——同形體的字記錄的可能不是同一個詞，誤將它們混淆。

2.1.1 形近致誤

例如「棘手」，用來形容事情難辦或難以對付，在口語中很是常用。如：李季《玉門兒女出征記·什麼是科學》：「頭幾個月還順利，越往後越**棘手**問題越多。」⁴ 肖淑惠、朱小美譯《嘉莉妹妹》第 21 章：「經理吃了一驚，被這問題擊中了，因為這問題比她的問題還要**棘手**。」其中的「棘手」與「棘手」同義。這個「棘手」，其實就是「棘手」。⁵

「棘手」的字面意思是荊棘刺手，比喻事情難辦或難以對付。據現有資料，這個詞明代後期已見。如明陳宏緒《寒夜錄》卷上：「董蘿石，年六十有八，……近從吳觀察秋圃處得其遺詩一帙，孤高秀逸，足稱其人，七言律尤佳。……而詠豆芽云：『蕪蕪亭後得褒封，金甲銀鉤奪化工。……』如此**棘手**題，大見工力。」明張岱《陶庵夢憶》卷 8「范與蘭」條：「後見王本吾琴，大稱善，盡棄所學而學焉。半年學《石上流泉》一曲，生澀猶**棘手**。王本吾去，旋亦忘之，舊所學又銳意去之，不復能記憶。」明錢謙益《牧齋初學集》卷 31「陶仲璞《遯園集》序」：「仲璞之集，稱心而言，指事而論，無薄喉**棘手**之艱，無東塗西抹之飾。」清代十分流行。如《龔自珍集》卷 5「在禮曹日與堂上官論事書」：「署中因循，憚於舉事。若再積數年，難保案牘無遺失者，他日必致**棘手**。」現代漢語仍常用。如沙汀《還鄉記》9：「羅敦玉覺得沒一樁不容易，就祇這件多少有點**棘手**。」

漢語中亦早有「棘手」一詞，始見於宋元，有「辦事剛嚴猛烈之人及能手或老手」和「厲害或毒辣的手段」兩個基本的意思，與「棘手」無關。如明朱長祚《玉鏡新譚》卷 1「敘略」：「雖爾，豈似今日魏忠賢之逆惡滔天也？忠義之士，感奮一言，立為齏粉，矯殺無算。即宮闈妃御，亦不能相保，懿親國戚，悉遭**棘手**。」明楊士聰《玉堂薈記》卷 3：「十一月之薦，以熊、姜、楊、廖諸人，漸有異議，將資巴縣之**棘手**以為援也，上皆從之。」明羅懋登《三寶太監西洋記通俗演義》第 35 回：「這番奴一連接了我兩枝箭，今番還他一個**棘手**，他才曉得。」清丁治棠《仕隱齋涉筆》卷 4：「君工經妍妙，不落凡響。生與友結詩社，拈題限韻，必擇**棘手**者為之。」

⁴ 為了節省篇幅，本文除常用古籍，在引例中但凡用到現代人的文學作品，亦不注明版本信息。下同。

⁵ 利用百度網搜索「更多關於棘手的問題」(2009/5/3)，共得到 2,133 項結果。其中第 1 頁 10 篇文章有以下的標題：「Vista 極棘手的問題，求教真正的高手來解決」「一件棘手的機關人際關係問題」「有會計事務高手嗎？棘手啊，幫幫忙」「比較棘手的共用上網問題」「電腦高手進來哦！問題很棘手，答錯很嚴重」「極品飛車 9 的安裝一個非常棘手的問題」。

「棘手」原無棘手義，從「棘手」原有的意義也難以「引申」出「棘手」的意義。顯然，義同棘手的這個「棘手」就是「棘手」——「棘手」用同「棘手」，是將「棘手」之「棘」誤認作「辣」之後的結果。⁶

2.1.2 同形致誤

相信很多人都會從醫生嘴裡聽到以下兩個病名，一個是「/qiángzhí/」，一個是「/yǔchǐ/」。前者指因肌肉萎縮引起的肢體活動障礙，後者即俗稱的「蟲牙」。看到漢字，才知道他們說的病應是「/jiāngzhí/」和「/qǔchǐ/」。⁷「/qiángzhí/」這個病名的出現，就是「同形致誤」的結果。

所謂「/qiángzhí/」，其書面形式是「強直」。「強直」在《素問·至真要大論》中已有其用例：「諸燥狂越，皆屬於火。諸暴強直，皆屬於風。」其中的「強」，是僵硬的意思，與「直」同義連文。「強」單用的例子如《世說新語·文學》：「殷仲堪云：三日不讀《道德經》，便覺舌本間強。」這個「強」，《廣韻》音「居良切」，現代普通話讀音是「/jiāng/」，本字當是「彊」。

「強」是一個多音多義字。其本義是米蟲。《說文解字·虫部》：「強，蛭也。從虫，弘聲。」張舜徽先生(1983:25-57)解釋說：「孔廣居曰：『弘聲不諧，不但秦刻石從口，即漢隸亦皆從口。愚意強，米中蛀蟲也。從虫口，會意；弓省聲。』舜徽按，籀文從彊聲，則強當從彊省聲明矣。《玉篇》『強』下云：『米中蠹小蟲。』此本義也。此蟲形小色黑而堅，米穀中多有之，湖湘間名之為鐵蠹牛，喻其強有力也。」從「強有力」的意義引申出堅強、強壯、強大諸義。「彊」，《說文解字·弓部》記錄的意義是「弓有力」。「強」和「彊」讀音和意義都相近，應是同源詞，二者可以互換。在古籍裡，「強」和「彊」混用，二者互為異體字。《玉篇·歹部》：「彊，死不朽也。」人或動物死了，屍體會僵硬，引申出「僵硬」之義。「彊」字不見《說文》，但當與另一個字「僵」有關。《說文解字·人部》有「僵」，釋作「僨」，就是「仆」。段玉裁注云：「僵謂仰倒。……今人語言乃謂不動不朽為僵。」就是以「僵」為「彊」。「僵」「彊」同源，二者

⁶ 在北京大學 CCL 語料庫裡，有這樣一段出自《人民日報》的語料：「在日常生活中，常常聽人把『瞻養』讀成『瞻養』，把『衛戍』讀為或寫為『衛戍』。……有的法官把『酗酒』讀為『凶酒』，有的醫生把『臀部』讀為『殿部』，有的老師把『愛迪生』讀為『愛由生』，把『忤（僂）悔』讀作『千悔』，有的領導幹部把『棘手』讀為『辣手』，把『柑桔』讀作『柑吉』，有的管科技的幹部把『造詣』讀作『造脂』。」（何漢聲：漢字的委屈與報復，《人民日報》1994/10/11）

⁷ 筆者的妻子、兒子及岳父母都是醫生。據他們說，各地的醫生都這麼說。若此，在特定的人群中已經「約定俗成」了。

和「彊」聲符相同，之間可形成通假關係，即該寫作「彊」（「僵」）的地方寫成「彊」，進而又寫成「強」，「強直」就是一個例子。因為不瞭解「強直」之「強」就是「彊」（「僵」），將其中的「強」的讀音誤為「/qiáng/」，創造出了一個新的「說法」。

相比之下，「/yǔchǐ/」產生的原因可能稍複雜一些。它固然也可以歸為「認錯字」，但致誤的原因不僅僅是字形，更在於字音，因此歸入「讀錯音」可能更準確。

2.1.3 讀錯音

「/yǔchǐ/」這個詞的書面形式是「齲齒」。「齲齒」作為病名使用的時代不晚於漢代，如《史記·扁鵲倉公列傳》：「齊中大夫病齲齒。」其中的「齲」，《說文解字》已見，《廣韻》音為「驅雨切」，正確的現代普通話讀音是「/qǔ/」。

按《說文解字·齒部》的說法，「齲」是形聲字，「齒」為義符，「禹」為聲符。隨著語音的變化，「禹」已經不能正確地表示「齲」的讀音。將「齲」/qǔ/讀成/yǔ/，是「讀半邊字」型的錯誤。這樣的錯誤在日常生活中是十分普遍的，導致「/yǔchǐ/」這個「新詞」的出現。

另一個例子是「/yǎnlián/」。人（動物）的眼部俗稱「眼皮」的部位，正式的名稱是「眼瞼」。這個詞的正確讀音是「/yǎnjiǎn/」。但在指稱同一部位時，我們常常聽到「/yǎnlián/」的說法。如 2008 年上半年，筆者應邀在新加坡國立大學中文系講授《訓詁學》。在當地電視台反覆播放的一則化妝品廣告中，作為模特兒的著名影星鞏俐就將眼皮叫作「/yǎnlián/」。不明就裡的人還以為「/yǎnlián/」是眼皮兒的另一種說法，其實就是「眼瞼」。為什麼會將本該讀作「/jiǎn/」的「瞼」讀成「/lián/」？

「瞼」是一個形聲字。以「僉」為聲符的字的聲母後來主要分化為兩個系統，一是「來母」，如「臉」「殮」「斂」等；一是「見母」，如「瞼」「撿」「劍」等。將見母的「瞼」讀為來母的「/lián/」，可能由於在目治時，將「眼瞼」之「瞼」與「臉」相混；當然，也有可能始作俑者不知道「瞼」的讀音，遂根據更常見的「臉」「斂」等字的讀音來讀「瞼」。

「認錯字」以及相關的「讀錯音」都是日常言語生活中常見的「言語錯誤」。這種類型的錯誤使指稱對象的名稱發生非合理改變，從而產生新的名稱。

2.2 望文生義

「望文生義」也是日常生活中常見的言語錯誤。致誤者簡單地從字面上理解詞義，往往導致所謂「新義」的產生。

「望洋興歎」一語或用於感歎與他人相比自愧弗如，或用於感歎想做事卻力不從心。如元劉壘《隱居通議》卷10「詩歌五·劉玉淵評論」：「少陵讀五車，倒三峽，吐驚人泣鬼之句；岱嶽岳陽，真能籠乾坤萬里於一詠之內。千古吟人，**望洋興歎**。」但今天有人將其解釋為望著大海感歎。如林語堂編著《當代漢英詞典》（香港：香港中文大學出版社，1972:813）「望」字下有「望洋興歎」，解釋為：「view with despair the vast ocean or vast work to be done（絕望地望著一片浩渺無垠的大海或一堆做不完的工作）」。北京外國語大學英語系《漢英詞典（修訂版）》（北京：外語教學與研究出版社，1997:1281）將這個成語的字面意義譯作「lament one's littleness before the vast ocean（面對巨大無比的海洋哀歎個人的渺小）」。⁸ 這是典型的望文生義。⁹

「望洋興歎」出自《莊子·秋水》：「秋水時至，百川灌河，涇流之大，兩涘渚崖之間，不辨牛馬。於是焉河伯欣然自喜，以天下之美盡在己。順流而東行，至於北海，東面而視，不見水端。於是焉河伯始旋其面目，望洋向若而歎曰：『野語有之曰「聞道百以為莫己若」者，我之謂也。』」晉郭象《莊子注》：「眊，莫剛反，又音『旁』，又音『望』，本亦作『望』。『洋』音『羊』。『眊洋』猶『望洋』，仰視貌。」又唐陸德明《經典釋文·莊子音義》：「『眊』莫剛反，又音『旁』，又音『望』。本亦作『望』。洋：音『羊』。司馬、崔云：『『眊洋』猶『望羊』。仰視貌。』」根據郭象和陸德明的注，可知他們當時看到《莊子》，「望洋」寫作「眊洋」，而「望洋」，似非通行的寫法。不論是「望洋」還是「眊洋」，記錄的是同一個單純詞（聯綿詞），而非合成詞，「望洋」二字只是記音而已。¹⁰

⁸ 《外研社現代漢英詞典》（北京：外語教學與研究出版社，2001年）是後者的改編本，同。

⁹ 王還主編《漢英雙解詞典》（北京：北京語言文化大學出版社，1997:901）「望洋興嘆」條的解釋有改進，沒有將「洋」解釋成「海洋」：「(former meaning) feel small and negligible in the face of sth. Great; (present meaning) bemoan one's inadequacy in the face of a great task（原指在偉大的事物面前感嘆自己的藐小，後比喻做一件事力量不足或條件不夠，而感到無可奈何）」，但卻將「望」解釋成「看」：「望洋：raise one's head to look at sth.」。按：此條的解釋語本中國社會科學院語言研究所《現代漢語詞典》。對於其中的「望洋」，該詞典（第五版，北京：商務印書館，2005:1412）僅僅解釋為「抬頭向上看的樣子」，似未妥。

¹⁰ 朱起鳳《辭通·七陽》（上海：上海古籍出版社重印本，1982:868）「望洋」條下除了「望洋」（《釋名》），還收錄了古籍中所見的其他一些寫法，如「望羊」（《史記》）「望陽」（《論衡》）「眊

「望文生義」型錯誤的另一個例子是現代對「罄竹難書」的誤用。這個成語形容不好的事情極多，難以盡載，始見於明代。明楊一清《平西夏頌》：「此賊可滅，臣死不避。十有七事，臣得其粗，窮其罪狀，**罄竹難書**。」明周起元《周忠潛奏疏》卷上「題為摘陳漕河吃緊要務以裨國計事疏」：「漕弊亟宜湔除，莫過於陋規。……此其弊，誠**罄竹難書**。」明朱國禎《湧幢小品》卷14「緒帖」：「再照編審之弊，不能盡言。病根積於『在圖還圖，在甲還甲』兩言，重以漏丁不查，報頂不允，而民之生理盡矣。夫甲止數人，若繫貧難，別無援救，貧者日貧，……其偏一也。……凡此五偏，猶其大略。至於瑣屑，**罄竹難書**。」《明史·鄒維璉傳》：「忠賢大奸大惡，**罄竹難書**。」清周西生《醒世姻緣傳》第54回：「作的那孽，**罄竹難書**。」

明清兩代，這個成語所適用的描寫對象都是罪惡或錯誤，其貶意十分強烈。然而，由於這種貶義在字面上看不到，現代人或望文生義，誤用此成語。鄒韜奮《抗戰以來》23：「淪陷區的同胞在抗戰中所表現的奇跡，真是所謂**罄竹難書**。」如果說作者在「罄竹難書」前面加了「所謂」一語，或表示其有意為之，下面的例子就不能說不是誤用了。

「關於老子，古多傳說，孰為史實，孰為謠言，今天已經難以分辨。加上老子又是道教的神，關於他的故事就更是**罄竹難書**了。比較樸實的還是司馬遷的記載。」（徐剛《古文源流考》，北京：北京大學出版社，2007:218）¹¹

我們把以上這類錯誤叫做「望文生義」。「文」指文字，指對那些字面義和實際義不一致的詞語，僅從字面義去理解和使用的語用錯誤。這種錯誤會導致語義的改變，甚至新詞的出現。

洋」（《莊子》）「茫洋」（唐孫樵《罵僮志》）等。又唐賈公彥《周禮正義》卷4：「『豕盲眊而交睫，腥。』……杜子春云：『「盲眊」當為「望視」。』」清佚名《周禮正義》卷8：「『豕盲眊而交睫腥』者，『盲眊』，《內則》作『望視』。『眊』『視』古今字，詳《太宰疏》。……案：《左》哀十四年傳云『陳豹……望視。』杜注云：『目望陽。』孔即本杜義。《家語·辨樂篇》王注云：『望羊，遠視也。』《莊子·秋水篇》『眊眊向若而歎。』彼《釋文》云：『本亦作「望洋」』，司馬、崔云：『「眊眊」猶「望羊」，仰視貌。』《莊子》『眊』與此『盲』字正同，『揚』『陽』『羊』『洋』字並通。段玉裁云：『「盲』『望』同音段借。子春不云『讀為』，云『當為』者，謂為聲之誤也。』」

¹¹ 《咬文嚼字》（上海文化出版社）2001年第3期載劉乾坤〈「關懷」怎用「罄竹難書」〉，對將這個成語用為褒義提出了質疑。該雜誌2003年第4期刊文〈「罄竹難書」的另一面〉，認為這個成語的貶義用法是一般用法，褒義用法是特殊用法，並不算錯。其實，歷時地看，特殊的用法正是一般用法的變異。如果再進一步問，這特殊的用法是如何產生的？不是由於「言語錯誤」又是由於什麼呢？

2.3 通俗詞源

通俗詞源 (folk etymology) 現象在所有語言中都存在。但在使用表意文字的漢語中，應有其獨特之處。以漢語為母語並受過教育的人，聽到一個新詞時，往往會思考這個詞的內部形式是什麼，如果寫下來，該如何寫。如果對其內部形式的解釋出現了錯誤，就會發生通俗詞源現象；如果再進一步用漢字記錄下來，一個新詞就產生了。¹²

例如，這些年，當人們談論考試的話題時，我們往往會聽到「/pāsi/」這個說法。不同的是，有些人通過考試時說，有些人考試失敗時說。例如：

(A) 問：考得如何？答：不錯，/pāsi/ 了。

(B) 問：考得如何？答：倒楣，/pāsi/ 了。

例 (A) 中的「/pāsi/」的意思是通過，顯然，這是源於英語「pass」的借詞；例 (B) 中的「/pāsi/」意思是不通過，它是哪個詞呢？原來，當會說英文的人借用「pass」一詞來表示「通過」義時，不懂英文的人既不瞭解這個詞的來源，也不瞭解這個詞的確切含義。他們根據這個詞的使用環境知道其與考試有關，從語音上推斷這個詞的內部形式可能是「*趴/拍死」，因此用在考試失敗的時候。值得一提的是，不論是哪個意思的「/pāsi/」，目前似乎還沒有完整的漢字書寫形式。¹³

另一個有意思的用例是「/máidān/」或「/mǎidān/」。近十多年，在北方的大小餐館中，流行「/máidān/」或「/mǎidān/」的說法，意思一樣，同為結賬或付賬，但內部形式不一樣。「/máidān/」直接來自粵語，在香港多寫成「埋單」，「埋」是個記音字，記錄的是粵語方言詞，意思是招呼侍者將帳單拿過來。¹⁴ 傳入北方後，除了用「埋單」外，還為這個詞創造出了另一種寫法「買單」（購買

¹² 對於漢語中的通俗詞源現象，前人已有許多精彩的研究。張紹麒 (2000) 是新近的重要成果。作者認為：通俗詞源「是語言變異研究的重要材料」。

¹³ 在百度網上，「考試 pass」的字樣很常見，都指考試通過；我們嘗試用「考試 pa 死」搜索，得到一個結果：「大三考研是不是可以報考，參加考試？報考時會不會被 pa 死吧？！」再用「考試 pasi」來搜索，得到一個結果：「可惜我此生與軍隊無緣，高考報志願的時候首先就把軍校 pasi 了。」再用「被 pa 死」搜索，得到一個結果：「我知道發上來可能會被 PA 死」。但沒有「趴死」。檢索時間：2009 年 4 月 8 日 10:00-12:00。

¹⁴ 這是中山大學周小兵教授告訴筆者的。在筆者的提示下，有香港本地學生說粵語口語裡有一個詞「mailai」，意思是叫人過來。又陳國球教授告粵語中「mai」有結算、會賬的意思，「maidan」應是結算賬單的意思。這樣看來，「埋」是一個記音字。在百度網 (<http://zhidao.baidu.com/question/7984412>，2012 年 1 月 20 日檢索) 上，有一些討論，可參看。

帳單)。更有人將「埋單」與結賬時侍者常常在遞賬單給付款人時，把單放在盤子裡，正面朝下「埋起來」，以免被客人看到數額的行業作法聯繫起來。這無疑就在北方話中創造出兩個新詞。

2.4 理據丟失

不少詞語，或因其字面晦澀難懂，造成理據（包括意義結構和語法結構）丟失。理據丟失往往造成對結構和語意誤解，引發「言語錯誤」。類似的錯誤可以叫做「理據丟失」。

「差強人意」是漢語古今常用的成語。《現代漢語詞典》（第五版，北京：商務印書館，2005:140）的解釋是：「大體上還能使人滿意。」所舉出的用例是：「那幾幅畫都不怎麼樣，只有這一幅梅花還**差強人意**。」

但在當代的言語活動中，這個成語有了一個新的意義。沈洪保的〈『差強人意』用錯多〉一文說：「最近幾年，許多記者在報導中都喜歡用『差強人意』這個詞。如『《七劍》首映上座率**差強人意**，只賣出十幾張票』、『這場演唱會，實在有些**差強人意**』、『台開案開庭，趙建銘的態度仍**差強人意**』、『《開創世紀》堪稱大手筆，但後四十集**差強人意**』、『上半時表現**差強人意**（六投一中，好像不會打球），但下半時起開始發威，數次以三分球命中，將比分拉開』……從這些例子看，他們都把『差強人意』用錯了，他們誤以為『差強人意』的詞意是『不能令人滿意』、『不盡如人意』、『不如人意』、『不盡人意』。在體育報導中用的頻率較高，凡比賽中『表現不佳』、『發揮不好』的，常用上一個『差強人意』。有人說該詞的錯用是幾年前某電視台體育頻道的主持人先用起來的，因為他們經常誤用『差強人意』，觀看體育頻道的青年人多，以訛傳訛，這詞的詞義就走向『反面』了。」（《溫州讀書報》2007-03-07）

我們姑且不論這個新用法是什麼時候發生的，始作俑者為何人，先來看看為什麼這個成語的意義會「走向反面」？最可能的原因是發生「言語錯誤」者不瞭解這個古代成語中的關鍵成分「差強」的內部語義形式，不瞭解「差強」的準確含義。而且，由於同樣的原因，這個「差強人意」在現代還出現了一個錯誤的形式變體「強差人意」。

在2008年12月，香港明愛醫院發生了一起病人在醫院門外因失救而不幸死亡的事故。香港各大媒體都報導了這件事，並轉述特區政府食物及衛生局局長周一嶽對此事的批評。如2008年12月23日的《星島日報》：「心臟病人上週六在明愛醫院門口失救，明愛醫院前天仍『死不認錯』，食物及衛生局局長周一嶽昨

批評醫院處理事件『強差人意』，直指醫院職員硬性跟從指引是一個『負累』。」同天的《頭條日報》以「周一嶽轟強差人意，醫管局明愛即道歉」為標題報導此事：「食環局局長周一嶽昨天指院方處理手法『強差人意』，又不滿有關職員死守指引辦事，要求院檢討。」¹⁵ 顯然，香港報章上所說的「強差人意」就是「差強人意」，因不瞭解「差強」的意思，使用者無意中「顛倒其辭」了。

有的時候，一個詞語的內結構並不複雜，卻也會發生類似的顛倒其辭現象。如 2011 年 11 月 15 日早上，筆者在鳳凰衛視中文台由吳小莉主持的「生命相髓·兩岸三地骨髓捐受者相見歡特別節目」中，第一次聽到「奔相走告」這個詞兒，用來形容在 2006 年蘇州發生的一起造血幹細胞配型成功（由台灣捐獻者提供）時，人們的興奮之情。「奔相走告」顯然是「奔走相告」的顛倒之形。上網查閱，才發現這並不是主持人的原創。¹⁶ 這個錯誤的發生，相信仍是與使用者對「奔走相告」的理據不十分清楚有關。

2.5 錯誤類推

前面說過由於「望文生義」類型的「言語錯誤」而使成語「望洋興歎」的內部形式發生了變異。但這個變異還沒有到此結束。

張中行《文言與白話》第 16 章「展望·文化遺產問題」：「不管《資治通鑑》《全唐詩》價值多高，你不會文言，就祇好望書興歎。」王蒙《堅硬的稀粥》：「十一點多，堂妹這一組佔著火做飯，由於挾爺爺之資深威重，別人祇能望火興歎。然後爸爸，然後叔叔。然後我能做飯時已是下午二時，祇好不做先去上班，然後晚飯同樣是望灶興歎。」梁曉聲《感覺日本·初識日本人》：「色情文化在日本不再是有豐利可圖的『文化商品』。即或色情描寫細膩到『官能小說』的地步，也還是望青年而興歎。」韓寒《三重門》第 16 章：「女性中如范曉萱之類頭髮短得可以讓喜歡扯住女人頭髮施威的暴君無處下手望頭興歎。」

以上用例中的「望書興歎」「望火興歎」「望灶興歎」「望青年而興歎」「望頭興歎」都產生於源自成語「望洋興歎」的固定結構「望 X 興歎」。這個結構是開

¹⁵ 筆者特別注意到，所有的媒體報導都作「強差人意」，可見不是某一媒體的筆誤。在北京大學 CCL 語料庫裡，「強差人意」的用例有 2 條，「差強人意」的用例有 63 條。在百度網上，與「強差人意」「相關網頁約 85,000 篇」，與「差強人意」「相關的網頁約 19,400,000 篇」。可見這個「言語錯誤」不僅僅發生在香港。檢索時間同上。進一步的討論將在另文做出（參看朱慶之：從「差強人意」到「強差人意」——對一個正在發生語言變異實例的初步觀察，載李向玉主編《澳門語言文化研究 2011》，澳門理工學院，2012 年，238-261 頁——作者補記）。

¹⁶ 在谷歌網查閱，「奔相走告」共有 58,700,000 條結果（2012 年 2 月 1 日）。

放的，從理論上講，「X」可以是任意一個能夠「望（看）」的名詞。而這個在現代漢語裡十分能產的表達結構，正來源於對成語「望洋興歎」的誤讀。

當一個錯誤發生後，有可能成為下一個「錯誤」的起點，類推出許許多多的相似的說法。這或許可以稱之為「錯誤類推 (mis-analogy)」，「望 X 興歎」同樣是一個較好的例子。

以上是我們在現實活的言語活動中看到的部分語言變異現象。類似的變異每天都在發生，稍加留意，不難觀察得到。由於有內省的能力，我們較容易地知道變異發生了，並且知道這些變異是如何發生的，原因是什麼。孤立地看，造成這些變異的原因可能千差萬別，但集中起來，規律十分清楚，歸納並不困難。如果進一步抽象，可以用一個詞來表述，這就是「言語錯誤 (speech-mistake)」，即語言的使用者在其日常言語活動中，對某些語言成分作了「錯誤理解 (misunderstand)」並導致「錯誤使用 (misuse)」。

3. 推衍：古代漢語變異例釋

由各類「言語錯誤」導致的語言變異為現代之獨有，抑或在古代同樣發生？按照社會語言學的理論，回答當然是肯定的。在古代，類似現代漢語中的「言語錯誤」同樣會不時發生，也正是許多變異的根本原因。以下主要討論成語中的一些例子。同樣先從「認錯字、讀錯音」說起。

3.1 認錯字、讀錯音

比如「不識一丁」，字面意思是連一個「丁」字都不認識。形容人沒有文化，未受教育。這個成語語本《新唐書·張宏靖傳》：「天下無事，爾輩挽兩石弓，不如識一丁字。」宋岳珂《鄂國金佖續編》卷1「高宗皇帝宸翰撫遺·紹興七年」：「於維紹興，扶危支傾。握校蒞戎，**不識一丁**。」元吳澄《吳文正集》卷42「舒卷堂記」：「察其為人，稽其行事，胸蟠萬卷之儒，或不如目不識一丁之夫，何哉？讀而不知其所以讀也。」明邵景瞻《覓燈因話》卷1「桂遷夢感錄」：「公不見吳之張萬戶、李都赤，**不識一丁**，而食祿千石，是何人也！」

但宋王楙《野客叢書》卷21「一丁字」條說：「今文人多用『不識一丁字』，祖《唐書》『挽兩石弓不如識一丁字』。出處僕考之，乃『个』字，非『丁』字。按，《續世說》書此，『个』字。蓋『个』字與『丁』字相似，傳寫誤焉。後又觀張翠微《考異》，亦謂『个』字，乃知《世說》之言為信。」「个」是

「箇」的俗寫，與「丁」形近而被認錯是完全有可能的。「不識一丁字」本是「不識一个字」。

由此又生發出「目不識丁（字）」的說法兒。元汪克寬《環谷集》卷5「西南林壑樓記」：「出入榛棘叢，血濺趾蹙；目不識丁字，罔知詩書。」明馮夢龍《警世通言》卷17：「他兩個祖上也曾出仕，都是富厚之家，目不識丁，也頂個讀書的虛名。」

另一個「認錯字」的例子是「雞口牛後（寧爲雞口，不爲牛後）」，意思是「寧可在局面小的地方自主，不願在局面大的地方聽人支配」。

《史記·蘇秦列傳》：「且大王之地有盡而秦之求無已，以有盡之地而逆無已之求，此所謂市怨結禍者也，不戰而地已削矣。臣聞鄙諺曰：『寧爲雞口，無爲牛後。』今西面交臂而臣事秦，何異于牛後乎？夫以大王之賢，挾強韓之兵，而有牛後之名，臣竊爲大王羞之。」《史記》的這段話可能取自《戰國策》，今本《戰國策·韓策一》「寧爲雞口，無爲牛後」云云相同。

但北齊顏之推《顏氏家訓·書證》說：「太史公曰：『寧爲雞口，無爲牛後。』此是刪《戰國策》耳。按延篤《戰國策音義》曰：『尸，雞中之主；從，牛子。』然則『口』當爲『尸』，『後』當爲『從』，俗寫誤也。」¹⁷ 如果顏之推不錯，這就是一則典型的「魯魚亥豕」式的言語錯誤。

以上都是「形近致誤」。下面的例子應是「同形致誤」。

「分庭抗禮」這個成語形容「（各方）地位相等，可以相互抗衡」。南朝陳姚最《續畫品·序》：「矯然獨步，終始無雙。有若神明，非庸識之所能效；如負日月，豈末學之所能窺！荀衛曹張，方之蔑矣；分庭抗禮，未見其人。」宋徐夢莘《三朝北盟會編》卷153「炎興下帙」：「本朝郎官出使序位在轉運之上，凡此蓋欲尊重天朝，習民於上下之分也。故事宰相坐待漏院，三衙軍官於簾外倒杖，聲喏而退。今見在分庭抗禮矣。」在現代漢語中，「分庭抗禮」進一步發展出搞分裂和鬧獨立的用法，如張一弓《趙鏞頭的遺囑》6：「這是不是與集體經濟分庭抗禮？是不是挖社會主義的牆腳？」

一般認爲，這個成語源自《莊子·漁父》：「萬乘之主千乘之君，見夫子未嘗不分庭抗禮，夫子猶有居傲之容。」成玄英疏：「伉，對也。分處中庭中，相對設禮，位望相似，無階降也。」是知「伉禮」指賓客和主人分別站在庭中的兩邊，相對行禮，以平等的地位相待。爲什麼後來意思從「平等」變爲「抗衡」？

¹⁷ 張守節《史記正義》說：「《戰國策》云：『寧爲雞尸，不爲牛從。』延篤注云：『尸，雞中主也。從，謂牛子也。言寧爲雞中之主，不爲牛子之從後也。』」參看鄭慧生：說「寧爲雞尸，無爲牛從」，《文史知識》2007年第9期，北京：中華書局。

原來，後人寫別字（通假字），將「伉」寫作「抗」，如《史記·貨殖列傳》：「〔子貢〕所至，國君無不**分庭與之抗禮**。」以致「抗禮」取代了「伉禮」，遂成爲通行寫法。¹⁸如宋謝良佐《上蔡語錄》卷 1：「孟子強勇，以身任道，後車數十乘，從者數百人，所至王侯**分庭抗禮**，壁立萬仞，誰敢正覲者？」元辛文房《唐才子傳》卷 8「陳搏」條：「太宗征赴，戴華陽巾，草屨垂條，與萬乘**分庭抗禮**，賜號希夷。」明彭大翼《山堂肆考》卷 68「築館招賢」條：「宋余玠安撫四川，築招賢館以禮士。時播州人冉璉與弟璞聞之，遂謁玠，玠與**分庭抗禮**，待之甚至。」其中的「抗禮」都是「伉禮」。「抗」成了「伉」的假借字後，就爲「同形致誤」創造成條件。後人不知此「抗（伉）」非彼「抗（對抗）」，最終導致這則成語意義的改變。

「讀錯音」而導致新詞新義在古代亦有，比如「好自爲之」和「箕風畢雨」。

《淮南子·主術》：「君人者不任能，而**好自爲之**，則智日困而自負其責也。」其中的「好自爲之」的意思是喜歡親自做事。「好」《廣韻》音「呼到切」，義「愛好」，是動詞，今讀去聲。「好自爲之」後來成爲成語。時間晚一點兒的用例如清楊懋建《京塵雜錄》卷 1「長安看花記」：「天下事惟在有心人**好自爲之**，豈獨繪事小道爲然？」清李伯元《文明小史》第 20 回：「劉學深於演說會，欲賺二八扣頭，可見上海地方，固隨在皆可賺錢，亦在人**好自爲之**耳。」《梁啟超文集》卷 6「霍布士學案」：「我父雖**好自爲之**，而我則未嘗預其事也。」

不過，大約從元代開始，這個成語有了另一種用法，勸人自己好好幹下去。《新元史·客烈亦王罕桑昆傳》：「桑昆睒然而出。王罕呼使反曰：『吾老矣，但思聚骸骨於一處，汝乃喋喋不已。**好自爲之**，毋貽吾憂可也。』」桑昆決意殺太祖。」《明詞匯編》王屋《前調（擬上堂語）其三》：「佛不插公話事，請君**好自爲之**。」清宣鼎《夜雨秋燈錄》卷 2：「此去仍爲僧，仰爲儒，均自便，非阿翁所能計及也。**好自爲之**，前途鄭重。」清許奉恩《里乘》卷 4「姁兒」條：「君幸努力自愛，**好自爲之**，何必抵死與人爭鬪體哉！」知新室主人譯《失女案》（載《新小說》第 2 年第 4 號）：「又謂奇山曰：『你**好自爲之**，爾命懸於吾掌中矣。』」《紅岩》第 27 章：「望你**好自爲之**，不負黨國重托。」劉震雲《一地雞毛》5：「爹娘沒本事，送你到居委會幼兒園，你以後就**好自爲之**吧。」但不同的是，其中的「好」《廣韻》「呼皓切」，義「善也，美也」，後來轉爲副詞，今讀上

¹⁸ 《說文解字·手部》：「抗，扞也。」今按，查先秦典籍，「伉禮」出現過三次，字均作「伉」而非「抗」。這個詞在《史記》中也出現三次，字均作「抗」。或許司馬遷是始作俑者。

聲，其實變成了一個新成語。分析這個「好自爲之」產生的原因，顯然也是一種言語錯誤，即將其中「好」的詞性誤解讀錯，意思也發生了變異，一直用到現代。

「箕風畢雨」猶言好風好雨，風調雨順。是對統治者的恭維話，現代已經鮮用。《藝文類聚》卷 7 南朝梁吳均《八公山賦》：「神基巨鎮，卓犖荆河，**箕風畢雨**，育嶺生峨。」南宋衛涇《後樂集》卷 8「表狀·瑞慶節賀表」：「臣某甲中賀恭惟皇帝陛下，道契兩儀，功高八極，掩井鉞參旗之耀，調**箕風畢雨**之和。」元王義山《稼村類稿》卷 21「代湖南倉司謝賜曆日表」：「璧月珠星，驗漢史太初之奏；**箕風畢雨**，叶洛書五紀之疇。」

但這個成語產生的原因也是使用者讀錯音。一般認為，「箕風畢雨」源自《尚書·洪範》：「庶民惟星，星有好風，星有好雨。」孔傳：「箕星好風，畢星好雨。」「好」也應是去聲，爲動詞，意思是人們的好惡各不同。後將「好」誤讀爲上聲，變成好壞之好，遂造「箕風畢雨」。

3.2 望文生義

古代的「望文生義」型的言語誤用，同樣可以從「望洋興歎」談起。

前邊說過，「望洋」是一個聯綿詞，其書寫形式可以是「望洋」，也可以是「眎洋」「望羊」「茫洋」「滂洋」「望佯」等。今人望文生義，將「望洋」解釋成「望著海洋」，古人也有同樣的錯誤。清桂馥《札朴》卷 4「望羊」條：「《春秋元命苞》：『武王柔肩望羊。』《晏子春秋》：『杜扃望羊待於朝。』《大戴禮》：『武王望羊。』哀公十四年《左傳》：『長而上僂，望視。』杜注：『目望羊。』字又作『洋』。《莊子·秋水篇》：『河伯始旋其面目，望洋向若而歎。』《釋文》云：『司馬、崔云：望洋猶望羊。仰視貌。』馥案：『正作「陽」。』《論衡·骨相篇》：『武王望陽。』」¹⁹ 桂馥「正作『陽』」是說「羊」「洋」「陽」三種寫法，只有「陽」才是正字，其他均爲假借字。我們知道，聯綿詞是單純詞，常常詞無定字，桂馥卻認為「望洋」一詞的各種寫法中，只有「望陽」是本字。這樣說的依據顯然來自「仰視貌」的意義，同樣是望文生義。

其實，對「望洋」的望文生義大概在宋代就已經出現了。宋釋道成《釋氏要覽·序》：「於戲！教海文繁，**望洋**懷怯，欲截其源而得其要脈者，難矣哉！」「教海」與「望洋」對文。宋鄭清之《可齋陳大卿政成之暇蒐討河渠爲鄉國長久

¹⁹ 《大戴禮》「武王望羊」，今本無，見《白虎通義》卷 7「聖人」條引；《左傳》注今本作「目望陽」。

慮開萬世利，非君侯其誰屬？因效一得以廣盛心焉》詩：「……東有錢湖浸，寒玉渺無價。西有它山源，盈科通晝夜。維此兩支邑，厥田俱下下。問之何因爾，水道無所舍。河伯空望洋，旱魃巧乘罇。緬思井田規，畎澮分激汙。……」韓澆《次韻王教授同寄昌甫器遠無爲臨漳》詩：「甌閩何處是，絕嶺更危谿。乘障消蠻貊，望洋蒐水犀。庾公情不淺，朱老世思齊。乘月胡床坐，角巾隨意低。」再如清屠紳《蟬史》卷 1「甲子城掘井得奇書」：「望洋知道岸雲遙，觀海覺文瀾甚闊。蕭閑歲月，非著書何以發微；浩淼煙雲，豈坐井而能語大。」其中「望洋」似乎都不是「茫然不知所措」或「仰視貌」的意思。而清人李慶辰《醉茶志怪》卷 1：「未幾，賊氛已退，榮議還鄉。買一巨舟，裝載行李，待風順起程。……適布帆翩翩作響，榮命家人登舟，中流擊楫，片刻已十餘里。生望洋興歎，無可奈何，恨不即生雙翼，飛過長河。」朱庭珍《筱園詩話》卷 2：「猶登山者一望昆侖，觀水者一朝南海，即侈然自足，而不知五嶽四瀆九江五湖三十六洞天奇之，天下尚別有無數妙境界也。則拘于方隅，必不能高涉昆侖之顛，遠航大海之外，徒自崖而返，望洋興歎已耳。」其中的「望洋興嘆」好像就是「望著茫茫大洋而發出無可奈何的感嘆」。以下再舉另一個例子。

黃庭堅《次韻德孺惠貺秋字之句》詩：「少日才華接貴遊，老來忠義氣橫秋。」「氣橫秋」形容老練而自負的神態，亦可形容蒼老挺拔的氣勢，褒義。由此形成了一個成語「老氣橫秋」。如宋樓鑰《攻媿集》卷 5「古體詩·題楊子元琪所藏東坡古木」：「東坡筆端遊戲，槎牙老氣橫秋。」清陳少海《紅樓複夢》第 55 回：「這棵老梅又生得古雅卓犖，老氣橫秋，我愛之欲死。」丁儀《詩學淵源》卷 8「李攀龍王世貞謝榛張佳胤」條：「近體五七言，悲壯蒼涼，老氣橫秋，更非於鱗世美兩家所及。」孫殿起《琉璃廠小志》第 1 章「琉璃廠匾額」條（出《春明古跡小識》）：「『茹古』爲翁叔平所寫，渾脫瀟灑，老氣橫秋。」

大約到了清代，「老氣橫秋」的意思發生了變化，有人望文生義，用來形容人擺老資格，或人沒有朝氣（尤其指年青人）。成爲一個貶意的詞語。如清許奉恩《里乘》卷 2「林妃雪」：「師雄僅托於夢，猶不敢公然唐突，惟有老逋無賴，判將一種清寒骨，老氣橫秋，硬呼我輩作妻，不尤令人噴飯耶？」吳趸人《二十年目睹之怪現狀》第 70 回：「眾人取笑了一回，見新人老氣橫秋的那個樣子，便紛紛散去。」陳森《品花寶鑑》第 17 回：「子雲道：『琪官性情剛烈，相貌極好。似欠旖旎風流。比他爲菊花罷。』高品道：『菊花種數不一，有白有黃，或紅或紫。白的還好，其餘似老氣橫秋。』」老舍《趙子曰》：「趙子曰老氣橫秋的用食指彈了彈煙灰，真帶出一些老大哥的派頭。」陳建功《皇城根》：「跟你們搞文藝的比，怎麼也不行。你們成天歡蹦亂跳的，走南闖北，有說有笑，一比，我

就顯得**老氣橫秋**。」王朔《我是你爸》：「還有，你們屁大的孩子，互相亂叫什麼『老李』『老張』的？小小年紀一個個**老氣橫秋**的，看著也不像啊。」

古樂府《雞鳴》：「桃生露井上，李樹生桃傍。蟲來齧桃根，李樹代桃僵。樹木身相代，兄弟還相忘。」詩中用桃李共患難比喻兄弟相愛相助。後來出現了一個成語「李代桃僵」，顯然源自《雞鳴》，但意思由褒變貶，比喻互相頂替或代人受過。宋劉攽《彭城集》卷 26「表·知襄州謝上表」：「果聞係踵之後人略取析豪之上，第涇以渭濁，故常畏於後生**李代桃僵**。」明賀復徵《文章辨體匯選》卷 117 馮琦「速賜考選以光聖治疏」：「況以前人之觸忤，杜後人之進用，疊受瓶恥，**李代桃僵**。」舊題明陳繼儒《真傀儡》：「古來史書上呵，知多少**李代桃僵**。」錢謙益《嘉興高氏家傳》：「寢殿圯，君代用抵罪，**李代桃僵**。」清李塨《周易傳注》卷 2：「夫邑人安居，自謂與人無患，與世無爭，何嘗有妄而橫被災禍，天下之**李代桃僵**者豈少哉！」發生如此變化的原因顯然是在脫離上下文的環境中出現了望文生義的言語錯誤。

3.3 理據丟失

前邊已經提到，《莊子》「每下愈況」的本義是愈往下，愈能比況出「道」的本質。但在宋代的文獻裡，這個成語意思大變，形容事物的境況一天不如一天。為什麼會有如此大的變化？原因應是「理據丟失」。

我們注意到，大概從宋代起，「每下愈況」被「顛倒其辭」，說成「每況愈下」。如宋李新《跨鰲集》卷 19：「昔人求禮於野人，求道於瓦礫，問迷於童子，**每況愈下**，謂愚者千慮，或有一得。」明葉子奇《草木子》卷 2：「莊子曰：道在秕稗，在瓦礫，在尿溺，**每況愈下**。蓋以道無乎不在也。」清黎元寬《豐城鼎建大正法寺碑記》：「……此猶是一家事，亦或不若正獲之間於監市履豨也，**每況愈下**者乎？」清洪吉亮《北江詩話》卷 3：「大抵釋氏書之精，皆莊、列之緒餘也；其至粗如『道在屎橛』等，釋氏亦竊之。南宋儒者似又竊釋氏緒餘，此即莊子所謂『**每況愈下**』也。」對此，清胡鳴玉編《訂譌襍錄》卷 8「每況愈下」條提出過批評：「《莊子·知北遊》『正獲之間於監市履豨也，每下愈況』，言每於至下之處，愈可比況他物。今用作『每況愈下』，非。」

「每下愈況」誤為「每況愈下」後，立即被按字面義來使用，產生了一個過渡性的新用法，即「用來比況的事物一個比一個低下」。用例如宋陸佃《埤雅》卷 14：「〔《詩》〕又曰『鴉鳴在桑，其子在梅，其子在棘，其子在榛』者，蓋先實者梅，後實者棘；先實者棘，後實者榛，故其序如此，亦其榛卑小於棘，棘卑小

於梅。詩以刺之，故**每況愈下**也。」宋胡仔《漁隱叢話後集》卷 26「東坡一」：「子瞻自言平生不善唱曲，故間有不入腔處，非盡如此。後山乃比之教坊司雷大使舞，是何**每況愈下**。蓋其謬也。」宋洪邁《容齋續筆》卷 8「蓍龜卜筮」條：「人人自以爲君平，家家自以爲季主，**每況愈下**，由是藉手于達官要人，舟車交錯于道路，毀譽紛紜而術益隱矣。」正是這種不顧語源的法，最後導致了新義的產生。其後的用例如元陸文圭《牆東類稿》卷 3「策·儒學吏治」：「宋人又愧唐人矣，人才不古，**每況愈下**，少知治體者，豈不寒心哉！」清黃宗義《外舅葉公改葬墓誌銘》：「自公云亡，**每況愈下**，諸張時文，啞鐘不打。」其中的「每況愈下」都是事情越來越糟的意思。

進一步，「每況愈下」的這個新義又「傳染」給了「每下愈況」。如清陳瀏《旬雅》卷上：「……客貨雖係疙疸釉而鮮妍古潤，聲價亦高。乾嘉以後，**每下愈況**，不堪入目。」清吳趸人《〈兩晉演義〉序》：「《三國演義》出，而膾炙人口，自士大夫以至輿台，莫不人手一篇。人見其風行也，遂競效之，然**每下愈況**，動以附會爲能，轉使歷史真象隱而不現彰。」

再比如「差強人意」。前面討論了這個成語在現代正在發生的語義和結構變異，但故事還沒有完，它的現代意義亦是古代一次語義變異的結果。根據《現代漢語詞典》，該成語的「標準意義」是「大體上還能使人滿意」。這個意義大約是唐代出現的。《唐大詔令集》卷 61「封魏伯玉陽成郡王制」：「上柱國芮國公魏伯玉，……累膺將帥之重，積有艱鉅之勳，殫力事邊，乃誠王室。伏波料敵每合我心，司馬理軍**差強人意**。」明丘浚《大學衍義補》卷 24「經制之義下」：「德宗爲君，楊炎爲相，無可取者，惟此一事**差強人意**，范祖禹謂炎知爲相之體，德宗知爲國之務，後世所當取法者也。」

但「差強人意」的最早出處是《後漢書·吳漢傳》：「諸將見戰陳不利，或多惶懼，失其常度，漢意氣自若，方整厲器械，激揚士吏。帝時遣人觀大司馬何爲，還言方修戰攻之具，乃歎曰：『吳公**差強人意**，隱若一敵國矣。』」對於「吳公差強人意」云云，古人的解釋是：差：甚；強：振奮；人意：士氣。意思是吳漢甚能提振將士的士氣。《魏書·于栗磾傳附于忠》：「元禧之謀亂也，車駕在外，變起倉卒，未知所之。忠進曰：『臣世蒙殊寵，乃心王室。臣父領軍付留守之重，計防遏有在，必無所慮。』世宗即遣忠馳騎觀之，而烈分兵嚴備，果如所量。世宗還宮，撫背曰：『卿**差強人意**。』賜帛五百匹。」用法相同。「人意」本是人的士氣或意志，卻被誤解爲「〔令〕人滿意」，應當是理據丟失的結果。

3.4 重新分析

《公羊傳·文公九年》：「〔經：〕冬，楚子使椒來聘。〔傳：〕椒者何？楚大夫也。楚無大夫，此何以書？始有大夫也。始有大夫，則何以不氏？許夷狄者，不一而足也。」注：「許，與也。足其氏，則當純以中國禮責之。嫌夷狄質薄，不可卒備，故且以漸。」意思是說既然楚已有大夫這樣的官職，為什麼不用「氏」來稱呼椒這個人？（因為）給未開化者（東西），不要因其做了一件事就讓他完全滿足。《漢書·陳湯傳》：「御史大夫貢禹博士匡衡以爲《春秋》之義『許夷狄者，不壹而足』，今郅支單于鄉化未醇，所在絕遠，宜令使者送其子至塞而還。」顏師古注：「言節制之，不皆稱其所求也。」其中的「不一而足」成爲一個成語。如宋鄭剛中《周易窺餘》卷 10「象曰或益之自外來也」條：「夫君之益臣，猶不一而足，彼賢智之士，能與二爲異乎？」

然而，不晚於宋代，這個成語意思發生了變化，從原先的「不要因其一而使其滿足」，變爲指同類的事物很多，不能一一列舉。如宋趙與時《賓退錄》卷 1：「周宣王，中興之賢君也。然考之於《詩》，曰箴，曰規，曰誨，曰刺，不一而足，第序《詩》者不能直書其事，故後世儒者無敢訾議。」明沈德符《萬曆野獲編·補遺》卷 3「刑部·重刊閩範序」：「上溯唐虞三代，下迄漢宋我朝，賢后哲妃，貞婦烈女，不一而足。」明凌濛初《二刻拍案驚奇》卷 34：「試看紅拂離了越公宅，紅綃逃了勳臣之家，此等之事，不一而足，可見生前已如此了。」清陳森《品花寶鑒》第 31 回：「近來卓然有些古怪，找他的不一而足。」

造成「不一而足」語義變化的原因，應是所謂的「重新分析 (reanalysis)」。
類似的變化還可以舉出「卑之無甚高論」和「安榮富貴」。

「卑之無甚高論」意思是所說見解一般，沒有什麼特別之處。宋胡次焱《梅巖文集》卷 1「山園後賦」：「子胡子喟然歎曰：文房四友，惟茲四子，卑之無甚高論，又何足與議？」明畢自嚴《度支奏議》卷 15「分委司官給散城守行糧疏」：「以上數款，雖卑之無甚高論，然使三軍果腹、萬口含哺。」又卷 16「遵奉聖諭議修屯政疏」：「以上十款皆卑之無甚高論。」明馬世奇《澹寧居文集》卷 10「與陳補思父母」：「以上數條皆卑之無甚高論，然地方切要，似無逾此。」清施克洪《畫學心法問答·跋》：「所見如是，所志如是，奚怪乎卑之無甚高論，而於繪事，何異買櫝而還珠乎？」

這個成語出自《史記·張釋之馮唐列傳》：「釋之既朝畢，因前言便宜事。文帝曰：『卑之，毋甚高論。令今可施行也。』」原意是要他談當前的事情，不要空發過高的議論。宋魏齊賢等編《五百家播芳大全文粹》卷 37 載熊子復《謝及第

啓》：「進言宜**卑之無甚高論**，寧姑爲升斗之計不敢犯咫尺之威？」仍是原來的用法。

「重新分析」大概是在宋代發生的。宋王楙《野客叢書》卷 2「卑之無甚高論」條：「今人以卑之無甚高論之語，卻所說之卑者，甚失當時之意。……所謂『卑之無甚高論』者，文帝懼釋之陳五帝三王上古久遠之事，無益於時，故令陳今可行之說。釋之遂言秦漢之事，文帝所以稱善。則『卑之，無甚高論』，自是兩句，今人作一句讀之，所以失當時之意也。」是知將兩個獨立的句子作了錯誤地重新分析，當作一個短語，並且將「卑」理解爲低下，約是宋人的創新。

「安富尊榮」指安於富貴的生活。《紅樓夢》第 2 回：「如今人口日多，事務日盛，主僕上下，都是**安富尊榮**，運籌謀畫的竟無一個。」

這個成語出自《孟子·盡心上》：「孟子曰：『君子居是國也，其君用之，則**安富尊榮**。』」「安富尊榮」本是並列結構。安，平安；富，富有；尊，地位高；榮，聲望大。宋呂本中《呂氏春秋集解》卷 4：「上使其君保**安富尊榮**之位也。」元景星《中庸集說啓蒙》卷上：「下面四句是**安富尊榮**之事，皆父母所願者，今皆得之。」而《紅樓夢》的用法，則將並列結構重新分析成動賓結構，不合原意。

3.5 省略失當

省略本來是常見的語用現象。但是如果省略導致語意改變，甚至導致新的語言形式出現，這種省略就是錯誤省略。

「安之若素」指面對某事安然而不驚慌，和往常一樣。清李寶嘉《官場現形記》第 38 回：「第二天寶小姐酒醒，很覺得過意不去。後來彼此熟了，見瞿太太常常如此，也就**安之若素**了。」《新青年》第 3 卷第 5 號載胡適文：「人皆知美爲自由之國，而俄爲不自由之國，而不知美爲最不爱自由之國，而俄爲最愛自由之國也。美之人已得自由，故其人**安之若素**，不復知自由代價之貴也。」指習以爲常。在近代，進一步引申爲麻木不仁。如姚靈犀《采菲錄》下編李增《遷安遵化天足會序》：「夫吾國號稱四萬萬人，因纏足而癱瘓者，不下兩萬萬人。楚毒同類，以爲風尚，天下可怪之端，未有過於此也。然竟習焉不察，**安之若素**。」梁啓超《知恥學會敘》：「民惟無恥，百人之中，識字者不及三十，**安之若素**。」

這個成語源自習語「安之若素～」，在「～」位置上的通常是動詞「有/無」或「習」。如宋曹勣《松隱集》卷 23：「竊見諸路州縣見行火甲以禦夜盜，民**安之若素**。臣嘗推行，備見其便。」《宋史·余玠傳》：「播州冉氏兄弟璵璣有文武才，隱居蠻中，前後閩帥辟召堅不肯起。聞玠賢，相謂曰：『是可與語矣。』」遂

詣府上謁。玠素聞冉氏兄弟，刺入即出見之，與分廷抗禮，賓館之奉，冉**安之若素有**。居數月，無所言。玠將謝之，乃爲設宴，玠親主之……」元王義山《稼村類稿》卷 19「講義」：「子文處人情甚不堪之時，而甚**安之若素所無者**。」「素有」猶言「平素所有者」，「素習」猶言「平素所習者」。

大約在明代，「安之若素～」開始出現省略的形式——「安之若素」，仍爲「安之若素有/習」之意。如明范允臨《輸寥館集》卷 4「林母徐孺人傳」：「晚歲，奉釋氏，食不二簋，啖茹衣單，**安之若素**。」明毛伯溫《毛襄懋文集》卷 4「藍雨田先生記」：「旋跡故里，以獎進後學爲己任，以燕翼孫子爲家謀。株守硯田，**安之若素**。」多指對窘境毫不介意，心情平靜得像往常一樣，有明顯的褒義。大約到了清代，這個成語也可用於一般義，指習以爲常，不以爲怪。這時的時間詞「素」就由原來修飾「有/習」的副詞，變爲名詞直接做動詞「若」的賓語。又引申指對於錯誤的言論或事物不聞不問，聽之任之。

4. 結論和餘論

4.1 以上所討論的漢語詞彙變異，不論古今，其發生的原因和機制，都可以歸結爲「言語錯誤」。

這裡所說的「錯誤」，是指語言使用者不按社會習慣或已有的規則使用語言的現象。錯誤主要發生在個人言語過程中，在無意識或被動情形之下，但也可能是故意或主動的，比如爲了修辭或追求新奇等的需要。「言語錯誤」在任何語言活動裡都會出現，具有跨越時空的普遍性。

「言語錯誤」分「個人言語錯誤」和「集體言語錯誤」。前者是「言語錯誤」的最初階段，錯誤還局限在始作俑者身上，沒有擴散開來；一旦個人錯誤被比較多的人接受，就變爲後者。後者的結果往往會導致一種新的意義或用法，甚至新詞語的產生。

需要強調的是，「言語錯誤」其實是語言變異的一個相當重要的原因，由「錯誤」而導致的語言變異與語言按照自身發展規律而產生的變異交織在一起。而造成「錯誤」的原因是各種各樣的，從理論上說，不同語言會有各自的特點。

雖然在以上討論中我們將不同的變異現象分別歸在不同的類型之中，其實變異往往是多種原因的結果，或者可以從不同的角度進行觀察。如將「望洋興歎」中的「望洋」理解爲「望著大海」，既與誤讀文字的「望文生義」有關，也與「通俗詞源」有關；「差強人意」從其本意到新義，既是理據丟失的結果，也可能是望文生義的產物，甚至是重新分析的創造，不一而足。有一個重要的特徵已

經讓我們能夠看清楚了，它們的共同點似乎是：發生在漢語中的「言語錯誤」多與漢字有關，進一步說，與漢語使用漢字作為文字並由此產生的民族文化心理有關。²⁰

2002 年筆者在韓國外國語大學教書期間，觀察到韓國人與（母語為漢語的）中國人在語言與文字關係之認識上的明顯差異，這是由於韓國人使用記音的諺文，而中國人使用表意的漢字。例如韓國有一個地方，名字的拉丁字母寫法是「osan」（音 [ou san]）。我問韓國人「osan」是什麼意思？他們大都一臉茫然，告訴我說這只是一個地名而已，不明白我為什麼要問這樣的問題。後來知道，它的漢字寫法是「烏山」。看到這個寫法，可以猜測它以山名得名，而山之所以名「烏」，可能與其山石的顏色有關。一個受過教育的中國人對這樣的問題會覺得很平常——當我們聽到某個地方叫「wūshān」時，一定會問：哪個「wū」？怎樣寫？當我們聽到一個人名時，也常會問：「怎麼寫？」也常聽到這樣的回答：「是干勾『于』」，「是口天『吳』」，等等。漢字不是記音文字，漢語的詞以單音節和雙音節居多。同音詞固然很多，但用來記錄同一音節的漢字同樣很多，使用時必須按照意義來進行選擇，這就造成講漢語的人對音、義和字、義關係的特殊敏感。當一個人聽到一個詞時，立即會想到什麼意思？怎麼寫？如果此時沒有人告訴你確定的寫法，我們就會根據自己的判斷來決定用哪個漢字。同樣，當一個人看到一個或一組字時，會根據字形（而非它所代表的音）來判斷它所記錄的是哪個詞（意義）。這些個過程反映的是我們對詞義做出了怎樣的解釋。如果理解與原意吻合，就不會發生「言語錯誤」；如果不吻合，就會發生「言語錯誤」。相關的問題有待進一步討論。²¹

4.2 本文涉及的漢語在當代的變異大都是言語 (speech) 性質的，也就是說正處在變異的過程中，能否從言語現象成為語言 (language) 現象，還很難說。但是正如社會語言學提倡的用語言的現在解釋語言的過去一樣，我們也可以用語言的過去預測語言的未來。這正是語言研究的終極目的之一。比如「差強人意」變異為「強差人意」是最近二三十年的事情，形式發生了變異，意義也從褒義產生出貶

²⁰ 根據譚力海 (2011) 介紹，他的研究團隊發現大腦的中國語言功能區有其特別之處，尤其是負責漢字的書寫和認讀的部分，而非以往西方學者研究所揭示的那樣，不同類型的語言和文字由同一語言中樞主管。如果這一成果得到證實，似可進一步推測漢語中的「言語錯誤」可能還有其生理學的原因。

²¹ 張紹麒 (2000) 說：「漢字在流俗詞源產生過程中可以起到重要的作用。聲訓（引按，俞敏用語）可以導致人們用不同的漢語記錄這個詞，俞先生稱這種字為假本字。……假本字寫常了，而且被大家接受了，流俗詞源也就形成了，並且固定了下來。」

義的用法，其情形與「每下愈況」的變異幾乎完全一樣。「每下愈況」因理據丟失而變異為「每況愈下」，又因望文生義而產生「越來越糟」的貶義。那麼，我們可以預測，「強差人意」極有可能從言語的詞變為語言的詞，而「差強人意/強差人意」的貶義用法也極有可能會從言語義變為語言義。²²

本文討論的「言語錯誤」現象集中在詞彙方面，其實，漢語語法甚至是語音的歷史演變，「言語錯誤」也是重要的原因。我們期待有興趣的同行參與討論。²³

引用文獻

- Crystal, David. 1997. *The Cambridge Encyclopedia of Language* (2nd edition). Cambridge & New York: Cambridge University Press.
- Crystal, David. (ed.) 2000. *Xiandai Yuyanxue Cidian* 現代語言學詞典 [*A Dictionary of Linguistics and Phonetics*] (4th edition), translated by Jiaxuan Shen (沈家煊). Beijing: The Commercial Press.
- Labov, William. 2001. *Principles of Linguistic Change*, Vol. 2: *Social Factors*. Oxford: Blackwell.
- Liu, Jiexiu (劉潔修). 1989. *Hanyu Chengyu Kaoshi Cidian* 漢語成語考釋詞典 [*The Dictionary of Chinese Idioms with Textual Criticisms and Explanations*]. Beijing: The Commercial Press.
- Luo, Zhufeng (羅竹風). (ed.) 1986-1993. *Hanyu Da Cidian* 漢語大詞典 [*Comprehensive Chinese Word Dictionary*]. Shanghai: Shanghai Lexicographical Publishing House.
- Northwest Normal University, Department of Chinese Language and Literature. (eds.) 1986. *Hanyu Chengyu Cidian* 漢語成語詞典 [*Dictionary of Chinese Idioms*]. Shanghai: Shanghai Education Press.
- Tan, Li-Hai (譚力海). 2011. Brain systems underlying Chinese language processing: insights from functional neuroimaging studies. Keynote speech delivered at the International Conference on Chinese Language Learning and Teaching in the Digital Age, November 25-27, 2011. Hong Kong: Hong Kong Institute of Education.
- Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. *Directions for Historical Linguistics*, ed. by Winfred P. Lehmann & Yakov Malkiel, 95-195. Austin: University of Texas Press.

²² 筆者調查了上世紀至今近一百年的語料，對「差強人意」到「強差人意」的變異情況作了精確的資料調查，結果將另有專文討論。

²³ 在 IACL-18 的討論中，汪維輝教授提示漢語語音史上一些特殊音變，即極有可能是「言語錯誤」造成的。

- Zhang, Shaoqi (張紹麒). 2000. *Hanyu Liusu Ciyuan Yanjiu* 漢語流俗詞源研究 [*Study on Folk Etymology in Chinese*]. Beijing: Beijing: Language and Culture Press.
- Zhang, Shunhui (張舜徽). 1983. *Shuowen Jiezi Yuezhu* 說文解字約注 [*A Brief Annotation to Shuowen Jiezi*]. Zhengzhou: Zhongzhou Books and Paintings Publishing House.

語料來源

- Zhongguo Jiben Guji Ku* 中國基本古籍庫 [*Database on Essential Ancient Chinese Literature*], ed. by Junwen Liu (劉俊文). Beijing: Erudition.
- Beijing Daxue Zhongguo Yuyanxue Yanjiu Zhongxin CLL Yuliaoku* 北京大學中國語言學研究中心 CLL 語料庫 [*Corpus of Center for Chinese Linguistic, Peking University*]. Beijing: Peking University.
- Dianziban Wenyuange Siku Quanshu* 文淵閣《四庫全書》電子版 [*E-edition of Wenyuange's Imperial Collection of Four*]. Shanghai: Shanghai People's Publishing House; Hong Kong: Digital Heritage Publishing Ltd, 1999.
- Guoxue Baodian* 國學寶典 [*E-database for Study on Ancient Chinese Civilization*], ed. by Xiaolin Yin (尹小林). Beijing.
- Zhushi Yuliaoku* 朱氏語料庫 [*Zhu's E-database for Linguistic Study*], ed. by Guanming Zhu (朱冠明). Beijing.

A Preliminary Study on the Effect of Speech-mistake Mostly Relating to Chinese Characters in the Historical Evolution of the Chinese Language Vocabulary

Qingzhi Zhu

*Hong Kong Institute of Education
Peking University*

Research on the history of the Chinese language needs to reveal not only the historical changes at different synchronic periods but also the underlying causes and rules. This is by no means an easy task as the historical changes of the Chinese language were rarely discussed in ancient documents currently available. Moreover, such ancient documents were largely written in *wenyan wen*, a classic written language of Ancient Chinese which shares a little features of spoken language. Consequently, it is almost impossible for us to understand the nature and processes of the historical changes of the Chinese language at different synchronic periods.

Relying on “the use of the present to explain the past”, a principle widely employed in the field of sociolinguistics, this article examines the processes and causes of the contemporary variations of the Chinese language, which were further compared to the variations of Ancient Chinese. It argues that speech-mistake, including misinterpretation/misannotation and misuse, mostly relating to Chinese characters appears to be one of the major causes of the evolution of both Ancient and Modern Chinese vocabulary.

Key words: historical Chinese linguistics, sociolinguistics, language change, speech-mistake, Chinese vocabulary, Chinese character

Frequency Effects in Chinese Morphology: Diachronic Evidence from a Synchronic Corpus^{*}

Chao-Jan Chen (陳超然)
National Chi Nan University

This paper explores the type frequency effect in Chinese compounding. We observe that the distribution of character productivity in compound words follows a power law. Based on relevant research on network connectivity, I suggest that the power-law distribution be regarded as the consequence of a positive feedback effect. Such a “rich-get-richer” effect is brought about by the type frequency effect of characters in compounding morphology during the diachronic development of Chinese compound lexicon. The study shows how diachronic evidence can be provided by a large synchronic corpus with statistical analysis. This work also provides empirical evidence for the usage-based hypothesis in the framework of Cognitive Grammar.

Key words: frequency effects, morphological productivity, compounding, power-law, network model

1. Introduction

The ideas of frequency and entrenchment play a fundamental role in the usage-based model of the framework of Cognitive Grammar (see Langacker 1987, 2000, 2008, and Barlow & Kemmer 2000). Different kinds of frequency effects have been explored in various fields of linguistic research, especially in morphology and phonology (e.g. Bybee & Hopper 2001, and Bybee 1985, 2007). Regarded as “a major factor determining the degree of productivity of a construction” (Bybee 1985, 2007), type frequency is supposed to play an important role in compounding. However, as Brinton & Traugott (2005:34) remarked, the original productivity in compounding might be obscured over time due to some phonological and other changes. Such obscurity of participant morphemes in words thus makes it difficult, if not impossible, to trace the morphological productivity of certain morphemes in the course of history.

^{*} This research was partially supported by Grants NSC (NSC 97-2410-H-260-036 and NSC 98-2410-H-260-044).

As a language extremely rich in compound words, Chinese is supposed to be an ideal target for the study of the type frequency effect in compounding. What is even more desirable for research on frequency effects is that the component morphemes in Chinese compound words generally remain transparent and thus perfectly discernable as units of *zi* (or Chinese characters¹) in the long course of history. However, quantitative studies about the overall morphological productivity of characters in compounding have rarely been the focus of research in Chinese morphology. For linguists who work under the generative paradigm, quantitative study involving frequency is regarded as a performance rather than competence issue and is thus neglected. On the other side, for computational linguists, who surely do quantitative research, the focus of morphological productivity is often placed on certain highly productive affixes (or fixed-placed roots). Therefore, researchers of Chinese morphology seldom pay attention to the productivity of “ordinary” morphemes in word formation, let alone study their productivity distribution pattern.

This paper thus aims to explore the type frequency effect in Chinese compounding. Instead of doing case studies on the morphological productivity for individual characters in forming compounds, we will focus directly on the overall distribution pattern of character productivity and explore its significance in diachronic linguistics under the theoretic framework of Cognitive Grammar, specifically in the usage-based model. Related research in dynamic network structures will be used to help build a usage-based model of compounding. The data we will use in this study is collected from a large corpus (containing five million word tokens) of contemporary Chinese, the Academia Sinica Balanced Corpus version 3.0 (abbreviated as ASBC). As two-character compounds are the only type of compound concerned in this paper, from now on, for the sake of simplicity, they will simply be referred to as “compounds”.

2. Theoretic framework

2.1 Usage-based model and compounding template

According to Langacker’s (1987, 2000, 2008) usage-based model, linguistic symbolic units can be schematized from usage events. In this theoretic framework, language structure is supposed not to be autonomous from the usage. Rather, the structure emerges from language use. Patterns can be abstracted from similar instances and entrenched through frequency to form schemas (or constructions in Construction Grammar’s term), which can later serve to form new instances.

¹ Throughout this paper, the term “character” will be used to refer to the monosyllabic semantic unit, almost always a morpheme, rather than a mere grapheme, in the word writing system of Chinese.

Adopting this dynamic viewpoint of constructions, we propose, in the case of Chinese compounding, a hypothesis that the production of compounds can be template-based: existing compounds can serve as templates (or “schemas”, to use Langacker’s term) to motivate the creation of new compounds, when related semantic constraints are satisfied. For example, an existing compound X-Y (with X and Y as component characters) can serve respectively as two templates [()-Y] and [X-()] to motivate the creation of new compounds such as W-Y or X-Z.

By the same token, the production of the compound X-Y can be motivated by template [()-Y], by template [X-()], or by both of them. Such a way of template-motivated compounding can also be regarded as *analogical creation* (Chen 2005) or *modification of prefabricated structures* (Barlow 2000).²

2.2 Type frequency effect and potential positive feedback

According to Bybee (2007), type frequency is “a major factor determining the degree of productivity of a construction”, which means “constructions that apply to a high number of distinct items also tend to be highly applicable to new items” (cf. Bybee 1985). While type frequency means the frequency of word types that conform to a schema, the productivity of a schema is argued to be a function of type frequency with the following hypothesis: “The productivity of a schema is a function of the type frequency of the instances of the schema.” (Bybee 1985:132-134, see also Bybee & Hopper 2001, and Langacker 2008).

Following Bybee’s hypothesis of the type frequency effect, in our usage-based model of Chinese compounding, we would expect that existing compounds serve as compounding templates, of which the type frequency will lead to the entrenchment of the compounding template (the schema in our case) as a frequency effect. To be more specific, we propose that the type frequency of the compounding template can determine, in a certain way, the probability of the derived compounds to be produced. A theoretical consequence of this proposal should then be positive feedback: the characters used more frequently in existing compounds will be more likely to be used in new compounds and thus become more productive afterward. In the following sections, we will use data in the Chinese corpus ASBC to test this hypothesis.

² Barlow (2000:318) argues that “the creativity or the expressive dimension to language comes in large part from the modification of prefabricated structures, rather than the novel combination of lexical categories.”

3. Character productivity: data and observation

3.1 Data collection

As our study of the type frequency effect is based on the framework of Cognitive Grammar (also abbreviated as CG) and focuses on two-character compounds in Chinese, we regard the compound word as a symbolic unit, combining two morpheme-functioning characters, which is well entrenched through language use. We are then allowed to avoid the problem, brought about by the traditional definition of compounds which involves confusing and controversial criteria about the distinction between bound and free morphemes, and between phrases and compound words in Chinese (see, for example, Dong 2002:25-27). As Li & Thompson (1981:45) remarked, “*no matter what criteria one picks, there is no clear demarcation between compounds and non compounds*”. In other words, there is no clear-cut between compounding and derivation in Chinese morphology.

The CG-based definition of compounds also allows us to use data in the ASBC corpus for our work, as the criteria of *wordhood* in ASBC generally are compatible with our criterion of compounds above (cf. the technical report no.95-02/98-04 of CKIP, the creating and maintaining group of the ASBC corpus). As multi-character morphemes are extremely rare in Chinese, virtually all the two-character words in ASBC can thus be considered as compounds. Since our study of character productivity in compounding is focused on two-character compound words only, we take as analysis data of compounds all the two-character words in the ASBC corpus. This simple criterion for *compoundhood* has its practical advantage. It allows us to do our corpus-based statistic analysis in a large-scale and automatic way and to avoid manual distinction of compounds among huge data, which is not quite feasible. Surely the two-character words in the corpus include few single-morpheme words. However, such relatively rare exceptions can be considered as noise technically tolerable in our statistical analysis because our task is mainly to observe an overall statistical pattern and therefore would not be sensitive to this trivial data noise.

3.2 An aristocratic world of productivity

Containing about five million word tokens, the ASBC corpus version 3.0 includes 66,722 (two-character) compound types, which are formed by 5,486 character types. Simple statistics show that the compounding productivity is extremely variable among the characters: while a great number of characters (1,070 or about 19.5%) only appear in one compound (see Figure 1, which lists only the partial distribution for the number of characters that appear in under 20 compound types), the most productive character

can appear in up to 607 compounds (occupying almost 1% of compound types) (see Table 1).

Since the *average* number of compound types formed by the characters is 24.42, we might expect that “normally” a character would appear in around 24 compound words. However, it is not the case. In fact, as we can see in Figure 1, the distribution of character productivity in compounding is not “normal” at all: it is not a bell-shaped distribution as that which is called *normal* distribution in statistics. Contrary to our expectations, the world of morphological productivity we face here is so “*aristocratic*” instead of “*equalitarian*” (terms borrowed from Buchanan 2002:119): while most of the characters are “impoverished” in compounding productivity, the top ranking ones are extremely “rich”.

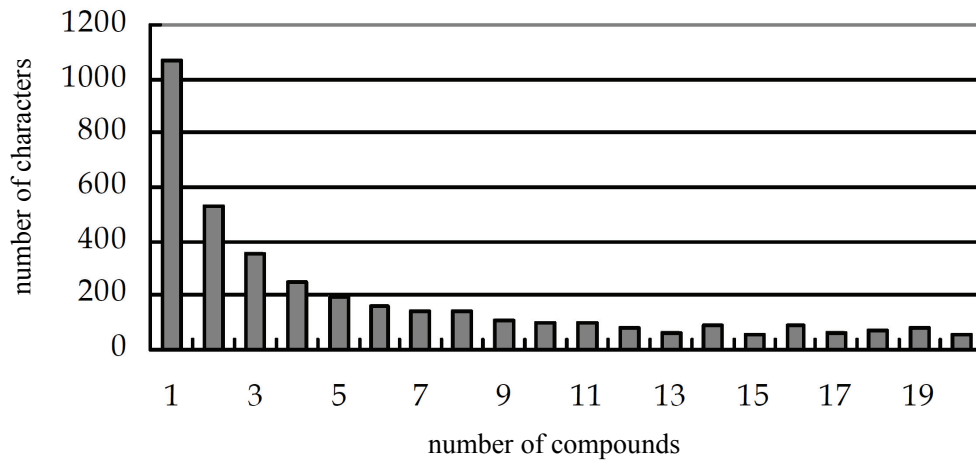


Figure 1: Distribution of character productivity (i)

Table 1: Distribution of character productivity (ii): Top 10 and Bottom 10
(#comp: number of compounds, #char: number of characters)

Top 10				Bottom 10			
#comp	#char	#comp	#char	#comp	#char	#comp	#char
607	1	415	1	10	99	5	191
553	1	392	1	9	108	4	244
513	1	383	1	8	142	3	351
482	1	340	1	7	142	2	527
418	1	322	1	6	162	1	1,070

4. Power-law distribution and its significance

4.1 A power-law distribution

If we present the productivity distribution of the characters in ordinary coordinates, we can see its decline tendency through the curve in the scatter plot (see Figure 2).

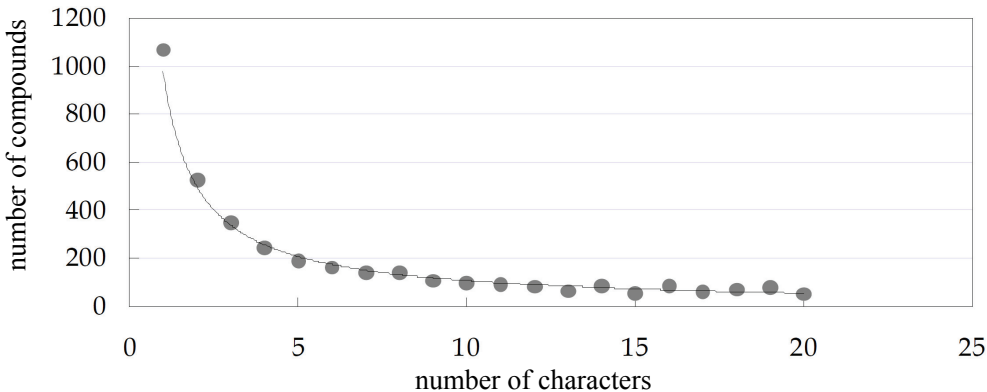


Figure 2: Distribution of character productivity (iii)

If we present this distribution in another way, namely in a plot in log-log coordinates, we will see that the productivity distribution of the characters turns out to be a straight line, instead of a curve, which is the decline tendency of the curve in Figure 2 (see Figure 3).

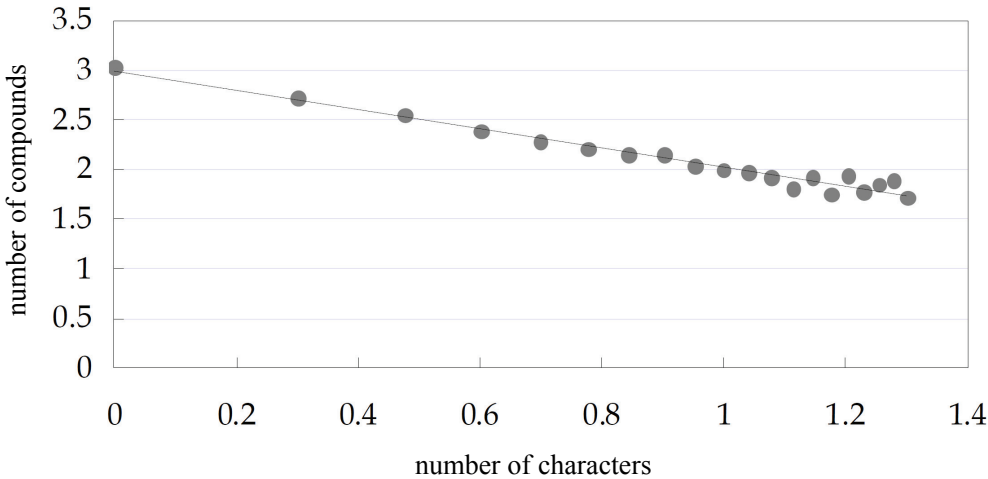


Figure 3: Distribution of character productivity (iv) (log-log)

By the method of least-square regression, we obtain a straight line as its best fitting curve:

$$(1) \log N(k) = 3.028 - 1.003 \log k \quad (r = -0.9846, r^2 = 0.9694),$$

where $N(k)$ is the number of characters participating in k compounds

The formula of productivity distribution can also be represented as follows:

$$(2) N(k) = 1066 k^{-1.003}$$

This formula clearly shows that the character productivity in compounding follows a power-law distribution.³

4.2 Significance of power-law distribution

In fact, the power-law distribution has been found in many kinds of real or virtual scale-free networks, including the World Wide Web, academic coauthor networks, the food chain web, airport networks, and the synonym web (see also Buchanan 2002, Barabási 2003, Barabási et al. 2003, Steyvers & Tenenbaum 2005, and Newman et al. 2006). In such networks, the number of nodes N is a function of the number of links according to the following formula:

$$(3) N(k) = C k^{-a}$$

According to Barabási & Albert (1999), a power-law distribution can be regarded as the signature of two properties in a network: (1) incremental growth of nodes and links, and (2) preferential attachment. In other words, the network following a power-law distribution is evolving with growing nodes and links; new nodes are linked to existing nodes by the principle of preferential attachment. That means the more links a node already has, the more likely it will get new links, which brings about a “rich-get-richer” effect. For example, a website linked by more websites will be more likely to be linked by new websites, a more frequently cited paper will be more likely to be cited again by other new papers, and so on.

³ According to the definition of power-law distribution, when $Y(x) = C X^{-a}$, we have a power-law distribution. We can see a straight line with a slope $-a$ in its corresponding log-log plot. As $\log Y = \log C - a \log X$, such a straight line in a log-log plot is thus regarded as the most salient feature of a power-law distribution.

5. A network model of compounding and its linguistic significance

In fact, as proposed in our previous approach (Chen 2005), we can regard compounding as a link: for a compound X-Y, there exists a link between X and Y. Therefore such compounding connectivity forms a network, in which characters are nodes and compounding relations are links, as shown in Figure 4.

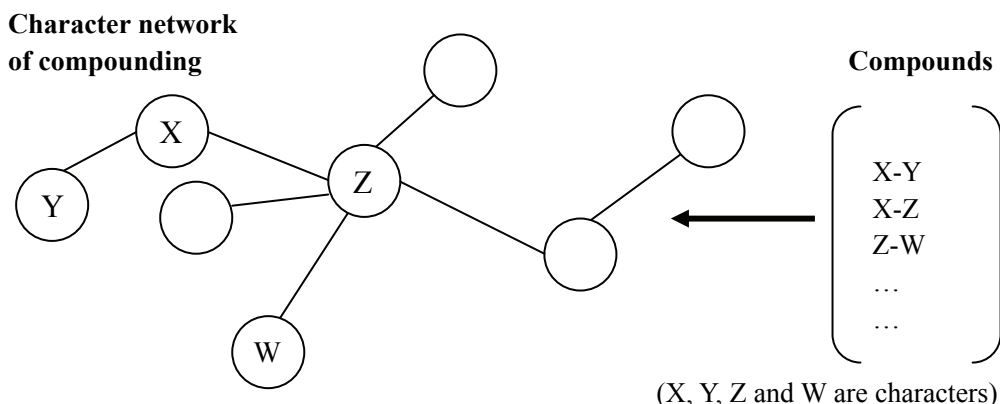


Figure 4: Compounding as links between characters

Viewing compounding as a network of characters, we can also find in this network the two critical factors accounting for the power-law distribution: the growth of the network and preferential attachment. Evidently, according to the diachronic development of a compound lexicon in Chinese, the modeling network should be growing. What is more important, the preferential attachment principle holds as a theoretical consequence under the Cognitive Grammar's hypothesis: the more used characters become more likely to be used again, thus being more productive in compounding (a "rich get richer" effect), which is exactly what the type frequency effect tells. In other words, when a character is more 'productive' in existing words, it is also more 'productive' for potential words.

However, in our compounding network model for two-character compounds, only the links grow, while the nodes (the characters) remain the same over time, which is slightly different from the dynamic mathematical models for network growth proposed in previous research such as that of Barabási & Albert (1999) (see also Newman et al. 2006 and Durrett 2007 for related mathematical models). Our previous research (Chen 2012) therefore sets up a stochastic model of compounding based on this feature. In this dynamic model, the probability of a compound X-Y being produced is proportional to the type frequency of template [X-()] and to that of template [()-Y]. The compounds thus produced can go on to contribute to the type frequency of related templates. The results

of the simulated evolution of a compound lexicon show that a power-law distribution can also be observed in such a model, as expected of other dynamic models with preferential attachment.⁴

Therefore, our evolving network model for Chinese compounding shows that the power-law distribution of character productivity can be accounted for by the type frequency effect proposed by the usage-based framework of Cognitive Grammar. Such a significant distribution pattern would not be otherwise satisfactorily explained linguistically. In other words, the power-law distribution of character productivity strongly suggests that the morphological mechanism of Chinese compounding is under the type frequency effect, which provides evidence for the usage-based framework of Cognitive Grammar.

6. Concluding remarks

In this paper, we have tried to show the effect of type frequency on Chinese compounding. To account for the fact that the distribution of character productivity in forming compounds follows a power law, we propose a network model of compounding, which is composed of characters and links between every pair of component characters in compounds. We have argued that such a network model can give rise to a power-law degree distribution, which is, according to related research in the structure and dynamics of various networks in virtual or real worlds, considered as the signature of a “rich-get-richer” growing network with preferential attachment. In the dynamic model we proposed, preferential attachment is actually the effect produced by a stochastic mechanism applying the hypothesis of type frequency effect.

Though our model is rather simplistic, it still allows us to capture the essence of the positive feedback effect in the diachronic development of the Chinese lexicon and to account for the power-law distribution we observed in character productivity. A broader theoretical significance of our work is that the effect of type frequency on compound production provides evidence for the usage-based model in the theoretical framework of Cognitive Grammar. The results of our study also suggest that compounding can be an analogical creation that is motivated by templates abstracted from existing compounds. Moreover, the study shows how diachronic evidence can be provided by a large synchronic corpus with statistical analysis. The application of such a global statistic method in diachronic linguistic research allows us to bypass the problem brought about by individual words whose historical development is unclear and difficult to trace.

⁴ As the main aim of this paper is to explore the linguistic explanation rather than to provide an elaborated mathematic model for the power-law distribution of character productivity, we are contented with such a simple and plausible model capturing the essence of the rich-get-richer effect without getting into mathematical details.

References

- Academia Sinica. Academia Sinica Balanced Corpus (<http://www.sinica.edu.tw/SinicaCorpus/>). Taipei: Academia Sinica.
- Barabási, Albert-László. 2003. *Linked: How everything is connected to everything else and what it means*. New York: Plume.
- Barabási, Albert-László, and Réka Albert. 1999. Emergence of scaling in random networks. *Science* 286:509-512.
- Barabási, Albert-László, Zoltán Dezső, Erzsébet Ravasz, Soon-Hyung Yook, and Zoltán Oltvai. 2003. Scale-free and hierarchical structures in complex networks. *Modeling of Complex Systems: Seventh Granada Lectures. AIP Conference Proceedings*, Vol. 661, 1-16. College Park: AIP.
- Barlow, Michael. 2000. Usage, blends and grammar. *Usage Based Models of Language*, ed. by Michael Barlow & Suzanne Kemmer, 315-345. Stanford: CSLI.
- Barlow, Michael, and Suzanne Kemmer. (eds.) 2000. *Usage Based Models of Language*. Stanford: CSLI.
- Brinton, Laurel J., and Elizabeth Closs Traugott. 2005. *Lexicalization and Language Change*. Cambridge & New York: Cambridge University Press.
- Buchanan, Mark. 2002. *Nexus: Small Worlds and the Groundbreaking Science of Networks*. New York: W.W. Norton & Company.
- Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam & Philadelphia: John Benjamins.
- Bybee, Joan L. 2007. *Frequency of Use and the Organization of Language*. Oxford & New York: Oxford University Press.
- Bybee, Joan L., and Paul Hopper. (eds.) 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam & Philadelphia: John Benjamins.
- Chen, Chao-Jan. 2005. *Modélisation de la sémantique des verbes composés chinois de type V-V*. Paris: University of Paris VII dissertation.
- Chen, Chao-Jan. 2007. Character-sense association: a study on automatic sense determination for Chinese V-V compounds. *Cahiers de Linguistique – Asie Orientale* 36.1: 3-36.
- Chen, Chao-Jan. 2012. Power-law distribution in morphological productivity: a statistical analysis of Chinese compounds. In *Search of Grammar: Experimental and Corpus-based Studies*, ed. by James Myers, 97-118. Taipei: Institute of Linguistics, Academia Sinica.
- Dong, Xiufang. 2002. *Cihuihua: Hanyu Shuangyinci de Yansheng he Fazhan* [Lexicalization: Origin and Evolution of Chinese Disyllabic Words]. Chengdu: Sichuan Nationalities Publishing House.

- Durrett, Rick. 2007. *Random Graph Dynamics*. Cambridge & New York: Cambridge University Press.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, Vol. 1: *Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2000. A dynamic usage-based model. *Usage Based Models of Language*, ed. by Michael Barlow & Suzanne Kemmer, 24-63. Stanford: CSLI.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford & New York: Oxford University Press.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Newman, Mark, Albert-László Barabási, and Duncan J. Watts. (eds.) 2006. *The Structure and Dynamics of Networks*. Princeton: Princeton University Press.
- Peyraube, Alain. 2004. Grammaire diachronique et cognition: l'exemple du chinois. *La linguistique cognitive*, ed. by Catherine Fuchs, 135-154. Paris: Ophrys.
- Steyvers, Mark, and Joshua B. Tenenbaum. 2005. The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science* 29.1:41-78.

Improper Subjects: Towards a Philosophical Grammar of Existence Predicates in Classical Chinese

Christoph Harbsmeier (何莫邪)
University of Oslo

Both Aristotelian logic and modern predicate logic require a subject of which a predicate can be asserted. Thus, impersonal sentences without logical or even grammatical subjects have invited much philosophical and linguistic discussion since antiquity. This paper is a first step towards analyzing impersonal constructions of Classical Chinese, comparing them with those known from contemporary and ancient Indo-European languages. For this purpose it distinguishes four types of subjectless sentences: (1) absence of logical subject (in Chinese without the presence of a grammatical dummy subject), (2) lexicalized omission of a lexically determinate subject, (3) lexicalized omission of contextually determinate subject and (4) absence of a specific subject, equivalent to a zero non-referential generic pronoun ‘one’, ‘you’. The paper focuses especially on different kinds of feature placing predicates (type (1)), looking at transitive and intransitive meteorological verbs, verbs for states of the world or society, as well as existence predicates, examining in detail the linguistic, logical and cognitive implications of classical Chinese using sentences without grammatical subjects as opposed to the use of a dummy grammatical subject such as German ‘es’ or English ‘it’.

Key words: impersonal constructions, subjectless predication, language and logic, Ancient Greek, Classical Chinese grammar

1. Introduction

Frédéric Lambert from Bordeaux has presented a truly fascinating magisterial account of the interpretation of impersonal verbs in classical Greek in which he has drawn my attention to the ways in which even a comedian like Aristophanes played around with the intractable problems of subjectless meteorological verbs.¹ In particular, Frédéric Lambert 2009 refers to a passage from *The Clouds* which I translate in my own way from the Greek below:

¹ The literature on impersonal constructions in the classical languages and elsewhere is rich. Particularly useful are Abraham & Leiss (2006), Chocheyras (1985), Comrie (1977), Cuzzolin & Napoli (2009), Desbordes (1991), Ernout (1909), Jacobi (1985), Lambert (2010), Maillard (1991), Ogura (1986), Pieroni (2000), Pinkster (1992), and Seriot & Berrendonner (2000).

Strepsiades:

ὁ Ζεὺς δ' ἡμῖν, φέρε πρὸς τῆς
γῆς, οὐλύμπιος οὐ θεός ἐστιν;

Come on for us now, and the earth be our witness:
is not Zeus, the Olympian, a god, then?

Socrates:

ποῖος Ζεὺς; οὐ μὴ ληρήσεις;
οὐδ' ἔστι Ζεὺς.

What kind of Zeus are we talking about? Don't
you give me this kind o' crap! In any case: he
doesn't exist!

Strepsiades:

τί λέγεις σύ;

What are you talking about?

ἀλλὰ τίς ὕει; τουτὶ γὰρ ἔμοιγ'
ἀπόφηναι πρῶτον ἀπάντων.

Who the hell do you imagine is raining then, eh?
Why don't you let me know before you go on to
anything else!

Socrates:

(370) αὗται δὴπου: μεγάλοις δέ
σ' ἐγὼ σημείους αὐτὸ διδάξω.

All right, then! As you say! With overwhelming
proofs I shall teach you this lesson, for sure.

φέρε ποῦ γὰρ πάποτ' ἄνευ
Νεφελῶν ὕοντ' ἤδη τεθέασαι;

Come on then: where have you ever seen him rain
in this world without clouds, eh?

καίτοι χρῆν αἰθρίας ὕειν αὐτόν,
ταύτας δ' ἀποδημεῖν.

And yet, you see, he ought to be raining along
while these clouds are off and away.

Strepsiades:

νῆ τὸν Ἀπόλλω τοῦτό γέ τοι δὴ
τῷ νῦν λόγῳ εὖ προσέφυσας:

By Apollo, you've made your point by your
present speech.

καίτοι πρότερον τὸν Δῖ ἀληθῶς
ᾧμην διὰ κοσκίνου οὐρεῖν.

And yet a moment ago I was convinced that Zeus
himself was pissing through some sieve when it
rained.

ἀλλ' ὅστις ὁ βροντῶν ἐστι
φράσον: τοῦτό με ποιεῖ
τετρεμαίνειν.

Let's have it, then: tell us who does the
thundering. That scares the shit out of me.

Socrates:

(375) αὐται βροντῶσι
κυλινδόμεναι.

These thunders just thunder as they roll along,
that's all!

Strepsiades:

τῷ τρόπῳ ὃ πάντα σὺ τολμῶν;

What way do you mean, brazen as you are?

Socrates:

ὅταν ἐμπλησθῶσ' ὕδατος πολλοῦ
κἀναγκασθῶσι φέρεσθαι,

Consisting of lots of water they cannot help being
carried along, those Clouds.

κατακρημνόμεναι πλήρεις
ὄμβρου δι' ἀνάγκην, εἶτα βαρεῖται

And when replete with rain

εἰς ἀλλήλας ἐμπίπτουσιν
ρήγνυνται καὶ παταγοῦσιν.

these Clouds bang into each other thus making
that terrific noise.

Strepsiades:

ὁ δ' ἀναγκάζων ἐστὶ τίς αὐτάς,
οὐχ ὁ Ζεὺς, ὥστε φέρεσθαι;

But ah! Who is it, if not Zeus, that forces them to
be carried along, eh?

Socrates:

(380) ἥκιστ' ἀλλ' αἰθέριος Δῖνος.

Not him in the least! The aetherial whirl!

Strepsiades:

Δῖνος; τουτί μ' ἐλελήθειν,
ὁ Ζεὺς οὐκ ὢν, ἀλλ' ἀντ' αὐτοῦ
Δῖνος νυνὶ βασιλεύων.

Oh damn! It never occurred to me
that Zeus being non-existing, that Whirl now
reigns supreme!

ἀτὰρ οὐδέν πω περὶ τοῦ πατάγου
καὶ τῆς βροντῆς μ' ἐδίδασκας.

That's all very well, but you haven't breathed a
word about that noise and thunder!

Socrates:

οὐκ ἤκουσάς μου τὰς Νεφέλας
ὕδατος μεστὰς ὅτι φημὶ

Didn't you hear me? The clouds, being full of
water,

ἐμπιπτούσας εἰς ἀλλήλας
παταγεῖν διὰ τὴν πυκνότητα;

knock against each other, and they make this noise
because they are so solid with water.

Strepsiades:

(385) φέρε τουτὶ τῷ χρηῖ
πιστεύειν;

Oh, come on! Who would believe any of this?

Socrates:

ἀπὸ σαυτοῦ ἴγώ σε διδάξω.

Very well: let me teach you through self-observation, then!

ἤδη ζωμοῦ Παναθηναίοις
ἐμπλησθεὶς εἶτ' ἐταράχθης

At the big Panathenaic Festival, stuffed with that broth, haven't you felt disturbed?

τὴν γαστέρα, καὶ κλόνος
ἐξαίφνης αὐτὴν
διεκορκορύγησεν;

In your tummy I mean, that inward turmoil rumbling its way out as a fart?

Strepsiades:

νὴ τὸν Ἀπόλλω καὶ δεινὰ ποιεῖ
γ' εὐθύς μοι, καὶ τετάρακται

By God, you have a point! That broth has a terrific effect!

ᾧσπερ βροντὴ τὸ ζωμίδιον
παταγεῖ καὶ δεινὰ κέκραγεν:

Like the thunder it races through your belly doesn't it, and makes a terrific din!

(390) ἀτρέμας πρῶτον παπᾶξ
παπᾶξ, κάπειτ' ἐπάγει
παπαπαπᾶξ,

It sets out mildly, papax papax, then goes papapappax,

ᾧσταν χέζω, κομιδῇ βροντᾶ
παπαπαπᾶξ ᾧσπερ ἐκεῖναι.

and finally, as I shit it out, it thunders forth indeed. Gently, though, with care. Just like those Clouds.

The spirit of this irreverent dialogue is entirely alien to ancient Chinese literary conventions. And yet, it seems to me, that the content is directly relevant to some fundamental issues concerning the nature of Chinese philosophical thought and Chinese historical syntax.²

As the comedian Aristophanes was delicately and irreverently aware, subjectless sentences are something of a well-known logical teaser, just the kind the ancient Greek sophists would have liked to discuss. According to Aristotelian logic, and indeed modern predicate logic in general, one can only assert propositions which attribute a predicate to a subject, propositions which assert that a predicate holds of a subject. Now the

² See Cikoski (1978), Freundlich (1988), and Li (1976).

scandal is that even a proposition like *It is raining*, with its grammatical subject ‘it’ is capable of being asserted, while it is in no way obvious that it has a logical subject.

Ancient Greek grammarians focused on impersonal constructions like the ancient Greek *dei* ‘one must; one should’ which do not invite the specification of a logical subject. Latin grammarians have indeed discussed impersonal subjectless constructions in elaborate detail. The great logician Peter Abaelard (1079-1142) has devoted much careful, analytic attention to the logical construal of impersonal sentences in his important commentary on Aristotle’s *Peri hermeneias* (*De Interpretatione*).

Klaus Jacobi has published on the web a masterful philosophical interpretation of Abaelard’s logical reflections on the subject in his commentary on *De Interpretatione*.

2. Subjectless predication or feature-placing predicates

In this brief contribution I wish to combine philological and logical methods to consider the case of subjectless predication or feature-placing predicates in the terminology of my philosophical master Peter F. Strawson in classical Chinese.³

Subjectlessness in classical Chinese poses a great number of entertaining problems of grammatical analysis that go beyond what is current in Latin and Greek.

In the context of syntactic analysis in my *Thesaurus Linguae Sericae* I distinguish between four importantly different types of absence of a subject:

1. **V0**
Absence of a **logical** subject.
Example: ‘It is raining.’
2. **V[0]**
Lexicalized omission of a **lexically determinate** subject.
Example: ‘[I] Thank you!’
3. **V(0)**
Lexicalized omission of a **contextually determinate** subject.
Example: ‘(It is) OK!’
4. **VØ**
Absence of a specific subject which can be read as the presence, in Chinese, of the zero non-referential generic pronoun **Ø**, translatable as ‘one; you’.

³ The philosophical literature on the subject includes Jacobi (1985), Moore (1936), Myhill (1997), Pears (1967), Pears & Thomson (1963), Strawson (1974), and Williams (1981).

In a comparable vein, I also distinguish between three types of object omissions:

1. **Vi**

Absence of a logical object.

Example: 'He *sleeps* [NOT an understood cognate object: a sleep.]'

2. **Vt[oN]**

Idiomatic omission of lexically determinate object.

Example: 'He *drinks* [(excessive amounts of) alcohol].'

3. **Vt(oN)**

Idiomatic omission of a contextually determinate object.

Example: 'I *agree* (with whatever is contextually determinate).'

In the present paper, I shall also concentrate on a few simple cases of omitted subjects.

We properly understand the 'it' in *it is raining* when we remain completely uncommitted as to what, if anything, this word refers to in the given context. Cases of this sort are not simple. Take the French interjection *chouette*: there may be no overt subject here, but it stands to reason that what occasions this interjection is also the logical subject of the predicate *chouette*.

In *thank you*, on the other hand, one only understands the phrase properly if one construes the performative verb as having a logical subject, that subject being the speaker using the phrase *thank you*. (It remains an open question in what sense *hereby* may be said to be "understood" in a phrase like *thank you*).

The identity of the omitted subject of *thank you* may be said to be determinate from the context. But there is a very neat lexicalized rule according to which the subject is retrieved: it is the internalized lexicon and not the context which tells us that the subject of *thank* is the current user of this word.

Things seem significantly though not radically different in the case of *OK*. All the lexicon tells us with this phrase is that we must look in the context for a suitable, logical subject, and that if we do not find one there we will be deemed to not have understood the utterance *OK*. Although, of course, while having failed to grasp its pragmatic force in the given context, we may well be deemed to have understood the lexical force of the expression *OK*.

In classical Chinese there is good reason to try to distinguish between these three types of subject absences for all subjectless words as used in a given context: We shall see that there is a surprisingly wide range of feature-placing predicates of the type **V0**, there is a fairly limited range of verbal expressions with idiomatically omitted lexically predictable verbs **V[0]**, and there is, of course, a disconcerting tendency of verbal

expressions **V(0)** to occur where modern Chinese as well as other modern languages would encourage the explicit specification of the subject involved.

It has only occurred to me very recently that there is in Chinese a fourth type of verbal expression with an idiomatically omitted subject, as in *rù xiāng suí sú* 入鄉隨俗 where it would seem unidiomatic if not ungrammatical to write any explicit subject, and where it sounds curiously misleading to speak of any particular lexical item that is omitted, but where it is entirely wrong to construe 入鄉 and 隨俗 in any of the senses mentioned so far.

There is a logical subject for both 入鄉 and 隨俗. Indeed, both verbal expressions have to be construed as having the same subject for the overall expression to be understood. And the Chinese is uncommitted as to whether the phrase is descriptive or imperative.

隨俗 is probably imperative, as is suggested in the English *When in Rome, do as the Romans do* (with its interesting *xiē hòu yǔ* 歇後語 variant *When in Rome...*). To specify the logical subject of this phrase as being the second person pronoun *you* commits an uncommitted phrase to a specificity of reference which is alien to it.

One might suspect that what is omitted here is what, in modern Mandarin, might come out as the idiomatically generic *nǐ* 你, in French as *on*, and in English as the generic pronoun *you* (not the second person pronoun). One might simply say that the French *on* comes out as zero \emptyset in Chinese. Yet such an interpretation seems to commit the Chinese where in fact it is uncommitted.

Contrast *it is your turn*, and also the Chinese *zhuǎn dào nǐ le* 轉到你了 / *lún dào nǐ le* 輪到你了 TURN TO YOU. There is no sense in which one needs to retrieve any grammatical subject in order to understand the phrase. And note that we say *hái méi yǒu zhuǎn dào nǐ le* 還沒有轉到你了.

LY 1.1 *xué ér shí xí zhī, bù yì yuè hū* 學而時習之, 不亦說乎? brings out the immense problems of specifying subjects for Chinese verbs very neatly: how must we construe the subject structure of the verbal expressions 學 and 習之?

The following possibilities present themselves for the analysis of this phrase:

1. **V(0)**

It is ‘we’ who study and then exercise something, is our exercise such a delight?

2. **V \emptyset**

The statement is general about “one”: ‘If one studies...’.

3. **V0**

The statement is indeed without subject: ‘To study..., is not that a pleasure?’

The reading **V[0]** does seem excluded: 學 and 習 cannot plausibly be taken to be

lexically subcategorised for idiomatic use with a certain subject.

Traditional Chinese philology has thrived in blessed indifference to alternatives like those I have lined up here. The fact is that Confucius is clearly uncommitted as to the distinctions we make. The phrase we are concerned with is underdetermined with respect to the distinctions we shall always try in vain to impose on it.

The theoretically important question that arises, then, is this: is this semantically undercommitted and underdetermined grammatical style with regard to the distinctions above pervasive to the classical Chinese language, or is it that classical Chinese does lexicalise these distinctions where they are deemed relevant?

This is an empirical question which depends entirely on the discovery of lexical entries which either are subcategorised themselves for specified omitted subjects of the type I have outlined, or which subcategorise other verbs they are in construction with.

I shall begin by considering verbs without logical subjects. Three types of these must be carefully held apart:

1. Feature-placing with dummy grammatical subject present, as in ‘*it* is raining’ (contrast the following Chinese phrases which have neither a logical nor a dummy subject: *yǔ* 雨 ‘it is raining’, *xià yǔ* 下雨 ‘it is raining’, and *yǔ xuě* 雨雪 ‘it is raining snow > it is snowing’.).
2. Feature-placing without dummy grammatical subject, as in ‘damn!’ *zāo gāo le* 糟糕了!
3. Subjectless grammatical construal of an action as a feature occurring, in a feature-placing subjectless sentence, as in *Hier wird geschlafen*. ‘this is where the sleeping goes on’.⁴

Compare the impersonal *il s’agit d’argent* ‘money is the question’ and the German *es geht um’s liebe Geld*. The radical defocusing of the subjects *il* and *es* comes out

⁴ *Hier wird von ihnen/einem/mir/Euch geschlafen* ‘here is slept by you’ are all unacceptable. And yet, there is no doubt that the sleeping is construed as the sleeping by someone. The sleepers are defocused, that is all. Cases of this sort need to be discussed together with such constructions as the German *es klopft (an der Tür)* ‘There is knocking (on the door).’ and *mir graust es* which translates only uncomfortably into ‘I feel unsettled’ because the German construes the matter logically subjectlessly, with a dummy grammatical subject, and with the affected “logical” subject in the dative. Consider also *Es grünt so grün, wenn Berlins Blüten blühen* ‘It greens so greenly when the flowers of Berlin are in flower.’ There is no doubt that in this case there is something that is green, but what is happening here is presented first in a subjectless poetic mode, and only thereafter something else is referred to in subject-predicate terms, with an undefocused subject.

when one tries to ask *qu'est ce qui s'agit d'argent?* or *was geht da um's Geld?*, *Was ist es, das da um's Geld geht*. In fact, it turns out that in this radical defocusing the subject is irretrievable, and cannot be asked for. The term “defocusing” becomes logically misleading: the subject is no longer only defocused, it is absent.

There are degrees of defocusing, and there are also degrees of the kind of conviction with which one states that the subject is absent.

What happens in the notorious German impersonal passives is a systematic defocusing not only of agents but of all subjects. And when this defocusing becomes not only systematic but radical then what we seem to have is the subjectless construal of situations that do involve subjects. And it is important to realize that in the case of the German *es klopft* the knocking may be an event without an agent, and even without even a subject that does the knocking: the knocking may be conceived like thunder, “impersonally”.

In classical Chinese it does not in fact rain “cats and dogs”, but the transitive use of the 雨 is worth recording.

Let me begin with a late example:

- (1) *Lùnhéng* 論衡, ed. TLS LH 19.6.10
雨粟陳、蔡 ‘If it had rained grain in Chén and Cài...’

One notes that it was not a case of ‘Chén and Cài being grain-rainy’.
In 624 it rained bees in Sòng:

- (2) *Chūnqiū* 春秋, ed. TLS CQ 6.3.5.1
雨螽於宋 ‘It rained bees in the state of Sòng.’
- (3) *Zhànguó cè* 戰國策, ed. TLS ZGC 11.1.3
雨血沾衣 ‘It was raining blood, which soaked the clothes.’

What soaked the clothes was not an understood subject of 雨.

- (4) *Chūnqiū* 春秋, ed. TLS CQ 2.8.5
雨雪 ‘It was raining snow.’
- (5) *Chūnqiū* 春秋, ed. TLS CQ 5.10.7
大雨雪 ‘There was great snowfall.’
- (6) *Lùnhéng* 論衡, ed. TLS LH 18.6.4; 18.6.7
雨穀 ‘It was raining grain.’

Contrast the possibility of the subject Heaven:

- (7) *Lùnhéng* 論衡, ed. TLS LH 19.12.12
天雨穀 ‘Heaven raining grain.’
- (8) *Lǚshìchūnqiū* 呂氏春秋, ed. TLS LS 2.1.1.2
雨水 ‘It was raining water.’ (!)

In what follows I shall consider a series of apparently subjectless intransitive predicative expressions in Chinese. Many of these seem to have no semantic parallel in Greek. Related to the meteorological subjectless verbs is the case of *bīng* 冰:

- (9) *Mòzǐ* 墨子, ed. TLS MO 19.4.2
夏冰 ‘When there is ice in summer...’

The verbal interpretation of *bīng* 冰 here is unaffected by the fact that the negation involves the nominal construction *wú bīng* 無冰, as in:

- (10) *Chūnqiū* 春秋, ed. TLS CQ 8.1.3.1
無冰。 ‘There was no ice.’
- (11) *Chūnqiū* 春秋, ed. TLS CQ 9.28.1.1
二十有八年春， ‘In the spring of the 28th year
無冰。 there was no ice.’
NOT: ‘The spring of the 28th year was ice-free.’

Shuǐ 水 ‘flood’ is current in the same kind of impersonal meaning:

- (12) *Chūnqiū* 春秋, ed. TLS CQ 2.1.4
秋，大水。 ‘In autumn there were great floods.’ is not to be
interpreted as ‘the autumn was highly watery.’
- (13) Oracle bone inscriptions, ed. TLS HEJI 33354
其水 ‘There will be a flood.’
不水 ‘There won’t be a flood.’

The *Annals* of the state of Lǚ 魯 provide a host of examples, sometimes so short that they invite the objection that they may be written in a telegraphic Chinese. Nonetheless, non-telegraphic cases do exist:

- (14) *Chūnqiū* 春秋, ed. TLS CQ 3.20.2
 夏， ‘In summer,
 齊大災。 **there was a great disaster from fire** in Qi.’
- (15) *Zuǒzhuàn* 左傳, ed. TLS ZUO 18.10.3.1
 夏五月， ‘In summer, in the fifth month,
 火始昏見。 the Huǒ star made its first appearance at dusk.
 丙子， On Bingzǐ
 風。 **there was wind.**’
- (16) *Zuǒzhuàn* 左傳, ed. TLS ZUO 5.15.4.3
 晉饑 ‘In Jin **there was famine.**’
 NOT: ‘Jin was famished.’

Consider now the following familiar opening of a book of the *Analects*.

- (17) *Analects*, ed. TLS LY 13.1
 子曰：「如有王者， ‘The Master said: “If there arises a true king
 必世而後仁。」 then **a generation has to pass** before **there is**
Goodness.”’

If we take *shì* 世 to be subjectlessly verbal here, both the presence of *bì* 必 ‘necessarily’ and of *érhòu* 而後 are explained. One might still wonder whether we might take the word 世 differently, as a denominal adverb. The question then arises what such a decision would logically entail. From a logical point of view the notion of an adverb is so ill-defined that it provides very little information.

Rén 仁 seems used as an impersonal verb meaning ‘goodness/humaneness prevails’.

- (18) *Analects*, ed. TLS LY 16.5
 蓋均無貧， ‘Presumably, when there is even-handedness **there is**
no poverty,
 和無寡， when there is harmony **there is no dearth,**
 安無傾。 when there is peace **there is no toppling of**
leadership.’

Or take this passage from Mencius in which it is not necessarily right to supply a grammatical subject:

(19) *Mencius*, ed. TLS MENG 2.1.2.25

治亦進，	‘If there is good order (in one’s state) then take office,
亂亦進，	if there is disorder (in one’s state) then likewise take
	office,
伊尹也。	that was Yī Yīn.’

The logically decisive question is whether it makes any logical difference to construe 治 and 亂 as abstract topics here “as for there being good order”. For the logical interpretation of these topics would lead us back to the very verbal construction we were trying to avoid by the assumption of a grammatical topic.

Zhuāngzǐ 莊子 is full of fine pieces of feature-placing.

(20) *Zhuāngzǐ* 莊子, ed. TLS ZHUANG 14.1.4

天有六極五常，	‘Heaven has Six Extremes and Five Constants.
帝王順之則治，	When the Sovereigns and Kings follow these there is
	good order ,
逆之則凶。	and when they go against it there will be inauspicious
	disaster .’

Examples of feature places can also be found in *Hánfēizǐ* 韓非子.

(21) *Hánfēizǐ* 韓非子, ed. TLS HF 51.1.7

是廢常上賢，	‘Thus when disregarding regular practise one honours
	moral talent
則亂；	then there will be political chaos ,
舍法任智，	and when setting aside the law one employs the
	competent,
則危。	then there will be political danger .
故曰：	Therefore it is said:
上法而不上賢。	One should honour the law more than moral talent.’

(22) *Hánfēizǐ* 韓非子, ed. TLS HF 20.7.6

有爭則亂，	‘When there is struggle then there will be chaos .’
-------	--

I shall revert in some detail to the existential first part of the sentence.

(23) *Hánfēizǐ* 韓非子, ed. TLS HF 2.3.3

夫一戰而不勝，	‘If one fails to win in one battle,
則禍矣。	then disaster will ensue .’

- (24) *Hánfēizǐ* 韓非子, ed. TLS HF 23.30.1
 雨十日， ‘**It rained** for ten days,
 夜星。 but at night **stars were visible**.’

Guǎnzǐ 管子 is full of good examples:

- (25) *Guǎnzǐ* 管子, ed. TLS GUAN 8.4.1
 夏行春政，風。 ‘[Calendar] in the summer, if government [suitable only to] spring is carried out, there will be winds.
 行冬政，落。 If government [suitable only to] winter is carried out, [plants] will droop.
 重則雨雹。 If this is repeated, it will rain hail.
 行秋政，水。 If government [suitable only to] autumn is carried out, there will be floods.’

The *Book of Changes* is another rich source for impersonal constructions of many kinds:

- (26) *Xící* 繫辭, ed. TLS XICI 1.2
 日月運行， ‘Sun and moon take their regular paths,
 一寒一暑。 and at one time **it is cold**, at another **it is hot**.’

Other impersonal constructions can be found in the *Lǚshì chūnqiū*:

- (27) *Lǚshì chūnqiū* 呂氏春秋, ed. TLS LS 17.7.1.3
 故一則治， ‘Where there is unity, order results,
 異則亂； where there are differences, chaos ensues;
 一則安， where there is unity, security results;
 異則危。 where there are differences, danger ensues.’

Emotions like sadness can be placed like features, as in *Yànzǐ chūnqiū* 晏子春秋:

- (28) *Yànzǐ chūnqiū* 晏子春秋, ed. TLS YAN 1.14.4
 … 義失則憂。 ‘...when rectitude is lost then there is worry.’

One can of course insist on reading this as ‘then people will be worried’. And it is only against the background of all the other impersonal usages, some of which I have documented above that it actually becomes quite plausible to also take this verb *yōu* 憂 in an impersonal way.

The *Shǐjì* 史記 provides a precious example involving the pest:

(29) *Shiji* 史記, ed. TLS SJ 6.2.3

天下疫。

‘There was the pest everywhere.’

NOT, I think: ‘All under Heaven was pestilential.’

3. The grammatical construal of existence predicates

Alain Peyraube became famous for asking necessary questions such as: “Where exactly does the object go in the sentence?” My teacher Angus Graham, on the other hand, always insisted: ‘Wait a minute: what exactly is the subject?’ Take the ubiquitous phrase *yǒu zhī* 有之, which means something like ‘There is such a thing/person’ or more idiomatically ‘True enough!’ Where’s the Chinese subject? If it is not there, what subject is omitted? If no subject is omitted and 之 must be taken to be that subject, why does that subject take the form of a grammaticalized and specialized *object* pronoun? What is going on here, logically, structurally, and grammatically?

(30) *Analects*, ed. TLS LY 4.6.1

蓋有之矣，我未之見也。

‘Perhaps there is such a person, but I have never seen such a person.’

之 recurs as an object.

(31) *Mencius*, ed. TLS MENG 1.2.3.1

齊宣王問曰：

‘King Xuān of Qí asked:

「交鄰國有道乎？」

“As for cultivating good relations with the neighboring states, is there any method (of doing it)?”

孟子對曰：「有」。

Mencius answered (politely): “Yes, there is.”

The question now is whether what is omitted in the reply 有 is the subject or the object of that verb.

There is is as opaque as *il y en a*. We need to know who or what *il* refers to, and what the logical structure is that the *there is* encodes idiomatically. Immanuel Kant famously insisted that ‘exist’ is not a predicate. But grammatically there is no doubt that in “unicorns exist”, *exist* looks and functions very much like a grammatical predicate. Logically, the claim is that the set of unicorns in the universe of discourse is not empty. So, then. If we take our inspiration from Aristophanes, we might go on to write a play about who does the 有-ing in classical Chinese.

Suppose now that 有 places as a feature what follows it, thus creating what works semantically like a feature-placing (existential) predicate. *There are ghosts* is then a systematically misleading expression, whereas the German *es spukt* ‘there are ghosts’ conveys the same idea in the manner closer to the one that I am attributing to the Chinese. Thus the existence of ghosts is expressed in German in the impersonal manner that existence is predicated in Chinese.

Then *Sòng yǒu guǐ* 宋有鬼 would not say ‘Song contains ghosts’, but ‘in Song *spukt es*’. This sentence, then, places an occurrence-feature somewhat in a French style of *il y a*, in which no one can sensibly ask what *il* is taken to refer to. Chinese uses the Chinese verb for *avoir*, and not any copula, as one does in English, and no concept of “giving” as in the German *es gibt*, literally ‘it gives’ or in the very interesting Russian passive *imeetsja* ‘is being had’ used to express existence.

Avoir is transitive, and the thing that is said to exist comes as the object of that verb, in French. And the subject of existential *avoir* is exactly the same as in *il pleut*; it is a dummy subject. It is the kind of subject that flies in the face of Angus Graham when he asks “Wait a minute: where exactly is the subject?” just as the *it* does in *it is raining*. And one notes that this sort of use of *it* is easily extended to the use of the subject in other words: *it never rains but pours*.

The logical subjectlessness then of *It rains cats and dogs* is like the subjectlessness of *il y a beaucoup de monde*. Classical Chinese does not use logically misleading expressions like dummy subjects; it does not do so when it places meteorological features in the physical world like rain or many, many other features. Nor does it do so when it places features in the universe of discourse by the transitive subjectless verb *yǒu* 有. The question then that French and Chinese (and Russian: *imeetsja*) pose in common is why a subjectless verb for TO HAVE is used to place features in the universe of discourse. The logical motivation that suggests itself is this: for a feature to be placed it must be contained in the domain one might call the universe of discourse. Indeed, logically speaking, to exist is always “to be in”, “to be contained in” a domain, that domain being unspecified and in the abstract case of existence radically unspecifiable. The modern Chinese *cúnzài* 存在 brings this out. The seemingly intransitive expression *cúnzài* 存在 delicately suggests an unsayable object that is dimly understood.

Thus *yǒu* 有 might be said to place a feature on some ontological map much as *yǔ* 雨 places a feature on a physical map. Generally, one asks of *yǔ* 雨 what *it* was raining, and usually the answer is *yǔshuǐ* 雨水. Similarly of *yǒu* 有 one asks, curiously, *qu’est-ce qu’il y a?*, and not *qu’est-ce qu’a?*. The French brings out the feature of placement in the use of “y” just as English automatically brings it out in *there*, whereas in Russian *imeetsja* and in the dialectal German *es hat* this placement features is missing.

Cháng yǒu shén guài 常有神怪 ‘there constantly are supernatural and strange features’ (*Lùnhéng* 論衡, ed. TLS LH 9.11.3) works like *dà yǔ hēi xuě* 大雨黑雪 ‘it greatly rained dark snow’: one asks of such verb phrases not “what?” but “where?”.

Just as we understand *yǔ xuě* 雨血 ‘it rained blood’ without asking whether the raining is “by Jove” or “by Heaven”, so we should understand *yǒu guǐ* 有鬼 ‘there are ghosts’ without asking who does the having/containing as if it said ‘it contains > manifests ghosts/ghosting’. Idiomatic English translation is not much help. What we need is not translation but structural deciphering, to use a splendidly untranslatable Chinese expression for our purpose, we need *pòyì* 破譯.

There is a verb *aristotelizein* in Greek, meaning ‘to do an Aristotle’. What I am suggesting here, light-heartedly, is that to say *Aristotle exists* in some languages might work a little like saying *aristotelizetai* ‘it Aristotles (somewhere)’. For ‘Plato existed’ one could then use the form *peplatōnizeto* ‘it has Platoed (somewhere)’ of the existing verb *platōnizesthai*. ‘There will be popes’ might be unpacked to become ‘it will pope (somewhere, some time)’ for a language like Chinese. And to consider this possibility is neither to pledge allegiance to Benjamin Whorf nor to claim that existence predicates must be construed like this from the point of view of some feature-placing logic. Nor is it a way of saying that the world of existing objects was, to the Chinese, like a world of disembodied wafting odors or atmospheres manifesting themselves in places. We are just quietly considering whether some current existence statements might not be structurally isomorphic with those statements that are so common in classical Chinese, which involve **logical** subjectlessness in transitive verbs.

Spirits don’t exist sounds as if it refers to spirits and goes on to claim that these do not exist. It is a systematically misleading expression because it first refers to something and then claims that there never was anything to refer to. This is a much rehearsed point. The classical *wú shén* 無神 ‘...there are no spirits’ (LNZ 1.7.7), when read as I propose, will raise no such logical problem because one can read its logic right off the surface structure as ‘it never contains > manifests spirits (anywhere)’.

Socrates existed similarly, will then have to be diagnosed as a systematically misleading expression. *On the mountain there was a holy man* must count as a logically opaque expression. And the classical Chinese *shān shàng yǒu shén rén yān* 山上有神人焉 ‘on the mountain it contains > manifests a holy man’ might turn out to be a logically less misleading and less opaque expression.

Consider the modern Chinese word *nào* 鬧: *Sòngguó nàole jīhuang* 宋國鬧了饑荒 must be construed impersonally as ‘in Sòng there was a famine’. *Suǒmǎliyǎ nàole hǎidào* 索馬利亞鬧了海盜 is impersonal ‘in Somalia there were robbers’, and the phrase is structurally different from *Suǒmǎliyǎ nàbiān hǎidào duō* 索馬利亞那邊海盜多 or *hǎidào zài Suǒmǎliyǎ nàoluàn* 海盜在索馬利亞鬧亂.

4. Conclusion

Some of the classical Chinese evidence might be taken to suggest to us a natural folk logic of the cognitive system of existence as feature-placings, which are NOT construed as predications of anything, and therefore NOT really predicates after all, just as Kant famously maintained long ago as a general philosophical point. The varied ethnography of folk logic that I advocate, which looks systematically for culture-specific differences and nuances, is subtly different from the study of some presumed biologically general “natural logic” applicable everywhere. As an ethnographer of folk logic one is, for example, quite ready to contemplate the possibility that there is nothing splendidly systematic or successful about construals of existence in human language and that different languages tinker differently with their systems to accommodate tricky notions like that of existence. Maybe existence, grammatically *as well as* existentially, is one of those many areas of conceptual tinkering where we just have to say: *on se débrouille*. Perhaps it is this process of structural and conceptual tinkering that increases what Wilhelm von Humboldt felicitously focused on as *die Verschiedenheit des menschlichen Sprachbaus* ‘the structural diversity of human language’.

Clearly, man is born with a biologically innate and universal ability for learning what are superficially observed to be very different languages. There are some general principles that man applies, and there are many widely varying and highly structured local conventions which man is biologically ready to learn. But beyond that, there are also areas of language which constitute articulatory dilemmas where what seems to be at work is not systematically structured local conventions, but rather linguistic *bricolage*, *Bastellei*, or structural tinkering.

Needless to say, I obviously have **not** solved the vastly complex philosophical problems around existence as a predicate, as it were, by a single sinological or philological sleight of hand. My modest purpose in this paper has been no more than simply to try to tease out some of the ethnography of certain logical sensibilities that might be inherent in classical Chinese ways of trying to construe existence linguistically. Professional logicians and philosophers of language may safely dismiss all of this as what it is: merely analytical ethnography of logical sensibilities, not real philosophy — philological day-dreaming. The complex logical problems surrounding statements of existence will indeed never be solved by any facile philological or sinological sleight of hand. But the patient ethnography of logical sensibilities remains a sublime pleasure and will continue to be so as long as it tries to deal with what in French is called *la logique vécue*. Long live Alain Peyraube, who has enabled and inspired so many of us to practice *la linguistique vécue*.

References

- Abraham, Werner, and Elisabeth Leiss. 2006. Personal and impersonal passives: definite vs. indefinite diatheses. *Transactions of the Philological Society* 104.2:259-296.
- Chocheyras, Jacques. 1985. *Autour de l'impersonnel*. Grenoble: ELLUG.
- Cikoski, John S. 1978. An analysis of some idioms commonly called "passive" in classical Chinese. *Computational Analyses of Asian and African Languages* 9:133-208.
- Comrie, Bernard. 1977. In defense of spontaneous demotion: the impersonal passive. *Syntax and Semantics*, Vol. 8: *Grammatical Relations*, ed. by Peter Cole & Jerrold M. Sadock, 47-58. New York: Academic Press.
- Cuzzolin, Pierluigi, and Maria Napoli. 2009. An overview of impersonal verbs in Indo-European. *Protolanguage and Prehistory: Akten der 12. Fachtagung der Indogermanische Gesellschaft in Krakau*, ed. by Rosemarie Lühr & Sabine Ziegler, 75-81. Wiesbaden: Reichert.
- Desbordes, Françoise. 1991. L'impersonnel d'après les textes théoriques de l'Antiquité. *L'impersonnel*, ed. by Michel Maillard, 11-18. Grenoble: LIDILEM, Université de Grenoble III.
- Ernout, M. A. 1909. L'emploi du passif latin à l'époque républicaine. *Mémoire de la Société de Linguistique de Paris* XV:1-61.
- Freundlich, Jerzy. 1988. *Nominal Functions and Nominalisation in Classical Chinese*. London: SOAS, University of London dissertation.
- Jacobi, Klaus. 1985. Diskussionen über unpersönliche Aussagen in Peter Abaelards Kommentar zu *Peri hermeneias*. *Mediaeval Semantics and Metaphysics: Studies Dedicated to L. M. de Rijk on the Occasion of His 60th Birthday*, ed. by Egbert P. Bos, 1-63. Nijmegen: Ingenium. (Artistarium: Suppl.; 2).
- Lambert, Frédéric. 2010. Peut-on parler d'impersonnel en grec ancien? *L'impersonnel: La personne, le verbe, la voix*, ed. by Didier Bottineau & Françoise Daviet-Taylor, 157-176. Actes du colloque International de linguistique, 8-9 décembre 2006, Université d'Angers, Presses Universitaires de Rennes.
- Li, Charles N. (ed.) 1976. *Subject and Topic*. New York: Academic Press.
- Maillard, Michel. 1991. *L'impersonnel*, Actes du colloque de Grenoble, 17-19 mai 1990. Grenoble: CEDITEL.
- Moore, George E. 1936. Is Existence a Predicate? *Aristotelian Society Supplementary Volume* 15:175-188.
- Myhill, John. 1997. Towards a functional typology of agent defocusing. *Linguistics* 35.5:799-844.
- Ogura, Michiko. 1986. *Old English 'Impersonal' Verbs and Expressions*. Copenhagen: Rosenkilde & Bagger.

- Pears, David F. 1967. Is Existence a Predicate. *Philosophical Logic*, ed. by Peter F. Strawson, 97-102. London: Oxford University Press.
- Pears, David F., and James F. Thomson. 1963. *Is Existence a Predicate?* London: Aquin Press.
- Pieroni, Silvia. 2000. Agents in Latin impersonal passives. *Mnemosyne* 53.3:288-301.
- Pinkster, Harm. 1992. The Latin impersonal passive. *Mnemosyne* 45.2:159-177.
- Seriot, Patrick, and Alain Berrendonner. 2000. *Le paradoxe du sujet*. Lausanne: Université de Lausanne.
- Strawson, Peter. 1974. Is Existence Never a Predicate? *Freedom and Resentment, and other Essays*, 208-217. London: Methuen.
- Williams, Christopher J. F. 1981. *What is Existence?* Oxford & New York: Oxford University Press.

The Causative in Old and Middle Chinese

Tsu-Lin Mei (梅祖麟)

Cornell University

The first part of this paper traces the evolution of the causative category from Old Chinese to Middle Chinese in four stages: (1) the causative *s- prefix, (2) the voiced/voiceless alternation in simplex/causative verbs, which is a reflex of the causative *s-, (3) the periphrastic causative formed by placing a verb meaning ‘to send’ or ‘to command’ before an intransitive verb, and (4) the V-RC (verb-resultative complement) construction. (1) and (2) are in Old Chinese and (3) and (4) are in Middle Chinese. The second part uses comparative evidence from Sino-Tibetan languages to establish the existence of the causative *s- in Proto-Sino-Tibetan and ends up by proposing a morphological criterion for membership in the Sino-Tibetan family.

Key words: Sino-Tibetan, Tibeto-Burman, Old Chinese, phonological, morphology
causative, reconstruction

1. Introduction

Some time ago, I opined that the rise of V-RC (verb-resultative complement) construction in Middle Chinese was due to the loss of Old Chinese affixes such as *s-, *-s, *-r- and the decline of the voiced/voiceless alternation in intransitive/transitive verbs (Mei 1991). I now wish to amend the statement as follows:

- (i) It is well-known that Old Chinese like other Sino-Tibetan languages has a prefix *s- (Yakhontov 1960).
- (ii) The voiced/voiceless alternation refers to pairs such as:
 - 敗 OC *brads > MC *bwai* ‘ruined, defeated’
 - OC *prads > MC *pwai* ‘to ruin, defeat’
 - 折 OC *djat > MC *ǰjat* ‘broken (of sticks, branches)’
 - OC *tjat > MC *tǰjat* ‘to break, to bend’
 - 現 OC *gians > MC *yien* ‘to appear’
 - 見 OC *kians > MC *kien* ‘to see’

These voiced/voiceless pairs were generated by the *s- causative:

敗 OC *brads	>	MC <i>bwai</i>	‘ruined, defeated’
OC *s-brads	> *s-prads > *prads	> MC <i>pwai</i>	‘to ruin, defeat’
折 OC *djat	>	MC <i>žjat</i>	‘broken (of sticks, branches)’
OC *s-djat	> *s-tjat > *tjat	> MC <i>tšjat</i>	‘to break, to bend’
現 OC *gians	>	MC <i>yien</i>	‘to appear’
見 OC *s-gians	> *s-kians > *kians	> MC <i>kien</i>	‘to see’

That is to say, the morphological function of the *s- causative is to form causatives, i.e. ‘to defeat X’ is ‘to cause X to be defeated’ or ‘to cause defeat in X’. Its phonetic effect is to devoice the immediately following voiced stop. Let *b-* represent voiced obstruents, and *p-* voiceless obstruents, then the devoicing effect of *s-* may be stated as follows:

(SDO) *s-* devoicing (obstruents): *s-b* > *s-p* > *p-*
s- is lost after devoicing the immediately following voiced obstruent.

According to this analysis, the voiced/voiceless alternation should be described in terms of its origin, as initial devoicing by the causative *s-. A further consequence of this analysis is that the causative *s- and the voiced/voiceless alternation should not be regarded as two distinct morphological processes, as my earlier statement implies, but rather the latter should be seen as a reflex of the former.

(iii) In Early Old Chinese, the *s- causative and the voiced/voiceless alternation were both active. In Late old Chinese (around the Warring States period) prefix *s- disappeared but its reflex—the voiced/voiceless alternation in pairs of verbs—remained in use until about the beginning of Middle Chinese (601 CE).

(iv) As Old Chinese morphological processes went into decline, two new processes arose to take their place. The first is to prefix 使 *shi* < *srjəgx ‘to send’ or 令 *ling* < *rjings ‘to command, to order’ to an intransitive verb. From the 1st to the 6th century, we have constructions of the following type:¹

V (+N) + *shi* + V or V (+N) + *ling* + V

¹ See Liu Cheng-hui (2002) and Wei Pei-chuan (2000) for more examples. Examples (1), (2), (3) are from the period before the Common Era, and examples (4) to (10) are from the period between Eastern Han (25 CE-220 CE) and Nanbeichao (420 CE-589 CE).

- (1) 民可使富也。(Mengzi 孟子, Jin xin shang 盡心上)
 min ke shi fu ye
 people can cause rich PART²
 ‘People can be made rich.’
- (2) 弱者使之強，短者使長，貧者使多糧(糧)。
 ruo zhe shi zhi qiang, duan zhe shi chang,
 weak NOM cause him strong short NOM cause long
 pin zhe shi duo liang
 poor NOM cause much grain
 ‘The weak is made strong, the short is made long, and the poor is made plentiful in grain.’ (Mawangdui Hanmu Boshu 馬王堆漢墓帛書, Shiwen 《十問》)
- (3) 見廩之粟積，義積之，勿令敗。
 jian yuan zhi su ji, yi ji zhi, wu ling bai
 see insect go grain pile should repile it do-not cause ruined
 ‘When (you) see insects going into piles of grain, you should repile the grain, and do not let the grain go bad.’ (Shuihudi, Qin Lü Shiba Zhong Cang Lü 睡虎地《秦律十八種倉律》)
- (4) 物穢者拭其上使明也。(Shiming 釋名, Shi yan yu 釋言語)
 wu hui zhe shi qi shang shi ming ye
 thing dirty NOM wipe GEN top cause bring PART
 ‘When things get dirty, wipe its surface to make it clean.’
- (5) 細切肉令散。(Shiming 釋名, Shi yin shi 釋飲食)
 xi qie rou ling san
 minutely cut meat cause shreds
 ‘Minutely cut the meat to make (it into) shreds.’
- (6) 發，撥也，撥使開也。(Shiming 釋名, Shi yan yu 釋言語)
 fa (*pjat) bo (*pat) ye, bo shi kai ye
 throw-out spread-out PART spread-out cause open PART
 ‘To throw out (an arrow) is to spread out; to spread out is to cause (it) to open.’

² The abbreviated grammatical glosses used in the examples are: GEN (genitive), LOC (locative), NOM (nominalizer), PART (particle).

- (7) 搗之如粉，以牛膽和之，煮之令燥。(Baopuzi 抱朴子, Huang bai 黃白)
 dao zhi ru fen, yi niu dan huo zhi, zhu zhi ling zao
 pound it like powder with ox gall mix it boil it cause dry
 ‘Pound it into powder, mix it with ox gall, and boil it until it becomes dry.’
- (8) a. 作魚醬法。去鱗，淨洗，拭令乾。
 zuo yu jiang fa: qu lin, jing xi, shi ling gan
 make fish sauce method remove scale clean wash wipe cause dry
 ‘Recipe for making fish sauce: remove the scales, wash it clean, and wipe to make it dry.’ (Qi Min Yao Shu 齊民要術, Zuo jiang fa 作醬法)
- b. 夜煮細糠湯，淨洗面，拭乾，以藥塗之。
 ye zhu xi kang tang, jing xi mian, shi gan,
 night boil fine chaff soup clean wash face wipe dry
 yi yao tu zhi
 with medicine smear it
 ‘At night boil fine chaff soup (mixed with water) into a soup; wash (your) face clean, wipe it dry, and smear with medicine.’ (Qi Min Yao Shu 齊民要術, Zhong hong lan hua zhi zi 種紅藍花梔子)
- (9) a. 黍熟，以淨席薄攤令冷，塊大者擘破，然後下之。
 shu shou, yi jing xi bao tan ling leng,
 millet cooked with clean mat thin spread cause cook
 kuai da zhe bo po, ran.hou xia zhi
 lump big NOM break.by.hand break then lower it
 ‘When the millet (cakes) are cooked, spread them out evenly on a clean mat and let them cool. Break the big clumps before putting down (the yeast).’ (Qi Min Yao Shu 齊民要術, Ben qu bing jiu 笨麴并酒)
- b. 於甕中搗黍令破，瀉甕中，復以酒杷攪之。(See above)
 yu weng zhong nuo shu ling po, xie weng zhong,
 LOC jar middle grasp millet cause break pour jar middle
 fu yi jiu ba jiao zhi
 again with wine rake stir it
 ‘In the jar, crush the millet (cakes) to break them up, pour into (another) jar, and again stir with a wine rake.’
- c. 擘破塊，內著甕中。(See above)
 bo po kuai, nei zhu weng zhong
 break.by.hand break lump enter LOC jar middle
 ‘Break up the lumps by hand and put into the jar.’

- (10) a. 尋使工師作七鐵丸，燒令極赤。
 xun shi gong shi zuo qi tie wan, shao ling ji chi
 subsequent cause work master make even iron ball burn cause very hot
 ‘Subsequently ask a blacksmith to make seven iron balls, fire them until they are very red.’ (Northern Wei, *Xian Yu Jing* 賢愚經, Da zheng cang 大正藏 V.372)
- b. 猶如鐵丸，猛火[火]燒赤，取而吞之。
 you ru tie wan, meng huo shao chi, qu er tun zhi
 just like iron ball fiery fire burn red take and swallow it
 ‘Just like an iron ball, burn to red hot with fiery fire, and then pick up and swallow it.’ (Later Qin, *Chu Yao Jing* 出曜經, Da zheng cang 大正藏 V.668)

The second process is the formation of V-RC compounds. What is written as V-RC can just as well be written as $V_1 + V_2$; with the proviso that V_1 is a transitive verb and V_2 is an intransitive verb. As we have just seen in (8a) and (8b), in (9a), (9b), and (9c), and in (10a) and (10b) “ $V_1 + ling$ 令 + V_2 ” give rise to $V_1 + V_2$ (which is the same as V + RC) when *ling* ‘to order, to command’ is deleted.

There is a second type of “V-RC”, as shown in Cao Guangshun’s (1995) classic study, in which the RC consists of grammaticalized transitive verbs, e.g.

~卻 <i>que</i>	< *khjak	‘to get, to ward off’
~得 <i>de</i>	< *tək	‘to expel, to capture’
~將 <i>jiang</i>	< *tsjang	‘to bring, to take’
~取 <i>qu</i>	< *tshjuqx	‘to take’

These verbs through grammaticalization lost their transitivity and became RC (resultative complement) and PC (phase complement) some time between Early Middle Chinese and Late Middle Chinese. Especially noteworthy is the phase complement ~卻 *que*, which, together with the phase complement ~了 *liao*, led to the rise of the perfective suffix 了 *le*.

The rise of “V-RC” and markers for periphrastic causative (such as 使 *shi* and 令 *ling*) are large and important topics, and I have only scratched the surface. But the main focus of this paper will be on the causative *s- and the voiced/voiceless alternation in simplex/causative pairs of verbs. These are morphological processes which Old Chinese inherited, directly or indirectly from Proto-Sino-Tibetan.

2. Initial devoicing by *s- causative in Tibeto-Burman languages

The theory of initial devoicing by *s- causative was first proposed for Written Tibetan by Betty Shefts Chang (1971) and extended to Tibetan and Liang-shan Lolo by Betty Shefts Chang & Kun Chang (1976). The theory was rediscovered by Gong (2002: 187) for Old Chinese through his comparison of WT pairs of verbs with OC pairs of verbs. I became convinced of its validity through reading *Han-Zangyu Gailun* (*An Introduction to Sino-Tibetan Languages*, 1991, 2003). In what follows, page number refers to the 2003 edition of *Han-Zangyu Gailun*. Similar accounts of individual languages are available in Graham Thurgood & Randy J. LaPolla eds., *The Sino-Tibetan Languages* (2003).

2.1 Dulong

In Dulong (*HZYGL*, pp.311-331), intransitive verbs can be made transitive or causative by the addition of the causative prefix *sǎ*, e.g.

Dulong	<i>īp</i>	‘to sleep’	<i>sǎ īp</i>	‘cause to sleep’
	<i>dām</i>	‘full’	<i>sǎ dām</i>	‘to fill’
	<i>tāt</i>	‘thick’	<i>sǎ tāt</i>	‘to thicken’
	<i>bə̀ɬ</i>	‘fat’	<i>sǎ bə̀ɬ</i>	‘to fatten’
	<i>dāt</i>	‘broken (of string)’	<i>sǎ dsāt</i>	‘to break (string)’

2.2 Yi or Liang-shan Lolo

In Yi or Liang-shan Lolo (*HZYGL*, p.431), causative pairs have voiced initial in the non-causative and Voiceless initial in the causative, e.g.

Yi (Lolo)	non-causative	causative
	<i>gu</i> ³³ ‘to hear’	<i>ku</i> ³³ ‘to cause someone to hear’
	<i>ga</i> ⁵⁵ ‘to wear clothes’	<i>ka</i> ⁵⁵ ‘to cause someone to wear clothes’
	<i>bi</i> ⁵⁵ ‘to come out’	<i>pi</i> ⁵⁵ ‘to cause to come out’
	<i>dzi</i> ³³ ‘to melt’	<i>dzi</i> ³³ ‘to cause to melt’
	<i>ge</i> ³³ ‘to break’	<i>khe</i> ³³ ‘to cause to break’
	<i>dzo</i> ³³ ‘to collapse’	<i>teho</i> ³³ ‘to cause to collapse’
	<i>ŋi</i> ³³ ‘to sit’	<i>ŋi</i> ³³ ‘to cause to sit’

In the case of Dulong, the PTB (Proto-Tibeto-Burman) *s- had acquired an epenthetic vowel, and being separated by this vowel, *s-* in *sə* could not affect the voiced initial of the intransitive verb. In Yi, on the other hand, *s- did devoice the voiced initials, yielding the voiceless initials of the cognate causative, thus:

<i>gu</i> ³³	‘to hear’	/	*s-g > <i>ku</i> ³³	‘to cause to hear’
<i>bi</i> ⁵⁵	‘to come out’	/	*s-b > <i>pi</i> ⁵⁵	‘to cause to come out’
<i>dzi</i> ³³	‘to melt’	/	*s-dz > <i>tei</i> ³³	‘to cause to melt’
<i>ŋi</i> ³³	‘to sit’	/	*s-ŋ > <i>ŋi</i> ³³	‘to cause to sit’

It should perhaps be noted that in the last pair, *s- devoices the (voiced) nasal initial of the noncausative ‘to sit’, and as a result, we have a voiceless nasal initial in the causative ‘to cause to sit’. The rule governing this sort of change is as follows. Let *m* represent voiced sonorants, and *ṃ* represent voiceless sonorants. Then we have a slightly different version of SD:

(SDS) *s-* devoicing (sonorants): *s-m* > *s-ṃ* > *ṃ*

As we will see, both SD and SDS are valid for Lolo-Burmese and Old Chinese, but only SD is valid for written Tibetan.

2.3 Written Tibetan

There are two ways to form the causative in Written Tibetan (*HZYGL*, pp.138-139). One is to prefix *s-* to the base of the verb; the other is by way of the voiced/voiceless alternation in the initials of non-causative/causative verbs.³

WT	non-causative	causative
s- + verb	<i>laŋ</i> ‘to rise’	<i>s-laŋ</i> ‘to cause to rise’
	<i>log</i> ‘to return’	<i>s-log</i> ‘to cause to return’
	<i>nyal</i> ‘to lie down; to sleep’	<i>s-nyal</i> ‘to lay down; to bed a person’
	<i>nub</i> ‘to sink; to decline’	<i>s-nub</i> ‘to cause to perish; to destroy’

³ The list for WT non-causative/causative verbs is assembled from Betty Shefts Chang (1971) and Huang Bufan (1981).

vcd / vcl	<i>bye</i>	‘to open by itself’	<i>phye</i>	‘to cause to open’
	<i>bud</i>	‘to come out, e.g. my tooth came out’	<i>phud</i>	‘to pull out’
	<i>bral</i>	‘to get separated’	<i>phral</i>	‘to cause to separate’
	<i>dul</i>	‘soft, tame, gentle’	<i>thul</i>	‘to tame, to conquer’
	<i>don</i>	‘to come out’	<i>thon</i>	‘to expel’
	<i>gum</i>	‘to die’	<i>khun</i>	‘to kill’ (Jaschke: ‘eleg.’)
	<i>grol</i>	‘loose, to become untied’	<i>khrol</i>	‘to release, to untie’

For bases with lateral or nasal initials, the causative is formed by prefixing *s-* to the base. For bases with voiced-stop initials, *s-* is seldom present in the causative. Instead we have alternations of voiced (*b d g*) and voiceless aspirated (*ph th kh*) consonants. It is not hard to deduce that **s-* was present, but disappeared after devoicing the immediately following voiced stop, thus:

$$\begin{aligned}
 s-b &> s-p > p > ph \\
 s-d &> s-t > t > th \\
 s-g &> s-k > k > kh
 \end{aligned}$$

The last step follows Li’s Law, discovered by Li Fang-kuei in 1933, which says voiceless unaspirated stops in absolute position change to voiceless aspirated stops.

The above is a simplified account of one of the features associated with Written Tibetan verbal morphology. For a more accurate and complicated account, consult Betty Shefts Chang (1971).

I will now cite, from *The Sino-Tibetan Languages (STL)*, passages describing the causative morphology in Kathmandu Newar, Tani, Hayu (Kiranti), and Tangut. The purpose of these citations is to show that all four languages have the voiced/voiceless alternation in non-causative/causative verbs, and that **s-* is their common source.

2.4 “Kathmandu Newar”, by David Hargreaves (*STL*, 371-376)

“Nepal Bhasa (Kathmandu Newar), spoken by approximately 690,000 speakers, is the language of the Newars of the Kathmandu Valley of Nepal.” (*STL*, 371-376).

“There are two morphological processes marking causativization. One process, confined to a restricted set of verbal-forms, is a non-productive reflex of the Proto Tibeto-Burman causative prefix **s-* (Malla 1985:99). Examples include: *den-* ‘lie’ > *then-* ‘lay’; *gya-* ‘afraid’ > *khya-* ‘frighten’. (*STL*, 371-376).

“A second causative process, fully productive, is marked by the causative morpheme *-k-/-kɔl-* suffixed to the verbal stem...” (*STL*, 371-376).

2.5 “The Tani Languages”, by Jackson T.-S. Sun (*STL*, 456-459)

“Tani refers to a compact cluster of Tibeto-Burman languages situated at the eastern end of the Himalayas, in an area skirted on four sides by Tibet, Assam, Bhutan, and Burma. The Tani languages are spoken by about 600,000 indigenous people of Arunachal Pradesh and northern Assam, including the Adi (many subtribes), Nishi(ng)-Bengni, Hill Miri, Tagin, and Apatani tribes of East Kameng, Lower Subansiri, Upper Subansiri, West Siang, East Siang, and Dibang Valley districts of Arunachal Pradesh, as well as the Mising people of Assam.” (*STL*, 456-459)

“The familiar Tibeto-Burman voicing alternation in the formation of simplex-causative pairs is preserved in some measure by certain languages, such as Padam-Mising” (Lorrain 1907):

<i>dir</i>	‘(as of sticks) to break’
<i>tir</i>	‘to break (sticks)’
<i>ben ~ bet</i>	‘(as of ropes) to break’
<i>pen ~ pet</i>	‘to break (ropes)’

Reduplication is an important morphological device favored by Tani languages to increase lexical bulk...”

2.6 “Kiranti Languages”, by Karen H. Ebert (*STL*, 505-509)

“The area where the thirty-two Kiranti languages are spoken stretches over eastern Nepal roughly from the Likhu river in the west into Darjeeling and Sikkim in the east. Hayu and Sunwar are situated a bit further west”. (*STL*, 505-509)

“Some western languages and Limbu also form causatives by devoicing and/or aspiration: Hayu *bok-* ‘be born’ / *phok-* ‘give birth’, Thulung *get-* ‘come up’ / *khet-* ‘bring up’, Limbu *tæks-* ‘be torn’ / *thæks-* ‘tear’.” (*STL*, 505-509)

2.6.1 “Hayu”, by Boyd Michailovsky (*STL*, 518-523)

“The Hayu (Nepali *hayu*, Hayu *wa:ju*) inhabit an area between 50 and 100 kilometers southeast of Kathmandu in the valley of Sun Kosi and across the Mahabharat range to the south.” (*STL*, 518-523)

Non-productive derivational morphology

Hayu preserves traces of the well known Tibeto-Burman causative morphology in the form of word-families, some thirty-five pairs of related verb roots. This morphology is not productive. In most pairs, the (roughly) non-causative member has a voiced stop initial and the causative member a plain or aspirated unvoiced one.

Examples :

<i>gik</i> (vtt) ‘to tie’	vs.	<i>kik</i> (vt, vtt) ‘to put on a belt’
<i>dzol</i> (vi) ‘to fall’	vs.	<i>tsol</i> (vt) ‘to save the life of’
<i>bok</i> (vi) ‘to rise, to get up’	vs.	<i>kok</i> (vt, vtt) ‘to rouse, to raise’
<i>dok</i> (vi) ‘to fall’	vs.	<i>tok</i> (vt, vtt) ‘to drop’ / ‘to cause to fall’
<i>dat</i> (vi) ‘to run out’	vs.	<i>that</i> (vtt) ‘to use up’
<i>bek</i> (vi) ‘to enter’	vs.	<i>phek</i> (vtt) ‘to bring/take in’ (pp.520-522)

Roots are classified as ‘intransitive’ (vi), ‘transitive’ (vt or vtt or both)... Transitive verbs which use the *k-* suffixes are identified as vt; those which use the *t-* suffixes as vtt. Many transitive verbs use both, and they are identified as vt, vtt.

2.7 “Tangut” by Gong Hwang-cherng (*STL*, 602-605)

“‘Tangut’ (also known as the Xixia language) is an extinct Tibeto-Burman language that was spoken in the Xixia empire that existed from 1038 to 1227 in northwest China.” (*STL*, 602-605)

Alternations of initial consonants

The alternation between voiced and voiceless aspirated consonants is relevant to the change in syntactic category. In the following examples the voiced initial consonants represent intransitive verbs, whereas the voiceless aspirated initial consonants represent transitive and/or causative verbs.

- | | |
|---------------|--------------------------------------|
| <i>bie</i> 2 | ‘to release, to open’ (vi) |
| <i>phie</i> 2 | ‘to release, to open, to untie’ (vt) |

2. *bja* 2 ‘end, cut off’ (vi)
phja 1 ‘cut off’ (vt)
3. *dwər* 2 ‘to burn, to blaze’ (vi)
thwər 1 ‘to burn, to cause to blaze’ (CAUS)
4. *dźjwi* 1 ‘to clean, be pure, melt’ (vi)
śhjwi 1 ‘make clean, to purify, melt’ (vt)

2.8 The historical development in Tibeto-Burman

Dai Qingxia of Central University of Minorities said the following in ‘The Historical Development of the Causative Category in Tibeto-Burmese Languages’ (in Chinese, 2001):

“The causative category in Tibeto-Burman languages has many inflectional forms.”

What are their relationships in developmental history. Through comparative studies, we can determine that the *s- causative is the earliest, and that the various phonetic alternations and prefixes which appeared later all came from this single source. The nature of *s- should be regarded as the first element of a consonant cluster and therefore part of the initial, which is to say it is not a free standing syllable. The *s- differs in phonetic characteristics from the syllabic *să which developed later, and therefore the ways they influence the development of phonetic expression of the causative category also differ. This distinction is very important, because it is related to the correct understanding of the rules governing the phonetic development in Tibeto-Burman languages.

The causative *s- took two different paths in its later development. The first path is for s- to fuse with the immediately following C, and thus the consonant cluster sC changes through this fusion into a single-consonant initial. If the noncausative is voiced in its initial, the causative formed from this fusion would be voiceless, thus generating the voiced/voiceless alternation which expresses the causative. This is an important development. Modern Tibeto-Burman languages which retain in the voiced/voiceless alternation to express the noncausative/causative contrast in pairs of verbs, in which the voiced goes with the noncausative, and the voiceless the causative, but there is a difference in frequency of occurrence of such pairs in different languages, for example:

Lolo (Yi)	<i>dze</i> ³³	‘to burn’	vs.	<i>tee</i> ³³	‘to cause to burn’;
Pumi (Prinmi)	<i>bze</i> ¹³	‘to break’	vs.	<i>phze</i> ¹³	‘to cause to break’;
Zhaba (札巴)	<i>a</i> ³³ <i>ji</i> ⁵⁵	‘to melt’	vs.	<i>a</i> ³³ <i>ei</i> ⁵⁵	‘to cause to melt’;
Nu (Nung)	<i>g.nu</i> ⁵⁵	‘to break’	vs.	<i>kh.nu</i> ⁵⁵	‘to cause to break’ etc.

Voiced/voiceless alternation as a way to express the noncausative/causative contrast is the most important method of using phonetic modification to signify the causative category in Tibeto-Burman, and it should be recognized that this type of alternation was the dominant process historically.

A different path was taken when *s- was separated from the following C and with the addition of a vowel became a free standing syllable. Thereupon prefixing *să to the noncausative verb expresses the corresponding causative verb. Once *să became separated from the noncausative verb, it could no longer influence the initial. This is another important development. Because the syllabic *să can be prefixed to any noncausative verb, the morphological function of the causative category is greatly expanded. After *să became a free standing syllable, the possibility of forming *sC (to express the causative) was cut off, that is to say, the path to generate alternation in pairs of verbs became impassable.

In living languages it can be clearly seen that those which possess a syllabic causative prefix, such as Jinghpo with its $să^{31}$, $t\check{s}^{31}$, $f\check{s}^{31}$ or Dulung with its $sə^{31}$ all lack alternation in the initial consonants. On the other hand, those which lack a syllabic causative prefix (such as Achang and Zaiwa (Atsi)) are more developed in using phonetic modification as a means to express the causative.

In the passage just cited, we find Michailovsky mentioning “the well-known Tibeto-Burman causative morphology”, Jackson Sun speaking of “the familiar Tibeto-Burman voicing alternation in the formation of simplex-causative pairs”, and Hargreaves referring to this well-known causative morphology as “a non-productive reflex of Proto Tibeto-Burman causative prefix *s-.” The same view is expressed by Dai Qingxia. The question naturally arises as to whether Old Chinese also had this familiar causative morphology. When we turn to the relevant sections of *Han-Zangyu Gailun*, or *The Sino-Tibetan Languages*, no clear answer can be found. Baxter & Sagart (1998) and Pulleyblank (2000) both mention the *s- prefix but neither paper has anything specific to say about its causative function. In the next section, I hope to fill the lacuna.

3. The causative morphology in Old Chinese

The causative morphology in Old Chinese consists of the following simplex/causative pairs of verbs.⁴

⁴ Wang Li (2000:469-488) and Zhou Zumo (1966) assembled an accurate collection of philological notes concerning voiced/voiceless alternation in simple/causative verbs such as 敗 *b- / 敗 *p-. Baxter & Sagart (1998) give a list of these pairs in English.

1. 敗	*brads > <i>bwai</i> *prads > <i>pwai</i>	‘ruined, defeated’ ‘to ruin, defeat’
2. 別	*brjat > <i>bjät</i> *prjat > <i>pjät</i>	‘to be different, to leave’ ‘to divide, to separate’
3. 斷	*duanx > <i>duan</i> *tuans > <i>tuan</i>	‘severed’ ‘to cut off’
4. 折	*djat > <i>žjat</i> *tjat > <i>tšjat</i>	‘broken (of sticks, branches)’ ‘to break, to bend’
5. 屬	*djuk > <i>žjwok</i> *tjuk > <i>tšjwok</i>	‘conjoined, connected’ ‘to connect, to join’
6. 長 _{長短} 長 _{生長}	*drjang > <i>djang</i> *trjangx > <i>tjang</i>	‘long, tall’ ‘grow tall, increase’
7. 解	*grigx > <i>yai</i> *krigx > <i>kai</i>	‘to crack, break open’ (vi) ‘to cut up (an ox)’
8. 現 見	*gians > <i>yien</i> *kians > <i>kien</i>	‘to appear’ ‘to see’
9. 繫	*gigs > <i>yiei</i> *kigs > <i>kiei</i>	‘linked, joined’ ‘to tie’
10. 會	*gwæds > <i>yuai</i> *kwæds > <i>kuai</i>	‘united, conjoined’ ‘to combine, to aggregate (the account)’
11. 滅 威	*mjat > <i>mjat</i> (*smjiat) > *hmjiat > <i>kjwat</i>	‘to drown, extinguish, destroy’ ‘to cause x to be destroyed’

In all pairs, the non-causative member has a voiced initial and the causative member a voiceless initial. By our previous argument, these pairs are a reflex of the causative *s-.

Noncausative	Causative
1. 敗 *brads ‘ruined, defeated’	*s-b > *s-p > *prads ‘to ruin, defeat’
4. 折 *djat ‘broken (of sticks)’	*s-dj > *s-tj > *tjat ‘to break, bend’
8. 解 *grigx ‘break open’ (vi)	*s-g > *s-k > *krigx ‘to cut up’
11. 滅 *mjat ‘to extinguish, destroy’	*s-m > *s-m̥ > *hmjiat 威 ‘to cause x to be destroyed’

The last example requires a word of explanation. Below I cite Ode 192 with Karlgren’s (1950) translation.

燎之方揚	‘When the fire is just flaming high
寧或滅之	how can anyone extinguish (*mjiat) it?
赫赫宗周	The majestic Tsung Chou
褒姒威之	(lady) Si of Pao has destroyed (*smjiat) it.’

In my earlier paper (Mei 1989:36) I argued that the last line should be read as ‘Lady Si of Bao has caused it to be destroyed’ and that 威 *smjiat has a causative sense, that is ‘cause to be destroyed’. Since we know that *s-* can devoice the immediately following nasal, Yakhontov’s (1960) formula $*sm- > x-$ can be rewritten as $*sm- > *ṁ- (*hm-) > x-$, where *hm-* is just a typologically convenient way to write the voiceless labial nasal *ṁ*. It should now be apparent that the *s-* that devoices *m-* is the same *s-* that devoices *b-*.

SDS: $*s-m > *s-ṁ > x-$	
滅 $*m- > m-$: 威 $*s-m > *ṁ > x$
‘extinguished, destroyed’	‘cause to be destroyed’

SD: $*s-b > *s-p > p-$	
敗 $*b- > b-$: 敗 $*s-b > *p > p$
‘defeated, ruined’	‘caused to be defeated’

Yakhontov’s (1960) list of cognate pairs should be added as examples exhibiting the denominative function of **s-*, that is, to convert a noun into a derived form which is not a noun.

The denominative function of **s-* prefix:

吏 $*rjəgs > lī$	‘official’	使 $*s-rjəgs > sī$	‘to send, employ’
林 $*rjəm > ljəm$	‘woods’	森 $*s-rjəm > sjəm$	‘woodsy’
墨 $*mək > mək$	‘ink’	黑 $*s-mək > xək$	‘dark, black’

4. Concluding remarks

Randy LaPolla (2004:46) sounded the theme of this paper when he said:

“Hopper (1991:24) gives the example of the past tense in English, where we have at least three layers: (a) Periphrasis: *We have used it* (newest layer), (b) Affixation: *I admired it* (older layer) and (c) Ablaut: *They sang* (oldest layer)’. In Tibeto-Burman we find a similar phenomenon in terms of causative marking.

The vast majority of Tibeto-Burman languages show evidence of either an *s- prefix or at least a difference in voicing/aspiration of the initial to mark a causative verb, though in almost all the languages this has ceased to be productive, so in an analytical causative, often formed using a verb meaning ‘do’, ‘make’, or ‘cause’, has developed. The older forms did not disappear, though, and can often be used together with the newer form of the causative, sometimes with variant shades of meaning.” (See LaPolla 2000 for example of Rawang.)

What LaPolla said about Tibeto-Burman languages is equally true of Old Chinese, for we have just seen that the causative is successively expressed by (1) the causative *s- prefix, (2) the voiced/voiceless alternation in simplex/causative pairs such as 敗 *brads ‘defeated’ / *prads ‘to defeat’; 折 *djat ‘broke’ / *tjat ‘to break, to bend’, (3) analytic causative (or periphrasis) formed by adding 令 *ling* ‘command, make’ or 使 *shi* ‘send, cause’ to the verb, for example, 使富 *shi fu* ‘to make rich, to enrich’ 令冷 *ling leng* ‘to make cool, to cool’.

The same three stages development is found in Yi (Lolo): (1) The causative *s- prefix, (2) the voiced/voiceless alternation in simplex/causative verbs, for example, *bi*⁵⁵ ‘come out’ / *pi*⁵⁵ ‘cause to come out’; *ga*⁵⁵ ‘wear clothes’ / *ka*⁵⁵ ‘cause to wear clothes’, (3) adding the causative marker *bi*⁴⁴ ‘give’ ... -*su*³⁴, for example:

- (12) a. *la*²¹ *bu*³³ *zi*³³ *dzu*³³ (non-causative)
 cow grass eat
 ‘The cow eats grass.’
- b. *la*²¹ *bu*³³ *bi*⁴⁴ *zi*³³ *dzu*³³-*su*³⁴ (causative)
 cow CAUS grass eat
 ‘Make (or let) the cow eat grass.’

There are three ways to express the causative in Daofu (道孚), a conservative Qiangic dialect in northwestern Sichuan (Huang Bufan 1991:27-28; *HZYGL*, p.280).

1. The s- prefix, e.g.

<i>dʒə-brə</i>	‘got scared, afraid’	<i>də-s-prə</i>	‘to startle’
<i>də-bja</i>	‘(bamboo) split’ (vi)	<i>də-s-pja</i>	‘split, break’ (vt)
<i>də-qə</i>	‘(fire) go out’	<i>nə-s-qə</i>	‘put out (fire)’
2. The voicing alternation in simplex/causative verbs, e.g.

<i>nə-brɛ</i>	‘(string) broken’	<i>nə-phrɛ</i>	‘to break (string)’
---------------	-------------------	----------------	---------------------

<i>də-dzoqa</i>	‘to leave’	<i>də-tchoqha</i>	‘to separate’
<i>də-brə</i>	‘become loose’	<i>yə-phrə</i>	‘to untie’

3. Adding *sphrə* ‘to cause’ or *və* ‘to make’ after the noncausative verb, e.g.

<i>rje</i>	‘to rise’	<i>rje sphrə</i>	‘to cause, to rise, to arouse’
<i>scçer</i>	‘to be afraid’	<i>scçer sphrə</i>	‘to scare, frighten’
<i>zəvzəv</i>	‘broken (into pieces)’	<i>zəvzəv və</i>	‘to crush, smash’
<i>rla</i>	‘destroyed, ruined’	<i>rlalə və</i>	‘to destroy’

In Daofu there is the layering effect noted by Hopper & LaPolla, that is, the simultaneous presence of the analytic causative and two forms of affixial causative. On the other hand, Early Old Chinese (1300-800 BCE) had both the causative *s-* and the voicing alternation, and Late Old Chinese (800-200 BCE) had the voicing alternation and the analytic causative. In this respect, Late Old Chinese is exactly like present-day Yi (Lolo).

In Mei (1980), I attempted to show that Old Chinese departing tone comes from the suffix **-s*, and gave the following cognate pairs as part of the evidence.

OC		WT	
Verb	量 _平 <i>*grjang</i> ‘to measure’	<i>N-grang</i>	‘to number, to count’
Noun	量 _去 <i>*grjangs</i> ‘to measure’	<i>grangs</i>	‘number’
Verb	織 _入 <i>*thək</i> ‘to weave’	<i>N-thag</i>	‘to weave’
Noun	織 _去 <i>*thəks</i> ‘stuff made of colored silk’	<i>thags</i>	‘texture, web’

In WT, adding *-s* to a verb converts it into a noun. In Old Chinese, the change from non-departing tone to departing has the same nominalizing effect. WT has *-s* but no tone; Late Old Chinese departing tone but no **-s*. **-s* and departing tone are similar in morphological function and in complementary distribution. Therefore, the departing tone of Old Chinese comes from Sino-Tibetan **-s*.

As final consonants **-s* and **-ʔ* evolved into tones in Old Chinese, the nominalizing effect of **-s*, and its successor, the departing tone, also declined. It came to pass that two new nominalizers, 所 *suo* and 者 *zhe*, arose to take their place.

In general, Early Old Chinese still had vestiges of the PST affixial morphological system, which includes the causative **-s-*, the nominalizing **-s*, **N-* and **r-*. With the disappearance of these affixes and their reflexes, the Chinese languages of the Han dynasty (206 BCE-220 CE) became a truly isolative language.

Finally, I would like to propose, in the spirit of Conrady (1896), a morphological criterion for membership in the Sino-Tibetan family.

Criterion of ST membership:

If a language or its proto-form has the causative *s*- and/or its reflex — the voicing alternation in simple/causative verbs — then it must be a member of the Sino-Tibetan family.

According to the arguments given in this paper, Old Chinese, Classical Tibetan, Lolo-Burmese, Tangut, Kiranti, Tani, Jinghbo, Dulong as well as Qiangic languages are all members of the Sino-Tibetan family.

Phonological Appendix

The theory of *s*- devoicing consists of two rules, one for obstruents and the other for sonorants.

SDO:	<i>s</i> - devoicing (obstruents)	$s-b > s-p > p$
SDS:	<i>s</i> - devoicing (sonorants)	$s-m > s-\eta > m$

SDO was discovered by Betty Shefts Chang (1971) for Classical Tibetan and applied to Tibetan and Lolo-Burmese by Betty Shefts Chang & Kun Chang (1976). It deserves to be called “Shefts’ Law”.

SDA was discovered for Tibeto-Burman languages by Jin Peng (1983), Dai Qingxia (1985), and Huang Bufan (1991). *Hang-Zangyu Gailun* (2003) has all the relevant data on $sm > \eta$, $sn > \eta$, $sn > \eta$, $sl > l$ on pps. 16-19, 25, 37-38, 204-205, 377-378. As these scattered page references indicate, a systematic account is implicit in the book.

Li Fang-Kuei’s (1971, 1980) reconstruction of Old Chinese was first published in 1971. At the time no one in Chinese historical phonology knew anything about the devoicing effect of *s*-. Consequently, in Li’s system there are hypothetical phonological rules which contradict either SDS or SDO. Below I list our preferred reconstruction according to SDO or SDS on the left and Li’s reconstruction on the right.

*sm- > *m- > x-	*sm- > s-
*荒 smang > xwang	*喪 smang > sang
*亡 mjang > mjwang	*亡 mjang > mjwang
*sd- > *t- > t-	*sd- > dz-
*斷 s-duans > *tuans	*摧 sdəd > dzuâi
*sdj- > *tj- > tsj-	*sdj- > zj-
*折 s-djat > *tjat	*詞 sdjəg > zji
*sg- > *k- > k-	*sg- > dz-, *sgj- > zj-
*絲 s-gigs > *kigs	*造 sgəgwh > dzâu
*俗 sgjuk > zjwok	

In Gong Hwang-cherng’s (2002) revised version of Li’s system, the rules on the right disappeared and are replaced by the rules on the left. This is also our position.

Pulleyblank (1973:114) proposed that the OC counterpart to WT *a-chung* is the cause for the voiceless/voiced alternation associated with transitive/intransitive pairs. At first the OC counterpart was thought to be a voiced laryngeal spirant *h*. When it is realized that WT *a-chung* was prenasal *N-*, its OC counterpart is changed accordingly to a prenasal **N-*. In Baxter & Sagart (1998:43, 46) we have the latest version of the “voicing” theory:

Transitive		Intransitive	
*敗 *p- > p-	‘to defeat’	*N-p > b-	‘defeated, ruined’
*打 *tj- > tsj-	‘to break’	*N-tjp- > ʒj-	‘broken (of sticks, branches)’
*系 *k- > k-	‘to tie’	*N-k- > ɣ-	‘linked, conjoined’

But the *a-chang* prefix is not an intransitivizer (LaPolla 2003:23). WT prenasal *N-* can aspirate a following voiceless unaspirated stop, e.g. *N-p* > *N-ph* (Betty Shefts Chang 1971:685, Gong 2002:208), but cannot voice it. Finally, we have a better account of the OC voicing alternation associated with simplex/causative pairs in term of causative **s-*, which is in accord with the morphological function of WT *s-*. For all theses reasons, the “N-voicing” theory should be replaced by the “s-devoicing” theory.

References

- Baxter, William H., and Laurent Sagart. 1998. Word formation in Old Chinese. *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, ed. by Jerome L. Packard, 35-76. Berlin & New York: Mouton de Gruyter.
- Cao, Guangshun. 1995. *Jindai Hanyu Zhuci* [Grammatical Particles in Late Middle Chinese]. Beijing: Language and Culture Press.
- Chang, Betty Shefts. 1971. The Tibetan causative: phonology. *Bulletin of the Institute of History and Philology Academia Sinica* 42.4:623-765.
- Chang, Betty Shefts, and Kun Chang. 1976. The prenasalized stop initial of Miao-Yao, Tibeto-Burman and Chinese: a result of diffusion or evidence of a genetic relationship? *Bulletin of the Institute of History and Philology Academia Sinica* 47.3:467-502.
- Conrady, August. 1896. *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten: ein Beitrag zur vergleichenden Grammatik der indochinesischen Sprachen, insonderheit des tibetischen, barmanischen, siamesischen und chinesischen*. Leipzig: O. Harrassowitz.
- Dai, Qingxia. 1985. Achangyu de qingbiyin [Voiceless nasals in the Burmish dialect of Achang]. *Minzu Yuwen* [Minority Languages of China] 1985.2:11-15.
- Dai, Qingxia. 2001. Zang-Mian yuzu shidong fanchou de lishi yanbian [The historical development of the causative category in Tibeto-Burmese languages]. *Journal of Chinese Linguistics* 29.1:1-11.
- Ebert, Karen H. 2003. Kiranti languages: an overview. *The Sino-Tibetan Languages*, ed. by Graham Thurgood & Randy J. LaPolla, 505-517. London: Routledge.
- Gong, Hwang-cherng. 2002. *Collected Papers on Sino-Tibetan Linguistics*. Taipei: Institute of Linguistics, Academia Sinica.
- Gong, Hwang-cherng. 2003. Tangut. *The Sino-Tibetan Languages*, ed. by Graham Thurgood & Randy J. LaPolla, 602-620. London: Routledge.
- Hargreaves, David. 2003. Kathmandu Newar (Nepāl Bhāṣā). *The Sino-Tibetan Languages*, ed. by Graham Thurgood & Randy J. LaPolla, 371-384. London: Routledge.
- Hopper, Paul J. 1991. On some principles of grammaticalization. *Approaches to Grammaticalization*, Vol. 1, ed. by Elizabeth Closs Traugott & Bernd Heine, 17-35. Amsterdam & Philadelphia: John Benjamins.
- Huang, Bufan. 1981. Gu Zangyu dongci de xingtai [Classical Tibetan verbal morphology]. *Minzu Yuwen* [Minority Languages of China] 1981.1:1-13.
- Huang, Bufan. 1991. Daofuyu [Horpa-Shangzhai]. *Zang-Mianyu Shiwu Zhong* [Fifteen Tibeto-Burman Languages], 1-46. Beijing: Yanshan Press.

- Jin, Peng. 1983. *Zangyu Jianzhi [A Sketch of the Tibetan Languages]*. Beijing: Nationalities Press.
- Karlgren, Bernhard. 1950. *The Book of Odes*. Reprinted from *Bulletin of the Museum of Far Eastern Antiquities* 16 and 17. Stockholm: Museum of Far Eastern Antiquities.
- LaPolla, Randy J. 2000. Valency-chaining derivations in Dulong/Rawang. *Changing Valency: Case Studies in Transitivity*, ed. by R. M. W. Dixon & Alexandra Y. Aikhenvald, 282-311. Cambridge & New York: Cambridge University Press.
- LaPolla, Randy J. 2003. Overview of Sino-Tibetan morphosyntax. *The Sino-Tibetan Languages*, ed. by Graham Thurgood & Randy J. LaPolla, 22-42. London: Routledge.
- LaPolla, Randy J. 2004. On nominal relational morphology in Tibeto-Burman. *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on his Seventieth Birthday*, 43-73. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Fang-Kuei. 1933. Certain phonetic influences of the Tibetan prefixes upon the root initials. *Bulletin of the Institute of History and Philology Academia Sinica* 4.2:135-157.
- Li, Fang-Kuei. 1971. Shangguyin yanjiu [Studies on Old Chinese phonology]. *Tsing Hua Journal of Chinese Studies*, New Series 9.1-2:1-61.
- Li, Fang-Kuei. 1980. *Shangguyin Yanjiu [Studies on Old Chinese Phonology]*. Beijing: The Commercial Press.
- Liu, Cheng-hui. 2002. *Hanyu Dongbu Jiegou Lishi Fazhan [Historical Development of the Verb-Resultative Complement Construction in Chinese]*. Taipei: Hanlu Books.
- Lorrain, J. Herbert. 1907. *A Dictionary of the Abor-Miri Language*. Shillong: Eastern Bengal and Assam Secretariat Printing Office.
- Ma, Xueliang. (ed.) 1991[2003]. *Han-Zangyu Gailun [An Introduction to Sino-Tibetan Languages]*. Beijing: Nationalities Press. (The book was originally published in 1991 by Peking University Press.)
- Malla, Kamal P. 1985. *The Newari Language: A Working Outline*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Mei, Tsu-Lin. 1980. Sisheng bieyi zhong de shijian cengci [Chronological strata in derivation by tone change]. *Zhongguo Yuwen [Studies of the Chinese Language]* 1980.6:427-443.
- Mei, Tsu-Lin. 1988. Neibu nigou Hanyu san li [Three examples of internal reconstruction in Chinese]. *Zhongguo Yuwen [Studies of the Chinese Language]* 1988.3:169-181.
- Mei, Tsu-Lin. 1989. The causative and denominative functions of the *-s- prefix in Old Chinese. *Proceedings of the Second International Conference on Sinology: Section on Linguistics and Paleography*, 33-51. Taipei: Academia Sinica.

- Mei, Tsu-Lin. 1991. Cong Handai de 'dong, sha', 'dong, si' lai kan dongbu jieyou 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.
- Mei, Tsu-Lin. 2000. *Mei Zulin Yuyanxue Lunwenji* [Collected Papers on Linguistics by Tsu-Lin Mei]. Beijing: The Commercial Press.
- Michailovsky, Boyd. 2003. Hayu. *The Sino-Tibetan Languages*, ed. by Graham Thurgood & Randy J. LaPolla, 518-532. London: Routledge.
- Pulleyblank, Edwin G. 1973. Some new hypothesis concerning word families in Chinese. *Journal of Chinese Linguistics* 1.1:111-125.
- Pulleyblank, Edwin G. 2000. Morphology in Old Chinese. *Journal of Chinese Linguistics* 28.1:26-51.
- Sun, Jackson T.-S. 2003. Tani languages. *The Sino-Tibetan Languages*, ed. by Graham Thurgood & Randy J. LaPolla, 456-466. London: Routledge.
- Thurgood, Graham, and Randy J. LaPolla. (eds.) 2003. *The Sino-Tibetan Languages*. London: Routledge.
- Wang, Li. 1965[2000]. Gu Hanyu zidongci he shidongci de peidui [The pairing of simplex and causative verbs in Old Chinese]. *Wang Li Yuyanxue Lunwenji* [Collected Papers in Linguistics by Wang Li], 469-486. (The article was originally published in *Zhonghua Wen Shi Luncong* 6 (1965)).
- Wei, Pei-chuan. 2000. Shuo Zhonggu Hanyu de shicheng jieyou [The causative constructions in Middle Chinese]. *Bulletin of the Institute of History and Philology Academia Sinica* 71.4:807-856.
- Yakhontov, S. E. 1960. Consonantal combinations in Archaic Chinese. Paper presented by the USSR delegation at the 25th Congress of Orientalists, Moscow. Moscow: Oriental Literature Publishing House.
- Zhou, Zumo. 1966. Sisheng bieyi shili [Examples of derivation by tone change]. *Wenxue Ji*, 81-119. Beijing: Zhonghua Book Company.

A Syntactic Analysis of Modal *bì* 必: Auxiliary Verb or Adverb?

Barbara Meisterernst (梅思德)

Humboldt-University Berlin

In this paper a syntactic analysis of the different words written with the character *bì* 必 in Han period Chinese is at issue with particular regard to its modal employments. The paper shows that there is conclusive evidence for an analysis of *bì* 必 as both an auxiliary verb and a modal adverb. Although the evidence for a verbal analysis of *bì* 必 is stronger, its syntactic constraints with regard to aspecto-temporal adverbs and — to a certain extent to negative markers — argue for an additional analysis of *bì* 必 as a modal adverb. Consequently, it is hypothesized that two different, but etymologically related words written *bì* 必 have to be assumed in the Han period literature according to the syntactic constraints *bì* 必 is subject to; one appears as a modal verb, employed independently or as auxiliary verb ‘must, need to’, and one appears as a modal adverb ‘certainly’.

Key words: modal auxiliary verbs, modal adverbs, *bì* 必, hierarchy of adverbs, Han period Chinese

1. Introduction

In this paper a syntactic analysis of the different words written with the character *bì* 必 in Han period Chinese is at issue with particular regard to its modal employments. The character *bì* 必 can write a full verb with the meanings ‘secure, ascertain, enforce; be certain’, and, according to the approach adopted in the linguistic literature, an adverb and/or an auxiliary verb and a conjunction. Yang Shuda (1928[1959:1, 10]) categorizes *bì* 必 as a verb which can be employed transitively or intransitively and as an adverb, both the verb and the adverb express certainty. Dobson (1959:237) distinguishes between *bì* 必 as a full word ‘must’ and as a grammatical word ‘must be’ in a determinative sentence (roughly corresponding to a sentence with a nominal predicate) expressing “Emphasis or Restriction upon the Unqualified Categorical: A rhetorical rather than a material quality is given to the predication (which remains unqualified and categorical)” (1959:112). In his study of Late Han Chinese (Dobson 1964:22), he additionally discusses a preverbal *bì* 必 in the sense of ‘decidedly, inevitably, of a certainty,’ etc. which can be negated by *wèi* 未 and by *bù* 不. The exact grammatical function of *bì* 必

is not specified. According to Shadick (1968:755ff), *bì* 必 ‘necessarily, always, certainly’ belongs to the *Predicate adjuncts* (PA) which resemble adverbs, but are distinguished from them “because they can also modify nominals, as adjunct to a nominal predicate”. More particular, *bì* 必 belongs to the PA of quality which “express the degree of certainty with which a predication is made or the intrinsic inevitableness of what is said”. To the same class belong the adverbs (PAs in Shadick’s terminology) *gù* 固 ‘certainly, actually’; *chéng* 誠 ‘truly’; *xìn* 信 ‘indeed, truly’; *shí* 實 ‘actually, in fact’; *yì* 亦 ‘indeed, in fact’; *guǒ* 果 ‘after all, in the end’, and others. Additionally Shadick (1968:765) shows that *bì* 必 can function as a conjunction in a second clause, referring to “all possible occasions” in the meaning ‘necessarily, always’. Gassmann (1997:196) qualifies *bì* 必 as a modal verb ‘müssen (must)’ comparable to *néng* 能 ‘can’; equally the Thesaurus Linguae Sericae categorizes *bì* 必 as a (auxiliary) verb in different syntactic constellations. Unger analyses *bì* 必 as a ‘Modalpartikel’ (modal particle), indicating assertion mostly in future contexts, in contrast to von der Gabelentz who categorizes it as an auxiliary verb.¹ Unger argues for his analysis with the fact that quantifiers ‘Distributiva’ usually follow *bì* 必. As an adverb *bì* 必 can, according to Unger, precede a nominal predicate (1987:100), a verbal predicate (1992:42ff), and a prepositional predicate (1989:245), and it can appear in subordinate and in matrix clauses. What Unger labels as ‘Modalpartikel’ corresponds in general to what the author labels ‘modal adverb’ in accordance with Pulleyblank (1995:18) who lists *bì* 必 as an adverb (1995:99) and as belonging to one of those “sentence adverbs that can be used in verbless noun predicates”. In the *Gǔdài Hànyǔ xūcí cídiǎn* (2000:23) *bì* 必 is analyzed as an adverb and a conjunction. However, *bì* 必 can be distinguished from many other adverbs by a number of syntactic criteria, and — amongst others — by the fact that it can occur on its own without a verb following it. This short summary clearly shows the difficulties in assigning a particular syntactic status either as a modal (auxiliary) verb or as a modal adverb to *bì* 必.

As has already been noted in Alleton (1972:23) and in a recent discussion on Modality in English and Chinese (Li 2004:111 and elsewhere), this is a general problem in the analysis of adverbs and modal expressions in Chinese.² The following discussion

¹ With the exception of initial and final particles all particles in Unger’s framework correspond to adverbs in the present author’s framework.

² Alleton (1972:23) states that adverbs and modal verbs mostly appear in the same position: “Dans la majorité des contextes, ainsi, le verbe modal se trouve dans la même position que l’adverbe.” But, according to her, they can be distinguished according to their capacity of forming an independent utterance “les verbes modaux sont des formes libres (susceptibles de constituer un énoncé), alors que les adverbes ne le sont pas,” although she concedes that this distinction is often difficult to obtain.

will try to solve the puzzle of the syntactic status of *bì* 必 in demonstrating that it clearly differs in its syntactic constraints from many other lexical items which traditionally have been analyzed as adverbs. It will show that *bì* 必 evidently displays syntactic constraints which argue for both its status as an auxiliary verb and its status as an adverb. Since the investigation of the syntactic hierarchy of adjuncts, including modal, aspecto-temporal and manner adverbs, in Classical and Han period Chinese is still in its infancy, results of this study have to be considered preliminary. The following discussion will focus on the syntactic constraints of *bì* 必, a semantic analysis of its modal values will be postponed to a separate study. The present study is only one of the first steps of an investigation into the precise syntactic and semantic constraints modal adverbs and auxiliary verbs in Classical and Han period Chinese are subjected to, an investigation which — with a few exceptions — has not been executed yet for the periods mentioned.³ The study is based on data predominantly taken from the *Shiji* 史記 (ca. 100 BC), the *Hànshū* 漢書 (1st c. AD) and the *Lùnhéng* 論衡 (1st c. AD).

2. The syntactic constraints of *bì* 必

In order to determine the syntactic status of *bì* 必 the different syntactic constraints it can be subject to will be discussed in this section. It will be hypothesized that the character *bì* 必 actually represents a full verb, a modal auxiliary verb, and an adverb. The syntactic tests applied in order to distinguish between its function as an auxiliary verb and an adverb concern the position of *bì* 必 with regard to *wh*-words, the relative particle *suǒ* 所, the relative position of negative markers, and the position of *bì* 必 in relation to other adverbs. According to the preliminary hierarchy of adverbs established in Meisterernst (2009) on the basis of a study of aspecto-temporal adverbs in Han period Chinese, modal adverbs by default precede aspecto-temporal and manner adverbs.⁴

First, examples for *bì* 必 as an independent verb will be presented and in the following the syntax of *bì* 必 in a subordinate position will be discussed.

2.1 *Bì* 必 as an independent verb

As an independent verb, *bì* 必 refers to a situation which is certain to happen, a notion which extends to its modal notion indicating certainty and necessity. In

³ See in particular Peyraube (1999, 2001), Liu (2000), and Meisterernst (2008b, 2008c).

⁴ This may provide some evidence for their analysis as belonging to the category of adverbs which Cinque (1999:11f) labels ‘higher’ (Sentence) adverbs. But a clear distinction between sentential adverbs and adverbs exclusively operating on the level of V still has to be established.

Pulleyblank (1991:34, 1995:175) *bì* 必 is not glossed as a full verb, only its adverbial meanings are given.

In the first example presented *bì* 必 appears as an intransitive verb with the meaning ‘secure’ following the modal verb *kě* 可.

- (1) 且漢王不可必，身居項王掌握中數矣… (*Shiji*: 92; 2622)

Qiě Hàn wáng bù kě bì, shēn jū Xiàng wáng
furthermore Han king NEG can secure person dwell Xiang king
zhǎngwò zhōng shù yǐ
jurisdiction middle few FIN⁵

‘Furthermore, the king of Han cannot be secured (his loyalty cannot be secured = he cannot be trusted), it has [already] been several times that he has personally been under the jurisdiction of the king Xiang, ...’

In example (2) *bì* 必 occurs as the main predicate, as an adjective with the meaning ‘be certain, sure’, marked by the final particle *yǐ* 矣 which usually indicates a change of state situation. In this construction *bì* 必 is quite frequently attested.

- (2) 吾聞秦軍圍趙王鉅鹿，疾引兵渡河，楚擊其外，趙應其內，破秦軍必矣。 (*Shiji*: 7; 305)

Wú wén Qín jūn wéi Zhào wáng Jùlù,
I hear Qin army encircle Zhao king Julu
jí yǐn bīng dù Hé,
quick conscribe army cross river

Chū jí qí wài, Zhào yīng qí nèi, pò Qín jūn bì yǐ
Chu attack its outside Zhao react its inside destroy Qin army certain FIN
‘I have heard that Qin has encircled the king of Zhao in Julu, but if we quickly take the army and cross the river, and if Chu attacks from outside and Zhao reacts from inside, it will be certain that we will destroy the army of Qin.’

In the following example *bì* 必 is employed as a transitive verb with the meaning ‘ascertain, make sure’, followed by a formally nominalized complement clause.

⁵ The abbreviations for the grammatical glosses in the examples are: CON connector; FIN final particle; FUT future; MOD modality; NEG negator; NEG_{asp} aspectual negator; NEG_{mod} modal negator; OBJ object pronoun; PREP preposition; PRT particle; REL relativizer; SUB subordinator; YI weak transitive verb ‘take, assume’ (Meisterernst 2008b:104f).

- (3) 今媾，郝又以不能必秦之不復攻也。(Shǐjì: 76; 2372)

Jīn gòu, Shí yòu yǐ bù néng bì Qín zhī
 now peace.treaty Shi again so NEG can secure Qin SUB
 bú fù gōng yě
 NEG again attack FIN

‘If we will have a peace treaty now, Shi will then still not be able to make sure that Qin does not launch another attack.’

In example (4) the syntactic analysis of *bì* 必 is less straightforward. It obviously appears in the meaning ‘must be’ followed by a prepositional phrase, a meaning which often also seems to be adequate if *bì* 必 is followed by a nominal predicate. Nevertheless, if *bì* 必 precedes nominal predicates, it is analyzed as a ‘Copula of unqualified categorical predication’ in Dobson (1959:237) and in Unger (1987:100) as a particle, i.e. an adverb, of assertion, identically to its employment in a sentence with a verbal predicate, and not as a proper verb. According to Unger (1989:245) the same adverbial function has to be assumed for *bì* 必 preceding a prepositional predicate; other adverbs which are attested in this position are the modal adverb *qí* 其 ‘perhaps, possibly’, the quantifier *gè* 各 ‘each’ and the negative marker *bù* 不 ‘not’. The fact that the negative marker *bù* 不 negates predicative prepositional phrases argues against their analysis as being syntactically identical to noun phrases; accordingly an apparent syntactic difference exists between nominal predicates and prepositional predicates in predicative position.⁶ If *bì* 必 is syntactically identical to the afore mentioned adverbs, the occurrence of the evident adverbs *qí* 其 and *gè* 各, in the same syntactic position as *bì* 必 could argue against the analysis of *bì* 必 as a verb in this position. However, since these adverbs differ considerably in their syntax and their semantics, only two of them, i.e. *qí* 其 and *bì* 必 have in common that they can precede both nominal and verbal predicates,⁷ they do not necessarily have to be assumed to be syntactically identical to *bì* 必. Accordingly, *bì* 必 will be analyzed as a verb with the meaning ‘must be’ in this example.

- (4) 蓋天好陰，祠之必於高山之下，小山之上，命曰「峙」…(Shǐjì: 28; 1367)

Gài tiān hào yīn, cí zhī bì yú gāo shān zhī xià,
 PRT heaven love Yin sacrifice OBJ must.be at high mountain SUB bottom

⁶ An analysis of prepositional phrases in predicative position is not at issue in this paper and has to be postponed to a separate study, but according to the examples presented in Unger, those with prepositional phrases apparently negated by *bù* 不 rather have to be analyzed as cases of ellipsis (with a deleted verb) than as cases of predicative prepositional phrases.

⁷ See Harbsmeier (1981:85) “Obviously, *ge* would be unthinkable in sentences with nominal predicates...”

xiǎo shān zhī shàng, mìng yuē: zhì...
small mountain SUB height name say altar...

‘Well, since heaven loves the *yin*, when sacrificing to it, this must be [performed] at the foot of a high mountain or on top of a small hill, the [place’s] name must be “Altar”...’

2.2 Modal *bì* 必: adverb or verb, some syntactic evidence

In this section the syntactic arguments for the classification of *bì* 必 as an adverb or a modal auxiliary verb will be discussed.

According to the data presented in Unger for Classical Chinese and the data discussed in this article, *bì* 必 evidently shows syntactic constraints which are different from most other adverbs, whether modal or aspecto-temporal. These concern in particular the position of negation markers, *wh*-words, in particular the *wh*-word *hé* 何, and the aspecto-temporal adverbs *jiāng* 將 and *qiě* 且 which are the aspecto-temporal adverbs mostly attested in combination with *bì* 必 in the Han period texts under consideration. The almost exclusive restriction to these adverbs is not surprising, since both predominantly serve to locate a situation in the future, identically to *bì* 必 which also usually qualifies situations in the future. In the following, a few examples for *bì* 必 in combination with the afore mentioned syntactic elements will be presented to reveal the syntactic status of *bì* 必. Unfortunately, many of the criteria which have been proposed for the syntactic distinction e.g. between auxiliary verbs, adverbs and full verbs in Modern Mandarin (for a summary and evaluation of these criteria see Li 2004, particularly p.130) are not applicable to Classical and Han period Chinese.⁸

2.2.1 *Bì* 必 in combination with *wh*-words

As a general rule *wh*-words follow modal and aspecto-temporal adverbs, but precede auxiliary verbs.⁹ As the following examples will show, if *bì* 必 appears in combination with a *wh*-word, the *wh*-word always precedes *bì* 必. This evidently contrasts *bì* 必 with

⁸ The main features which are not applicable to Classical and Han period Chinese — since they do not yet exist during these periods — are ‘reduplication’ and the employment of ‘aspectual suffixes’ which are both not permitted with modal auxiliary verbs (Li 2004:130) in Modern Mandarin.

⁹ Interrogation is listed as one of the criteria to determine the linguistic status of auxiliary verbs in Modern Mandarin, but since *wh*-questions and yes-no questions in Modern Mandarin differ considerably from those in Classical and Han period Chinese, the criteria established for Modern Mandarin are not necessarily applicable to earlier stages of the language.

the modal adverb *yì* 亦, the syntactic status of which as an adverb is much more evident than that of *bì* 必 and which is usually followed by the *wh*-word. In examples (5) to (7), *bì* 必 is followed by a noun phrase and not by a verb in which case it is analyzed as a full verb in analogy to those instances where it precedes a prepositional phrase as in example (4).

- (5) 名何必湯武，俗何必成康！(Shìjì: 112; 2957)
 Míng hé bì Tāng Wǔ, sú hé bì Chéng Kāng
 name why must.be Tang Wu custom why must.be Cheng Kang
 ‘Regarding your name, why must it be [that of] Tang or Wu, and regarding your customs, why must they be [those of] Cheng or Kang!’
- (6) 觀所以得尊寵及所以廢辱，亦當世得失之林也，何必舊聞？
 (Shìjì: 18; 878)
 Guān suǒyǐ dé zūn chǒng jí suǒyǐ fèi rù, yì dāng
 view whereby get honour favour and whereby abandon shame also at
 shì dé shī zhī lín yě, hé bì jiù wén
 generation get loose SUB multitude FIN why must.be old question
 ‘If one considers how one gains honour and favours and why one becomes abandoned and humiliated, in our times, too, there are manifold occasions for getting lost, why does one need old questions?’
- (7) 奉職循理，亦可以爲治，何必威嚴哉？(Shìjì: 119; 3099)
 Fèng zhí xún lǐ, yì kě yǐ wéi chí,
 serve office follow principle also can take make govern
 hé bì wēi yán zāi
 why must.be awe stern FIN
 ‘Those who fulfil their functions and follow the principles can also be employed to govern, why must there be awe and sternness?’

Contrastively to example (5) to (7) where *bì* 必 is followed by a noun phrase, in examples (8) to (11) it is followed by a verb in which case it is analyzed as an auxiliary verb; one of the main criteria to characterize an auxiliary verb is the fact that they have to be followed by another verb (Li 2004:113, Peyraube 1999:28f) with which they establish a Verb+Object (complement) structure (Peyraube 1999:29, Pulleyblank 1995:39) in contrast to adverbs which modify the verb they precede. In adverbial modification, the modified verb constitutes the centre, whereas with auxiliary verbs, it is the auxiliary verb which constitutes the centre of the phrase (Peyraube 1999:39).

- (8) 子路曰：「有民人焉，有社稷焉，何必讀書然後爲學！」
(*Shǐjì*: 67; 2212)
Zǐ Lù yuē yǒu mín rén yán, yǒu shè-jì yán,
Zi Lu say there.are people man there there.are spirits.of.land.and.grain there
hé bì dú shū ránhòu wéi xué
why must read book afterwards become learn
‘Zi Lu said: “There are the people and there are the spirits of land and grain,
why must one read books in order to be considered learned afterwards?”’
- (9) 宮殿中可以避世全身，何必深山之中，蒿廬之下。」(*Shǐjì*: 126; 3205)
Gōngdiàn zhōng kě yǐ bì shì quán shēn, hé bì
palace middle can yi avoid world perfect personality how must
shēn shān zhī zhōng, hāo lú zhī xià
deep mountain SUB middle waste.away hut SUB below
‘In the middle of a palace one can avoid the world and perfect one’s per-
sonality, why must one go deep into the mountains or waste away in a hut?’
- (10) 如恐左右之見，懷屏隱匿之處，足以使蜚不見，何必食之？
(*Lùnhéng*: 20.3.51)
Rú kǒng zuǒyòu zhī jiàn, huái bǐng yǐn nì zhī chù,
if fear entourage SUB see worry screen hide hide SUB place
zú yǐ shǐ zhì bú jiàn, hé bì shí zhī
suffice yi make leech NEG see why must eat OBJ
‘If one fears that the entourage might see it, one may take care to screen it off
in a hidden place, and this is enough to make sure that the leech will not be
seen, why must one eat it?’
- (11) 夫宋人父子脩善如此，神報之，何必使之先盲後視哉？(*Lùnhéng*: 20.7.5)
Fú Sòng rén fù zǐ xiū shàn rú cǐ, shén bào zhī,
fu Song man father son cultivate good be.like this spirit reward OBJ
hé bì shǐ zhī xiān máng hòu shì zāi
why must make OBJ first blind afterwards see FIN
‘Well, if these men from Song, father and son, cultivated the good like this
and the spirits rewarded them, why must they first be made blind and after-
wards be made seeing [again]?’

2.2.2 Modal *bì* 必 preceded by *suǒ* 所

Occasionally, modal *bì* 必 immediately follows the relativizer *suǒ* 所. Apparently not many adverbs are allowed in this position, amongst them none of the above

mentioned modal adverbs including the adverb *yì* 亦 ‘also, indeed, in fact’. Only aspecto-temporal adverbs, in particular *cháng* 常 ‘habitually’ and *cháng* 嘗 ‘once (in the past)’, and manner adverbs seem to be licensed in the position between *suǒ* 所 and the verb. The fact that *bì* 必 is permitted in this position, although it is only infrequently attested, argues for its analysis as an auxiliary verb at least in these examples. However, it must be conceded that in Classical and Han period Chinese not all auxiliary verbs are attested in the *suǒ* 所 — phrase; however, if a modal auxiliary verb is present, it usually follows *suǒ* 所. Although in example (12) the negative marker follows *bì* 必, this does not necessarily — at least in the present construction — argue for its analysis as an adverb; the negative marker has scope only over the second verb, the complement of *bì* 必, and not over the entire predicate, a construction which is also possible with the modal verb *kě* 可 as has been shown in Meisterernst (2008c) and other auxiliary verbs.

- (12) 死者，人之所必不免也。(Shǐjì: 79; 2407)
 Sǐ zhě, rén zhī suǒ bì bù miǎn yě
 die NOM man SUB suo must NEG avoid FIN
 ‘Dying is what man cannot avoid.’

- (13) 此明聖所必加誅也。(Hànshū: 64B; 2818)
 Cǐ míng shèng suǒ bì jiā zhū yě
 this bright wise suo must add punishment FIN
 ‘... this is what the wise and virtuous must inflict a punishment upon.’

2.2.3 Modal *bì* 必 in combination with negative markers

By default, negative markers follow modal and aspecto-temporal adverbs, but precede auxiliary verbs when the entire modal proposition is negated.¹⁰ Accordingly, a negative marker preceding one of these adverbs can — under certain conditions — argue for a change in the analysis of the supposed adverb from adverb to auxiliary verb.¹¹ There are a few exceptions to this rule, particularly concerning the aspectual and modal negative marker *wèi* 未 ‘not yet, never’ which regularly precedes the aspecto-temporal adverb *cháng* 嘗 ‘once (in the past), habitually (in the past)’ and some synonymously employed adverbs. The aspectual negative marker *wèi* 未 and the aspecto-temporal adverbs in

¹⁰ As already mentioned, they can also follow a modal auxiliary verb; in this case only the second verb is negated.

¹¹ Alleton (1972:37, 43) excludes the position of negative markers as a criterion to establish adverbs as a syntactic category, since in Modern Mandarin, *bù* 不 can appear in different positions in relation to the adverb and, additionally, negative markers can appear independently (1972:43).

general evidently share the same position within the verb phrase which accounts for the differences in position of *wèi* 未 and the other negative markers. If the combination *wèi cháng* 未嘗 is again combined with the neutral negative marker *bù* 不, *bù* 不 always appears in its default position following *cháng* 嘗. Additionally to this regular exception with the negative marker *wèi* 未, the negative marker *bù* 不 can either precede or follow the modal adverb (particle expressing assertion) *guǒ* 果 ‘really’ in Classical Chinese (Unger 1992:39f). But according to the examples presented in Unger, the adverb *guǒ* 果 apparently rather has to be analyzed as a manner adverb when preceded by the negative marker *bù* 不 and not as a modal adverb.¹² Furthermore some other adverbs, for instance, the modal adverb *yì* 亦 ‘also, indeed, in fact’¹³ — equally to e.g. the aspecto-temporal adverbs *jì* 既 and *yǐ* 已 ‘already’ — can occasionally be preceded by the negative marker *bù* 不, but this construction is confined to marked cases, i.e. to rhetorical questions which are in general followed by the final particle *hū* 乎.

2.2.3.1 A negative marker follows the modal adverb

First, some examples for the default position of negative markers in combination with modal and aspecto-temporal adverbs will be presented. The most frequently attested negative marker following modal *bì* 必 is the neutral negative *bù* 不, but also the modal negative markers *wù* 勿 ‘(do) not’ and *wú* 無/毋 ‘(do) not’ are attested, although the latter predominantly has to be analyzed as a verb ‘not have’ when preceded by *bì* 必. The following examples represent the employment of the different negative markers following *bì* 必. The adverbial analysis of *bì* 必 in all examples is quite evident.

- (14) 朔不肯，曰：「子必不絕趙祀，朔死不恨。」 (Shiji: 43; 1783)
 Shuò bù kěn, yuē zǐ bì bù jué Zhào sì,
 Shuo NEG willing say you certain NEG interrupt Zhao sacrifices
 Shuò sǐ bú hèn
 Shao die NEG anger
 ‘Shuo was not willing and said: “You will certainly not interrupt the sacrifices of Zhao, so there will be no problem when I die.”’

¹² Manner adverbs seem to be less confined regarding their syntax, apparently they can be either preceded or followed by negation markers. But since their syntactic constraints have not yet been studied in detail, a precise analysis cannot be given in this context.

¹³ This multifaceted adverb exhibits many different functions (Harbsmeier 1981:136ff) and certainly deserves a separate study which would exceed the framework of the present study.

- (15) 寡人雖不肖，必不能事秦。(Shǐjì: 69; 2253)
 Guǎ rén suī búxiào, bì bù néng shì Qín
 lonely man even.if unworthy certain NEG be.able serve Qin
 ‘I may be unworthy, but I will certainly not be able to serve Qin.’
- (16) 諸學仙術，爲不死之方，其必不成，猶不能使冰終不釋也。
 (Lùnhéng: 24.25.22)
 Zhū xué xiān shù, wéi bù sǐ zhī fāng,
 all learn immortal technique make NEG die SUB method
 qí bì bù chéng, yóu bù néng shǐ bīng zhōng bú shì yě
 MOD certain NEG complete as NEG can make ice finally NEG melt FIN
 ‘All this studying the techniques of immortality and practising the methods
 of how not to die will certainly not achieve anything, just as one cannot
 cause ice not to melt eventually.’
- (17) 闔廬且死，告其子夫差曰：「必毋忘越。」(Shǐjì: 41; 1739)
 Hé Lú qiě sǐ, gào qí zǐ Fúchāi yuē bì wú wàng Yue
 He Lū FUT die tell his son Fuchai say certain NEG_{mod} forget Yue
 ‘When He Lū was about to die, he told his son Fuchai: “You must certainly
 never forget Yue.”’
- (18) 君必勿聽也。(Shǐjì: 76; 2370)
 Jūn bì wù tīng yě
 prince certain NEG_{mod} listen FIN
 ‘You certainly must not listen to him.’

2.2.3.2 A negative marker precedes modal *bì* 必

Different negative markers are attested in this position. These are the aspectual and modal negative marker *wèi* 未 ‘not yet’ which almost never follows *bì* 必 in the Classical and Han period literature,¹⁴ the neutral negative marker *bù* 不 ‘not’ and the negative marker *fēi* 非 ‘is not, unless’ which usually, but not exclusively, negates nominal predicates.¹⁵ First, examples with *wèi* 未 will be discussed.¹⁶ Since *wèi* 未 ‘not yet’ functions

¹⁴ Two instances of *bì wèi* 必未 are attested in the *Mòzǐ* (3.3/21/27, 3.3/22/1), one in the *Zhànguó cè* (461/226/27), one in the *Huáinán zǐ* (10/83/16) preceding *wèi cháng* 未嘗, and several instances in the post-Han historical literature, in all instances an analysis of *bì* 必 as a modal adverb is evidently most conclusive.

¹⁵ Additionally, the negative quantifier *mò* 莫 ‘no one’ precedes *bì* 必 in one instance (Shǐjì 23; 1164) in which *bì* 必 clearly functions as a full verb.

¹⁶ For a detailed discussion of the negative marker *wèi* 未 in Han period Chinese see Meisterernst (2008a).

as an aspectual adverb in contrastive distribution with the aspecto-temporal adverbs *jì* 既 and *yǐ* 已, it is not surprising that it differs syntactically to a certain extent from the other negative markers. However, since it belongs to the class of aspecto-temporal adverbs, it is supposed to follow a modal adverb and accordingly its position preceding *bì* 必 argues for an analysis of *bì* 必 as a modal (auxiliary) verb. As has been demonstrated in Meisterernst (2008a, 2008c) *wèi* 未 frequently serves to modify modal auxiliary verbs, usually expressing a strong categorical denial (Meisterernst 2008c). In the *Gǔdài Hànyǔ xūcí cídiǎn*, the combination *wèi bì* 未必 is analyzed as an adverbially employed modifier-head construction with — according to the categorization of *bì* 必 (2000:23) therein — an adverb as its head. It seems to be of younger origin; it is not attested yet in the *Zuǒzhuàn*, the *Guóyǔ*, *Lúnyǔ* and in *Mèngzǐ* (Unger 1992:50f), the earliest examples presented in the *Gǔdài Hànyǔ xūcí cídiǎn* are from the *Zhuāngzǐ*. Although Unger discusses it under the label of *bì* 必 as a particle, he does not exclude that *bì* 必 originally has to be analyzed as a verb when combined with the negative marker *wèi* 未. However, according to him it is already grammaticalized as a combination of two function words in the examples presented. In example (19) *bì* 必 is followed by a complement clause which includes the modal auxiliary verb *néng* 能 ‘be able to’, whereas in (20) and (21) it is followed by a transitive VP.¹⁷ In example (19) *bì* 必 evidently has to be analyzed as a full verb ‘make sure’.

- (19) 今遣少子，未必能生中子也，而先空亡長男，奈何？(*Shǐjì*: 41; 1753)
 Jīn qiǎn shǎo zǐ, wèi bì néng shēng zhōng zǐ yě,
 now send younger son NEG_{asp} certain be.able live middle son FIN
 ér xiān kōng wáng zhǎng nán, nài hé
 CON first vain perish older man what about
 ‘Well, if you now send the younger son, before you have made sure (but you have not made sure yet) that the middle son can live, and let the older son perish in vain, what could be the use of that?’
- (20) 江河之決皆天事，未易以人力爲彊塞，塞之未必應天。(Shǐjì 29; 1409)
 Jiāng Hé zhī jué jiē tiān shì, wèi yì yǐ rén lì
 Jiang He SUB opening all heaven affair NEG_{asp} easy with man strength
 wéi qiáng sài, sài zhī wèi bì yīng tiān
 make strong boundary boundary SUB NEG_{asp} certain correspond heaven
 ‘The breaches [in the banks] of the Jiang and the He are all heavenly work, and it is not easy to build strong bulwarks with human strength, and the bulwarks may not conform with Heaven[’s will].’

¹⁷ In Meisterernst (2008a:140f) *bì* 必 in a similar construction has been analyzed as a modal adverb, an analysis which according to the data presented here cannot be maintained.

The combination *wèi bì* 未必 is frequently attested in the *Lùnhéng*.

- (21) 案呂后之崩，未必遇雷也。(Lùnhéng: 23.10.21)

Àn Lǚ hòu zhī bēng, wèi bì yù léi yě
 regarding Lǚ lady SUB pass.away NEG_{asp} must meet thunder FIN
 ‘When Lady Lǚ died, she need not have been struck by a thunderbolt.’

In example (22) *bì* 必 is followed by the aspecto-temporal adverb *chū* 初 which evidently modifies the complement of *bì* 必. Equally to the complement VP of the modal auxiliary verb *néng* 能 ‘can, be able’ the complement VP of *bì* 必 can be modified by an adverb.

- (22) 案古人君臣困窮，後得達通，未必初有惡，天禍其前；卒有善，神祐其後也。(Lùnhéng: 21.17.28)

Àn gǔ rén jūn chén kùn qióng, hòu dé
 regarding antiquity people ruler subject distress exhaust afterwards get
 dá tōng, wèi bì chū yǒu è, tiān huò qí
 reach.to penetrate NEG_{asp} must first have bad heaven misfortune its
 qián; zú yǒu shàn, shén yòu qí hòu yě
 before eventually have good spirit bless its after FIN
 ‘If amongst men of antiquity rulers and their subjects were distressed, but later achieved success, this need not be, because there was first something bad and Heaven sent misfortune to their predecessors and then eventually there was something good and the spirits sent blessings to their successors.’

In example (23) *bì* 必 is followed by the aspecto-temporal adverb *cháng* 常; a predicate modified by *cháng* 常 or *cháng* 嘗 is always negated by *wèi* 未 preceding the aspecto-temporal adverb, but in this example *wèi* 未 modifies the modal verb ‘must’ and *cháng* 常 modifies the second verb. In example (24) *bì* 必 is followed by the quantifier *jiē* 皆 ‘all’ which usually quantifies the subject.¹⁸

- (23) 如據太平，太平之帝，未必常致鳳鳥與河圖也。(Lùnhéng: 28.14.21)

Rú jù tài píng, tài píng zhī dì, wèi bì cháng
 if rely great peace great peace SUB emperor NEG_{asp} must regularly
 zhì fèng niǎo yǔ Hé tú yě
 bring phoenix bird and River map FIN
 ‘If they rely on the great peace, the emperors of the [times of] great peace need not always have brought on the Phoenix and the Map of the River.’

¹⁸ For a discussion on the quantifier *jiē* 皆 see Harbsmeier (1981:78).

- (24) 當時鄉縣之吏，未必皆伏罪也。(Lùnhéng: 49.2.25)

Dāng shí xiāng xiàn zhī lì, wèi bì jiē fù zuì yě
 at time district county SUB official NEG_{asp} must all admit guilt FIN
 ‘At that time it was not yet necessary that the officials of the districts and
 counties had all admitted their guilt.’

Without being able to give a final account of the syntactic constraints of *jiē* 皆 in this context, it can be stated that it by default follows the negative marker *wèi* 未 but can precede or follow the negative marker *bù* 不. It belongs to the adverbs which are attested modifying the complement VP of the modal auxiliary verb *néng* 能 ‘can’; and apparently it generally precedes aspecto-temporal adverbs and manner adverbs.

The following example represents one of the not very frequent cases of double negation with *bì* 必 in the Classical and Han period literature: i.e. *bì* 必 is preceded by the negative marker *wèi* 未 and followed by the negative marker *bù* 不. This combination, although not very frequently attested, is by no means unique; contrastively, double negation with *bù* 不 as both the first and the second negative marker is almost non-existent. Double negation is proposed in the linguistic literature as one of the criteria to distinguish auxiliary verbs from adverbs (according to Liu 1960:2, cf. Li 2004:116) in Modern Mandarin, but in Classical and Han period Chinese this argument does not hold for all adverbs.¹⁹ According to Li (2004:126), who argues against double negation as a criterion to identify auxiliary verbs, in Modern Mandarin *bì* 必 is excluded from double negation.

- (25) 有求而不得之者矣，未必不求而得之者也。(Lùnhéng: 3.5.31)

Yǒu qiú ér bù dé zhī zhě yǐ,
 have search CON NEG get OBJ NOM FIN
 wèi bì bù qiú ér dé zhī zhě yě
 NEG_{asp} must NEG search CON get OBJ NOM FIN
 ‘If there is someone who searches without getting anything, this must not
 imply that there is someone who gets something without searching.’

Whereas in the examples (19) to (25) the aspectual and modal negative marker *wèi* 未 precedes *bì* 必, in the following examples it is the neutral negative marker *bù* 不 which by default follows adverbial *bì* 必 in not additionally marked sentences. Accordingly, the negative marker *bù* 不 in this position, too, argues for an analysis of *bì* 必 as a modal (auxiliary) verb. In the following examples (26) and (27) *bì* 必 is followed by a

¹⁹ In the combination *wèi cháng bù* 未嘗不 *cháng* 嘗 certainly has to be analyzed as an adverb.

noun phrase. Since noun phrase predicates are usually negated by the negative copula *fēi* 非 and not by *bù* 不, the presence of *bù* argues for an analysis of *bì* 必 as the main verb in this construction. Furthermore, if a negated noun phrase predicate is additionally modified by the assertive modal adverb *bì* 必, this adverb precedes the negative copula *fēi* 非.²⁰

- (26) 故禮也不必一道，而便國不必古。(Shǐjì: 43; 1810)
 Gù lǐ yě bú bì yī dào,
 therefore rites FIN NEG must.be one way
 ér biàn guó bú bì gǔ
 CON serve state NEG must.be antiquity
 ‘Therefore, regarding the rites, there must not be [merely] one way, and to serve the state, there must not be [the example of] antiquity.’
- (27) 馬有千里，不必騏驎之駒；鳥有仁聖，不必鳳皇之鷄。
 (Lùnhéng: 50.20.6)
 Mǎ yǒu qiān lǐ, bú bì qílín zhī jū; niǎo yǒu
 horse have thousand mile NEG must unicorn SUB colt bird have
 rén shèng, bú bì fènghuáng zhī chú
 benevolence wisdom NEG must phoenix SUB fledgling
 ‘A horse may run a thousand miles, however, it must not be the colt of a unicorn; a bird may be benevolent and wise, but it must not be the fledgling of a phoenix.’

In examples (28) and (29) *bì* 必 is followed by a verbal complement and accordingly can be analyzed as a modal auxiliary verb.

- (28) 漢得之不必待以全。(Shǐjì: 53; 2016)
 Hàn dé zhī bú bì dài yǐ quán
 Han get OBJ NEG must.be wait in.order.to complete
 ‘Han may get hold of them, but it must not rely on them in order to be complete.’
- (29) 簡子見之，若寢夢矣，空虛之象，不必有實。(Lùnhéng: 15.5.28)
 Jiǎn zǐ jiàn zhī, ruò qǐn mèng yǐ, kōng xū zhī xiàng,
 Jian zi see OBJ be.like sleep dream FIN empty empty SUB semblance

²⁰ Unger (1987:100) presents two examples for this construction, one from *Xúnzǐ* and one from *Lǚshì chūnqiū*.

bú bì yǒu shí

NEG must have reality

‘When Jian zi saw them, it was like in a dream, they were empty semblances and need not have had any reality.’

In examples (30) and (31) *bì* 必 is preceded by the negative copula *fēi* 非 which is not confined to nominal predication. Modifying a verbal predicate, it negates the entire proposition and not the predicate alone. According to Harbsmeier (1981:19) who provides a comprehensive discussion of preverbal *fēi* 非, it expresses “a negative judgement, rather than a report that something did not happen or a straightforward description that something is or was not ‘thus’.” *Bì* 必 in the instances presented certainly has to be analyzed as a verb; in example (30) it is followed by a formally nominalized full complement clause with an overt subject, and in (31) it is followed by a complement clause with an transitive verb and without an overt subject.

- (30) 田乞及常所以比犯二君，專齊國之政，非必事勢之漸然也，蓋若遵厭兆祥云。(Shǐjì: 46; 1903)

Tiānqǐ jí Cháng suǒ yǐ bǐ fàn èr jūn, zhuān Qí guó
Tianqi and Chang REL with compare revolt two prince usurp Qi state
zhī zhèng, fēi bì shì shì zhī jiàn rán yě,
SUB government NEG certain serve opportunity SUB gradual be.like FIN
gài ruò zūn yàn zhào xiáng yún
rather be.like follow satiate omen auspicious FIN

‘[The reason] why Tianqi and Chang can be compared with the two revolting princes is that they usurped the government of the state Qin, and this need not be the case because the circumstances gradually became like that, but they rather followed and satisfied the prognostications.’

- (31) 陳王起蕲，至陳而王，非必立六國後。(Shǐjì: 89; 2575)

Chén wáng qǐ Qí, zhì Chén ér wàng,
Chen king rise Qi arrive Chen CON become.king
fēi bì lì liù guó hòu
NEG necessary establish six state descendant

‘The king of Chen started his uprising in Qi, and when he arrived in Chen he became king, it is not the case that one must enthrone a descendant of the Six Kingdoms.’

2.2.4 Modal *bì* 必 in combination with different kinds of adverbs

In this section, the position of modal *bì* 必 with respect to other adverbs, modal, and aspecto-temporal adverbs is analyzed. As already mentioned, in general, modal adverbs precede aspecto-temporal and manner adverbs in the hierarchy of adverbs. This hypothesis is based on the study of the position of aspecto-temporal adverbs within the verb phrase (Meisterernst 2009) — which most likely have to be generated within TP — which has to be confirmed with regard to the different modal adverbs and their combination and interrelation with other adverbs.

2.2.4.1 Modal adverbs preceding *bì* 必

Since *bì* 必 as an adverb evidently belongs to the category of modal adverbs, its position with respect to other modal adverbs does not provide any evidence for the discussion on the syntactic status of *bì* 必. Nevertheless, it can be stated that *bì* 必 by default follows the modal adverb *qí* 其 ‘perhaps, possibly’ which expresses different modal values, and the adverb *yì* 亦 ‘also, indeed, in fact’. But, since these two adverbs also regularly precede modal auxiliary verbs, their position cannot argue for either the status of *bì* 必 as an adverb or as a modal auxiliary verb. The same also accounts for the modal adverb *dài* 殆 ‘nearly, almost’, expressing probability and related meanings. Since it follows most other modal adverbs, *bì* 必 — if analyzed as an adverb in this position — appears quite low in the hierarchy of modal adverbs.

- (32) 天子曰兆民，諸侯曰萬民，今命之大，以從盈數，其必有眾。

(*Shǐjì*: 39; 1642)

Tiān zǐ yuē zhào mín, zhūhóu yuē wàn mín,
heaven son say million people feudal.lord say ten-thousand people
jīn mìng zhī dà, yǐ cóng yíng shù,
now name OBJ grand in.order.to follow full number
qí bì yǒu zhòng
MOD certain have multitude

‘The son of Heaven is named “A Million People”, the feudal lords are named “Ten-Thousand People”, one calls them grand now, and in order to conform to the correct proportion he will certainly possess the masses.’

- (33) 而周武王克紂後，成王封叔虞于唐，其地阻險，以此有德與周衰並，亦必興矣。(Shǐjì: 42; 1758)

Ér Zhōu Wǔ wáng kè Zhòu hòu, Chéng wáng fēng Shúyú
CON Zhou Wu king conquer Zhou after Cheng king enfief Shuyu

yú Táng, qí dì zhǔ xiǎn, yǐ cǐ yǒu dé yǔ
 at Tang its land inaccessible dangerous with this have good.term and
 Zhōu shuāi bìng, yì bì xīng yǐ
 Zhou decline common also certain rise FIN
 ‘And after king Wu of Zhou conquered Zhou, king Cheng gave Shuyu Tang
 as his fief, its land was inaccessible and dangerous, but it was on good terms
 with them and they shared Zhou’s decline, they will certainly also rise.’

- (34) 高子泣血，殆必有之。(Lùnhéng: 26.10.10)
 Gāo-zǐ qì xuè, dài bì yǒu zhī
 Gao-zi cry blood almost certain have OBJ
 ‘That Gao-zi shed bloody tears has almost certainly been the case.’

2.2.4.2 Aspecto-temporal adverbs following *bì* 必

A comprehensive study of the syntax and the semantics of aspecto-temporal adverbs has demonstrated that they appear by default below modal adverbs in the hierarchy of adverbs (Meisterernst 2009). In those cases where a lexical item — which usually functions as a modal adverb — follows an aspecto-temporal adverb, it generally has to be interpreted differently, i.e. either as an (auxiliary) verb or as a manner adverb. Manner adverbs by default appear in the position closest to the verb and accordingly follow both modal and aspecto-temporal adverbs. Aspecto-temporal adverbs in general precede modal auxiliary verbs, only very few of them are attested in the position following a modal auxiliary. In this case the complement of the modal verb usually consists of an entire clause including an adverbially modified verb as in examples (22) and (23). According to the syntactic constraints of aspecto-temporal adverbs and auxiliary verbs, *bì* 必 evidently has to be analyzed as an adverb when it is followed by an aspecto-temporal adverb. Due to the fact that *bì* 必 predominantly appears in future contexts, it is most frequently attested in combination with the aspecto-temporal adverbs *jiāng* 將 and *qiě* 且 which locate a situation in the future, but it is not confined to them. Modal values are only secondary with these adverbs (Meisterernst 2004).

- (35) 惠、懷無親，外內棄之；天未絕晉，必將有主，主晉祀者，非君而誰？(Shǐjì: 39; 1662)
 Huì, Huái wú qīn, wài nèi qì zhī, tiān wèi
 Hui Huai not.have close outside inside abandon OBJ heaven NEG_{asp}
 jué Jìn, bì jiāng yǒu zhǔ, zhǔ Jìn sì zhě,
 cut.off Jin certain FUT have master master Jin sacrifice NOM

fēi jūn ér shuí

NEG prince CON who

‘Hui and Huai have no one close, outside and inside one has abandoned them; but Heaven has not yet cut off Jin [from its sacrifices], and it will certainly have a ruler, and the one ruling over the sacrifices of Jin, who else could it be than you, My Lord?’

- (36) 足下必將戰勝然後略地，攻得然後下城，臣竊以爲過矣。

(*Shǐjì*: 89; 2575)

Zúxià bì jiāng zhàn shèng ránhòu lüè dì, gōng dé
sir certain FUT fight win afterwards annex land attack get
ránhòu xià chéng, chén qiè yǐwéi guò yǐ
afterwards subdue city subject take.liberty consider mistake FIN

‘You, Sir, will certainly have to fight and win before you annex the country, to attack and conquer before you subdue the cities, but I consider this wrong.’

- (37) 王雖東取地於越，不足以刷恥；必且取地於秦，而後足以刷恥於諸侯。(Shǐjì: 40; 1726)

Wáng suī dōng qǔ dì yú Yuè, bù zú yǐ shuā chǐ;
king even.if east take land PREP Yue NEG suffice yi wipe.off disgrace
bì qiě qǔ dì yú Qín, ér hòu zú yǐ shuā
certain FUT take land PREP Qin CON afterwards suffice yi wipe.off
chǐ yú zhūhóu
disgrace PREP feudal.lord

‘In the east Your Majesty may take land from Yue, but it will not suffice to wipe off the disgrace; you certainly will have to take land from Qin, and only then will it be enough to wipe off the disgrace among the feudal lords.’

In example (38) the auxiliary verb *yù* 欲 ‘which, want’ follows the modal and the aspecto-temporal adverbs *bì* 必 and *qiě* 且.

- (38) 秦不遺餘力矣，必且欲破趙軍。(Shǐjì: 76; 2371)

Qín bù yí yú lì yǐ, bì qiě yù pò Zhào jūn
Qin NEG leave surplus strength FIN certain FUT wish destroy Zhao army
‘Qin does not save himself from any additional pains, it will certainly wish to destroy the army of Zhao.’

Example (39) represents one of the very infrequent instances in which an aspecto-temporal adverb different from *jiāng* 將 and *qiě* 且 follows the modal adverb *bì* 必, i.e.

the adverb indicating the resultative and completive *yǐ* 已 ‘already’ which, although referring to a completed situation is not excluded from future tense contexts. The aspecto-temporal adverb is followed by the negative marker *bù* 不 which is its default position with regard to aspecto-temporal adverbs.

- (39) 後當復出見之鳳皇騏驎，必已不與前世見出者相似類。(Lùnhéng: 50.24.16)

Hòu dāng fù chū xiàn zhī fènghuáng qílín, bì
afterwards ought again go.out visible SUB phoenix unicorn certain
yǐ bù yǔ qián shì xiàn chū zhě xiāng sìlèi
already NEG with before generation visible go.out NOM mutually resemble
‘A phoenix or a unicorn which will become visible later do certainly not
already resemble those which became visible in former generations.’

2.2.4.3 Other than aspecto-temporal adverbs following *bì* 必

In this section, quantifiers and manner adverbs following *bì* 必 will be discussed. Quantifiers following *bì* 必 have already been alluded to in the section on *bì* 必 in relation to negative markers. In the combination [NEG + *bì* 必 + Quantifier] *bì* 必 has been analyzed as a verb, and consequently the same analysis for *bì* 必 will be assumed for affirmative sentences with a quantifier following *bì* 必.

- (40) 太子焉用孔惺？雖殺之，必或繼之。(Shǐjì: 37; 1601)

Tàizǐ yān yòng Kǒng Kuī? Suī shā zhī,
crown.prince how use Kong Kui? Even.if kill OBJ
bì huò jì zhī
certain someone continue OBJ
‘How can the crown prince make use of Kong Kui? Even if he were to kill him, it is certain that someone would replace him.’

- (41) 馮將以秦韓之兵東卻齊宋，馮因擣三國之兵，乘屈丐之弊，南割於楚，故地必盡得之矣。(Shǐjì: 46; 1896)

Píng jiāng yǐ Qín Hán zhī bīng dōng què Qí Sòng, Píng yīn tuán
Ping FUT with Qin Han SUB army east attack Qi Song Ping then bundle
sān guó zhī bīng, chéng Qū Gài zhī bì, nán gē yú Chǔ,
three state SUB army utilize Qu Gai SUB ruin south cut.off PREP Chu
gù dì bì jìn dé zhī yǐ
therefore land certain complete get OBJ FIN
‘... With the armies of Qin and Han Ping will attack Qi and Song in the East,

and then Ping will unite the armies of the Three Kingdoms and, profiting from the ruin of Qu Gai, cut land off from Chu in the South, and regarding the land, it is certain that he will get it completely.’

In the following example *bì* 必 is followed by the manner adverb *dà* 大 ‘great’ which does not block the analysis of *bì* 必 as an adverb, since modal adverbs regularly precede manner adverbs.

- (42) 今王已絕於齊而責欺於秦，是吾合秦齊之交而來天下之兵也，國必大傷矣。(Shiji: 40; 1724)

Jīn wáng yǐ jué yú Qí ér zé qī yú Qín,
now king already break.off PREP Qi CON blame deception PREP Qin
shì wú hé Qín Qí zhī jiāo ér lái tiān xià
this I join Qin Qi SUB relation CON come heaven below
zhī bīng yě, guó bì dà shāng yǐ
SUB soldier FIN state certain great damage FIN
‘Now, Your Majesty has already broken off with Qi, and You blame Qin for its deception, this means that, if we enter into good relations with Qin and Qi and cause the armies of the empire to come, then the state will certainly suffer great damage.’

2.2.4.4 Aspecto-temporal adverbs preceding *bì* 必

The following is one of the infrequent examples with an aspecto-temporal adverb preceding *bì* 必.²¹ In most of the instances in which the order *jiāng bì* 將必 or *qiě bì* 且必 is attested, both *jiāng* 將 and *qiě* 且 may not be analyzed as aspecto-temporal adverbs, but a different analysis, i.e. the nominal analysis for *jiàng* 將 and the analysis as a conjunctive adverb for *qiě* 且, has to be assumed. However, even if evidence for this order is quite sparse, it evidently exists and provides additional arguments for the existence of an (auxiliary) modal verb *bì* 必 in the Classical and the Han period literature.

- (43) 主君之子且必有代。(Shiji: 43; 1788)

Zhǔ jūn zhī zǐ qiě bì yǒu Dai
Ruler prince SUB son FUT must have Dai
‘The son of Your Rulership must have Dai in the future.’

²¹ Unger (1992:48, 52) presents a few examples for *jiāng bì* 將必 in the Classical literature.

3. Conclusion

As the examples have demonstrated the modal employments of *bì* 必 whether as a modal auxiliary verb or as a modal adverb are clearly closely related to its employment as an independent verb indicating certainty. Syntactically it differs considerably from many of the other modal adverbs and evidently shows constraints which argue for its status as an auxiliary verb. The arguments for an analysis of *bì* 必 as an auxiliary verb or an adverb will be summarized in the following according to the tests applied.

- Auxiliary verbs are characterized by the following syntactic constraints:
 - a) *Wh*-words precede auxiliary verbs.
 - b) The relativizer *suǒ* 所 precedes an auxiliary verb.
 - c) Negative markers (other than *wèi* 未) precede auxiliary verbs.
 - c') The negative marker *wèi* 未 precedes auxiliary verbs.
 - d) Aspecto-temporal adverbs (usually) precede auxiliary verbs.
 - e) Quantifiers precede or follow auxiliary verbs.
- Modal (and aspecto-temporal) adverbs are characterized by the following syntactic constraints:
 - a) *Wh*-words by default follow modal and aspecto-temporal adverbs.
 - b) The relativizer *suǒ* 所 usually does not precede a modal adverb, but it can precede aspecto-temporal adverbs.
 - c) Negative markers other than *wèi* 未 usually follow modal and aspecto-temporal adverbs (in rhetorically un-marked sentences).
 - c') The negative marker *wèi* 未 by default precedes the aspecto-temporal adverb *cháng* 嘗 and its synonyms.
 - d) Aspecto-temporal adverbs follow modal adverbs.
 - e) Quantifiers follow modal adverbs and precede aspecto-temporal adverbs.

As the following two tables demonstrate, there is conclusive evidence for an analysis of *bì* 必 as both an auxiliary verb and a modal adverb. Although the tests provide more arguments for the verbal analysis, its syntactic constraints with regard to aspecto-temporal adverbs and — to a certain extent to negative markers — conclusively argue for an additional analysis of *bì* 必 as a modal adverb. Consequently, two different, but etymologically related words both written as *bì* 必 have to be assumed in the Han period literature according to the syntactic constraints *bì* 必 is subject to; one appears as a modal verb, employed independently or as auxiliary verb ‘must, need to’, and the other appears as a modal adverb ‘certainly’. However, it has to be conceded that fuzzy

cases remain in which — because of the lack of additional syntactic evidence — no clear syntactic distinction can be made between auxiliary verb and adverb.

Table 1: Arguments for *bì* 必 as a modal auxiliary verb

	<i>Bì</i> 必 precedes	<i>Bì</i> 必 follows
Precedes <i>wh</i>-words	+	
Follows <i>suǒ</i> 所		+
Follows NEG <i>bù</i> 不, <i>wú</i> 無, <i>fēi</i> 非		+
Follows NEG <i>wèi</i> 未		+
Follows aspecto-temporal adverbs²²		(+)
Precedes or follows quantifiers	+	

Table 2: Arguments for *bì* 必 as an adverb

	<i>Bì</i> 必 precedes	<i>Bì</i> 必 follows
Follows <i>wh</i>-words		–
Precedes <i>suǒ</i> 所	–	
Precedes NEG <i>bù</i> 不, <i>wú</i> 無, <i>fēi</i> 非	+	
Precedes NEG <i>wèi</i> 未	(+)	
Precedes aspecto-temporal adverbs	+	
Precedes or follows quantifiers	+	

Although the data provides some evidence for a precise analysis of the syntactic status of *bì* 必 it simultaneously confirms the difficulties in distinguishing precisely between adverbs and auxiliary verbs in (Han period) Chinese in general. This is partly due to the fact that the precise picture of the syntax and the semantics of adjuncts in Han period Chinese has not yet been achieved, but is also due to the lack of any morphological and only a few syntactic distinctions between auxiliary verbs and adverbs in general, which both occupy a position to the left of the verb. Nevertheless, the following order of syntactic elements, which still has to be confirmed by more evidence, can be established with regard to adverbs and modal auxiliary verbs with *bì* 必 being permitted in both the position of ADV_{mod} and the position of AUX :

$$[ADV_{mod} (= bì) \textit{ suo} \textit{ NEG}_{asp} \textit{ ADV}_{asp} \textit{ WH-Op} \textit{ NEG} \textit{ AUX} (= bì) \textit{ V}].$$

²² Brackets indicate that instances for this syntactic position are extremely infrequent.

References

- Alleton, Viviane. 1972. *Les adverbos en chinois moderne*. Paris: Mouton.
- Chinese Academy of Social Sciences, Institute of Linguistics. 1999. *Gudai Hanyu Xuci Cidian* [Dictionary of Function Words in Ancient Chinese]. Beijing: The Commercial Press.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford & New York: Oxford University Press.
- Dobson, W. A. C. H. 1959. *Late Archaic Chinese*. Toronto: University of Toronto Press.
- Dobson, W. A. C. H. 1964. *Late Han Chinese*. Toronto: University of Toronto Press.
- Gassmann, Robert H. 1997. *Grundstrukturen der antikchinesischen Syntax*. Bern & New York: Peter Lang.
- Harbsmeier, Christoph. 1981. *Aspects of Classical Chinese Syntax*. London & Malmö: Curzon Press.
- Li, Renzhi. 2004. *Modality in English and Chinese: A Typological Perspective*. Boca Raton: Dissertation.com.
- Liu, Jian. 1960. Lun zhudongci [On auxiliary]. *Zhongguo Yuwen* [Studies of the Chinese Language] 1960.1:1-4.
- Liu, Li. 2000. *Xianqin Hanyu Zhudongci Yanjiu* [A Study of Auxiliary Verbs in Pre-Qin Chinese]. Beijing: Beijing Normal University Press.
- Meisterernst, Barbara. 2004. The future tense in Classical and Han period Chinese. *Meaning and Form: Essays in Pre-Modern Chinese Grammar*, ed. by Ken'ichi Takashima & Shaoyu Jiang, 121-146. München: Lincom Europa.
- Meisterernst, Barbara. 2008a. The negative *wei* 未 in Han period Chinese. *Journal of Chinese Linguistics* 36.1:121-154.
- Meisterernst, Barbara. 2008b. Modal verbs in Han period Chinese, Part I: the syntax and semantics of *kě* 可 and *kěyǐ* 可以. *Cahiers de Linguistique – Asie Orientale* 37.1:85-120.
- Meisterernst, Barbara. 2008c. Negative markers in combination with the modal auxiliary verbs *kě* 可 and *kěyǐ* 可以. *Cahiers de Linguistique – Asie Orientale* 37.2:197-222.
- Meisterernst, Barbara. 2009. *Temporality in the Shiji*. Manuscript.
- Peyraube, Alain. 1999. The modal auxiliaries of possibility in Classical Chinese. *Selected Papers from the Fifth International Conference on Chinese Linguistics*, ed. by H. Samuel Wang, Feng-fu Tsao & Chinfa Lien, 27-52. Taipei: Crane.
- Peyraube, Alain. 2001. On the modal auxiliaries of volition in Classical Chinese. *Sinitic Grammar: Synchronic and Diachronic Perspectives*, ed. by Hilary Chappell, 172-188. Oxford & New York: Oxford University Press.

- Pulleyblank, Edwin G. 1991. *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese and Early Mandarin*. Vancouver: UBC Press.
- Pulleyblank, Edwin G. 1995. *Outline of Classical Chinese Grammar*. Vancouver: UBC Press.
- Shadick, Harold Ernest, and Chien Ch'iao. 1968. *A First Course in Literary Chinese*. Ithaca: Cornell University Press.
- Unger, Ulrich. 1987-97. *Grammatik des Klassischen Chinesisch*. Münster (unpublished).
- Yang, Shu-da. 1928[1959]. *Ci Quan [Annotation of Words]*. Taipei: The Commercial Press.

說“定知”： 古漢語中的一種長距離副詞移位*

李 明

中國社會科學院

“定知”是魏晉南北朝至唐五代時期的一個常用語詞。通過考察魏晉南北朝至唐五代時期的文獻，我們發現：“定知”的形成，多是由於“定”從“知”後的小句長距離移位至“知”前而形成的。由於頻繁使用，“定知”在唐代可能成詞。成詞之後，在人們的理解中，“定”或“知”可能失去原有的意義。

關鍵詞：定知，長距離移位

1. 引言

“定知”是魏晉南北朝至唐五代時期的一個常用語詞。以前的研究主要有以下三種看法：

(一) 蔣禮鴻(1997:527-535)花了很大篇幅，論證了唐五代“知”可以用作“語助詞，沒有意義”。摘引幾例如下：

- (1) 王今伐吳，定知自損。(變文，《伍子胥變文》)
- (2) 王若用宰彼(懿)之言，吳國定知除喪。(同上)
- (3) 雖知世尊是親兄弟，且不肯出家。(變文，《難陀出家緣起》)
- (4) 慈父已生於天上，終朝快樂逍遙；母身墮在阿鼻，日日惟知受苦。(變文，《目連緣起》)

例(1)(2)兩例爲“定知”，例(3)(4)“知”一附於連詞“雖”後，一附於副詞“惟”後。

* 本文初稿曾提交“漢語歷史詞彙與語義演變研討會”(杭州，2008年8月)，承董志翹、方一新、高列過、何亞南、沈家煊、松江崇等先生提出批評意見，謹此致謝。

(二) 江藍生 (1998:105) 認為蔣先生所舉的例子中，“有的可能確實是沒有什麼意義的語助詞，但其中也有不少難以用語助詞解釋”。江先生認為例 (1) 一類的“知”表示“依情理可能出現某種情況”，是個表示將然的助動詞，相當於現代漢語中“植物無水會乾死”的“會”。“定知自損”義為“必將會使自己受到損害”。她的證據是：“知”、“解”、“會”在“理解、領會、知曉”義上是同義詞，它們都可以引申為表示可能出現某種情況。¹

(三) 王雲路 (1999:151-152) 舉到六朝詩歌中“定知”的例子，摘引兩例如下：

- (5) 今年喜夫婿，新拜羽林郎。定知劉碧玉，偷嫁汝南王。（《先秦漢魏晉南北朝詩·北周詩》卷二，庾信《結客少年場行》）
- (6) 千年與昨日，一種併成塵。定知今世土，還是昔時人。焉能取他骨，復持埋我身。（同上，卷六，釋亡名《五盛陰》）

她認為：“‘定知’謂料想，有揣測之義。也有的可釋為‘才知道’。‘定’本身即有料想的意思。”同時王雲路先生還指出：“定知”一詞在唐詩中習見。

“定知”到底是什麼意思？是怎樣形成的？本文擬就此進行專門的考察。同時，本文還將就“定知”所牽涉到的句法問題嘗試著進行分析。

2. “定”的移位

通過考察魏晉南北朝至唐五代時期的文獻，我們發現：“定知”的形成，多是由於“定”從“知”後的小句移位到了“知”前。比較：

- (7) 王即自知定死不久。（元魏·吉迦夜共曇曜譯《雜寶藏經》4/484c）²
- (8) 爾時王子婆奴面上色相出現，於七日內必當命終。……時辟支佛，教授威儀，過六日已，至其七日，定知命終。（隋·闍那崛多譯《佛本行集經》3/931a）
- (9) 相師占雨，有五因緣，不可定知，使占者迷惑。云何為五？一者雲有雷電，占謂當雨，以火太多故，燒雲不雨。是為占師初迷惑緣。……（後秦·佛陀耶舍共竺佛念譯《長阿含經》1/137a）

¹ 類似的分析還可參看江藍生 (1993[2000])。

² “4/484c”分別代表《大正藏》的冊數、頁碼、欄次。下引譯經材料同。

例(7)爲“知定……”，“定”未移位；例(8)、(9)都爲“定知”，但例(8)“定知命終”其實就是“知定命終”的意思，這時“定”從包孕小句移位到了主句動詞之前，而例(9)義爲“有五種因緣不可確知”，“定”本來就生成於“知”前，而非移位而來。³

再看“必知”的例子：

- (10) 時婆羅門聞女所說，知必賢能。（元魏·慧覺等譯《賢愚經》4/399b）
- (11) 復有釋言：視其形相，不驚不畏，以此推之，必知是實。（元魏·吉迦夜共曇曜譯《雜寶藏經》4/497b）
- (12) 王告彼人：吾誦馬相，聞馬母聲，必知其駒好惡。（姚秦·竺佛念《出曜經》4/713a）

這三例可分別同上三例類比：例(10)爲“知必……”，“必”未移位；例(11)、(12)都爲“必知”，但例(11)“必知是實”就是“知必是實”，“必”有移位，而例(12)“必”本來就在“知”前，沒有移位。

再比較以下的例子：

“必定”：

- (13) 此菩薩瑞相，悉同過去佛，以是知菩薩，必定成正覺。（劉宋·求那跋陀羅譯《過去現在因果經》3/639c）[“必定”未移位]
- (14) 我昔既聞阿私陀說，及眾相師，并諸奇瑞，必定知汝不樂處世。（同上，3/632b）[“必定”移位]

“決定”：

- (15) 仙人既見太子智慧，又觀志意堅固不虧，知決定成一切種智。（同上，3/638b）[“決定”未移位]
- (16) 我觀大王太子諸相皆得其所，又極明顯，是以決定知成正覺。（同上，3/627c）[“決定”移位]

³ 例(8)“定知命終”如果理解爲“時辟支佛，教授威儀，過六日已，至其七日，肯定知道（王子）將命終”，也就是把“定”視爲本來就在“知”前，則上下文變得不太順暢，因爲這樣理解則“定”顯示敘述者的判斷，句子好像是說敘述者斷定辟支佛知道王子將命終。在這裡，敘述者判斷的加入顯得非常突兀。從上下文看，顯然是“辟支佛知道王子定將命終”而不是“辟支佛肯定知道王子將命終”的意思。

“決”：

- (17) 又復象馬諸畜生，并諸幢鬘等來至，星速急疾向菩薩，決知當作佛世尊。又復一切淨居天，持其清淨莊嚴體，曲躬頂禮於仁者，知仁決作佛世尊。（隋·闍那崛多譯《佛本行集經》3/774b）[前一個“決”移位，後一個未移位]

從以上的分析，可以得出以下三個結論：一、有兩個“定知”，其中一個“定”有移位，另一個“定”沒有移位。二、“定”的移位並不是強制的，它可移可不移。三、不僅僅是“定”可以非強制移位，表必然義的副詞都可以移位。

再看一組例子：

- (18) 若人愚癡，不信三寶，誹謗毀道，應墮畜生。為病所困，唯得伏臥，不得偃側，不喜善言。左右定知此人必死。（元魏·慧覺等譯《賢愚經》4/378c）
- (19) 昔摩羅國有一剎利，得病極重，必知定死。（蕭齊·求那毘地譯《百喻經》4/551b）
- (20) 爾時頻毘娑羅王及諸大眾，聞優樓頻螺迦葉說此偈言，心大歡喜，於如來所深生敬信，決定得知如來必成於一切種智，審知迦葉是佛弟子。（劉宋·求那跋陀羅譯《過去現在因果經》3/651a）

這三個句子，“定”、“必”、“決定”移位之後，原來的位置上又分別出現了“必”、“定”、“必”，類似於話題化之後又用代詞復指。

以上用的是翻譯佛經的材料，翻譯佛經不一定體現真實的語言面貌。下面以中土文獻來映證：

- (21) 今年喜夫婿，新拜羽林郎。定知劉碧玉，偷嫁汝南王。（《先秦漢魏晉南北朝詩·北周詩》卷2，庾信《結客少年場行》）[同例(5)]
- (22) 千年與昨日，一種併成塵。定知今世土，還是昔時人。焉能取他骨，復持埋我身。（同上，卷6，釋亡名《五盛陰》）[同例(6)]
- (23) 步搖金翠玉搔頭，傾國傾城勝莫愁。若逞仙姿遊洛浦，定知神女謝風流。（《全唐詩》，卷317，武元衡《贈佳人》）
- (24) 寒樹雪晴紅豔吐，遠山雲曉翠光來。定知別後無多日，海柳江花次第開。（同上，卷536，許渾《題陸侍御林亭》）

以上 4 例“定知”，都可以理解爲是“知 [(NP) 定 VP]”中的“定”移位到了“知”之前。其意義其實就是：知道劉碧玉一定會偷嫁給汝南王，知道今世土肯定還是昔時人化成的，知道神女定會遜色不少，知道別後不久海柳江花定會依次開放。下面是兩例沒有移位的例子：

- (25) 江村搖落暮蟬鳴，執手臨歧動別情。古岸相看殘照在，片帆難駐好風生。牽吟一路逢山色，醒睡長汀對月明。況是清朝至公在，預知喬木定遷鶯。（《全唐詩》卷 748，李中《送夏侯秀才》）
- (26) 與爾情不淺，忘筌已得魚。玉臺挂寶鏡，持此意何如？坦腹東床下，由來志氣疏。遙知向前路，擲果定盈車。（同上，卷 175，李白《送族弟凝之滁求婚崔氏》）

再看“必知”的例子：

- (27) 聞錢自華知舉，自知必不中第，遂東歸。（《太平廣記》，卷 278，皇甫弘）[“必”未移位]
- (28) 人知必死，郡縣無以應命。（同上，卷 269，宋昱、韋儼）[“必”未移位]
- (29) 龍朔二年，司列少常伯楊思玄恃外戚貴，待選流多不以禮而排斥之。為選者夏侯彪所訟，御史中丞郎餘慶彈奏免官。時中書令許敬宗曰：“必知楊吏部之敗。”或問之，對曰：“一彪一狼，共著一羊，不敗何待！”（同上，卷 185，楊思玄）[移位]
- (30) 臣自日日沙堤上，見丞相所乘，皆良馬也。是以必知通馬經。（同上，卷 250，黃幡綽）[移位]
- (31) 青蚨似蟬而狀稍大，其味辛，可食。每生子，必依草葉，大如蠶子。人將子歸，其母亦飛來，不以近遠，其母必知處。（同上，卷 477，青蚨）[未移位]
- (32) 開成中，有龍復本者無目，善聽聲揣骨，每言休咎，無不必中。凡有象簡竹筭，以手捻之，必知官祿年壽。（同上，卷 224，龍復本）[未移位]

通過以上的說明，基本可以斷定：“定知”的形成，一個來源是“定”的移位。下面重新審視例 (1)、(2)：

- (1) 王今伐吳，定知自損。（變文，《伍子胥變文》）
(2) 王若用宰彼（詬）之言，吳國定知除喪。（同上）

“定知自損”源於“知定自損”，“知”既不必理解為助詞，也不必理解為表將然的“會”。那麼，“吳國定知除喪”該如何理解？我們認為這樣的結構是由“知吳國定除喪”經過兩次移位形成的：

第一步：“定”移位

“知吳國定除喪” > “定知吳國除喪”

第二步：子句主語話題化

“定知吳國除喪” > “吳國定知除喪”

上面舉的很多“定知”的例子，都只經歷第一步，比如例(21)-(24)。⁴ 第二步在古今漢語中都較常見，例如：

- (33) 我保證他沒事 > 他我保證沒事
(34) 我要修於佛果，汝須速上天宮；莫將諸女獻陳，我家當知不受。（變文，《維摩詰經講經文(五)》）

上例“我家當知不受”即“當知我家不受”。

例(1)“定知自損”一類句子，由於子句主語省略，既可以補足為“定知王自損”，也可以補足為“王定知自損”。

“定知”所涉及到的移位，可以與下面的例子類比：

- (35) 必知菩薩，問得維摩，二空之理既同，七辯之詞不異。（變文，《維摩詰經講經文(四)》）
(36) 朕兵已到江南，宋之君臣必知恐懼。（元·陶宗儀《南村輟耕錄》卷一，“獨松關”條）[引自蔣禮鴻(1997:534)]

例(35)“必”移位，但子句主語沒有話題化；例(36)“必”移位，同時子句主語話題化。

⁴ 這裡分析“吳國定知除喪”為：

(i) “知吳國定除喪” > “定知吳國除喪” > “吳國定知除喪”

有沒有可能是下面(ii)的方式？

(ii) “知吳國定除喪” > “吳國知定除喪” > “吳國定知除喪”

(i) 中“定”的移位先於子句主語的話題化，(ii) 中“定”的移位後於子句主語的話題化。但從現有的材料來看，(i) 比 (ii) 的可能性要大得多，因為我們還沒有發現“知定”的實例。

3. “定知”的成詞

上一小節對“定知”的分析可能還存在問題。

“定知”在唐代可能已經成詞，有兩個證據：

(一) 頻率統計顯示，唐·道世撰集的《法苑珠林》中，“定知”有 12 次，“必知”1 次，“必定知”0 次，“決定知”1 次，“決知”0 次。《全唐詩》中，“定知”出現了 60 餘例。

(二) “定知”不僅出現於詩歌之中，如上引例 (23) (24)，還進入了當時的口語：

- (37) 周靜樂縣主，河內王懿宗妹。懿妹短醜，武氏最長，時號大哥。縣主與則天並馬行，命元一詠。曰：“馬帶桃花錦，裙銜綠草羅。定知幃帽底，儀容似大哥。”則天大笑，縣主極慙。（唐·張鷟《朝野僉載》卷 4，引自《太平廣記》卷 254，張元一）⁵

此例“定知”出現於張元一戲謔靜樂縣主的打油詩。這首打油詩中的“底”，一般視為結構助詞“底”（“的”的前身）的較早例證（“幃帽底”即“戴幃帽的”之義），同時詩中出現了“大哥”這個稱呼，可證這首詩用的是當時的“大白話”。

如果“定知”在唐代或更早已經成詞，那麼，我們再看下面的例子：

- (38) 王若用宰彼（詬）之言，吳國定知除喪。（變文，《伍子胥變文》）
[同例 (2)]
- (39) 腰帶定知今瘦小，衣衫難作遠裁縫。（《全唐詩》，卷 416，元稹《酬樂天得稹所寄紵絲布、白輕庸製成衣服以詩報之》）

根據上文的分析，這樣的例子涉及兩次移位：先是“定”移位，然後子句主語移位。可是在實際運用、習得過程中，很難想像人們（包括說者和聽者）會有這樣繁雜的操作過程。這樣的例子可能並不涉及移位的問題：人們可能就把“定知”當作一個表必然的副詞，其中的“知”無義，也就是蔣禮鴻先生所說的語助詞。

可是在某些結構中，需要作完全不同的理解。請看例句：

⁵ 此引《太平廣記》據中華書局 1961 年排印本，實際“懿妹短醜”應作“懿宗短醜”，否則上下文難通。

- (40) 若人愚癡，不信三寶，誹謗毀道，應墮畜生。為病所困，唯得伏臥，不得偃側，不喜善言。左右定知此人必死。（元魏·慧覺等譯《賢愚經》4/378c）[同例(18)]
- (41) 汝今若其不肯出家，我定知汝命終之後，必墮惡道。（隋·闍那崛多譯《佛本行集經》3/930c）
- (42) 車匿既見太子剃除鬚髮身著法服，定知太子必不可回。（劉宋·求那跋陀羅譯《過去現在因果經》3/634b）
- (43) 若是汝之祖父已來所有衣者，應當解著，云何顛倒？用上為下，以下為上，以不解故。定知汝衣必是偷得，非汝舊物。（唐·道世撰《法苑珠林》53/699b）
- (44) 如今小兒新長成，明年聞道又徵兵。定知此別必零落，不及相隨同死生。（《全唐詩》，197卷，張謂《代北州老翁答》）
- (45) 由是定知必有舍利，乃聚眾掘之。（《太平廣記》，卷379，劉薛）

對於以上例子，上面分析例(18)-(20)時曾說：可以認為是“定”移位後再用“必”填補。可是，在這種“知”前後各有一表必然的副詞的格式中，“定知……必”佔絕對優勢。這促使我們懷疑：也可能人們已經把“定知”當作了一個動詞，“定”在這裡沒有意義，因此在後面又用一個表必然的副詞。如果這種分析成立，這時，“定”也不再涉及移位的問題。

總之，“定”從子句移位上來，這是一種非常特殊的句法現象（詳見下文）。換言之：“定”出現在了不該出現的位置上。由此造成的“定知”這個極其“彥扭”的組合，可能給人們的理解造成困難。同類現象還有“必知、必定知、決定知、決知”等，但只有“定知”偶然性地被挑選出來，高頻使用，逐漸成詞。在成詞過程之中，由於“定”、“知”本不是直接成分，難以捏合在一起，因此它的組合意義，有三種可能：

（一）人們可能仍然知道“定”、“知”本不是直接成分，“定”是從子句移位而來。這時“定”、“知”都有意義。如例(23)(24)。

（二）“定”有義，“知”無義，例(38)(39)的情況可能如此。這種現象可比方作“一葉障目”：在“定知劉碧玉，偷嫁汝南王”（即例(5)、(21)）這樣的句子裡，“定”移位上來後，像一片樹葉遮擋了人們的視線，使“知”的動詞義變得晦暗。一旦子句主語再移位上來，變為主句的話題，形成“吳國定知除喪”（即例(2)、(38)）這樣的句式，“知”前有了雙重障礙，“知”就更容易被理解為語助詞。

(三) “定”無義，“知”有義，例(40)-(45)的情況可能如此。此時“定”“形同虛設”。

現在我們重新審視開頭所提到的對於“定知”的三種觀點：

(一) 蔣禮鴻先生認為“定知”中的“知”是語助詞，部分情況下可能是成立的。

(二) 江藍生先生認為“定知”中的“知”相當於表將然的“會”。我們認為可能不是這樣。“解”、“會”發展為表示將然的“會”，源於表示技能，而“知”表示知識，意義很泛，並不是專用於表示技能；⁶“知”可能沒有“解”、“會”那樣的詞義變化。

(三) 王雲路先生認為“‘定知’謂料想，有揣測之義”，這是把“定知”理解為動詞義，部分情況下也是成立的。

綜上所述，一些“定知”的實例最初涉及“定”的移位。“定知”在成詞之後，就不再涉及移位的問題。但“定知”到底在什麼時候成詞，難以確知。成詞之後，在人們的理解中，“定”或“知”可能失去原有的意義。

4. 其他意義的副詞的移位

上面談到的是必然義副詞“定、必、必定、決定、決”等的移位，其他意義的副詞有沒有這種情況？我們發現表“終究、確實”義的副詞也有這種情況，不過例子不多。

首先看“審”、“信”：

(46) 爾時頻毘娑羅王及諸大眾，聞優樓頻螺迦葉說此偈言，心大歡喜，於如來所深生敬信，決定得知如來必成於一切種智，審知迦葉是佛弟子。（劉宋·求那跋陀羅譯《過去現在因果經》，3/651a）[同例(20)]

(47) 令其勿犯北門，竟爾自惑，信知仙界不可妄至也。（《太平廣記》，卷36，李清）

(48) 於是歎異，信知聖教不虛，遂加崇信焉。（同上，卷99，李大安）

⁶ “知”表示技能的例子不是沒有，但少見。例如江藍生(1998:105)所引《世說新語》的一例：
小兒知談，卿可與語。（《世說新語·排調》）

又如：

能作白酒，而不知作麴釀。（《三國志·魏志·烏丸鮮卑東夷傳》注引《魏書》）

再看變文中的幾例“定知”：

- (49) 朕得舞筵（筵）之內，忽占面色憂文。定知與我相離，所以適來惆悵。（變文，《歡喜國王緣》）
- (50) 忽地夫人氣色昏，淚流如線莫能勝。定知玉貌終歸土，爭忍夫人化作塵。（同上）
- (51) 每想夫人辭家出，夜夜尋看房臥路。玉貌定知歸那裡？且喜恩沾說修持。（同上）
- (52) 兒與娘娘今日別，定知相見在何年？（變文，《大目乾連冥間救母變文》）

這幾例“定”其實是“終究、究竟”義。⁷ 例(49)義為：知道終究與我相離。“定”移位到了“知”前。例(50)義為：知道夫人終究歸於泥土。這時子句還有一個與“定”同義的“終”，這個“定知”更易單理解為“知”，而“定”不表義（參看上文對例(40)-(45)的分析）。例(51)義為：夫人究竟去了哪裡？與例(50)相比，此例子句主語已移位作主句的話題，“定知”更易單理解為“定”，而“知”不表義（參看上文對例(38)、(39)的分析）。例(52)義為：究竟相見在何年？這個“定知”似乎也是“定”有義，而“知”已無義。

5. 其他動詞

前面的內容都涉及動詞“知”，那麼，其他動詞有沒有這種現象？我們發現“恐”也存在這種現象。請看例句：

定恐：

- (53) 爾時淨飯王聞是語已，如是思惟：大仙私陀言不虛妄，定恐太子舍家出家，我今更可增益五欲，令其染著，勿使出家。（隋·闍那崛多譯《佛本行集經》3/725a）

必恐：

- (54) 吾鉢已受持數十年，借汝必恐損之。（《太平廣記》，卷 94，華嚴和尚）

⁷ “定”至遲從六朝起，就有“終究、究竟”義。比如：

鄧艾口吃，語稱“艾艾”。晉文王戲之曰：“卿云‘艾艾’，定是幾艾？”對曰：“‘鳳兮鳳兮’，故是一鳳。”（劉宋·劉義慶《世說新語·言語》）

- (55) 梁武帝多策事。因有貢徑寸栗者，帝與沈約策栗事：帝得十餘事，約得九事。及約出，人問今日何不勝。約曰：“此人忌前，不讓必恐羞死。”（同上，卷 197，沈約）

多恐：

- (56) 謾排酒饌應難吃，久坐時多恐損人。（變文，《捉季布傳文》）

將恐：

- (57) 今主君德薄，不足聽之。聽之，將恐有敗。（《韓非子·十過》）
(58) 房稽首曰：將恐今之視古，亦猶後之視今也。（《世說新語·規箴》）
(59) 我今無嗣，種姓將恐斷絕。（《祖堂集》卷一，第七釋迦牟尼佛）

終恐：

- (60) 今既見兄，勳不負諾矣，終恐此靈物亦非兄所有。（《太平廣記》，卷 230，王度）
(61) 願嫁得前時柳生足矣。必不允，某與外兄終恐不生全。（同上，卷 342，華州參軍）

以上例子，都應當理解為副詞從“恐”後包孕小句中移位到“恐”之前，比如例(53)“定恐太子舍家出家”實為“恐怕太子一定會舍家出家”。餘此類推。⁸

6. 進一步的句法分析

本文上面談到的副詞的移位，具體地說：

- (一) 它是從一個非論元位置移到另一個非論元位置，因此是 A' 移位 (A'-movements)。
(二) 它移出了包孕子句，是長距離移位。

⁸ 即使把這些“恐”理解為表示推測的副詞，也同樣存在移位的問題。比如“定恐”是“定”從“恐”後移位到了“恐”前，因為只能說“恐怕一定（會下雨）”而不能說“一定恐怕（會下雨）”。只不過這時候，移位是局部移位，而不再是移出小句的長距離移位。

這種長距離的移位對主句動詞、移位的副詞有什麼限制呢？

關於副詞長距離移位對主句動詞的限制，一個可能是敘實性 (factivity) 與非敘實性 (non-factivity) 的問題。根據 Kiparsky & Kiparsky (1970)，只有非敘實動詞才能進行否定提升 (neg-raising) (詳細討論見下文)。否定提升涉及否定詞的長距離移位，所以可以設想：副詞的長距離移位是否也與敘實與非敘實的問題有關？

“知”是敘實 (factive) 動詞，涉及副詞的長距離移位；但是，“恐”是非敘實 (non-factive) 動詞，同樣可以有這種現象。

另外，同是非敘實動詞的“想”，卻沒有這種現象發生。我們沒有見到移位上來的“想必”的例子，只有“想……必”、“想必”：

(62) 想君小時，必當了了！（劉宋·劉義慶《世說新語·言語》）

(63) 正禮元子，致有志操，想必有以殊異。（《三國志·吳書·劉繇傳》）

“想必”後來還成為詞；“知”就沒有成詞為“知定”而是成詞為“定知”。

由上可知，本文提到的這些副詞的移位，跟敘實、非敘實的區分沒有關係。這其中的限制，由於牽涉到的動詞太少，一下子看不清楚。

對於移位的副詞，涉及到的有表必然的“定、必、必定、決定、決、多”，表終究、確實義的“審、信、定、終”，表將來的“將”。可以看出，它們都是句子副詞 (sentential adverbs)，如果我們接受漢語副詞為附加語 (adjuncts) 的觀點（參看 Tang 2001），那麼，可以長距離移位的副詞，就只限於 IP 的附加語，而不能是 VP 的附加語，比如表示方式的副詞就沒有發現可以進行長距離移位的實例。

另外，還有一個現象值得注意。先請看例句：

(64) 江村搖落暮蟬鳴，執手臨歧動別情。古岸相看殘照在，片帆難駐好風生。牽吟一路逢山色，醒睡長汀對月明。況是清朝至公在，預知喬木定遷鶯。（《全唐詩》卷 748，李中《送夏侯秀才》）[同例 (25)]

(65) 與爾情不淺，忘筌已得魚。玉台挂寶鏡，持此意何如。坦腹東床下，由來志氣疏。遙知向前路，擲果定盈車。（同上，卷 175，李白《送族弟凝之滁求婚崔氏》）[同例 (26)]

(66) 少年輕會復輕離，老大關心總是悲。強說前程聊自慰，未知攜手定何時。公才屈指登黃閣，匪服胡顏上赤墀。想到長安誦佳句，滿朝誰不念瓊枝。（同上，卷 271，竇叔向《酬李袁州嘉祐》）

這三例都是“定”未移位的例子。值得注意的是，這三例“知”前已分別有表時間的副詞“預”、“遙”以及否定詞“未”。

以上談到的兩點現象讓我們推測：“定”等的移位，可能是因為它們作為句子副詞，與子句主要動詞的關係相對比較鬆散，因此，在一定情況下，它可以受主句動詞的吸附，移位至主句動詞之前，形成一個雙音節的組合，組成一個在韻律上非常穩定的單位。這樣的推測有以下兩個佐證：一、主句動詞前如已有副詞與之構成雙音節的組合，如上三例，“定”等不再移位；二、“必定”、“決定”等本已是雙音節的副詞雖也有移位成“必定知”、“決定知”的實例，但非常少見。馮勝利 (2009) 甚至把韻律視為一種形態，指出漢語的韻律在構詞、促發句法移位等方面的作用。本文討論的副詞移位現象，可能主要也是由韻律引發的。

7. 古漢語中的否定提升

否定提升例如：

- (67) a. John doesn't think Peter will come.
- b. John thinks Peter will not come.

上兩例意思相同，都義為“John 認為 Peter 不會來”，但 (67a) 否定詞由子句提升到了主句。

否定提升與本文所說副詞的移位有相似之處，但否定詞與一般的副詞還是不同。否定詞到底是附加語 (adjuncts) 還是自成一個功能語類，存在爭議。因此本文分開討論。

何莫邪 (Harbsmeier 1981:§1.2) 談到古漢語中“使”、“令”存在否定提升的例子：

- (68) 為國家者，見惡如農夫之務去草焉，芟夷蕞崇之，絕其本根，勿使能殖，則善者信矣。（《左傳·隱公六年》）
- (69) 天生民而立之君，使司牧之，勿使失性。有君而為之貳，使師保之，勿使過度。（《左傳·襄公十四年》）
- (70) 且比化者無使土親膚，于人心獨無悛乎？（《孟子·公孫丑下》）

上三例並不是祈使句，卻用了通常認為是表示禁止的“勿”、“無”，一個合理的解釋就是：“勿”、“無”是從後面的子句移位上來的；在“使”後的子句中，由於不是直陳事實，所以否定要用“勿”、“毋（無）”而不是“不”、“弗”（詳參呂叔湘 1942[1995], 1942-1944[1990:241-242]）。又如：

- (71) 寡君不忍，使群臣請於大國，無令與師淹於君地。（《左傳·成公二年》）

此例明顯宜理解為“令輿師無淹於君地”。

類似的現象，還出現於元明時代的“勿請”一詞（參看袁賓 1990:20）：

- (72) 邵曰：“吾思已決，勿請驚疑！”（《清平山堂話本·死生交范張雞黍》）
(73) 來日宋江梯己已聊備小酌，對面論心一會，勿請推卻。（《水滸傳，62回》）

“勿請”都應理解為“請勿”。

同樣，變文中有“不請”一語，項楚(1984)釋為“請勿、不必”。例如：

- (74) 直須認取速行行，不請無端戀意情。（變文，《維摩詰經講經文(三)》）
(75) 莫生憂慮，不請疑猜；師兄便望收留，弟子今朝佈施。（變文，《維摩詰經講經文(五)》）

我們推測“不請”也是由於否定詞的提升而形成的。

8. 結論

本文討論“定知”一語，推及副詞及否定詞的長距離移位問題。主要結論如下：

一、大多數“定知”其實是由於“定”從子句移位而形成的。由於頻繁使用，“定知”在唐代可能成詞。成詞之後，在人們的理解中，“定”或“知”可能失去原有的意義。

二、涉及這種長距離移位現象的主句動詞除“知”外，還有“恐”。由於牽涉到的動詞太少，其中的限制不是很清楚。發生移位的副詞都是句子副詞，移位可能是由於韻律引發的。

三、古漢語中同時存在過少量的否定提升現象。

引用文獻

- Feng, Shengli (馮勝利). 2009. Lun Hanyu yunlü de xingtai gongneng yu jufa yanbian de lishi fenqi 論漢語韻律的形態功能與句法演變的歷史分期 [On morphological function of prosody and the chronologies of syntactic changes in Chinese]. *Lishi Yuyanxue Yanjiu* 歷史語言學研究 [*Historical Linguistics Study*] 2:11-31. Beijing: The Commercial Press.
- Harbsmeier, Christoph (何莫邪). 1981. *Aspects of Classical Chinese Syntax*. London: Curzon Press.
- Hu, Jianhua (胡建華). 2007. Fouding, jiaodian yu xiayu 否定、焦點與轄域 [Negation, focus, and scope]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2007.2:99-112.
- Jiang, Lansheng (江藍生). 1993[2000]. Xiangguan yuci de leitong yinshen 相關語詞的類同引申 [Extended meaning found in related words]. *Jindai Hanyu Tanyuan* 近代漢語探源 [*Searches for the Origin of Modern Chinese: Papers by Jiang Lansheng*], 309-319. Beijing: The Commercial Press.
- Jiang, Lansheng (江藍生). 1998. Yanyifa yu Jindai Hanyu ciyu kaoshi 演繹法與近代漢語詞語考釋 [Deductive method in textual research of Early Modern Chinese words]. *Yuyanxue Luncong* 語言學論叢 [*Essays on Linguistics*] 20:99-107. Beijing: The Commercial Press.
- Jiang, Lihong (蔣禮鴻). 1997. *Dunhuang Bianwen Ziyi Tongshi* 敦煌變文字義通釋 [*A Comprehensive Glossary of the Special Terms Found in the Genre of Dunhuang Bianwen*]. Shanghai: Shanghai Ancient Books Publishing House.
- Kiparsky, Paul, and Carol Kiparsky. 1970. Fact. *Progress in Linguistics: A Collection of Papers*, ed. by Manfred Bierwisch & Karl Erich Heidolph, 143-173. The Hague: Mouton.
- Lü, Shuxiang (呂叔湘). 1942[1995]. Lun wú yu wù 論毋與勿 [Differences between wú and wù]. *Hanyu Yufa Lunwenji* 漢語語法論文集 [*Collected Papers of Chinese Grammar*], 73-102. Beijing: The Commercial Press.
- Lü, Shuxiang (呂叔湘). 1942-1944[1990]. *Zhongguo Wenfa Yaolue* 中國文法要略 [*Brief Introduction to Chinese Grammar*]. Beijing: The Commercial Press.

- Tang, Chih-Chen Jane. 2001. Functional projections and adverbial expressions in Chinese. *Language and Linguistics* 2.2:203-241.
- Wang, Yunlu (王雲路). 1999. *Liuchao Shige Yuci Yanjiu* 六朝詩歌語詞研究 [*Research on Terms Found in Poems of the Six Dynasties*]. Haerbin: Heilongjiang Education Press.
- Xiang, Chu (項楚). 1984. Bianwen ziyi ling shi 變文字義零拾 [Selected words from Dunhuang manuscripts]. *Zhuming Zhongnian Yuyanxuejia Zixuanji: Xiang Chu Juan* 著名中年語言學家自選集・項楚卷 [*Selected Papers by Xiang Chu*], 2-43. Hefei: Anhui Education Press.
- Yuan, Bin (袁賓). 1990. *Chanzong Zhuzuo Ciyu Huishi* 禪宗著作詞語彙釋 [*Compilation and Explanations of the Colloquial Terms Found in Zen Writings*]. Nanjing: Jiangsu Ancient Books Publishing House.

On *Dingzhi* ‘know certainly’: A Type of Long Distance Movement in Medieval Chinese

Ming Li

Chinese Academy of Social Sciences

The word *dingzhi* ‘know certainly’ was commonly used from the Wei, the Jin and the Northern and Southern Dynasties until the Tang and the Five Dynasties. Based on documents of that period, it can be hypothesized that the formation of *dingzhi* ‘know certainly’ comes from the long distance movement of *ding* from a small clause following *zhi*. Due to its frequent occurrence, it became a lexical word during the Tang Dynasty but once it was formed, the meanings of *ding* and *zhi* may have disappeared.

Key words: *ding zhi* ‘know certainly’, long distance movement

中古以來表患病類詞語的搭配組合

方一新 王雲路

浙江大學

詞彙是一個系統。本文以中古近代漢語時期表示疾病、治療和病愈的詞語組合為例，從系統的角度，考察這幾類詞語的組合關係和習慣表達方式，探討此類詞語的發展過程。

關鍵詞：中古漢語，詞彙史，患病，組合

1. 詞語的搭配與組合

漢語史幾個相關領域研究中，詞彙史相對薄弱。美國歷史語言學家馬爾基耶爾教授在回答北京大學徐通鏘教授關於“在過去，歷史語言學集中研究語音的發展，很少研究詞彙和語法，這是什麼原因”問題時說：“首先，音位的數目在典型的語言中是很有限的，只有 20 個，或 30 個，或 35 個，或 40 個，很少超過這個數目。所以語音學的研究容易駕馭。……至於詞彙，你要對付的或許也是一種系統，但它包含上萬個單位。要把這堆龐大的材料組織到一個緊湊的系統中去是很困難的。”“詞彙單位太多，看起來沒有嚴密的結構，就是說，一種變化不一定觸發別的變化，如此等等，預見性比較弱。”¹

詞彙是一個系統，這已經為人們所公認。但要確定其系統性是比較困難的。新詞與舊詞，口語詞、俗語詞和文言雅詞，實詞與虛詞，單音詞與複音詞，諸如此類的詞彙成分、詞彙組織中間究竟具有何種關聯，有何異同，龐大的詞彙組織如何構成一個系統，都是具有挑戰性的研究課題。如果我們從詞族聚合的角度逐類理清，或許是看清詞彙組織構成的一個方法。本文擬以中古近代漢語時期表示疾病、治療和病愈的詞語組合為例，探討此類詞語的發展過程。

關於詞語如何表示概念，即關於詞語的搭配、組合，國內外學者都有一些論述，如：

¹ 見《語言學論叢》第 13 輯，210-211 頁。

克魯斯 (D. Alan Cruse) 說：

“搭配 (collocation) 這個名稱用來指習慣上共同出現的詞項，每個詞項都是一個語義要素，因此整個語義是完全明朗的。”

“如果搭配的一個成分的意義要求在直接語境中有一個特定的詞項，那麼語義的粘著尤為緊密（所有成分都固定的情況似乎並不出現）。”²

蔣紹愚 (1989:49-50) 指出：

“詞的組合關係，簡單地說就是詞的搭配關係。詞的組合關係的歷史變化，主要表現在以下兩個方面：

- 1) 同一個詞，詞義基本不變，但在不同的歷史時期組合關係有所不同。
- 2) 在某一時期的語言平面中，既有繼承前一時期的舊詞和舊語法成分，又有新產生的新詞和新語法成分，新舊的組合關係往往不同。”

這裡探討一組漢語（列舉漢語部分）表示疾病、治療和病愈的詞語搭配的實例並酌作討論。

2. 古代表示患病的詞語搭配

在中國古代，生產力低下，生活和醫療水準不高，人們容易感染疾病。故在古代詞語的搭配和組合中，表示感染疾病類的很多。

張詒三 (2005:1-4) 曾對“病”表示疾病義的搭配情況作了考察，指出：以“病”為中心語素而產生的詞語搭配，古今漢語有別。先秦時期，單音節詞語相對少，可以和“病”搭配的動詞主要有“有”、“遇”、“為”；至魏晉南北朝，和“病”搭配的動詞大大增加，除了保留“遇”外，新增了“臥”、“被”、“發”、“得”、“生”、“致”、“抱”、“中”、“寢”、“嬰”、“起”、“成”、“滯”、“結”等，而現代漢語中可以和“病”搭配

² 見 D. A. Cruse 《詞彙語義學》第二章〈詞彙單位的橫組合定界〉第九節“成語和搭配”。

的動詞主要有“有”、“生”、“患”、“染”、“得”、“犯”等，有了明顯的變化。其中，“有”、“生”二詞和“病”的搭配比較穩定，自先秦、六朝一直到現代都有例子。但魏晉南北朝的“被病”、“遇病”、“中病”、“抱病”、“寢病”、“嬰病”、“滯病”、“起病”、“成病”、“結病”等搭配在現代漢語中已經不用；現代漢語的“染病”、“患病”、“犯病”等搭配在先秦、六朝典籍中未見用例。³

根據我們的初步考察，張說大體正確，但還有可補充之處。六朝時期，還有其他能和“病”結合的單音詞，如“感”、“患”、“邁”、“染”、“害”等。

通常說“感病”或“感……”。《南齊書·張欣泰傳》：“元徽中，興世在家，擁雍州還資，見錢三千萬。蒼梧王自領人劫之，一夜垂盡。興世憂懼感病卒。”（881頁）《南史·褚澄傳》：“豫章王感病，高帝召澄為療，立愈。”（756頁）南朝梁寶唱撰《比丘尼傳》卷二《南皮張國寺普照尼傳》：“宋元嘉十八年十二月，因感勞疾，雖劇，而篤情深信，初自不改。”（50/938c）還有“感氣病”、“感疾”、“感心疾”、“感腳疾”等說法，用例很多。也有“感患”同義連文者，指患病，染病。《高僧傳》卷十三《釋僧護》：“有建安殿下感患未瘳，若能治剡縣僧護所造石像得成就者，必獲平豫。”（491頁）⁴

近代漢語中“感（…）疾”一類的組合仍習見。如：宋邵伯溫《邵氏聞見錄》卷二〇：“熙寧十年夏，康節先生感微疾，氣日益耗，神日益明。”（221頁）莊綽《雞肋編》卷上：“長孫順德喪息女，感疾甚，唐太宗薄之。”（15頁）羅大經《鶴林玉露》丙編卷五“范雲”條引《南史》范雲事有“感寒疾”一語（見該書320頁）。《張孔目智勘魔合羅》第一折：“（正末云）‘這那裡便有賊？老的，我如今感了風寒，一臥不起，只望老的你便寄個信與俺渾家，教他來看我。’”“感了風寒”，即患了感冒。《老殘遊記》第十五回：“誰知這個女婿去年七月感了時氣，到了八月半邊，就一命嗚呼哀哉死了。”“感了時氣”，即得了傳染病。

中古以來未見“患病”之例，但“患”常與帶有染上疾患的身體部位和所患疾病的名稱結合。有“患氣”，指罹患呼吸道疾病，《全晉文》卷二五王羲之《雜帖》：“昨得熙廿六日書，云患氣，懸情。”“患面”，指面部得了癰腫，《全晉文》卷二七王獻之《雜帖》：“患面疹腫，脚中更急痛，兼少下。”“患眼”，指眼睛得病，南朝梁元帝蕭繹《金樓子》卷二《后妃篇》：“繹始學弱

³ 參看張詒三《詞語搭配變化研究：以隋唐若干動詞與名詞的搭配變化為例》，齊魯書社，2005。

⁴ 參看方一新《東漢魏晉南北朝史書詞語箋釋》43-45頁，黃山書社，1997。

年，患眼之始，衣不解帶，冬則不近炎火，夏則不敢風涼。”有“患嗽”、“患手”，分別指得了咳嗽病和手上染疾，唐張文成《游仙窟》：“十娘曰：‘兒近來患嗽，聲音不徹。’下官答曰：‘僕近來患手，筆墨未調。’”（《近代漢語語法資料匯編·唐五代卷》，6頁）此外還有類似組合：稗海本《搜神記》卷八“司勳張員外”條：“司勳忽患舌腫，須臾出於口外，其大如斗。”（《搜神後記》附，110頁）五代靜、筠《祖堂集》卷二《慧可禪師》：“及至禮師，不稱姓名，云：‘弟子身患風疾，請和尚爲弟子懺悔。’”（唐五代卷，451頁）宋孫光憲《北夢瑣言》卷一四：“忽患腳瘡，痛不可忍。”（109頁）“患舌腫”、“患風疾”、“患腳瘡”和“患嗽”相似，都是指出所得的病症和部位。

“患”加病名，組成“患～”，是動賓結構，如：有“患利”，指瀉肚。《隋書·王勇傳》：“我爲患利，不脫衣臥。”也可作“利患”，則屬於偏正結構，謂拉肚子病。《宋書·謝莊傳》：“利患數年，遂成痼疾，吸吸惛惛，常如行屍。”是其例。近代以來有“患疾”，泛指生病，是動賓結構。⁵《警世通言》四〇卷：“經由故國，知主上患疾，特來顧之。”《東周列國志》八三回：“（趙）鞅偶患疾，使無恤代將以往。”也說“疾患”。《清平山堂話本·楊溫攔路虎傳》：“不被他打得疾患，也得你不識李貴。”（宋代卷，423頁）這裡似乎可以看做動詞並列。這種“AB”“BA”式均可的構詞法值得注意。

漢魏六朝以來，又有由“遘”+“疾、病”組成的詞語。⁶

有“遘疾”，漢張衡《周天大象賦》：“或含丹而舉兵，或隕黑而遘疾。”《三國志·蜀志·諸葛亮傳》：“事臨垂克，遘疾隕喪，朕用傷悼，肝心若裂。”《南齊書·高帝紀下》：“遘疾彌留，至於大漸。”

有“遘病”，《庾子山集》卷一五《周驃騎大將軍開府侯莫陳道生墓志銘》：“擐甲遘病，死於轅門，春秋五十一。”唐張鷟《朝野僉載》卷二：“後於並州營幢子，未成遘病，臨終歎曰：‘夫生死者，人之大分，如來尙所未免，但功德未成，以此爲恨耳！’”《太平廣記》卷八四“王居士”條（出《關史》）：“果有延壽坊鬻金銀珠玉者，女歲十五，遘病甚危，眾醫拱手不能措。”比較而言，“遘疾”使用早且用例多，而“遘病”使用晚（南北朝始用）且用例較少，未詳其故。⁷

⁵ “患疾”亦有疾病義，如《前漢書平話》卷中：“何將通衣服拽住：‘大夫，你早來念甚來？你何患疾？’”

⁶ 此承友生真大成博士提示，謹致謝忱。

⁷ “詬病”與“遘病”音同，容易混淆，或許是一個原因。

唐宋以來，說“染疾”。如宋孫光憲《北夢瑣言》卷一二：“柳玘出官瀘州郡，泊牽復，沿路染疾，至東川通泉縣求醫。”（95頁）

也說“害……”。《敦煌變文校注·燕子賦》：“雀兒被額，更害氣噴，把得問頭，特地更悶。”（378頁）“噴”，甲卷作“咽”。“更害氣咽”，加上患氣短咽塞（的毛病）。⁸稗海本《搜神記》卷二“王子珍”條：“眾共尋，乃見白雞在架牆上而坐，害左眼。珍見，思此物是我父之冤家也。……瞎左眼者，所射中也。”（82頁）“害左眼”謂左眼有病，和“患左眼”略同，也就是下文所說的“瞎左眼”。《古今小說·宋四公大鬧禁魂張》：“侯興一個兒子，十來歲，叫做伴哥，發脾寒，害在床上。”（宋代卷，491頁）元喬吉《兩世姻緣》第二折：“兒嚶，你害的是甚的病，怎麼這等憔悴了？”

這些詞語的選擇有什麼特點呢？上古的“有”、“遇”、“爲”是泛指動詞，可用於許多方面，說明當時並沒有指疾病感染的專門詞語。詞彙量少，則詞義兼職必然是普遍現象。

中古以來除了保留“遇”外，新增了“臥”、“被”、“發”、“得”、“生”、“致”、“抱”、“中”、“寢”、“嬰”、“起”、“戍”、“滯”、“結”、“感”、“患”、“邁”、“染”、“害”等。

這些動詞大致可以分爲這樣幾類：(1) 泛指單音節動詞與疾病名稱結合，“致”、“發”、“得”、“生”等即是。(2) 與疾病相關的單音節動詞，如表示遭遇義的“被”、“邁”等；表示纏繞不去的“嬰”、“結”、“滯”等；表示生病狀態的“寢”、“臥”等。(3) 後來逐步成爲專職表示疾病類的動詞，如“染”、“害”、“感”、“患”等。這些動詞數量多，說明當時詞彙相當豐富。另外，“患”使用頻率高，兼有動詞和名詞兩種用法，也是一個值得注意的現象。

現代漢語中可以和“病”搭配的動詞主要有“有”、“生”、“患”、“染”、“得”、“犯”、“感”等，其間有些差別。如“有”、“生”二詞和“病”的搭配比較穩定，細分“有病”表示已有的、持續的狀態；“生病”表示產生、初起狀態。與“有病”類似的是“患病”等，與“生病”類似的是“染病”、“得病”等；“感”則多表示相對輕微的、偶爾的疾病，故通常多稱“感疾”而極少稱“感病”，也表示外來的病源，如“感風寒”等；“染”表示受外

⁸ 黃征、張涌泉《敦煌變文校注》謂“噴”即“賁”之“增旁字”，“‘氣賁’爲病名。”又作“奔”。舉《諸病源候論》卷十三《賁豚氣候》“夫賁豚氣者，腎之積氣，起於驚恐憂思所生”、同卷《奔氣候》“……故氣奔急也”等例，參看《敦煌變文校注》406頁。中華書局，1997。

來的傳染，如“染病”；“犯病”則表示疾病重新發作。

關於病情發作，中古時有專門的動詞，稱“發”、“動”、或“發動”、“動發”：

動。《三國志·蜀志·蔣琬傳》：“會舊疾連動，未時得行。”（1059頁）

發。《後漢書·方術傳下·華佗》：“佗曰：‘此病後三朞當發，遇良醫可救。’”（2738頁）《世說新語·言語》第15則劉注引嵇紹《趙至叙》：“自痛棄親遠遊，母亡不見，吐血發病，服未竟而亡。”（41頁）

動發。《雜阿含經》卷十五：“云何良醫善知治病已，於未來世永不動發。謂良醫善治種種病，令究竟除，於未來世永不復起。是名良醫善知治病。更不動發。”《全唐文》卷四五〇齊映《爲蕭復讓宰相表》：“今逆臣殲夷，庶物忻暢，餘生獲泰，宿疾頓來。豈無諸醫？竟未有效；忪悸之疾，動發無時。”

發動。《抱朴子內篇·雜應》：“其尚盛者，則生諸疾病，先有疹患者，則令發動。”（271頁）《後漢書·方術傳下·華佗》：“乃戒之曰：‘後十八歲，疾當發動，若不得此藥，不可差也。’”（2739頁）

宋元以來稱“舉發”，如：《元典章·刑部四》：“委因舊患心風病證舉發，昏迷不省，不知怎生將喬老打死。”又《刑部十六》：“訊瘡舉發，才方保放。”

現代漢語則稱“發”、“發作”、“暴發”，等等。如“發病”、“疾病發作”、“當地痢疾暴發”等，滿足了與單音詞搭配、與雙音詞搭配或表示群體性疾病發作的不同需要。

這些動詞雖然是同義或近義關係，但還是有細微差別的，可以表達疾病狀態、來源或程度的差別，也可以滿足不同結構類型與音節韻律搭配的需要，這是漢語詞彙豐富的原因之一。

3. 中古以來表示身體不適的詞語搭配

詞語的形成、發展與人們的日常生活緊密相關。受社會整體發展水準的限制，我國古代醫藥並不發達，各種疾病嚴重威脅著人們的身體健康，有關疾病、治療、痊癒、去世等詞彙十分豐富，像二王法帖中，談論身體狀況和疾病的內容比比皆是。與此相應，也出現了很多表示身體狀況欠佳的詞語，均比較委婉。如中古用“不安”、“不和”、“不佳”、“不快”、“不平”、“不適”、“違和”，近代用“不佳”、“不快”、“不適”、“違和”、“違憂”，現代用“不適”、“欠安”、“欠爽”等。

不安。《漢書·王褒傳》：“其後太子體不安，苦忽忽善忘，不樂。”《三國志·魏志·文昭甄皇后》裴注引《魏書》：“時武宣皇后體小不安，後不得定省，憂怖。”（160頁）

不和。《全晉文》卷二五王羲之《雜帖》：“李母猶小小不和，馳情。伏想行平康。”《晉書·愍懷太子傳》：“十二月，賈后將廢太子，詐稱上不和，呼太子入朝。”（1459頁）

不佳。《北堂書鈔》卷一四四引《郭林宗別傳》：“林宗嘗不佳，夜間命作粥。”《全晉文》卷二六王羲之《雜帖》：“君頃就轉佳不？僕自秋便不佳，今故不善差。”《搜神後記》卷五：“及悅疾篤，導憂念時至，積日不食。忽見一人，形狀甚偉，被甲持刀。問是何人。曰：‘僕，蔣侯也。公兒不佳，欲為請命，故來爾。公勿復憂。’”

不快。《三國志·魏志·華佗傳》：“又有一士大夫不快，佗云：‘君病深，當破腹取。’”《全晉文》卷二三王羲之《雜帖》：“妹不快，憂勞，餘平安。”（商務印書館本 221 頁）《晉書·愍懷太子傳》：“二十九日早入見國家，須臾遣至中宮。中宮左右陳舞見語：‘中宮旦來吐，不快。’使住空屋中坐。”元關漢卿《竇娥冤》第二折賓白：“‘婆婆，你今日病體如何？’（卜兒云）‘我身子十分不快哩。’”

不平。《漢書·王嘉傳》：“天子以相等皆見上體不平，外內顧望，操持兩心。”《魏書·術藝·徐謩傳》：“體小不平，及所寵馮昭儀有疾，皆令處治。”

不適。《陸雲集》補遺《春節帖》：“三月十六日雲白：春節余不適，得示，知足下平安。”

違和。南朝梁沈約《齊禪林寺尼淨秀行狀》：“又於一時復違和，亦甚危困，忽舉兩手，狀如捧物，語旁人不解。問言：‘為何所捧？’答云：‘見寶塔從地出，意欲接之。幡花伎樂，無非所有。’於是疾恙豁然而除，都無復患。”

《大唐大慈恩寺三藏法師傳》卷七：“帝發京時，雖少違和，而神威睿慮，無減平昔。”五代靜、筠《祖堂集》卷六《洞山和尚》：“問：‘四大違和，還有不病者也無？’”（唐五代卷，511 頁）又卷十四《江西馬祖》：“院主問：‘和尚四體違和，近日如何？’”（同上，543 頁）《歐陽修集》卷一五〇《書簡》卷七：“見兒子，言尊候違和，豈非患腹臟邪？秋後，慎生冷為佳。”《五燈會元》卷十二《文公楊億居士》：“公因微恙，問環大師曰：‘某今日忽違和，大師慈悲，如何醫療？’”

以上爲中古常見用法。與之相同，近代表示身體不適，通常也採用委婉的說法，有“不佳”、“不快”、“不適”、“違和”等。

沿用中古詞語的如：

不佳：宋司馬光《涑水記聞》卷九：“扑見吏衣紫，誤以爲醫官，因引手案上，謂曰：‘扑數日來體中不佳，君試爲診之。’聞者傳以爲笑。”宋趙與峕《賓退錄》卷六：“朱希真《夢記》略云：‘紹興戊寅除夜，體中不佳，三更方得睡。’”

不快：《黃庭內景經》肺部章第九：“素錦衣裳黃雲帶，喘息呼吸體不快。”《喻世明言》卷三：“吳山道：‘我身子不快，不要點心。’”

不適：《續資治通鑒》卷九三：“攸別居賜第，一日，詣京，甫入，遽起，握父手爲診視狀，曰：‘大人脈勢舒緩，體中得毋有不適乎？’京曰：‘無之。’”《閱微草堂筆記》卷十九《灤陽續錄一》：“嘗見一人服松脂十餘年，肌膚充溢，精神強固，自以爲得力。然久而覺腹中小不適，又久而病燥結，潤以麻仁之類，不應。”

違和⁹：五代靜、筠《祖堂集》卷六《洞山和尚》：“問：‘四大違和，還有不病者也無？’”（唐五代卷，511 頁）又卷十四《江西馬祖》：“院主問：‘和尚四體違和，近日如何？’”（同上，543 頁）《五燈會元》卷一二《廣慧璉禪師法嗣·文公楊億居士》：“公因微恙，問環大師曰：‘某今日忽違和。大師慈悲，如何醫療？’”（727 頁）明楊爵《慰人心以隆治道疏》：“數年以來，因聖體違和，朝儀間缺，經筵未講。”

南北朝、唐宋以後產生的新的表達形式如：

違憂。唐韓愈《鳴雁》詩：“違憂懷息性匪他，凌風一舉君謂何？”錢仲聯《韓愈詩集釋》引陳景雲說：“違憂懷息，即有病求息義。”

欠安。《二刻拍案驚奇》卷三：“翰林道：‘昨日到宅，渴想妹子芳容一見，見說玉體欠安，不敢驚動。’”明無名氏《霞箋記·父子傷情》：“下官前日獲一小恙，身體欠安。”《明珠緣》第四七回：“不意聖躬欠安，客巴巴傳出信來，叫忠賢親往問安。見聖躬日漸清腹。”《紅樓夢》第二六回：“總是我沒造化，偏又遇著叔叔欠安。”

欠爽。本指無精打采，不爽快。《海上花列傳》第四八回：“子富獨在房中，覺得精神欠爽，意欲吸口鴉片烟。”也指身體欠佳、不適。《玉蟾記》第三三回：“彩鸞說：‘我又怕吃茶了。鈴兒，我精神欠爽，莫不是昨日在園中受些

⁹ “違和”已見於南北朝，如：南朝梁沈約《南齊禪林寺尼淨秀行狀》：“又於一時，復違和，亦甚危困。”

風露麼？你稟知夫人去。’鈴兒到夫人樓上說：‘小姐今日欠安。’夫人說：‘快去請徐先生診視。’”

由以上各例不難看出：表示身體感覺不適的詞語，通常用否定的方法，即由表示舒適義的單音節形容詞¹⁰加上否定副詞“不”，如“不安”、“不和”、“不佳”、“不快”、“不平”、“不適”等。另一種是動詞後加上表示舒適義的單音節形容詞，近代用“欠安”、“欠爽”或“欠佳”，語氣更委婉；“違和”、“違憂”則稍微直接些。而“違憂”是“違和”的類推，意謂“違和而憂慮”，從意義上講並不準確。

從以上例證可以看出，表示生病多為動賓結構，表示發病多用並列式動詞，表示不適多用否定式形容詞，共同點是多為雙音節。

4. 古代表示診脈的詞語搭配

探討詞語的發展，既有詞語如何表示概念，也有概念如何改變名稱的問題。對此，王力《漢語史稿》下冊第六十節已經論及，作者列舉了“腿”、“走”、“跑”、“錯”、“怕”、“偷”等多例詞語，考察這些概念由舊詞向新詞的演變，指出：隨著時代的不同，概念外延的擴大或縮小，也會引起名稱的變化。概念也有由專指（特殊化）向泛指（一般化）演化的趨勢，概念的一般化和特殊化，都和社會的發展有關。避諱和禁忌，是概念變更名稱的原因之一。1984年，王力在《漢語史稿》下冊的基礎上修訂完成了《漢語詞彙史》，對“概念是怎樣變了名稱的”一節作了一些增補。¹¹

有時候，有些概念在不同的時代有不同的說法，表現出有聯繫又有差別的組合關係。前面已經舉述，古代表示生病和身體不適的詞語很多，同理，表示醫病的詞語也相當繁多。古代治病都是中醫，中醫診病有“望、聞、問、切”四種方法，其中“切”指搭脈，是四法中直接和病人接觸的一法，起到最後的判斷作用，十分重要。《史記·扁鵲倉公列傳》：“問臣意：‘診病決死生，能全無失乎？’臣意對曰：‘意治病人，必先切其脈，乃治之。敗逆者不可治，其順者乃治之。心不精脈，所期死生視可治，時時失之，臣意不能全也。’”（2817頁）

圍繞著“切”（搭脈）這一診病方法，在漢語發展史上有不同的表達方法。以下以表示搭脈診療的複合詞“～脈”的組合關係為例。

¹⁰ 具體參見下文。

¹¹ 《漢語詞彙史》，商務印書館，1993。又收入《王力文集》第十一卷，山東教育出版社，1990。

先秦以來有“切脈”。《史記·扁鵲倉公列傳》：“越人之爲方也，不待切脈、望色、聽聲、寫形，言病之所在。”“切脈”後代也說，如：唐劉禹錫《因論·鑒藥》：“切脈、觀色、聆聲，參合而後言曰：‘子之病，其興居之節舛。’”宋孫光憲《北夢瑣言》卷一〇：“醫者意也，古人有不因切脈，隨知病源者，必愈之矣。”（79頁）

有“診脈”。《史記·扁鵲倉公列傳》：“扁鵲以其言飲藥三十日，視見垣一方人。以此視病，盡見五藏癥結，特以診脈爲名耳。（2785頁）又：“齊王太后病，召臣意入診脈。”（同上，2801頁）《世說新語·術解》：“（殷）浩感其至性，遂令舁來，爲診脈處方。”（383頁）唐宋以後也見：如宋邵雍《邵氏聞見錄》卷一七：“士人之妻孕，診其脈曰：‘六脈皆絕，反用子氣資養，故未死。子生，母即死矣。’已而果然。”（184頁）明葉權《賢博編》：“飾一妓爲女子，使一人爲之父，若農莊人，棹小船載魚肉酒果，俟無人，投寺中，乞僧爲女診脈，歷說病源，故爲痴態。”（7頁）清西周生《醒世姻緣傳》第四回：“只是有件毛病不好：往人家去，未曾看病，先要吃酒，掇了個酒杯，再也不肯進去診脈。”清崔象川《玉蟾記》第三三回：“小姐梳洗已畢，徐先生上樓來，請過夫人安，就替小姐診脈，說：‘小姐微有感冒。’”¹²

有“案脈”。《漢書·王嘉傳》：“前東平王云與后謁祝詛朕，使侍醫伍宏等內侍案脈，幾危社稷，殆莫甚焉！”（3492頁）《三國志·魏志·華佗傳》：“李將軍妻病甚，呼佗視脈，曰：‘傷娠而胎不去。’將軍言：‘聞實傷娠，胎已去矣。’佗曰：‘案脈，胎未去也。’”（802頁）

有“相脈”。《張家山漢墓竹簡·脈書》：“相脈之道，左□□□□案（按）之，右手直踝而簪之。”《馬王堆漢墓帛書·脈法》：“相脈（脈）[之道]，左□□□走而案（按）之，右[手直踝而簪之]。”¹³

有“視脈”。《三國志·魏志·華佗傳》：“故甘陵相夫人有娠六月，腹痛不安，佗視脈，曰：‘胎已死矣。’”（799頁）《三國志·魏志·華佗傳》：“故督郵頓子獻得病已差，詣佗視脈，曰：‘尙虛，未得復，勿爲勞事，御內即死。’”（800頁）《宋書·文九王傳·始安王休仁》：“范陽祖翻有醫術，姿貌又美，殷氏有疾，翻入視脈，說之，遂通好。”（1878頁）

有“看脈”。《三國志·魏志·華佗傳》裴松之注引《佗別傳》：“有人病

¹² 《原本老乞大》：“我有些腦痛頭眩，請大醫來診候脈息，看甚麼病。”（外研社本 58頁）“診候脈息”應該就是“診脈”的擴展形式。

¹³ 二例引文據《張家山漢墓竹簡》[二四七號墓]（釋文修訂本）245-246頁（63-65號簡，126頁），文物出版社，2001[2006]。

兩脚蹙不能行，舉詣佗，佗望見云：‘已飽針灸服藥矣，不復須看脈。’”（802頁）《五燈會元》卷二〇《楚安慧方禪師》：“張公會看脈，李公會使藥，兩個競頭醫，一時用不著。”

有“候脈”。《北齊書·方伎傳·馬嗣明》：“邢邵子大寶患傷寒，嗣明爲之診，候脈。”（680頁）宋周密《齊東野語》卷一四《針砭》：“若唐長孫后懷高宗將產，數日不能分娩，詔醫博士李洞玄候脈。”也作“候視脈”，《三國志·魏志·華佗傳》裴注引《佗別傳》：“佗使悉解衣倒懸，令頭去地一二寸，濡布拭身體，令周匝，候視諸脈，盡出五色。”（804頁）

有“驗脈”。唐張讀《宣室志·李生（二）》（《太平廣記》卷二二〇引）：“天寶中，有隴西李生自白衣調選桂州參軍，既至任，以熱病旬餘，覺左乳痛不可忍，及視之，隆若癰腫之狀。即召醫驗其脈，醫者曰：‘臟腑無他。若臆中有物，以喙攻其乳，乳痛，而癰不可爲也。’”

隋唐以後也產生了新的說法：

有“對脈”。《說郛》卷七四引令狐澄《大中遺事》：“唐宮中以診脈爲對脈。”宋周輝《清波別志》卷下：“稼軒在上饒，屬其室病，呼醫對脈。”¹⁴

有“評脈”。《太平廣記》卷四三五“續紳”條（出唐康駢《劇談錄》）：“咸通乾符中，京師醫者續坤頗得秦醫和之術，評脈知吉凶休咎。”（3536頁）《拜月亭》二【梁州第七】白：“郎中仔細的評這脈咱。”

有“准脈”。《新校元刊雜劇三十種·老生兒》一[點絳脣]白：“當日婆婆上席去來，我暗使人喚的個穩婆婆與小梅准脈來。”¹⁵

這些表示切脈的詞語大致有兩類：一是視覺類，如“診脈”、“相脈”、“視脈”、“看脈”、“候脈”；一是觸覺類，如“切脈”、“案脈”、“對脈”、“驗脈”等。（先是觸覺類，如“切脈”、“案脈”等。到了後代，則有兼含視覺與觸覺這二者的說法，如“驗脈”、“對脈”、“評脈”等。）中醫判斷病情有“望、聞、問、切”四種方法，切脈已經表明感官中這兩官是不能分開的了。這些都是同義詞，意義上並無差別，只是文言與白話的區別，或者是不同時代、不同地域的習慣叫法而已。

這些有關切脈的不同說法，或有遺漏，但也足以反映出漢語表達上的豐富多彩。同時，有不少組合只出現在歷史進程中的某個階段，並未保留到今天。在現代漢語普通話中，有關的診病手段說“切脈”、“診脈”；吳方言如紹興話說“看脈”，魯迅《狂人小說》四：“大哥說：‘今天請何先生來，給你診一

¹⁴ 參看《唐五代語言詞典》105頁，上海教育出版社，1997。

¹⁵ 參看《元語言詞典》440頁，上海教育出版社，1998。

診。’我說：‘可以。’其實我豈不知道這老頭子是劊子手扮的！無非借了看脈這名目，揣一揣肥脊。”上海話說“搭脈”、“號脈”；杭州話說“把脈”、“按脈”，其他的說法已經很少用了。¹⁶

5. 中古以來表示身體尚可或病情好轉的一組詞語

東漢末至魏晉南北朝時期政局險惡，戰爭頻繁，加之當時醫學尚不發達，疾病、瘟疫嚴重威脅人們的健康，百姓的生命十分脆弱，反映在文獻特別是書信中有不少詢問身體狀況，企盼健康的詞語。

第一類是報告身體安康或問候。在佛、道著作以及二王法帖中，表示身體好、健康，有“安”、“平安”、“平健”、“平康”、“平平”、“佳”、“佳善”、“可”等眾多詞語，反映了人們渴求健康的心理。

安：《全晉文》卷二二王羲之《雜帖》：“上下安也。和緒過，見之欣然。”（商務本 216 頁）

平安：《全晉文》卷二二王羲之《雜帖》：“餘粗平安，知足下情至。”（商務本 210 頁）《全晉文》卷二六王羲之《雜帖》：“阮生何如？此粗平安，數絕問為慰。”（商務本 253 頁）

平健：《經律異相》卷五引《法句經》卷三：“婦曰：‘若汝立死，食尚得，況今平健，欲望我食！’”（22 頁）

平康：《全晉文》卷二五王羲之《雜帖》：“夫人遂善平康也，足下各可不？”（商務本 244 頁）《高僧傳》卷七《釋僧含》：“後平康無疾，忽告眾辭別。”（276 頁）

平平：《太平經》卷四九：“其次疾病多而不得常平平，忿然往學，可以止之者。……是故吾書教學人，乃以天長壽之法，旦夕自力為之，才得且平平耳。”（161 頁）

佳：《全晉文》卷二二王羲之《雜帖》：“先生適書，亦小小不能佳，大都可耳。”（商務本 211 頁）

佳善：《全晉文》卷二五王羲之《雜帖》：“得告：承長平未佳善，得適適，君如常也。知有患者，耿耿。”（商務本 247 頁）

可：《全晉文》卷二二王羲之《雜帖》：“先生適書，亦小小不能佳，大都

¹⁶ 台灣學者曲守約《續辭釋》在列舉不少切脈詞語後說：“凡此，皆有關診脈之不同辭語，及其真正之意旨也。”可謂精彩。參看曲守約《續辭釋》413 頁，聯經出版事業公司，台北，1982。

可耳。”（商務本 211 頁）《全晉文》卷二五王羲之《雜帖》：“遠婦疾猶爾，其餘可耳。”（商務本 247 頁）

第二類是表示病情好轉等。表示病情經治療有好轉、治愈、康復的詞語也很多，如“平”、“平善”、“平復”、“平完”、“平豫”、“佳”等。

平：《風俗通義·怪神·世間多有見怪驚怖以自傷者》：“（杜）宣意遂解，甚夷憚，由是瘳平。”（328 頁）《三國志·魏志·方技傳朱建平》：“吾所苦漸平，明日雞鳴，年便五十。”（809 頁）《續談助》三引《漢武故事》：“天子至鼎湖，病甚。……上乃令發根禱之，即有應。上體平，遂迎神君會於甘泉，置之壽宮。”（《古小說鈎沉》458 頁）

平復：敦煌漢簡八八 DYTGC：五 A、B：“弟子王習，詣吳師門下，叩頭叩頭，煩勞遠苦，問之，甚厚意，大佳。元小疾未平復，叩頭叩頭，出，謁告弟子，皆如前，甚毋失期。有者，坐之。交故兩師循行，互處之。”¹⁷《道行般若經》卷九《薩陁波倫菩薩品》：“欲哀我者，使我身體平復如故。”《大明度經》卷六《普慈闍士品》：“闍士聞之，心大歡喜，即取刀自刺兩臂，以血與之，復割兩髀肉，又破骨以髓與之。……報言：‘天王哀我者，使身平復。願則如舊瘡愈，身強力氣逾前。’”《大寶積經》卷七九《富樓那會》第十七之三《大悲品》：“深心菩薩於求他利不貪身命，以淨心布施因緣，臂還平復，無有瘡癩。”

平完：《賢愚經》卷九《善事太子入海品》：“婦因自誓：我今一心，共相尊奉，無有他意，大如毛髮，若當實爾，至誠不虛，令汝一目，平完如故。”

平愈：《後漢書·光武帝紀下》李賢注引《東觀（漢）記》：“御坐廡下淺露，中風發疾，苦眩甚。……遂到章陵，起居平愈。”（68 頁）

平豫：《高僧傳》卷十三《釋僧護》：“有建安殿下感患未瘳，若能治剡縣僧護所造石像得成就者，必獲平豫。”（491 頁）

佳：《全晉文》卷二二王羲之《雜帖》：“知部兒不快，何所在？今已佳也？”（商務本 212 頁）又卷二三王羲之《雜帖》：“數得桓公問，疾轉佳也。”（商務本 220 頁）

這些詞語的特點是：或者是單音節形容詞，或者是並列式形容詞，均表達了美好的願望和期盼，使用頻率最高的是“平”，因為平和、平安、平常都是最正常的，也是最寶貴的。由此類單音節形容詞加上否定詞“不”，則表示身體不適，屬於由健康到生病的一個中間階段，例見前。

¹⁷ 據甘肅省文物考古研究所《敦煌漢簡》，中華書局，1991。此例蒙友生路方鵠同學提供線索，附記致謝。

從以上的初步考察可以看出：中國古代表示疾患、不適、治療的詞語很多，表示病情好轉、治愈的詞語也不少；而且隨著時代的發展有所變化，組合成不同的詞族。這反映了古代民眾的生活面貌，更豐富了漢語的詞彙系統，其中有不少詞語一直沿用（或略有變化）到今天，是漢語詞彙的寶貴財富。以上從生病、診病和表示身體狀況等方面分析了大量的同義詞和相關詞語，說明了詞語的構成搭配是很有規律的，無論結構方式還是語素運用，都很有條理，反映了人們共同的認知觀念。

引用文獻

- Chu, Shou-yueh (曲守約). 1982. *Xu Ci Shi 續辭釋* [*The Sequel of Annotation*]. Taipei: Linking Publishing.
- Cruse, D. Alan. 2009. *Lexical Semantics*. Beijing: World Publishing Corporation.
- Fang, Yixin (方一新). 1997. *Donghan Wei Jin Nanbeichao Shishu Ciyu Jianshi 東漢魏晉南北朝史書詞語箋釋* [*The Annotation for Words in History Book of the Eastern Han, Wei, Jin, Southern and Northern Dynasties*]. Hefei: Huangshan Shushe.
- Huang, Zheng (黃征), and Yongquan Zhang (張涌泉). 1997. *Dunhuang Bianwen Jiaozhu 敦煌變文校注* [*The Annotation of the Material of Dunhuang*]. Beijing: Zhonghua Book Company.
- Jiang, Lansheng (江藍生), and Guangshun Cao (曹廣順). 1997. *Tang Wudai Yuyan Cidian 唐五代語言詞典* [*The Lexicon of Language in the Tang and Five Dynasties*]. Shanghai: Shanghai Education Press.
- Jiang, Shaoyu (蔣紹愚). 1989. *Gu Hanyu Cihui Gangyao 古漢語詞彙綱要* [*The Outline of the Ancient Chinese Vocabulary*]. Beijing: Peking University Press.
- Li, Chongxing (李崇興), Shuxian Huang (黃樹先), and Zesui Shao (邵則遂). 1998. *Yuan Yuyan Cidian 元語言詞典* [*The Lexicon of Language in the Yuan Dynasty*]. Shanghai: Shanghai Education Press.
- Qi, Yucun (戚雨村). 1997. *Xiandai Yuyanxue de Tedian he Fazhan Qushi 現代語言學的特點和發展趨勢* [*The Characteristics and Development Trend of Modern Linguistics*]. Shanghai: Shanghai Foreign Language Education Press.
- Vendryes, Joseph (房德里耶斯). 1920[1992]. *Yuyan 語言* [*Le langage: Introduction linguistique a L'histoire*], translated by Qixiang Cen & Feisheng Ye. Beijing: The Commercial Press.

- Wang, Li (王力). 1980. *Hanyu Shi Gao* 漢語史稿 [*An Outline of Chinese Language History*]. Beijing: Zhonghua Book Company.
- Wang, Li (王力). 1990. *Hanyu Cihui Shi* 漢語詞彙史 [*The History of Chinese Vocabulary*]. Ji'nan: Shandong Education Press.
- Xu, Tongqiang (徐通鏘). 1984. Meiguo yuyanxuejia tan lishi yuyanxue 美國語言學家談歷史語言學 [The American linguists talk about historical linguistics]. *Yuyanxue Luncong* 語言學論叢 [*Essays on Linguistics*] 13:210-211. Beijing: The Commercial Press.
- Zhang, Yisan (張詒三). 2005. *Ciyu Dapei Bianhua Yanjiu: Yi Sui Qian Ruogan Dongci yu Mingci de Dapei Bianhua Wei Li* 詞語搭配變化研究：以隋前若干動詞與名詞的搭配變化為例 [*Study of Word-Collocating Varieties*]. Ji'nan: Qilu Shushe.

Combinations of Phrases Relevant to Illness from Medieval Times Onwards

Yixin Fang and Yunlu Wang
Zhejiang University

Vocabulary is a system. This article investigates the combination of relations and customary expression of these phrases which are relevant to disease, treatment and recovering in medieval times and modern times, and explores their process of development.

Key words: Medieval Chinese, history of vocabulary, disease, combination

Contact and Change in the History of the Chinese Language

Hsiao-jung Yu (遇笑容)

University of California, Santa Barbara

This article is a preliminary examination of whether language contact has influenced Chinese language development, and if so, which other languages might have been of relevance. Special attention is given to the period between the third and ninth centuries when large volumes of Buddhist sutras were translated from Sanskrit or Pali into Chinese. During the past few decades, a significant body of literature has been published on the subject of language contact and its influence on language change, but it deals mainly with synchronic studies of language contact and rarely or scarcely touches upon modern Chinese, much less Chinese language development. Chinese linguists long held the view that language change in Chinese has been due only to internal factors. However, external borrowing through language contact could play a crucial role in understanding Chinese language changes, especially given the fact that China has had large-scale language contact with non-Han peoples such as Central Asian monks, Mongols, Manchus, and so on. This study employs data collected from a large volume of documents and texts from three periods in Chinese history: the Medieval Period (3rd-7th c.), the Yuan dynasty (late 13th - early 14th c.) and the Qing dynasty (1644-1911). We isolate and examine specific linguistic features in the documents and texts representative of various time periods, and we discuss the social and historical conditions which gave rise to certain changes as a result of contact with non-Chinese languages, the mechanisms by which these linguistic features emerged in Chinese as well as the nature of the language that was influenced by non-Chinese languages and how it developed over time.

Key words: language contact, Chinese language development, translated Buddhist sutras

1. Introduction

One of the major questions in the study of grammatical development concerns the underlying mechanisms responsible for change — are they due to analogy and reanalysis (including grammaticalization), or are they caused by external factors such as borrowing or interference? Since the 1980s, linguists have increasingly turned to external factors such as language contact and borrowing as underlying forces that can motivate a language

to evolve. Important studies include Thomason & Kaufman (1988) on creolization and pidginization of languages, Harris & Campbell (1995) on historical linguistics, and more recently the work of Heine & Kuteva (2005) on the effect of language contact on grammatical structure. During the past few years, however, scholars have increasingly come to realize that external factors might have played a part in the development of the Chinese language.

This article is a preliminary examination of whether language contact has influenced Chinese language development, and if so, which other languages might have been of relevance. Special attention is given to the period between the third and ninth centuries when large volumes of Buddhist sutras were translated from Sanskrit or Pali into Chinese.

1.1 Language contact in the history of Chinese language

Throughout Chinese history, a number of non-Chinese (or non-Han) ethnic groups with languages belonging to various language families have resided in or occupied the northern Chinese territories. Two groups occupied all of China for significant periods of time. The following is a brief account of the major non-Han groups:¹

- Xiongnu: a Turkish-speaking group, founded the first steppe empire in the third century BC on the grasslands of northern China and were conquered by the Han in the first century BC and first century AD.
- Rouzhi: probably an Indo-European-speaking group, sought by Han China as allies against the Xiongnu in the second century BC.
- Xianbei: a Mongolian-speaking group of people who resided in eastern Mongolia in the third century AD and invaded China in the fourth century.
- Tobgach (Tuoba): largely Mongolian language speakers, founded the Northern Wei dynasty (AD 386-534) in North China.
- Turks (Tujue): a Turkish-speaking group who established an empire in China circa AD 581.
- Khitan (Qidan): a Mongolian-speaking group, founded the Liao dynasty in North China and adjacent areas, which ruled from AD 947 to 1125.
- Jurchen (ancestors of Manchus): a Tungusic language group, founded the Jin dynasty, which ruled North China from AD 1122 to 1234.
- Tanguts: a Tibetan-speaking group, founded the Xixia Kingdom in Northwest China, which lasted from AD 1038 to 1227.

¹ This brief account of the major non-Han groups is primarily based on Fairbank's list (1978:153), with minor revisions.

- Mongols: a Mongolian-speaking group, founded the Yuan dynasty and ruled over all of China from AD 1271 to 1368.
- Manchus: a Manchu-speaking group, founded the Qing dynasty and occupied all of China from AD 1644 to 1911.

Of the non-Han ethnic groups described above, the Khitan, Jurchen and Tanguts occupied areas in the north or northwest of China, while the Mongols and Manchus took over the entire country for 97 years and 267 years, respectively. This prolonged period of foreign contact resulted in extensive cultural convergences and exchanges with these peoples. Further, many integral aspects of Chinese culture (e.g., art, dance and music) emerged under the influence of non-Han cultures. Considering such a context, repeated large-scale linguistic contact throughout the history of China was inevitable, especially given the fact that Mongolian was the official language during the reign of the Yuan dynasty and Manchu was one of the official languages in the imperial court during the Qing dynasty.

However, the most significant and major influence on China in the areas of philosophy, religion, literature, architecture and language did not come from non-Han conquerors such as Mongols and Manchus but through the spread of Buddhism in China during the period from the first to the sixth centuries. This is regarded as one of the major events of Chinese cultural history and the single most important foreign influence on Chinese civilization before the impact of the West in the nineteenth and twentieth centuries. While contact under Mongol control and under the Manchu regime mark two of the three major language contacts in Chinese history, the Buddhist contact signifies the first and probably the most important language contact to influence the syntactic development of the Chinese language.

1.2 Contact through Buddhist works during the Medieval Period (3rd-7th c.)

Buddhist ideologies were imported into China primarily through the large-scale translation of Buddhist sutras from Sanskrit, Pali, or certain Central Asian languages into Chinese. Between the Han dynasty (206 BC-AD 220) and the Tang dynasty (AD 618-960), a vast number of Buddhist texts were translated into Chinese. The first documented translation of Buddhist sutras into Chinese dates from AD 148 as An Shigao, the Parthian prince-missionary, arrived at the Han Dynasty capital of Luoyang, where he set up a center for the translation of Buddhist texts. Kushan Lokaksema, Kumārajīva and several others were the main translators after An Shigao and none of them were native speakers of Chinese. From the late Han to the end of the Tang dynasty, over 170

translators were involved in the translation of 2,278 Buddhist sutras from Sanskrit or Pali into Chinese. Within these translated sutras, there are 7,046 *juan* ('fascicles') containing approximately 56 million characters.²

2. The nature of the contact

Of the three major language contacts introduced above, the Central Asian monks traveled to China in order to spread Buddhist teachings, while the Mongols and Manchus did so in order to control the land and its people. Their motivations varied, but all three groups underwent some process of learning the Chinese language. In the written texts left by these groups of non-native Chinese speakers, there appear certain special, unauthentic linguistic elements or features that had not previously existed in Chinese. These are the results of imperfect Chinese language learning due to interference from the foreigners' mother tongue. The Chinese these people used was essentially an intermediary language, or interlanguage.

2.1 The interlanguage in the translated Chinese Buddhist texts

It is widely accepted that the translated sutras or texts stimulated the production of a new Chinese lexicon. Yet understanding the extent to which these non-Chinese elements led to changes in Chinese syntax requires the careful study of the elements and features borrowed from the original Sanskrit as well as the interlanguage errors of the translators. We will return to this point with more discussion in §3 and §4 below.

2.2 The interlanguage in *Yuan baihua*

In order to foster communication with the Chinese, the Mongols used a mixture of Chinese and Mongolian known as *Yuan baihua*, or "Mongolian Chinese". While some linguists consider this language a pidgin or creole (see Zu 2005, 2007), it actually is more like an interlanguage (see Cao & Chen 2009). The following are two examples:³

² Taisho Tripitaka Vol. T55, No. 2154.

³ These two examples are taken from the paper "The Mongol empire in China and the Mongolian influence on Chinese language" presented by Cao Guangshun in the *Steppes and their suburbs* workshop in France in 2007.

- (1) 賚把行的聖旨與了也。
 lai ba xing de shengzhi yu-liao-ye.
 hold hold carry.out *de* imperial.edicts give-PERF-FP⁴
 ‘(The emperor) gave the imperial edict [to the monk] for him to carry out.’
- (2) 五嶽、五鎮、四海、四瀆咱每的神祈有。
 wuyue, wuzhen, sihai, sidu, zanmei-de shen qi-you.
 five.mounts five.areas four.seas four-lakes our god-LIG pray-AP
 ‘[You, the Taoist practitioners go to] worship the gods of ours [who are in charge of] the Five Mountains, Five Areas, Four Seas and the Four Lakes.’

In both examples, the text is written in Chinese characters and the words are Chinese, but the syntactic structure is typical of *Yuan baihua*; the Mongols had always refused to learn the Chinese language. Perhaps due to the fact that the Mongols did not rule over China for a very long time, *Yuan baihua* became extinct very soon after the Mongols lost their power over the Chinese.

2.3 Manchu, interlanguage and sinicization

The Manchus occupied China for almost 300 years. Although both Manchu and Chinese were official languages of the imperial court, there was a significant shift among the Manchus from being monolingual Manchu speakers (before the 1660s) to bilingual Manchu and Chinese speakers (1660s-1790s) and then to monolingual Chinese speakers (after the 1790s). During the transitional period, a form of interlanguage was used in that the written script and words were Chinese but the word order was Manchu. The *Compilation of the Qing Language* (*Qingwen huishu* 清文匯書), a well-known Manchu-Chinese dictionary compiled under the Yongzheng reign (1722-1735), has the following Chinese character equivalents for Manchu expressions:

- (3) Generengge akū
 去的沒有
 qu de meiyou
 go-NOM NEG
 ‘No one went.’

⁴ The following abbreviations are used in the examples: AP assertive particle; FP sentence final particle; LIG ligature marker; NEG negation; NOM nominalizer; and PERF perfective aspect.

(4) Ebubumbi

凡從高處使人諸物往下下

fan cong gaochu shi ren zhuwu wangxia xia

all from high.place make people things towards.downward move

‘As long as ordering people to move things down from a high place...’

Though Chinese is a subject-verb-object language, both of these examples consist of Chinese characters and lexical items in the object-verb word order of Manchu. According to the Qing document *Records of the Eight Banners* (*Baqi wenjin* 八旗文經), before and after the bannermen went into Beijing, “All the Manchus, Mongols, and Chinese armies [knew how to] speak the national language (Manchu)... A hundred years later, not everyone [could] speak the national language (Manchu).” When the Manchus first took over the northeast of China in the 1620s, the Manchu national founder Huang Taiji (1559-1626), also known as Nurhaci, created an official position known as “*Qixinlang*” (啓心郎 literally ‘clerk who clarifies thoughts’, meaning ‘interpreter’). Each official bureau had a “*Qixinlang*” to facilitate communication between the Manchus and the Chinese. However, this position had already become superfluous in 1672 when it was abolished by Emperor Kangxi due to the large number of Manchus who had become bilingual within this 56-year period. The number of Manchu speakers continued to decline from then on. Official Qing historiographical accounts record that in 1776, the fortieth year of the Qianlong reign, the emperor was shocked and infuriated by the fact that his Manchu subordinates could not speak their mother tongue,⁵ the decline of the Manchu language was inexorable by this point. During the Guangxu reign (1875-1908), “all of [the Manchus] follow[ed] the Han customs in speaking and writing,” and “there were only one or two out of a thousand Manchus” who could speak the Manchu language (*Hulanfu zhi*, vol. 10). Official documents before 1840 contained both Manchu and Chinese, placed side by side. After 1840, however, documents were written only in Chinese. Even the most powerful woman in the history of the Qing dynasty, the Empress Dowager, knew no Manchu whatsoever.

3. Externally induced influence on the Chinese language—taking examples from Chinese sutras translated during the Medieval Period

There are distinctive linguistic differences between the language in the Chinese sutras translated during the Medieval Period and that in native Chinese texts of the same

⁵ “Manchu rule over China and the attrition of the Manchu language”, paper presented in the “Steppes and their suburbs” Workshop, 2007.

period. These differences fall under two categories: the emergence of unusual, innovative features in Chinese and pre-existing Chinese features used with excessive frequency in translated Buddhist texts.

3.1 Examples of distinctive linguistic features in the language of translated Chinese sutras

(a) The postverbal *yi* 已 denoting completed action:

The translated Buddhist texts contain many examples of the “V (O) + *yi*” construction. In pre-Qin times and the former Han dynasty, the verb in this construction could only be a continuous verb (*chixu dongci* 持續動詞, e.g., *shi* 食 ‘to eat’), such as *shi yi* 食已 ‘have eaten.’ In cases where an instantaneous verb is used (*shunjian dongci* 瞬間動詞, e.g., *si* 死 ‘to die’), to indicate that something had come to an end involved placing *yi* before the verb, such as *yi si* 已死 ‘[already] died’. However this verbal constraint is not respected in the translated sutras. The translators used both types of verbs in the “V (O) + *yi*” construction without recognizing the fact that placing this second type of verb in the pre-existing Chinese construction violates a grammatical rule. This error is a form of interlanguage caused by the lack of such a distinction between the two types of verbs in the language of the translators. This new form of “V (O) + *yi*” construction is frequently found in the Buddhist sutras. The “mistake” of using *yi* with an instantaneous verb (*yi* later developed to *liao* > *le*₁) is an innovation resulting from language contact and the influence of Sanskrit⁶ on Chinese.

(b) The copular construction—“subject, noun *shi*” construction:

This type of copular construction in which a verb is placed in the final position of a sentence is most likely the influence of Sanskrit (see Jiang 2003 and Chen 2003).

(c) The special usage of the interrogative word *yunhe* 云何:

Yunhe, meaning ‘what’, has only one function in traditional Chinese texts: to serve as an interrogative word in a question. This interrogative word is used in Chinese sutras as a translation for the Sanskrit word *kim* ‘what’. In our previous studies (Yu 2003 and 2007), we compared the language in translated Chinese sutras with the language in corresponding Sanskrit texts and found that placing *yunhe* in a superfluous or redundant position at the beginning of a sentence in the translated sutras without it having any function as an interrogative is due to the Sanskrit influence.

⁶ Qiantan ‘Qiren bai wang, fu yi si liao’—Jianlun yuyan jiechu yingxiang Hanyu yufa fazhan de jizhong moshi, paper presented in Guizhou, 2006.

(d) The cause marker *gu* 故 used at the end of sentences:

Standard Chinese texts from the Medieval Period place *gu* in the beginning of the second clause, “..., *gu*...”. In the translated sutras however, *gu* is placed at the end of the sentence, “..., ... *gu*”. This post-positioned *gu* is the result of the Sanskrit influence (see Wang 2004, Yu 2007, 2008).

(e) The disposal construction “*qu* OV” (取 OV):

Chinese is an SVO language. The verbal final word order is very likely to be influenced by the word order of Sanskrit (see Cao & Yu 2000a, 2000b).

(f) The genitive marker *suo/xu* (所/許):

The usage of *suo/xu* in the Chinese translated sutras probably corresponds to the genitive case in Sanskrit (see Zhao 2009).

(g) Number:

In Classical Chinese, the words *deng* (等), *bei* (輩) and *cao* (曹) have various functions. One of which, when placed after a personal noun or personal pronoun, is to denote that ‘there is more than one’ (e.g., *wubei* ‘we’). In Sanskrit, however, numeration is quite complicated. There are singulars, duals, and plurals for nouns, pronouns and adjectives, and linguists have pointed out that, in many cases, the expanded use of *deng* and *cao* used as plural markers in Chinese sutras is due to Sanskrit influence (see Zhu 1992).

(h) Passive construction:

It has been widely acknowledged among scholars that the frequency of passive constructions in the Chinese of translated Buddhist sutras is noticeably higher than that in Chinese texts of the Medieval Period (see, for example, Wu Jinhua 1983 and Zhu Qingzhi 1993). The passive construction in Chinese, as in many East Asian languages such as Japanese, Vietnamese, and Thai, primarily denotes adversity. The use of the passive voice in Sanskrit is different. It can be used in sentences with transitive verbs, as in English. It can also be used in the so-called “impersonal construction,” where the verb can be intransitive. It is apparent that Sanskrit has a strong preference for the use of passive sentences in Sanskrit scripture. This may explain why there is a significantly higher frequency of passive sentences in the Chinese sutras.

3.2 Two axes in the language of translated Chinese Buddhist sutras

All of the above examples mark the different linguistic features in the language of

Chinese translations of Buddhist sutras. These distinctive phenomena can be observed on two axes:

- i. Synchronically, or on a “horizontal” differential axis, (i) there are special/unusual and new linguistic features that emerged in the language of Chinese-translated Buddhist sutras (zero > innovative features). Examples (a) through (f) above belong to this group. (ii) Certain linguistic elements have a much higher frequency of use in the translated sutras than in the Chinese texts (few > more). Examples (g) and (h) above are of this type. All of these linguistic phenomena are associated with the Buddhist influence via language contact regardless of whether it is a new feature or preexisting feature of high frequency.
- ii. Diachronically, or on a “vertical” evolutionary axis, while all of the aforementioned features appear in translated Chinese sutras, not all of them were accepted into the Chinese language system and used in standard Chinese works. Only a very few of them were accepted and triggered grammatical changes in Chinese.

3.3 Borrowing vs. interference

In our previous studies (Yu & Cao 2007), in examining the relationship between the distinctive linguistic phenomena found in the language of translated Chinese sutras, compared with that of their corresponding Sanskrit sutras, we divided the new phenomena found in the translated sutras into two categories: borrowing and interference. In the former category, linguistic elements of the original Sanskrit language were directly borrowed (i.e., transferred or copied) into Chinese. *Yunhe*, for example, is a typical case where the translators simply took the literal translation as well as the initial sentence position of *kim* ‘what’ and copied it into Chinese. As a result, the Chinese interrogative word *yunhe* ‘what’ is found appearing at the beginning of a sentence in translated Chinese sutras without having the function of an interrogative. The “V (O) + *yi*” construction is an example of interference. There was a clear grammatical rule in Chinese regarding the use of *yi* indicating the completion of an action but Sanskrit does not have such grammatical constraints. The error of not differentiating between the different syntactic behaviors of the two groups of verbs (continuous verbs vs. instantaneous verbs) when using *yi* is due to interference from the translators’ language: Sanskrit. We will come back to this point in the next section.

Borrowing and interference are two separate types of influence on the Chinese language; borrowing refers to linguistic features of other languages that are superimposed onto Chinese while interference involves changes in grammatical development induced

by other languages. The linguistic result of the two types of influence in language contact is different in Chinese. To our knowledge, to date only interference has had an impact on Chinese syntactic development.

4. Language contact and the development of the Chinese language

As mentioned above, language contact is a mechanism of language change, and throughout the history of Chinese language development, one question of major interest has always been whether or not Chinese has been influenced by this phenomenon. If there has been such an external influence, the question now is how it has impacted upon the Chinese language.

First of all, we must define the word ‘influence’. If we take the broad sense of the word, then whenever ungrammatical features in Chinese arise as the result of external contact, we may say that Chinese has been influenced by another language. However, this kind of influence is limited in scope. Ungrammatical elements might only appear within certain genres (e.g., Buddhist scriptures or Buddhist literature), and they may only remain for a short period of time. On the other hand, in the narrow sense of the word, influence causes changes in the development of the Chinese language. These changes spread throughout the Chinese language, they are accepted into the Chinese language and finally are incorporated as parts of the Chinese language system. The latter type of influence is an external mechanism that induces the development of the Chinese language.

4.1 The broad sense of influence of language contact

In the broad sense of influence of language contact, we could count every single ungrammatical linguistic element associated with foreign contact that has ever appeared in any Chinese written text throughout Chinese history. In doing so, then we would find that there are a lot of non-Chinese elements that appear in Chinese writings in various periods of Chinese history. Aside from the examples of the translations of Buddhist sutras given above, there are a number of other examples that can be seen in *Yuan baihua* as well, such as the SOV word order, the post-positioned object case marker *gendi* 根底, the suffix encoding causative *shang* 上 *shangtou* 上頭, the plural marker *mei* 每, the particle indicating direct quotation *modao* 麼道, and so on.

Even though there are many such distinctively unusual phenomena that can be found in writings, they are limited to certain documents or texts from certain time periods. They were not widespread, nor were they accepted into the Chinese language system, and they have never broadly appeared in the written language of the Chinese people.

For example, the sentence final cause marker *gu*, as examined in our previous study (Yu 2008), first appeared in sutras translated by Central Asian monks during the Medieval Period. It continued to be used in Buddhist sutras translated by Chinese monks during the Tang dynasty, and was later also found to have been used in Buddhist writings by Chinese around the ninth through eleventh centuries, including in *Excerpts of Recorded Sayings of the Zen Sect* (*Chanzong yulu*, 禪宗語錄). Yet from beginning to end, the sentence final *gu* only appeared in Buddhist writings and was used only within the limit of that particular religious domain. Similar situations are found with *Yuan baihua*. Cao Guangshun & Chen Dandan (2009) pointed out that many of these special linguistic phenomena only appear in documents or texts relating to the Mongols and these unusual or ungrammatical features fell out of use shortly after the Mongols lost their power to the Chinese.

4.2 The narrow sense of influence of language contact

The narrow sense of language contact influence refers to those unusual linguistic elements that have made an impact on Chinese language development. Based on what we know, very few grammatical changes in Chinese have been triggered by language contact. Studies have shown that the two best known cases of Buddhist influence on the Chinese language are the following: the postverbal *yi* denoting the completion of an action in the “V (O) + *yi*” construction and the disposal construction. In the following, we will take the *yi* construction as an example to demonstrate its development.

It is commonly known that the verb *yi* first appeared in pre-Qin texts meaning ‘complete, finish’. The verb was found to be used in Western Han (206 BC to AD 24) texts in the grammatical frame “V (O) + *yi*”, where V is a continuous verb and *yi*, a verb indicating a completed action.

From the Eastern Han (AD 25-220) to the Sui (AD 581-618) and the Tang (AD 618-907) dynasties, the “V (O) + *yi*” construction with continuous verbs was still not commonly used in Chinese texts, and there was no trace of instantaneous verbs used in this construction. Chinese texts where this structure appears include *Records of the Three Kingdoms* (*Sanguo zhi* 三國志), *History of the [Former] Han* (*Hanshu* 漢書), *History of the Latter Han Dynasty* (*Hou Hanshu* 後漢書), *New Account of Tales of the World* (*Shishuo xinyu* 世說新語), *Qimin yaoshu* 齊民要術, and *Luoyang qielan ji* 洛陽伽藍記. However, there are a large number of the “V (O) + *yi*” examples present in translated Chinese Buddhist sutras.

Yi was one of the most frequently used verbs in second position in the “V₁OV₂” construction during the Medieval Period where V₂ is a verb denoting a completed action, the other forms being *bi* 畢, *jing* 竟 and *qi* 訖. All of these forms share the same

meaning ‘complete’, ‘finish’, or ‘come to an end’. Since the late 1970s, linguists have noticed the usage of *yi* in the translated sutras suggesting this could be related to the influence of the Sanskrit language in the original Buddhist sutras. Important studies include Hung-Nin Cheung (1977), Christoph Harbsmeier (1989), Zhu Qingzhi (1993), and Seishi Karashima (1998 and 2000). In a comparison of the Buddhist sutras in Sanskrit with their corresponding Chinese translations, Karashima states that this kind of *yi* mostly corresponds to the absolutive or gerund in Sanskrit (Karashima 1998, 2000). Jiang Shaoyu takes Karashima’s point further and remarks that there are two types of *yi*: *yi*₁ and *yi*₂, with *yi*₁ being a verb denoting perfective action which can only take continuous verbs, and *yi*₂, being a “highly grammaticalized (高度虛化 *gaodu xuhua*)” element which takes instantaneous verbs. *Yi*₁ has been used in Chinese since the Pre-Qin Period, while *yi*₂ is a new feature from the translation of the Sanskrit absolutive or gerund.

The appearance of excessive numbers of the “V (O) + *yi*” construction and the emergence of instantaneous verbs in the construction are thus due to the Sanskrit influence and the interlanguage errors of the translators in learning Chinese. Imperfect learning of the grammatical constraints of the Chinese “V (O) + *yi*” construction expanded the scope of the use of *yi*. The crucial point we must notice here is that the new phenomena in the “V (O) + *yi*” construction ($V_{\text{instantaneous}} O yi$) was not found in any Chinese texts prior to the Tang dynasty, a fact indicating that the expanded construction was not accepted into the Chinese linguistic system at that time. During the Tang dynasty, the new verb member (instantaneous verbs) in the “V (O) + *yi*” construction is not only frequently found in *bianwen*, it also appears in the writings of Chinese literati. The development of *yi* did not stop there. It was later gradually replaced by another perfective verb *liao* and in the late Tang dynasty, the perfective verb *liao* developed into a perfective aspect marker, the modern Chinese *le*₁.

To date, it is still unclear to us how the new, expanded construction merged and was accepted into the Chinese language. Even though we don’t know the exact process involved, the evolution of *yi* nonetheless is a case demonstrating that external language contact caused linguistic changes in the Chinese language. From this example, we can see that the change of *yi* was triggered by the foreign translators’ lack of distinction between the two different groups of verbs in Chinese. Their incorrect use of *yi* blurred the rule of the “V (O) + *yi*” construction. The interference of this new usage of *yi* resulted in a chain of changes in the Chinese language. This interference is the starting point of the changes. In this snowball effect, even though language contact did not directly influence the development of the perfective aspect marker *le*, it did play a role in forming the beginnings of the change.

5. Some observations

There have been a good number of special linguistic features introduced through language contact in Chinese texts during various periods as discussed above. Why is it that only a very few have influenced Chinese grammatical development? One of the reasons is probably related to the constraints of social factors. We think that whether or not a contact-induced feature can trigger language change depends on a crucial key—the acceptance of the new feature by the speakers of that language. We can see from the above examples that those new linguistic features harmonizing with or similar to Chinese syntactic structure seem to be accepted more easily, while those distinct from Chinese structures are less tolerated. The postverbal case markers found in translated sutras such as *gu* and those that appeared in *Yuan baihua* like *gendi*, *shang*, and *shangtou* are typologically different from the structure of Chinese. These features were only short-lived in certain written texts and never came to be accepted in the Chinese language system in the end. The perfective action *yi*, on the other hand, merged into the Chinese language. It blurred the original constraining rule of the “V (O) + *yi*” construction, but the appearance of instantaneous verbs in the verb position did not break the structural frame of the Chinese language. A somewhat parallel situation can be observed in the disposal construction. Between the Han and Tang dynasties, the serial verb construction ($V_1O_1V_2O_2$) in Chinese underwent a series of developments and changes: when $O_1=O_2$, the repeated object can be omitted. This time period happens to be the period when sutras were translated into Chinese from Sanskrit. The translators, due to the influence of the Sanskrit SOV word order, omitted O_2 , which resulted in the new grammatical form of $V_1O_1V_2$. This new form developed into the narrow disposal construction where the full meaning of V_1 was bleached or grammaticalized (Cao & Yu 2000a, 2000b). If the serial verb construction did not undergo changes, the new formation of the disposal construction probably would not have become one of the most important grammatical constructions in the Chinese language. Like the case of the new form of the “V (O) + *yi*” construction, this new form was influenced by language contact, and it also emerged in Buddhist texts earlier than in native forms, with higher occurrence rates as well.

As mentioned above, the influence of language contact on the structure of a language is closely tied to social factors. And yet social factors often are unpredictable. When we study language contact and its influence on language change synchronically, speakers of the language and social factors may help us to pinpoint the principles or rules for change. But in diachronic studies, because we don’t have anyone who speaks Ancient Chinese, all we can do is use the corpus of linguistic facts drawn from written texts during various dynasties. When these linguistic facts serve as the sole basis for us to study language contact and its effect, certain rules we have proposed may, at times, become

inapplicable. Taking the “V (O) + *yi*” formation as an example, we proposed that the “V (O) + *yi*” construction was expanded by blurring the Chinese structural rule for *yi*. This new formation was later accepted by Chinese and further developed in being applied to *liao* 了, and then to *le*₁. It is a successful case where an error caused by imperfect learning of the Chinese language influenced Chinese language development. However, there are other cases where the Chinese structural rules were also blurred by the interlanguage errors of non-Chinese speakers and yet these features were not accepted into the Chinese language. For example, the words to indicate “plurality” of people in Classical and Medieval Chinese are *deng*, *bei* and *cao*. The frequency of the use of these words was significantly higher in translated sutras than in the Chinese texts of the same time since Sanskrit has a number category. After the Northern Song dynasty (AD 960-1126), *mei* 每 (or *men*) replaced the above forms as a plural marker for human beings. During the Yuan dynasty, there was a much higher frequency of the use of *mei*. In *Yuan baihua*, the use of *mei* was expanded beyond the marking of human beings to include animate nouns, such as *mamei* 馬每 ‘horses’, *ying-goumei* 鷹狗每 ‘eagles and [hunting] dogs’, and even to inanimate nouns such as *caodimei* 草地每 ‘grass’, and *chaitanmei* 柴炭每 ‘firewood-charcoal’. The expansion of the use of *mei* is associated with the language contact with the Mongols. The different system in marking plurality in Mongolian resulted in errors in overtly applying the plural markers to nouns beyond human beings. But the newly expanded usages of *mei* can only be found in *Yuan baihua* and later became extinct after the Mongols lost their power to the Chinese. It did blur the Chinese structure, but it was not permanently integrated into Chinese.

From the examples of the translated Buddhist sutras in the Medieval Period and those of *Yuan baihua*, it seems that structural similarity is a condition for a contact-induced feature to be accepted by Chinese language. But in reality, structural similarities do not always guarantee the acceptance of a new feature introduced via language contact. The prerequisite for a non-Chinese feature to emerge in the Chinese language system is that the feature has to be accepted by native Chinese language speakers. The new usage of *yi* and *mei* discussed above expanded Chinese grammatical structures, yet while the former triggered changes in Chinese grammar, the latter became extinct. If we examine the examples given in this paper from the perspective of typology, we can see that the three languages mentioned are of various language families, namely Sino-Tibetan (Chinese), Indo-European (Sanskrit), and Altaic (Mongolian). From the cases in Chinese, on the one hand, we can see that it is possible for grammatical changes to be triggered by a typologically different language in language contact situations. On the other hand, only very few language contacts have had a long-lasting impact. Social factors seem to play a more crucial/decisive role than syntactic comparability.

References

- Cao, Guangshun. 2007. The Mongol Empire in China and the Mongolian influence on Chinese language. Paper presented at the “Steppes and their Suburbs” Workshop, September 29-October 1, 2007. Porquerolles, France.
- Cao, Guangshun, and Dandan Chen. 2009. Yuan baihua teshu yuyan xianxiang zai yanjiu [Reexamination of the special features in Yuan baihua]. *Lishi Yuyanxue Yanjiu* [Studies of Historical Linguistics] 2:108-123. Beijing: The Commercial Press.
- Cao, Guangshun, and Hsiao-jung Yu. 2000a. The influence of translated Later Han Buddhist Sutras on the development of the Chinese disposal construction. *Cahiers de Linguistique – Asie Orientale* 29.2:151-177.
- Cao, Guangshun, and Hsiao-jung Yu. 2000b. Zhonggu yijing zhong de chuzhishi [The disposal construction of translated middle Chinese Buddhist sutras]. *Zhongguo Yuwen* [Studies of the Chinese Language] 2000.6:555-563.
- Chen, Xiulan. 2003. *Wei Jin Nanbeichao Wen yu Hanwen Fodian Yuyan Bijiao Yanjiu* [A Comparison Study on the Language of the Wei Jin Nanbeichao Period and the Language in Translated Chinese Buddhist Sutras]. Hangzhou: Zhejiang University dissertation.
- Cheung, Hung-nin Samuel. 1977. Perspective particles in the Bian Wen language. *Journal of Chinese Linguistics* 5.1:55-74.
- Fairbank, John K. 1978. *China*. Boston: Houghton Mifflin.
- Harris, Alice C., and Lyle Campbell. 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge & New York: Cambridge University Press.
- Harbsmeier, Christoph. 1989. The Classical Chinese modal particle *yi*. *Proceedings of the Second International Conference on Sinology, Section on Linguistics and Paleography*, 475-504. Taipei: Academia Sinica.
- Heine, Bernd, and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge & New York: Cambridge University Press.
- Jiang, Lansheng. 2003. Yuyan jiechu yu Yuan Ming shiqi de teshu panduanju [Language contact and the special copular construction during the Yuan and Ming periods]. *Yuyanxue Luncong* [Essays on Linguistics] 28:43-60. Beijing: The Commercial Press.
- Jiang, Shaoyu. 2001. *Shishuo Xinyu, Qimin Yaoshu, Luoyan Qielan Ji, Xianyu Jing, Baiyu Jing* zhong de ‘yi’, ‘jing’, ‘qi’, ‘bi’ [‘Yi’, ‘jing’, ‘qi’ and ‘bi’ in the *Shishuo Xinyu, Qimin Yaoshu, Luoyan Qielan Ji, Xianyu Jing* and the *Baiyu Jing*]. *Yuyan Yanjiu* [Studies in Language and Linguistics] 2001.1:73-78.

- Karashima, Seishi. 1997. Hanyi fodian de yuyan yanjiu [A study on the language of the translated Chinese Buddhist sutras]. *Suyuyan Yanjiu* [Studies of Vernacular Language] 1997.4:29-49.
- Karashima, Seishi. 1998. *Zheng Fahuajing Cidian* [A Glossary of Dharmaraksa's Translation of the Lotus Sutra]. Tokyo: International Research Institute for Advanced Buddhology, Soka University.
- Li, Yanji. 1751. *Qingwen Huishu* [The Compilation of the Qing Language]. Beijing: Yinghua Tang.
- Thomason, Sarah Grey, and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Wang, Jihong. 2004. *Jiyu Fan-Han Duikan de Fojiao Hanyu Yufa Yanjiu: Yi Apidamo Jushelun-Fenbiejie Pin Wei Li* [A Grammatical Study of the Buddhist-Chinese via the Comparative Use of Sanskrit and Chinese Buddhist Sutras]. Beijing: Peking University dissertation.
- Wu, Jinhua. 1983. 'R wei A suojian V' shi [The construction of 'R wei A suojian V']. *Zhongguo Yuwen* [Studies of the Chinese Language] 1983.3:207-210.
- Yu, Hsiao-jung. 2003. Shuo 'yunhe' [A discussion on the interrogative word *yunhe*]. *Kaipian* 22:48-50.
- Yu, Hsiao-jung. 2006. Qiantan 'Qiren bai wang, fu yi si liao' — jianlun yuyan jiechu yingxiang Hanyu yufa fazhan de jizhong moshi [A discussion on language contact and its influence on certain kinds of syntactic changes in Chinese]. Paper presented at the 12th National Pre-modern Linguistics Conference. Guizhou, China.
- Yu, Hsiao-jung. 2007. Manchu rule over China and the attrition of the Manchu language. Paper presented at the "Steppes and their Suburbs" Workshop, September 29-October 1, 2007. Porquerolles, France.
- Yu, Hsiao-jung. 2008. Lilun yu shishi—yuyan jiechu shijiao xia de Zhonggu yijing [Theory and fact: a study of the translated Buddhist sutras of the Medieval Period from the perspectives of language contact]. *Hanyushi Xuebao* [Journal of Chinese Language History] 7:121-127. Shanghai: Shanghai Education Press.
- Yu, Hsiao-jung, and Guangshun Cao. 2007. Zai tan Zhonggu yijing yu Hanyu yufashi yanjiu [Revisiting the language in translated Buddhist sutras of the Medieval Period and its significance in the studies of historical Chinese grammar]. *Han Zangyu Xuebao* [Journal of Chinese Language History] 1:1-15. Beijing: The Commercial Press.
- Zhao, Changcai. 2009. Zhonggu Hanyi fojing zhong de houzhici 'suo' he 'bian' [The postpositions *suo* and *bian* in Medieval translation of Buddhist scriptures]. *Zhongguo Yuwen* [Studies of the Chinese Language] 2009.5:438-447.

- Zhu, Qingzhi. 1992. *Fodian yu Zhonggu Hanyu Cihui Yanjiu* [*A Lexical Study on the Language in Buddhist Sutras and Medieval Chinese*]. Taipei: Wenjin Press.
- Zhu, Qingzhi. 1993. Hanyi fodian yuwen zhong de yuandian yingxiang chutan [A preliminary survey of the influence of the language in original Buddhist sutras on the language in translated Chinese sutras]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 1993.5:379-385.
- Zu, Shengli. 2005. Yuandai Mengguyu he beifang Hanyu yuyan jiechu de wenxianxue kaocha [A textual analysis of language contact between Mongolian and Northern Chinese during Yuan Dynasty]. *Menggushi Yanjiu* [*Studia Historica Mongolica*] 8:52-79.
- Zu, Shengli. 2007. Yuandai de Mengshi Hanyu ji qi shi ti fanchou de biaoda: yi zhiyiti wenxian de yanjiu wei zhongxin [On the Mongolian Pidgin Chinese of Yuan Dynasty with special reference to its tense and aspect category]. *Dangdai Yuyanxue* [*Contemporary Linguistics*] 2007.1:1-13.

元白話語言性質再研究

曹廣順

中國社會科學院

元白話是指元代文獻中在詞彙語法上呈現蒙古語特徵的資料，以往的研究常把這種資料視為元代的北方口語。本文根據對元代文獻的分類比較證明：元白話中蒙古語的影響十分有限，其中特殊語言現象的使用範圍有明顯的局限和差異。元白話是由不完全習得而產生的帶有蒙古語影響 TL2，以後的發展中元白話特殊語言現象進入漢語的為數甚少。所以，元白話不是元代北方漢語口語，並沒有記錄當時北方漢語口語的實際面貌，只是當時蒙古人學習漢語的過程中產生的一種中介語。

關鍵詞：元白話，語言接觸，語言性質

1. 元白話

元白話常常作為元代接近口語的文獻語言的統稱使用，但是引起研究者注意的，是其中經常出現的不同於漢語的詞彙語法特徵。雖然統稱為元白話，但在不同的文獻中，這些特徵出現的多寡有所不同。

下面是《元典章四十一·刑部卷三》中的一段文字：

承奉皇慶二年正月十三日本台官奏過事內一件：「昨前，省家俺根底與將文書來：『淮東廉訪司官人每文字裡說將來：“楊州宣慰使汪元昌小名的人，他的爺歿了呵，省家交他奔喪去麼道與將文書去呵，他不奔喪，將省家文書藏了，每日筵席有。完澤篤皇帝升天了呵，別了大體例，家裡喚得樂人來，交唱著，他自彈著，筵席有。更歿了爺不奔喪的罪過，遇赦免了也。”刑部擬著他除名不敘。』麼道。俺商量來：依著他每擬定的行呵，怎生？」奏呵，「那般者。」麼道，聖旨了也。欽此。咨請欽依施行。

這段文字是《元典章》中的文書詔令，元白話中包含特殊現象最多一類文獻，常常被稱為“蒙式漢語”，其中許多語法現象是漢語裡沒有，明顯帶有蒙古語影響痕跡的，如：

- (1) 使用借詞：格標記“根底”；
- (2) 大量使用複數詞尾“每”；
- (3) 使用句末動詞“有”；
- (4) 大量使用語氣詞“呵”；
- (5) 句末使用祈使語氣詞“者”；
- (6) 使用引語助詞“麼道”。

與元白話有關的另一個概念是“漢兒言語”。太田辰夫先生指出：“‘漢兒言語’，它在‘漢兒’中間使用，也在北方民族和‘漢兒’之間使用，這是不言而喻的。但令人吃驚的是，即使在母語各異的北方民族中間它也被作為共同語，並且連極僻遠的地方也可通行。”；“這一時代（指元代）的‘漢兒言語’似乎比以前更加普遍地在北方民族間通行。……那種漢語自然是卑俗的，或者是不合規範的漢語，這一點勿庸置疑。”；“不單‘漢兒’連蒙古人中懂‘漢兒言語’的也不少。”；“因為‘漢兒言語’作為中間媒介來使用，所以‘漢兒言語’就成為當時的共通語，這是顯而易見的。”¹ 祖生利先生進一步說：“我們所說的元代的‘漢兒言語’則用來特指元代後期在蒙古語（及諸色目人語）影響下形成的流行於大都等北方地區的漢語變體，它是元代蒙古語同北方漢語語言接觸的產物，有著顯著的蒙古語的干擾特徵。”² 祖生利認為，典型的“漢兒言語”是像古本《老乞大》那樣的語言。如：

恁這月盡頭到的大都那到[不]得？知他，那話怎敢道。天可憐見，身已安樂呵，也到得有。恁是高麗人，卻怎麼漢兒言語說的好有？俺漢兒人[上]學文書來的上頭，些小漢兒言語省的有。你誰根底[學]文書來？我在漢兒漢堂裡學文書來。

¹ 參閱：太田辰夫. 1954[1991].〈關於漢兒言語：試論白話發展史〉，《漢語史通考》（江藍生、白維國譯）。重慶：重慶出版社。

² 參閱：祖生利. 2009.〈試論元代的“漢兒言語”〉，《歷史語言學研究》第二輯。北京：商務印書館。

這段古本《老乞大》中仍有蒙古語影響的痕跡，如：

- (1) 使用借詞：方位詞的格標記“根底、上、上頭”；
- (2) 使用句末動詞“有”；
- (3) 大量使用語氣詞“呵”。

此外，也有些元白話文獻中，並沒有明顯的蒙古語影響的痕跡。如《直說通略》卷下：

帝又到中牟縣圍獵，踐踏了百姓田禾，中牟縣令當馬前諫說道：陛下為民父母，怎生損壞民食，餓死了百姓？帝怒教將起去殺了。敬新磨再拿來到馬前，責罵說道：恁做縣令，怎不理會得天子好獵，如何卻縱令百姓耕種，妨誤天子飛放，您罪過該死！帝笑教放了。

雖然也是元代較為接近口語的文獻，也是元白話的組成部分，但其中幾乎看不到受到蒙古語影響的特徵。

如果以上幾種文獻都是元白話，那麼元白話就是一個比較寬泛的概念，是元代比較接近口語的文獻語言的統稱，其中包含不同類型的文獻資料。其中部分文獻的語言受到蒙古語的影響，受影響的程度依文獻不同有所差別。

2. 元白話的語言性質

對於這樣一個比較寬泛的概念，元白話的語言性質也就難以界定了。最近的研究認為，元白話可以劃分為不同的層次，蒙古王公貴族詔書飭令等的翻譯（以上我們列舉的第一類），是一種蒙式漢語，存在大量蒙古語的語法、詞彙成分，明顯受到蒙古語的影響，但其使用僅限於翻譯蒙古貴族官僚的命令文書，日常生活中並沒有人使用。一般百姓的語言（我們列舉的第三類），則基本上看不到蒙古語的影響，詞彙語法完全是漢語的。有趣的是中間一類“漢兒言語”，太田辰夫 1954 提出：“把漢兒言語看作是廣義的漢兒間的共通語這樣的看法比較合適。”但“共通語”究竟是一種什麼樣的語言？是否就是共同語？這種“漢兒言語”在當時是否有口語基礎？什麼人會說這種語言？所有這些問題，仍有待我們進一步考察。

3. 元白話特殊語言現象的使用

3.1 從古本《老乞大》看元白話特殊語言現象的使用

既然古本《老乞大》是比較典型的“漢兒言語”，我們首先考察一下古本《老乞大》的語言特徵，也就是其中受到蒙古語影響的特殊語言現象的使用情況。概括起來說古本《老乞大》中特殊現象重要是以下幾類：

3.1.1 OV 句式（非“有”字）

OV 句式的出現是“漢兒言語”受到蒙古語影響最明顯例證，從漢語的 SVO 語序到蒙古語的 SOV 語序，改變了漢語的基本類型特徵。例如：

- (1) 咱每為父母心盡了，不曾落後。
- (2) 茶飯吃了呵，碗子傢俱收拾者。

但是，這些 OV 語序的句子很難與漢語的受事主語句分開，目前的研究中大家判定的依據多是根據語感，而語感實際上不足為據。即使是根據語感，可能的 OV 句式在古本《老乞大》也只有不足 20 例。這在整個古本《老乞大》中，約佔句子總數的 1.6%。

3.1.2 動詞“有”字句

動詞“有”字句，出現 132 句，“有”字在元白話中使用比較特殊，我們把它單獨拿出來統計。132 例“有”字句中 OV 語序的 18 句，VO 語序的 112 句，另有 2 例使用了兩種語序的疊加句式（如例(5)）。

- (3) 為什麼這般的歹人有？
- (4) 父母在生時，家法名聽好來，田產物業有來，孳畜頭匹有來，人口奴婢有來。
- (5) 恁這布裡頭長短不等。有勾五十尺的有麼？

“有”字句中 OV 語序的句子出現數量少，說明其使用有很大的隨意性，用不用 OV 語序來表達，是不固定的。

3.1.3 判斷句

判斷句，出現 148 句，“S+是+O”類 133 句，“S+O+是”類 9 句，“S+O+有”1 句，“S+是+O+有”5 句，特殊判斷句僅出現 15 句，佔總數的十分之一。

- (6) 一主兒的不是，這四箇伴當是四箇主兒。
- (7) 俺買呵，買一兩個，自穿的不是，一發買將去要覓些利錢。
- (8) 這言（個）伴當便是。
- (9) 離閣有一百步地向街，那北巷裡向街開雜貨鋪兒便是。

在判斷句中也有同樣的問題，是否使用受到蒙古語影響的特殊句式有很大的隨意性，同樣的句子，可以用，如例 (10a)；也可以不用，如例 (10b)。

- (10) a. 這蓼是新羅蓼有，也著中。
- b. 俺的是新羅蓼。

3.1.4 使用格標記“根底、上、行”

在古本《老乞大》中“根底”表示動作的對象，相當於漢語的“向/對”（4 例）：

- (11) 好媳婦別人根底去也。
- (12) 恁這等慣做買賣的人，俺一等不慣的人根底多有過瞞有。

用“上、行”表示動作的對象、處所，相當於漢語的“在”“到”等，也可以表示領屬，相當於漢語的結構助詞“的”。“行”4 例，表示動作的對象；“上”4 例，表示動作對象的 1 例，表示處所的 2 例，表示領屬的 1 例。

- (13) 每日學長將那頑學生師傅行呈著。
- (14) 鍋子上蓋覆了，休著出氣，燒動火，暫霎兒熟也。
- (15) 那般者，你買下飯去時，這間壁肉案上買豬肉去。
- (16) 兀的燈來也，壁子上掛者。
- (17) 你是姑舅弟兄，誰是舅舅上孩兒，誰是姑姑上孩兒？

上述用法在古本《老乞大》中的使用頻率很小，常見的仍是使用一般漢語語

法，也就是用介詞來表示對象、處所的句子。如：

- (18) 他在那個房子裡下？
(19) 小人姓王，在東京城裡闖北街東住。

3.1.5 假設複句

一般認為，元白話中把“呵、時”等置於假設條件分句的末尾，然後再接結果的假設複句是受到蒙古語影響產生的，古本《老乞大》中也使用這一類特殊的假設複句。如：

- (20) 咱每都去了時，這房子裡沒人，敢不中。
(21) 安樂時不快樂呵，真個呆人也。

這類假設複句有 39 句，是古本《老乞大》中使用最多的假設複句句式。另外古本《老乞大》中還有一些既含有漢語的假設連詞，又含有“呵、時”的句子。可以看作是漢語和蒙古語兩種語法規則雜糅的表現。其中假設連詞有“若、怕、不爭”等。如：

- (22) 若背不過時，教當直學生背起，打三下。
(23) 他每若是歹人，來歷不明呵，怎生能勾到這裡來？
(24) 怕你不信時，別個店裡試商量去。
(25) 有甚麼熟菜蔬，將些來與客人吃。怕無時，有蘿蔔、生蔥、茄子將來，就將些醬來。
(26) 不爭將去時，連其餘的馬都染的壞了。

這類句子共有 17 例。

漢語的假設複句通常是在一個分句的句首加連詞“若、如”，構成假設條件。古本《老乞大》中這樣的假設複句有 11 句。如：

- (27) 若做直身襖子，有剩。
(28) 如馬來處不明，賣主一面承當。

以上介紹的古本《老乞大》中假設複句似乎是以受到蒙古語影響的形式為主，但是這裡需要指出的是，用“呵、時”表示假設並非始見於元白話，宋代文

獻中已經可以見到類似的情況，只是元代使用的更多（孫梅青 2006）。從其發展看，或許元代只是推動了這種句型使用增加，而不是造成了一種新句型的出現。其他元白話特殊現象在古本《老乞大》中都只出現少量的個別例子，假設複句在古本《老乞大》中的使用情況也可以從另一個側面證明，這個句式可能不是一個受到蒙古語影響產生的新句式。

3.1.6 因果複句

古本《老乞大》中因果複句有很大一部分是由方位詞“上，上頭”放在原因分句的句尾構成，如：

- (29) （恁是高麗人，卻怎麼漢兒言語說的好有？）俺漢兒人□學文書來的上頭，些小漢兒言語省的有。
- (30) 那般呵，消化不得上頭，腦痛頭眩，不思飲食。
- (31) 恁偏不理會的，從年時天旱，田禾不收，饑荒的上頭，生出歹人來。

這種表示原因的後置詞“上，上頭”，研究者認為是“直譯中古蒙古語表示原因的後置詞 *tula* 和形動詞工具格附加成分 *-ar/-bar*”（祖生利 2004）。但同時也應該指出的是，方位詞“上”表示對象、原因的演變過程在元代之前就已經開始了，和假設複句一樣，元白話可能只是推動了這一過程的發展。

綜上所述，古本《老乞大》中元白話特殊語言現象有些很少使用，有些雖然使用較多，但常常是在元代以前就已經出現的語言現象，難以確定是在蒙古語影響下新產生的。

如果古本《老乞大》就是元代“漢兒言語”，即北方漢語的“共通語”的話，以上我們的分析說明，古本《老乞大》的語言是一種包含零星的元代受到蒙古語影響的特殊語言現象的漢語，這些特殊語言現象的出現是非強制性的，使用頻率與一般的漢語表達方式相比非常低。也就是說，元代“漢兒言語”並沒有一種語法和詞彙上相對獨立的系統，在文獻中看到的蒙式漢語和漢族人的一般漢語之間，古本《老乞大》的“漢兒言語”不是一種協商之後的簡化和妥協，而是基本上使用一般的漢語，同時夾雜個別的蒙式漢語特徵。這些蒙式特徵的出現，可能是與蒙古人交際習慣的流露，也可能是吸收的一些當時接受度較高的蒙式漢語特徵。同時，我們說到古本《老乞大》的時候，大家都習慣的強調這是當時的漢語教科書，應該接近當時的口語，但可能沒有注意到的是，古本《老乞大》是一本韓國人編輯的漢語教科書，在編輯過程裡，儘管他們會盡可能的忠實於漢語，

但其母語的影響不可能沒有。比較蒙語和韓語，對韓國人來說，蒙式漢語應該更容易接受，在其教科書中不自覺的使用一些蒙式漢語的特徵，是有可能的。

3.2 從元代其他文獻看元白話特殊語言現象的使用

古本《老乞大》的語言是一種包含零星受蒙古語影響產生的特殊語言現象的漢語，如果把它定義為“漢兒言語”的話，太田先生所說的“共通語”實際上就是當時的漢語。那麼，在元代有沒有一種為社會所接受的、帶有明顯蒙古語影響的漢語呢？以下我們對元代的幾種文獻分別進行考察。

3.2.1 從《元典章·刑部》看元白話特殊語言現象的使用

《元典章·刑部》是研究元代漢語的重要資料，其中記載的各類供詞，記錄了當時不同階層語言的使用狀況。

1. 卷三【燒烙前妻兒女】延佑三年十月

有後母韓端哥，不知主何情意，用鐵鞋錐于俺孫女郝醜哥舌頭上，烙訖三下；脊背上，烙訖七十二下。小廝郝罵兒也烙了七錐子。
你昨日城裡來的晚了。您兩個孩兒偷出小荳，客人處換梨兒吃。
我是換了五個梨兒吃來。

2. 卷七【品官妻與從人通姦】至元十八年十月

劉阿孫道：“劉提舉那廝，十二三年不曾來我行宿臥。我根你去。”

3. 卷四【取受被察推病，依例罷職】大德七年五月

京州行省的文卷刷去的監察每文書裡說有：“那行省姓郭的都事，為刷馬的上頭，徐知州小名的人根底取受了一個馬，阿難答的奧刺赤田亨的根底取受了十一定鈔。出錢的、過錢的人每根底問呵，明白指證的文字與了也。他根底喚呵，‘病’麼道推辭著，不肯出來對證。交醫人驗去呵，‘沒病’道有。又交喚呵，‘病’麼道，一個月不曾出來。”

以上三段文字，第一段是一般百姓的供詞，沒有出現受蒙古語影響的特殊語言現象；第二段是官員妻子的供詞，出現了一例後置詞“行”；第三段是有關案情的奏摺，幾乎每一句中都使用了元白話的特殊語言現象。

《元典章·刑部》中記錄的供詞基本上可以分爲這三類，平民百姓的供詞中基本上沒有出現過元代的特殊語言現象，一般官員零星使用，王公貴族集中使用。由此可以看出，一般百姓在日常的交際中是很少或不會使用這些特殊的語言現象的，他們在交流的過程中，即使面對的是審問他們的官吏（很可能是蒙古人），他們所說的話仍是“純漢語”。反之，地位越高的蒙古官吏越多的使用元白話特殊語言現象。在這個漸變的過程中，我們很難想像除了一般漢語之外，還有一個作爲“共通語”使用的“漢兒言語”。

3.2.2 從《元朝秘史》旁譯與總譯的差別看元白話特殊語言現象的使用

《元朝秘史》明朝初年被從蒙古語譯成漢語，翻譯者不是漢人，如果元代有一種流行於北方的“共通語”“漢兒言語”的話，《元朝秘史》的譯者一定會受其影響，在譯文中有所反映。以下我們從《元朝秘史》旁譯和總譯的對比來看元白話特殊語言現象的使用。

(32) 我是王罕。

必 王_中罕 備
我 人 名 有

(33) 答三十七下

_中忽臣 朵羅安 別_舌裏額思 莎餘額惕 必丹突兒 迓步_中灰
三十 七 條 子 教導了 咱 行 行的自的
班 別_舌兒客失顏 阿主兀。(d227)
行 作 難 有 來

(34) 聽得不兒罕山野物廣有。

不山而（罕）[_中罕]·（哈）[_中合]勒敦 訥 戈_舌劣額孫 戈_舌魯兀黎
山 名 的 野 物 可 捕
撒亦禿 _中合剌兒 撒因 客延。
好 有 的 地 好 麼 道

我們在《元朝秘史》總譯中看到，所有的判斷句在翻譯的過程裡都由原來的“S+O+有”改成了“S+是+O”式（如例(32)）。

在總譯中，用於句末做助詞或表示時體的“有”，用與不用兩可。出現的頻率，以“有來”爲例，在《元朝秘史》旁譯中出現 188 次，總譯中保留的有 20 次，保留的比例約 1/10（如例(33)）。

《元朝秘史》總譯中仍出現了一些受到蒙古語影響的 OV 等特殊語序（例 (34)），但數量較少，在全部句子中佔的比例很低。

《元朝秘史》兩種翻譯的對比表明，元白話中這些“特殊的語言現象”，確實有其蒙古語的根源，是從蒙古語到漢語的翻譯過程中，受蒙古語的影響而形成的。但是這些特殊語言現象的使用基本上是例外，在大多數情況下，翻譯者在翻譯過程中使用標準的漢語句式，而不會採用這些特殊的語言現象。

3.2.3 從《經世大典》與《元典章》的比較看元白話特殊語言現象的使用

《經世大典》也稱《元會要》，天曆二年（1329 年）冬，有旨命奎章閣學士院與翰林國史院編纂，歷時一年多，於至順二年（1331 年）五月纂修完成。成書時共有八百八十卷，目錄十二卷，公牘一卷，纂修通議一卷。現已全部亡佚，只在《永樂大典》殘卷和《元文類》中還保留了一些章節。另外還有部分散見於《元史》、《元史類編》、《元明事類鈔》等書的引文部分。《經世大典》成書晚於《元典章》，其中一些內容與《元典章》相同，是從《元典章》中摘引改寫的。同是元代的官修典籍，參與編撰的同樣是蒙古貴族高官和漢族官僚，內容也相同，但使用的文字卻有所不同，《經世大典》把《元典章》中一些明顯帶有蒙古語影響特徵的蒙式漢語，改寫成了一般的漢語，如：

《元典章·戶二》：

至元二十一年二月十七日，通政院官等奏：“我從女真田地裡來的時分，信州站戶每我行文字與著告有來：‘裡住的打捕鷹的忽都魯回還家去的時分，白日經過時，見在的羊肉與呵，不肯吃，俺行嗔責道不是，要訖活羊吃了也。俺每忽都魯根底攀著說有，若別個的似這般行的，見有的肉與呵，不吃，要活羊吃的多有，似這般啊，俺生受有。您使臣每根底告有。怎生上位奏知的，您識者。’道來。”麼道，奏呵。“呆站家與來也者。如今見在肉有呵，與見在肉者。若見在肉與呵，不吃的人每根底，水也休與者。”麼道，聖旨了也。

《永樂大典》卷一九四一八《經世大典·站赤》：

二十一年二月……十七日，通政院奏：“譯史阿散來自女真之地，據信州站戶告言：‘打捕鷹房官忽都魯者，由川路回過本站，見在之肉不食，必需活羊烹之。後來者若複仿此，實難供給。’”上說：“此站何遇而與之！今後見在有肉則與之，與而不食，雖水勿與可也。”

《元典章》和《經世大典》都是爲了治理國家的需要，把有關的法規政令、案例條文集中整理，可以想見，如果受到蒙古語影響的“漢兒言語”是當時流行於世的“共通語”，《經世大典》的編纂者一定會用“漢兒言語”來改寫《元典章》中的蒙式漢語。而我們看到的卻是，蒙式漢語被一般漢語所替代，其中的蒙古語特徵都被修改了。《經世大典》對《元典章》文字的修改，從另一個側面證明，不僅蒙式漢語在當時用於交流有問題，“漢兒言語”也不是普遍流行的用法，交流的有效工具，仍然是一般的漢語。

4. 簡短的結語

以上我們介紹了不同類型的元代文獻所反映出的語言面貌，如果我們把元白話視爲一個整體，那麼這個整體是由不同類型的語言構成的，最上層是蒙古貴族們使用的蒙式漢語，然後是或多或少受到蒙古語影響的官僚士紳，以至於市井生意人（如《老乞大》所反映的）使用的帶有蒙古語影響特徵的漢語，最後是一般老百姓使用的漢語。

不言而喻，和任何語言接觸造成的語言改變一樣，元白話中特殊語言的產生來源於蒙古人進入中原，成爲統治者。少數蒙古族統治者要治理這個國家，而絕大多數的人民，是以漢語爲母語的。在這個歷史背景之下，蒙古統治者爲了實現其統治，不得不使用漢語。於是，出現了這場漢語歷史上最廣泛的語言接觸，接觸的主流是作爲統治者的蒙古人學習漢語。像所有第二語言習得者一樣，他們也有一個從不完全習得到完全習得的學習過程，而這個過程依交流需要不同而出現差異：帝王貴冑最大程度的保留了自己的蒙古語特徵，官僚士庶保留程度多少不等。於是我們看到：

- (1) 在元代文獻中元白話特殊語言現象的使用數量較少；
- (2) 使用範圍主要局限於蒙古統治者；
- (3) 一些重要的典章制度在用元白話記錄之外，仍需改寫成標準漢語；
- (4) 顯然，元白話仍然不是當時社會廣泛流通的語言，它是一種作爲中介的“共通語”，而非普遍使用的“共通語”。

引用文獻

- Cao, Guangshun (曹廣順), and Dandan Chen (陳丹丹). 2009. Yuan baihua teshu yuyan xianxiang zai yanjiu 元白話特殊語言現象再研究 [Reexamination of the special features in Yuan baihua]. *Lishi Yuyanxue Yanjiu* 歷史語言學研究 [*Historical Linguistics Study*] 2:108-123. Beijing: The Commercial Press.
- Lee, Tae-Soo (李泰洙). 2003. *Laoqida Sizhong Banben Yuyan Yanjiu* 《老乞大》四種版本語言研究 [*Analysis on Language Usage Trends from Four Editions of Laoqida*]. Beijing: Language and Culture Press.
- Lee, Tae-Soo (李泰洙), and Lansheng Jiang (江藍生). 2000. Laoqida yuxu yanjiu 《老乞大》語序研究 [On word order in *Laoqida*]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 2000.3:71-82.
- 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.
- Sun, Meiqing (孫梅青). 2006. *Yuan Zhi Mingchu Hanyu Teshu Yufa Xianxiang Yanjiu* 元至明初漢語特殊語法現象研究 [*A Study of the Special Chinese Linguistic Features during the Yuan and Early Ming Periods*]. Beijing: Chinese Academy of Social Sciences dissertation.
- Yu, Hsiao-jung (遇笑容). 2004. Hanyu yufashi zhong de yuyan jiechu yu yufa bianhua 漢語語法史中的語言接觸與語法變化 [Language contact and grammatical changes in historical Chinese syntax]. *Hanyushi Xuebao* 漢語史學報 [*Journal of Chinese Language History*] 4:27-34.
- Zu, Shengli (祖生利). 2001. Yuandai baihua beiwen zhong fangweici de ge biaoji zuoyong 元代白話碑文中方位詞的格標記作用 [The usages of case marker of localizers in the vernacular inscriptions of Yuan Dynasty]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 2001.4:62-75.
- Zu, Shengli (祖生利). 2004. Yuandai zhiyiti wenxian zhong de yuanyin houzhici 'shang/shangtou' 元代直譯體文獻中的原因後置詞“上/上頭” [The special usage of causal postposition of *shang* and *shangtou* in the translated documents of Yuan Dynasty]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 2004.1:47-52.
- Zu, Shengli (祖生利). 2007. Yuandai de Mengshi Hanyu ji qi shi ti fanchou de biaoda: yi zhiyiti wenxian de yanjiu wei zhongxin 元代的蒙式漢語及其時體範疇的表達：以直譯體文獻的研究為中心 [On the Mongolian Pidgin Chinese of Yuan Dynasty with special reference to its tense and aspect category]. *Dangdai Yuyanxue* 當代語言學 [*Contemporary Linguistics*] 2007.1:1-13.

- Zu, Shengli (祖生利). 2009. Shi lun Yuandai de 'Han'er yanyu' 試論元代的“漢兒言語” [On the Han'er language of the Yuan Dynasty]. *Lishi Yuyanxue Yanjiu* 歷史語言學研究 [*Historical Linguistics Study*] 2:124-135. Beijing: The Commercial Press.

Revisiting the Characteristics of the Language in *Yuan Baihua*

Guangshun Cao

Chinese Academy of Social Sciences

Yuan baihua is a language that has lexical and syntactic features of Mongolian language recorded in the Yuan Dynasty texts. The previous studies often take it as a northern colloquial language used in China during the Yuan period. After having divided the Yuan documents and texts into categories, comparing and contrasting their linguistic features, however, we find that the Mongolian influence on the *Yuan baihua* is rather limited. These special, unauthentic linguistic elements only appear in documents or texts related to the Mongols. *Yuan baihua* is a TL2 resulting from the imperfect learning of Chinese by the Mongols. Only very few of these special elements of *Yuan baihua* remain to be sustained in Chinese. Therefore, *Yuan baihua* should not be considered as a northern Chinese colloquial language during the Yuan Dynasty, as it did not record the actual spoken language used in northern China at the time. Rather, it is only an interlanguage of the Mongols learning Chinese.

Key words: *Yuan baihua*, language contact, linguistic characteristics

也說“漢兒言語”^{*}

江藍生

中國社會科學院

本文回述了“漢兒言語”的興起與主要研究資料；歸納綜合了“漢兒言語”的十一項語法特點，認為這些特點的形成都是阿爾泰語語法對漢語的影響與滲透；通過與現代某些西北方言的對比，認為西北方言是“漢兒言語”的活化石，具有很高的研究價值。

關鍵詞：漢兒言語，阿爾泰語，西北方言，語言接觸

1. 前言

1954年，著名日本漢語學家太田辰夫先生發表了〈關於漢兒言語——試論白話發展史〉一文，對於漢語史研究特別是漢語白話發展史研究來說，這是一篇見解獨到，具有重要意義的學術論文。在這篇文章以及早一年發表的另一篇論文〈《老乞大》的語言〉中，太田先生考證了宋以後“漢人、南人”之別，“漢兒、漢兒言語”所指，自北朝以來“漢兒言語”的發展軌跡以及反映“漢兒言語”的歷史文獻等。其考據的審慎和眼光的敏銳，即使在時隔半個多世紀的今天也很令人欽佩。但是由於研究資料等條件的局限，當時對於“漢兒言語”面貌的了解還是比較粗疏的，本文擬在前人研究的基礎上對時賢與筆者個人研究的成果加以綜合歸納，以推動這項對語言接觸與漢語白話史研究有重要意義的課題持續、深入地進行下去。

^{*} 十年前，我應邀在法國國家科研中心東亞語言研究所訪問研究，與貝羅貝先生合作撰寫了有關“漢兒言語”的比擬式的論文。貝羅貝先生早年畢業於北京大學中文系，我們是校友兼同行，在學術研究上多所切磋交流，啟發良多。多年來貝羅貝先生在繁忙的學術行政領導工作之餘一直堅持漢語語法史研究，且續有創獲，跟他有相同經歷的我，很能體會他的堅韌，而作為一個外國人研究漢語史，他所付出的辛苦卻是包括我在內的常人所難以想像的，對此我深表欽佩。在貝先生六十五歲華誕之際，謹以此文表示敬意和祝賀。

1987年我在日本京都大學人文科學研究所做訪問學者一年，其間參加了太田辰夫先生在京都產業大學主持的中古漢語研讀會，太田先生是“漢兒言語”研究的先行者，在此也一併表達內心的敬仰與懷念。

2. “漢兒言語”的興起及其主要研究資料

據太田先生考證，宋以後“漢人”多與“南人”並舉，這種分別並非自元朝始，早在北朝時期就已有“漢兒”與“吳兒”的分別。“漢兒”一詞出現在反映北方社會生活、歷史的文獻中，如北朝《折楊柳歌辭》：“我是虜家兒，不解漢兒歌。”《北史》中也頻頻出現“漢兒”一詞。文獻資料顯示，北朝胡漢之別並不嚴格以血統論，而是泛指中國北方的漢人和漢化了的北方民族，與此相對，南方的漢人被稱作“吳兒”。這一分別唐代依然沿用，日本所謂“漢音、吳音”也基於自南北朝至唐五代間這種“吳、漢”的區別。“漢兒”與“漢人”所指相同，只有稱呼上的雅俗之別，“漢人”是文的稱呼，“漢兒”是口語的。元朝時把漢人分為兩種，所謂“漢人、漢兒”，狹義的指金國統治下的漢民族，廣義的指除了北方漢人之外，還包括契丹、高麗、女真、渤海等北方已經漢化了的民族在內。所謂“漢兒言語”，是在北方漢人和漢化了的北方民族之間通行的共同語。

“漢兒言語”在中國北方的推廣和流行，歷史久遠，不自金元始。我們知道，自東漢末年開始，中國陷入了戰亂頻仍的時期，大規模的、頻繁的遷徙加速了北方不同民族成分之間的融合，也促進了漢語與操阿爾泰語的北方各民族語言的接觸。經過五胡十六國的戰亂後，中國北方長期由異民族統治，建立後趙的石勒為羯人，北魏、東魏、北齊都是鮮卑人建立的。

北魏孝文帝推行禁止說“北語”（鮮卑語等其他少數民族語言）的政策，從國家制度層面加速了鮮卑等族與漢族的同化，推動了漢語在北方各地的推廣和通行。另一方面，與北魏不同，北齊、北周推重鮮卑語，漢人出於謀生出世的需要也要學習鮮卑語言和文化。鮮卑等異族人在學講漢語時不可避免地會把母語的某些特點帶進來，而漢人在跟北方異族人交際時也會反過來盡量向這種不標準、不地道的漢語靠近，久而久之，連北地漢人的語言也會被這種不地道的漢語同化。可以設想，這一時期在北方通行的口語應是一種不標準的、帶有阿爾泰語某些特點的漢語，可稱之為某種程度的混合語，它就是“漢兒言語”的最初形態。

社會的動亂變遷同樣也加速了江南與中原語言文化的接觸，北齊顏之推《顏氏家訓·音辭》篇說當時漢語的狀況是“南染吳越，北雜夷虜”。永嘉之亂後，北方士族南渡（317），南朝建都建業，一時間江南士族紛紛仿效洛陽風氣，學習起洛陽話來。《宋書》卷八十一《顧琛傳》載：“先時宋世江東貴達者會稽孔季恭、恭子靈符、吳興邱淵之及琛吳音不變。”陳寅恪據此判斷：“其餘士人，雖本吳人，亦不操吳音，斷可知矣。”（見〈東晉南朝之吳語〉）周一良具體說明

了僑居揚州的中原人與吳人語言相互影響的過程：“蓋揚州之僑人不自覺中受吳人熏染，於中原與吳人語言以外漸形成一種混合之語音。同時揚州土著士大夫求與僑人沆瀣一氣，競棄吳語而效僑人之中原語音。然未必能得其似。中原語音反因吳人之模擬施用，益糅入南方成分。此種特殊語音，視揚州閭里小人之純粹吳語固異，視百年未變之楚音（此指北方語言）亦自不同。”“這種吳人口中的北語，隔了多少年之後，連僑人也受到同化，一樣地說那種不南不北的吳化洛陽語了。這種吳化洛陽語相當於藍青官話，因為是官話，所以只行於士族間。”（見〈南朝境內之各種人及政府對待之政策〉）這種語言接觸文化交融的情況在唐詩中也時有反映，例如：劉禹錫《武陵抒懷》詩：“鄰里皆遷客，兒童習左言。”王維《送李判官赴江東》：“封章通左語，冠冕化文身。”賈島《送人南歸》：“雖然南地遠，見說北人多。”張籍《永嘉行》：“北人避胡皆在南，南人至今能晉語。”所謂“左言”即指江東吳地的語言。“南染吳越，北雜夷虜”的結果是，中原漢語因地域的分割、人文語言環境的改變，在長期的發展變化過程中發生了分化。夾雜“夷虜”之語的即通行於北方的“漢兒言語”，糅染“吳越”之音的應是建業一帶深受吳語影響的漢語，從今天的杭州話可見其一斑。

研究“漢兒言語”，追索其發展軌跡，必然要借助各代的有關資料，然而，北朝時期的資料極少，只在《北齊書》、《北史》、《隋書》中留下一些片言隻語。宋遼金時期較前略多，最重要的為北宋徐夢莘的《三朝北盟會編》，其中卷二十《宣和乙巳奉使金國行程錄》有一段文字很能說明漢語作為北地各民族中介語的作用：

第三十三程。自黃龍府六十里至托撒字董寨，府為契丹東寨，當契丹強盛時，擒獲異國人則遷徙散處於此。南有渤海，北有鐵離吐渾，東南有高麗靺鞨，東有女真韋室，東北有烏舍，西北有契丹、回紇、黨項，西南有奚，故此地雜諸國風俗，凡聚會處諸國人言語不通，則各為漢語以証方能辨之。

太田先生徵引了這段文字，他說：在這樣遠離中國本土的地方（據考黃龍府為今長春以北的農安縣），北宋末年早已通行漢語的情況是值得注意的。他並且斷言：那種漢語必定是極其鄙俗的口語。但是這種鄙俗的漢語其詞彙、句法特點到底是什麼樣的？由於資料缺乏而難以詳知（太田先生文中舉了一些零星的詞語，如第二人稱代詞“你”、以及“捎空、歹”等）。十多年前我在閱讀宋人筆記時看到一則很有用的資料，已轉引於〈重讀《劉知遠諸宮調》〉一文，現再舉於下：

宋洪邁《夷堅丙志》卷十八《契丹誦詩》云：“契丹小兒初讀書，先以俗語顛倒其文句而習之，至有一字用兩三字者。頃奉使金國時，接伴副使秘書少監王補每為余言以為笑。如‘鳥宿池中樹，僧敲月下門’兩句，其讀詩則曰：‘月明裡和尚門子打，水底裡樹上老鴉坐’，大率如此。補，錦州人，亦一契丹也。”

這段記載包含的信息很多，其一，它反映了契丹人重視學習漢文化，其後代要學習唐代詩歌。其二，這是契丹人按契丹語詞序念誦賈島詩句的生動例子（契丹語屬蒙古語族），其中“和尚門子打”（賓語前置）、“樹上老鴉坐”（處所補語前置）這種“顛倒其文句”的讀法正是 SOV 句型的反映，乃是當地“俗語”的通例，所以可以用來教小孩子。其三，王補身為契丹人，卻每每拿這種鄙俗的漢語當笑話，可知他本人已經完全漢化了，也可以看出這種北地通行的俗語跟中原漢語的區別是相當明顯的。

女真自十二世紀崛起東北，國號“金”，定都上京會寧（後遷至燕京、開封），1125 年滅遼，1127 年滅北宋，疆域直至淮河以北，統治中國長達 120 年之久。女真在統一中國北部的過程中，曾多次進行大規模移民，讓漢人、契丹人等充實京師。為防止異族人“懷二三之心”，又命女真人遷入原遼宋舊地，數量達幾百萬。這種大換班加速了民族的融合。女真人入主中原後，為了維繫和加強自己的政權，也廣泛接受了包括科舉制度在內的漢民族文化，迅速地走向漢化。金海陵王完顏亮漢化尤深，能做漢詩，他登位後停止以女真同化天下的政策。在漢字楷體和契丹字體制的基礎上創制的女真字使用不廣，即使皇家子弟讀書，也是“每日先教漢字”（《金史》卷九十八《完顏匡傳》）。世宗大定（1232）以後，漢語文字逐漸成為官方語言，以至後來女真人竟對本民族語言“或不通曉”了（《金史》卷三十九《樂志》）。

北方各族之間的文化交流影響是相互的，在女真統治中國淮河以北達一個多世紀的時間裡，漢人也受到女真文化的同化，特別是燕雲地區的漢人，經過遼金數百年統治，已呈“胡化”傾向。此間出使金國的南宋官員親眼目睹了這種變化。樓鑰《攻媿集》卷一一一《北行日錄上》記載：金世宗大定九年（宋乾道五年），樓鑰從使金國賀正旦，沿途所見，北宋故地“只是舊時風範，但改變衣裝耳”。及過白溝，“人物衣裝，又非河北。北男子多露頭，婦人多耆婆。把車人曰：‘只過白溝，都是北人，人便別也。’”（白溝在涿州境內，原是遼宋界河）陸游在一首詩中也記述了這種物是人非的滄桑之變：“上源驛中搥畫鼓，漢使作客胡作主，舞女不記宣和妝，廬兒盡能女真語。”（《得韓無咎書寄使虜時

宴東都驛中所作小闕》；廬兒：家奴，僕從。）

這個時期可以作為考察“漢兒言語”面貌的文學作品為金代的兩種諸宮調：《劉知遠諸宮調》和董解元《西廂記諸宮調》。拙文（江藍生 1999b）舉出前者有些句式可以看出阿爾泰語的影響。其一，方位詞“前、前面”用如格助詞，表示動作的對象或方向，如：“每番只是人前走踢行拳，凶頑無賴”（對他人拳打腳踢）、“九州安撫，三翁前面，捧盞跪勸香醪”（向三翁勸酒）；其二，出現“像”義動詞“似”後置的比擬式，如“紫玉似顏色”、“傾盆也似雨降”，這種用法此前未見，與阿爾泰語比擬式相合；其三，異常語序：賓語前置、處所補語前置，如“莫想清涼傘兒打”（打清涼傘兒）、“唐末龍蛇未辨，布衣下官家潛隱”（官家潛藏在布衣人中）。這兩句異常的語序跟契丹小兒念誦的俗語“月明裡和尚門子打，水底裡樹上老鴉坐”如出一轍。但是，像這樣的資料不僅是絕無僅有，而且由於是戲曲韻文，在文體上也必然有許多局限。

比較能觀察到“漢兒言語”廬山真面目的白話資料在元代。元代白話文獻大體可以分成兩大類：

甲、純漢語的資料，如元雜劇、散曲、南戲和講史平話等。其中元人雜劇和散曲中可以看到一些受蒙古語詞彙和語法影響的痕跡，而平話類作品中則有相當的文言成分。

乙、直講、直譯體白話。其中有典章吏牘體白話如《元典章》、《通制條格》和蒙語直譯體白話碑文等。直講體即大臣用當時的口語給皇帝講解漢文典籍的白話講章，如許衡《大學要略》、《大學直解》，貫雲石《孝經直解》，吳澄《經筵講義》等。最重要的元代白話資料是朝鮮很早以來就廣泛使用的漢語教科書《老乞大》、《朴通事》，它的語言跟直講體十分接近，比直講體還要口語化，更能反映當時北方漢語口語的真實面貌。據太田先生考證，二書最初成書於元代，作者不明，應是高麗僑民。《老乞大》的著者應是住在東北瀋陽或遼陽一帶的人，《朴通事》的著者可能是常住北京的人，兩書應是用通行於東北至華北一帶的漢語口語寫成的。元朝滅亡以後，“漢兒言語”勢力衰退，南北通行官話。成宗因二書語言“乃元朝時語也，與今華語頓異，多有未解處”，遂下令“且選其能漢語者，刪改《老乞大》《朴通事》”。（《成宗實錄》十一年十月十九日條）於是在 1480-1483 年間由兩名中國人修訂，其後（16 世紀前半葉）朝鮮語言學家崔世珍為修訂本加上了諺解。諺解本可作為元、明間的資料，在 1998 年之前，國內外研究者所看到和利用的就是這兩種諺解本。

值得萬分慶幸的是，1998 年初，韓國慶北大學教授南權熙在大邱發現了一種與諺解本內容有異的《老乞大》，次年初，我從韓國博士生李泰洙那裡看到這本

書的複印件，經反覆細看比較後認定，這就是崔世珍在為《老乞大》《朴通事》作注的《單字解》和《老朴集覽》中所稱的“古本”或“舊本”，其內容編寫於元代無疑。古本《老乞大》的發現對研究“漢兒言語”具有十分重要的意義，因為作為教科書，它的語言應該比較標準，應是一種實際存在、且有廣泛適用性的語言，能夠反映當時通行於北方的漢語的真實面貌。在為古本《老乞大》的發現而震驚與慶幸的同時，也深為對二書研究作出傑出成績的太田先生沒能看到而深感遺憾。由於古本《老乞大》（以下簡稱“古本”）的發現，使我們可以把它跟反映明初語言的諺解本（以下簡稱“諺解”）和刊行於清代乾隆年間的《老乞大新釋》（1761年，以下簡稱“新釋”）和《重刊老乞大》（1795年，以下簡稱“重刊”）加以比較，從而發現“漢兒言語”與明清時期的漢語的異同，歸納出其特點等。下文對“漢兒言語”特點的歸納很多都借助於這種比對。

3. “漢兒言語”的特點

“漢兒”一詞北朝已見，但是“漢兒言語”一詞卻是在《老乞大》裡才看到的。古本《老乞大》中凡四處六見，例如：

- (1) 恁是高麗人，卻怎麼漢兒言語說的好？俺漢兒（上）學文書來的上頭，些小漢兒言語省得有。(1a)
- (2) 如今朝廷一統天下，世間用著的是漢兒言語。我這高麗言語只是高麗地面裡行的，過的義州，漢兒田地裡來，都是漢兒言語。(2a)

由此可知，元朝“一統天下”後口語中通行的是自北朝以來在口語中自然形成的“漢兒言語”，北方各民族包括作為統治者的蒙古族在內都是用這種語言跟漢人或他們相互間進行交際的。那麼這種語言有哪些特點，它跟中原地區的正統漢語之間有些什麼區別？根據中外學者目前研究的結果，大家都承認這種漢語在語音、詞彙、語法各方面都或多或少地受到阿爾泰語的影響，關鍵是要釐清這些影響到底有多大？具體表現在哪些方面？為此，必須要從類型學上把握漢語與阿爾泰語這兩種語言的主要特徵，才能準確而敏銳地捕捉有關的語言現象。語言接觸最深度的影響體現在語法層面，在語法上，阿爾泰語與漢語的主要區別在於：

- i. 詞序不同。漢語是 SVO 型語言，賓語一般置於動詞後面；而阿爾泰語是 SOV 性語言，賓語置於動詞前面。

- ii. 漢語是孤立語，沒有嚴格意義的形態變化，主要靠前置的虛詞表達各種語法關係和語法意義。阿爾泰語是黏著語，有格範疇，用後置的格助詞表示語法關係和語法意義。可以說，元代白話資料中的種種異常現象幾乎都跟這兩種語言的上述根本區別有關。

根據前輩、時賢也包括筆者個人的研究，本文擬將“漢兒言語”的主要特點初步歸納為以下十一個方面：(一) 複數詞尾的特殊用法；(二) 第一人稱複數代詞區分包括式與排除式；(三) 方位詞充當各類語法標記；(四) 賓語前置；(五) 句末使用表示肯定的語助詞“有”；(六) 判斷詞後置與疊加式判斷句；(七) 動詞“有”表示人或事物所在位置、處所；(八) “像”義動詞後置的比擬式；(九) 副詞的異常位置；(十) 在表示行為動作方向或場所的句子中，處所名詞前置，不用介詞；(十一) 比較句中用狀中式替代述補式。由於有關論文和專著對上述各點多有詳考，為避重複繁瑣，本文對有些特點僅點到為止，有的略作舉例性的簡介，有的則稍加說明。

3.1 複數詞尾的特殊用法¹

在金元白話資料中，複數詞尾主要用在指人名詞後表示多數。除此之外，還可以用在以下三種場合，顯示出獨有的特點。

3.1.1 指示代詞之後加複數詞尾“懣、每”

指示代詞“這（的）、那（的）”之後加複數詞尾“懣、每”，指“這些人、那些人”，有的相當於“他們”（宋金用“懣”，元代用“每”，明代用“們”）如：

- (3) 李洪義、李洪信，如狼虎；棘針視、倒上樹，曾想他劣缺名目，向這懣眉尖上存住。（《劉知遠諸宮調》2.32 [中呂調·木罌綏]）。

南宋《揮塵錄餘話·王俊首岳侯狀》中也一見：“我去後，將來必共這懣一處。”

¹ 例句引自江藍生(1999a)、祖生利(2001, 2002)等。

金元資料中“那每、這的每、那的每”指“他們”的例子如：

- (4) 那每殷勤的請你，待對面商議。（《董解元西廂記諸宮調》卷三）
- (5) 這的每寺院裡、房舍裡使臣休安下者。（元代白話碑）
- (6) 俺眾人與南人每一處商量來，那的每也則這般說有。（《元典章·戶部》）

3.1.2 可用於動植物乃至無生命的事物

漢語的複數詞尾僅用於指人名詞之後，不能用於動植物，更不能用於無生命的事物。但是在“漢兒言語”資料裡，表示複數的“每、們”等卻無此限制。用於非指人名詞後最常見的是牲口，如“馬每、驢每、頭口每、驢馬每”（見於古本《老乞大》、元人散曲、《元朝秘史》等多種資料，《老乞大諺解》、《朴通事諺解》“每”作“們”）。此外還可以用於無生命的事物之後，例如“窗隔每、椅卓每”（錢霖《般涉調·哨遍》），“角頭每”（《朴通事諺解》），“經文每、印板每”（元代白話碑），“草地每”（《元典章·兵部》）等。

3.1.3 單數人稱代詞、名詞後加“每、們”²

- (7) 你不須提起蔡伯喈，說著他每忒歹。（《琵琶記》第三十八齣）
- (8) 小姐，你揀個好財主每、好秀才每，或招或嫁，可不好那？（元曲《鴛鴦被》一折）
- (9) （牢子云）你是甚麼人？（正末云）叔待，孩兒每是個莊家。（元曲《黑旋風》三折）

這種用法的“每、們”已不表示複數，虛化為純粹的語助詞。據莫超（2008）考證，在《蒙古秘史》中有這種用法，說明元雜劇中的此類用法並非沒有來歷。

3.2 第一人稱複數代詞區分包括式與排除式

“咱、咱每/們”為包括式 (inclusive)，包括交談對方在內；“俺、俺每、我們”等為排除式 (exclusive)，不包括交談對方在內。

² 據莫超（2008）。

3.3 方位詞充當各類語法標記

由於漢語沒有後置的格助詞，而方位詞的語法位置總是在名詞、代詞、動詞等類詞的後面，相同的句法位置，使得方位詞被借來充當阿爾泰語的各類格助詞。

3.3.1 表示行為動作對象的標記“～行（上）、～根底”等³

宋元明白話文獻後置詞“～行”有兩種句式：

A 式：“動/介+N 行（+VP）”（向咱行、向誰行宿）

B 式：“N 行+VP”（君王行奏、誰行借）

二式的區別在於前面是否用動詞或介詞。在金元以前的文獻中一般都用動/介詞（A 式），如：

(10) 若言無意向咱行，為甚夢中頻夢見？（柳永詞；無意到我這裡）

(11) 低聲問向誰行宿，城上已三更。（周邦彥詞；在哪裡過夜的）

但在金元白話文獻中前面不用動詞，介詞的 B 式很常見，例如：

(12) 動不動君王行奏。（《金錢記》三折；向君王奏）

(13) 我官人行說了。（《魔合羅》三折；我對官人說了）

(14) 嫂嫂母親行更加十分孝。（《替殺妻》二折；嫂嫂對母親更加孝順）

(15) 大師行深深的拜了。（《西廂記》一本二折；對著大師深深地拜了）

(16) 床頭金盡誰行借？（元·喬吉《山坡羊》散曲；向誰借）

(17) 你這等賊心賊肝，我行須使不得。（《水滸傳》十四回；對我可使不得）

以上各例不用表示對象的動詞或介詞“向、到”等，逕用方位詞“行（上）”表示動作行為的對象。這種用法跟《蒙古秘史》漢語對譯完全一致：

³ 拙文（江藍生 1998）考證“行”為“上”的輕讀音變。

- (18) 音譯：成吉思 合罕 亦禿兒堅 突兒 兀祿客 列列勒敦
對譯：太祖 皇帝 人名 行 不 說話
(成吉思汗不跟亦禿兒堅說話。“～行”對譯蒙語位格助詞)

《蒙古秘史》漢語對譯中還用“行”對譯與格助詞和賓格助詞，此不贅引。可以認為，A 式是漢語自古以來就有的，B 式是元代漢語受蒙古語格範疇影響而產生的新興句式。

除了用“行(上)”，《老乞大》中還用“根底”表示行為動作跟隨的對象，例如：你誰根底學文書來？(古本、諺解本 1a10、2a9)，而乾隆二本分別作：“你跟著誰學書來著？”(新釋 1b2)“你跟著誰學書來？”(重刊 2a9)

3.3.2 表示原因的標記“～上頭/～上”⁴

- (19) 癭驚上離了茲(慈)親，悞然地兩腳到您莊院。(金《劉知遠諸宮調》1.11 [商調·拋球樂])

意思是：因為心情鬱悶離家出走，不知不覺來到您家莊院。“癭驚上”的“上”即表原因。

請看《老乞大》四個本子中表示原因的不同說法：

- (20) 古本：從年時天旱，田禾不收，饑荒的上頭，生出歹人來。(8a6)
諺解：從年時天旱，田禾不收，饑荒的上頭，生出歹人來。(24a9)
新釋：因去年年成荒旱，田禾沒有收成的上頭，就生出這些歹人來了。(9a5)
重刊：因去年年成荒旱，田禾沒有收成，就生出歹人來了。(23b)

《古本》、《諺解》本中用後置於小句的“～上頭”表示原因，意思是“因為去年天旱沒有收成、鬧飢荒，所以生出壞人來”。但此句在《新釋》裡被改作“因 X 的上頭”，前面加上了介詞“因”。到了《重刊》本，刪除了“的上頭”，只在句前加了個“因”。其變化可圖示為：X 的上頭——因+X 的上頭——因 X。另一處情況相仿：

⁴ 據太田辰夫(1953)、江藍生(1999a)、李泰洙(2003)。

- (21) 古本：底似的漢兒言語說不得的上頭，不敢言語。(15a1；底似：抵死。用為程度副詞，相當於“十分”)

諺解：他漢兒言語說不得的，因此上不敢說語（話）。(46b7)

新釋：他不懂漢人說的官話，故此不敢說話。(17a3)

重刊：他不懂中國的話，故此不能說話。(47b2)

《古本》作“X 的上頭”，《諺解》本作“因此上”，《新釋》和《重刊》只用連詞“故此”。前後的變化為：X 的上頭——因此上——故此。其中“因此上”的結構為：“因 此上”，後來重新分析為“因此 上”。

由上可以看出元代“漢兒言語”用方位詞“上頭”或“上”表示原因（後置式），明初開始用“因+X 的上頭/上”（疊加式），清代基本改用了前置式“因”或乾脆另用連詞“故此”。

3.3.3 表示憑藉、工具的標記“～裡、～根底”⁵

- (22) 長生天氣裡，大福蔭護助裡皇帝聖旨。(《元代白話碑》51 頁)

此句意思是托著上天大福蔭保佑的皇帝的聖旨，“裡”用在名詞後，表示憑藉、托福的“托”。到了明初，一般在前面加上動詞“托著”，用疊加式“托著+X+裡”：

- (23) 孩兒在都，托著爺娘福蔭裡，身已安樂，不須憂念。(《朴通事諺解》10b)

古本：車子水裡去呵，水裡行不得，舡裡載著有。(31b4)

諺解：車子水裡去時，水裡行不得，須用船裡載著。(39b7)

《古本》作“舡+裡”（後置式），《諺解》作“用+船+裡”（疊加式）。

元人雜劇中也有用例：

- (24) 這手帕剪了做靴簷，染了做鞋面，撶了做鋪持。一萬分好待你，好覷你，如今刀子根底，我敢割得來分零麻碎。(元刊雜劇《詐妮子調風月》二折[上小樓])

“刀子根底”即“用刀子”，“根底”用在名詞後面表示工具格。

⁵ 據李泰洙(2003)、祖生利(2000)。

3.3.4 表示動作的起點、方向、場所的“～裡、～前”⁶

這些語法功能在漢語裡通常用介詞“從、從打、到、往、在”等表示，但在“漢兒言語”資料中大量出現不用介詞、僅用方位詞“裡、前”表示的用法。如：

- (25) 恁高麗田地裡將什麼行貨來？
（古本《老乞大》20b；你從高麗地面帶了什麼貨物來了？）
《諺解》同此，僅把“恁”改為“你”（2a）。
(26) 咱每則這後園裡大（去）淨手不好那？
（同上 11a；咱們就到這後園子裡解手豈不好？）
《諺解》同此，僅把“每”改為“們”，“則”改為“只”（33b）。
(27) 到晚師傅行撒簽背念書。
（同上 1b；晚上在師傅跟前撒簽背念書）
《諺解》同此，僅把“行”改為“前”（3a）。

3.4 賓語前置句⁷

如前所說，阿爾泰語的賓語通常都置於動詞之前：OV、SOV；漢語一般則是置於動詞之後：VO、SVO。賓語前置的情況在金元白話文獻中十分常見，前已舉《劉知遠諸宮調》例，下面再以《老乞大》各本為例：

- (28) 古本：布帳子疾忙打起者，鋪陳整頓者，房子裡搬入去者。鞍子轡頭，自己睡臥房子裡放者，上頭著披氈蓋者。那的之後，鑼鍋安了者，疾忙茶飯做者。（32a2）
諺解：把帳房忙打起來，鋪陳整頓了，搬到帳房。鞍子轡頭，搬到自己睡處放下，上頭把氈子蓋了。然後埋好了鑼鍋，疾忙做茶飯。（37b1）

稍加對比就可看出，《古本》的 OV 句在《諺解》本中大多改為 VO 句，有的改為“把”字句。如“布帳子疾忙打起”改為把字句“把帳房忙打起來”；“鑼鍋

⁶ 例句引自李泰洙 (2003)。

⁷ 例句引自李泰洙 (2003)。

安了”改爲“埋好了鑼鍋”；“疾忙茶飯做者”改爲“疾忙做茶飯”等。以下幾組例子都是《古本》爲OV句，其他各本改爲VO句的例子：

- (29) 古本：咱每爲父母心盡了，不曾落後。(31a)
諺解：咱們盡了爲父母的心，不曾落後。(38b)
- (30) 古本：咱每結相識行呵，休說那你歹我好，朋友的面皮休教羞了。
(32a7)
新釋：咱們會相與人，不要說你歹我好，不要羞了朋友的面皮。
(37b6)
- (31) 古本：為什麼這般的歹人有？(8a5)
諺解：為什麼有這般的歹人？(24a6)
新釋：為什麼有歹人呢？(9a4)
重刊：為什麼有歹人？(24a4)
古本：賣的好弓有麼？(27b10)
諺解/新釋/重刊：有賣的好弓麼？(27b4, 32b1, 29a2)

在否定判斷句中《古本》否定判斷詞“不是”位於句末，其他三本改爲前置：

- (32) 古本：俺買呵，買一兩個，自穿的不是，一發買將去要覓些利錢。
(36a5)
諺解：我買時，不是買自穿的，一發買將去要覓些利錢。(54a2)
新釋/重刊：我買去，不是自家穿的，……。 (41b9, 56a3)

黃征 (1996) 舉出《敦煌變文集》中所見蕃語痕跡：五毒嗔心便起 | 方便與舜 | 王子此度且放 | 李陵蕃中在 (另一處作“陵在蕃中”)。黃謂之“倒裝句”，其實爲OV句式。

3.5 句末使用表示肯定語助詞“有”⁸

《老乞大集覽》卷上“漢兒人有”條下云：“元時語必於言終用有字。如語

⁸ 據太田辰夫 (1953)、李泰洙 (2003) 等。

助而實非語助，今俗不用。”這種句末助詞“有”主要出現在陳述句、判斷句句尾，表示肯定語氣。用於陳述句的例子如：

- (33) 這鈔裡頭真假俺高麗人不識有。（古本《老乞大》37b）
- (34) 家裡都好嗎？都安樂好有。（同上 20b）
- (35) 一箇手打呵，響不得有；一箇腳行呵，去不得有。（同上 31b）
- (36) 在下的作亂呵，有罪過有；眾人中爭鬥呵，有傷損有。（《孝經直解·紀孝行章第十》）

用於判斷句中的更爲多見，例如：

- (37) 你的師傅是什麼人？是漢兒人有。（古本《老乞大》2b）
- (38) 這段子外路的，不是服地段子有。（同上 27b）
- (39) 孝道的勾當，是德行的根本有。（《孝經直解·開宗明義章第一》）
- (40) 金銀是鈔的本有。（《元典章·戶部·卷六·存留鈔本》）
- (41) 將他一箇懷孕的婦人拿住，問他：“你是什麼人氏有？”（《蒙古秘史》）

句末語氣助詞“有”在明代文獻中較少看到，《老乞大諺解》中有一例，《朴通事諺解》中見到兩例，應是明初修改未盡所遺。《朴通事諺解》中的兩例是：

- (42) 無職時有什麼事？律條裡明白有。（下 16a；律條裡明明白白寫著呢）
- (43) 《西遊記》熱鬧，悶時節好看有。（下 17a）

這種僅見於元代白話文獻的句末助詞“有”，用於陳述句的是語氣詞，而用於判斷句的則是蒙古語句末表判斷的助動詞的直譯，當它融進漢語時被重新分析爲句末語氣詞，崔世珍說“有”“如語助而實非語助”恐怕與“有”的兩種不同性質有關。

3.6 判斷詞後置與疊加式判斷句⁹

漢語的判斷句經過了從古代無係詞與有係詞兩類並存，到漢晉以後以有係詞

⁹ 據江藍生(2003)。

爲主的過程；係詞從“惟、爲、是”並用變化爲主要用“是”；“主語＋是＋表語”很早就成爲漢語判斷句的典型格式而固定下來。但是，在元人雜劇和元末明初白話小說《水滸傳》中有甲乙兩種用於自我介紹的特殊判斷句，其共同點是判斷詞“便是”位於句末。

甲式：“S+N（的）＋便是”：

- (44) 老夫，王員外便是。（元刊《小張屠》）
- (45) 貧道，呂岩便是。（元刊《鐵拐李》）
- (46) 貧道，陳搏先生的便是。（元刊《陳搏高臥》）
- (47) 小生孫華，小字蟲兒的便是。（《元曲選·殺狗勸夫》）

乙式：“S＋是＋N（的）＋便是”，可稱爲疊加式判斷句：

- (48) 貧道是司馬德操的便是了。（元刊《單刀會》）
- (49) 小人便是白虎山前莊戶孔亮的便是。（《水滸》五十八回）
- (50) 某乃宋江便是。（《元曲選·李逵負荊》）

這兩類特殊判斷句的出現跟蒙古語的判斷句句法有關。通過《蒙古秘史》的對譯和總譯可以看出，甲類是蒙古語判斷句的直譯：

- | | | | | | |
|------|-----|---|----|----|-----------|
| (51) | 音譯： | 必 | 王罕 | 備由 | （卷七 P446） |
| | 對譯： | 我 | 人名 | 有 | “S O 有” |
| | 總譯： | 我 | 是 | 王罕 | “S 是 O” |

“有”是蒙古語句末表示肯定的助動詞的直譯，對譯句“我王罕有”總譯爲“我是王罕”。所以元代“漢兒言語”中甲類特殊判斷句就是把“我王罕有”的“有”改爲漢語的“便是”而來的。用雙音節的“便是”結句更合漢語的韻律。

乙類疊加式判斷句是漢語判斷式“S 是 N”與蒙古語判斷式“SN 有”相疊加融合而成的。即：

S 是 N + SN 有——S 是 N 有

“S 是 N 有”判斷句的用例已在上文舉過，在使用中爲了向漢語靠攏，又把句末的助動詞“有”改爲漢語的判斷詞“便是”，這樣就產生了疊加式“貧道是司馬

德操的便是”這類特殊判斷句了。

3.7 動詞“有”表示人或事物所在位置、處所¹⁰

漢語存在義動詞爲“在”和“有”，其區別是表示人或物存在的處所、位置用動詞“在”（可帶處所賓語），如“老李在嗎”，“老李在家”；表示有無用“有”（不帶處所賓語），如“屋裡有人”，“河裡有隻小船”。但元代白話資料中該用“在”的地方多見用“有”的：

- (52) 大都有大小衙門裡官人每令使每要了肚皮的，交臺裡首者；外頭有的，按察司首者。（《元典章·刑部十》）

意思是：在大都（即後來的北京）的大小官吏中向令使們索賄的，由臺裡舉發；在大都之外的索賄官吏由按察司舉發。“大都有”指在大都的，“外頭有的”指在京城之外的。

如果問人在不在某處，應該用“在”不用“有”，而元代白話資料中多見用“有”者，例如：

- (53) 不免叫一聲，店主人有麼？（元曲《盆兒鬼》一折[金盞兒]白）
(54) 這店裡賣毛施布的高麗客人李舍有麼？（《古本》20a、《諺解》63b）

此句本是詢問“李舍在不在這店裡”，但清代《新釋》本卻按漢語理解爲“這店裡有沒有李舍”，因此改爲：這店裡卻有賣毛藍布的朝鮮客人李舍麼？(23a)

漢語回答“在不在”問句時應該用“在”或“不在”，而不用“有”或“沒有”，但是元代資料中多見用“有”回答的用例。試比較《老乞大》四個本子對處所問句作答的差異：

- (55) 古本：店在那裡？兀那西頭有。(19b)
諺解：店在那裡？那西頭有。(19b)
新釋：店在那裡？在那西頭有。(19b)
重刊：店在哪裡？在那西頭。(19b)

¹⁰ 例句引自李崇興等(1998)、李泰洙(2003)等。

四本中只有《重刊》用“在”回答，符合漢語通例。《新釋》用“在……有”疊加式作答，應是為遷就《古本》而作的變通，嚴格地說那種回答與問句不完全對應。

“有”表位置或處所還經常用在讓某人待在某處等著場合，如元雜劇中的“一壁有者”（在一旁等著）：

(56) 子龍，且一壁有者。（《隔江鬥智》二折白）

從《老乞大》各本的異文也能看出“有”是元代的特殊用法：

(57) 古本： 你都這裡有者，我稅契去。(25b)

諺解： 你都這裡等候著，我稅契去。(18b)

《新釋》和《重刊》都改為“你們在這裡等候著”(29a, 19a)。

3.8 “像”義動詞後置的比擬式¹¹

漢語的比擬式在從先秦到唐宋這段漫長的歲月裡，基本上都是“如 X 一般”這樣的句式（全式），也有少數省去“像”義動詞的“X 一般”句式（簡式）。全式、簡式的語法功能幾乎清一色做謂語，從五代開始，才見到為數有限的簡式做狀語的例子，可看作複句的緊縮。例如：

(58) 大人才見兩僧，生佛一般禮拜。（《祖堂集》4.38；可擴展為：像見了生佛一般，恭敬禮拜）

(59) 爛冬瓜相似變將去，土裡埋將去。（《五燈會元》395 頁；可擴展為：像爛冬瓜相似，一日日變將去）

也就是說，在十二、十三世紀以前，漢語的比擬式在形式和語法功能上都是比較單一的。

但是到了金元時期，白話文獻中出現了功能和形式都與前代不同的新的比擬式：X+似+NP/VP。其特點是：

¹¹ 據江藍生(1992, 1999a)。

其一，“X+似/也似”前面一律不用“像”義動詞，但後面一定要有 NP 或 VP。

其二，“X+似/也似”的語法功能是做修飾語，或修飾 NP 做定語，或修飾 VP 做狀語。例如：

- (60) 把山海似深恩掉在腦後。（金《董解元西廂記諸宮調》卷二）
- (61) 虎狼似惡公人，撲魯推擁廳前跪。（元曲《魔合羅》4折）
- (62) 做娘的剜心似痛殺殺刀攢腹，做爺的滴血似撲簌簌淚滿腮。（元曲《看錢奴》2折）
- (63) 絳雲也似丹臉若頻婆。（元曲《西蜀夢》3折）
- (64) 這鏝刀是俺親眷家的，不付能哀告借將來，風刀也似快。（古本《老乞大》6a）
- (65) 閻裡吉思戲言也似說來也者。（《元典章·刑部》）
- (66) 宋江探頭看時，一隻快船飛也似從上水頭搖將下來。（《水滸》37回）

金元比擬式與前代的區別在於：

(i) 句式結構不同

前代：“如+X+一般”（全式），

“X+一般”（簡式）；

金元：“X+似+NP/VP”：前面不用“像”義動詞；後面必須有中心語 NP 或 VP。

(ii) 語法功能不同

（全式）“D+X+Z”作謂語。

（簡式）“X+Z”作謂語，少數作狀語（複句緊縮而成）。

“X+似+NP” “X+似”作定語，不能作謂語。

“X+似+VP” “X+似”作狀語，不能作謂語。

金元“X+似+NP/VP”比擬式的來源與阿爾泰語如蒙古語的比擬表達式有關。蒙古語的比擬表達方式是在名詞、代詞或少數動詞後面加上後置詞 *metü*，然後再接中心語：

- (67) ral metü ularan tur
火 似 紅 旗 （像火一樣紅的旗子）

- (68) mori nisqu metü qaruluna
馬 飛 似 跑 (馬像飛一樣地跑)

至於“似”又作“也似”，拙文（江藍生 1992 等）認為是爲了避免歧義而添加的助詞。上面兩句如果直譯，就成了“火似紅旗”“馬飛似跑”。直譯有歧義，“似”前加上助詞“也”作“火也似紅旗”“馬飛也似跑”就不會有這種誤會了。

3.9 副詞的異常位置¹²

3.9.1 程度副詞遠離被飾成分

漢語的程度副詞修飾形容詞時，一般緊位於該形容詞前，而在反映“漢兒言語”的資料中，程度副詞卻不與被飾成分相鄰，例如：

- (69) 那般時，馬每分外喫的飽。（古本《老乞大》7b2；喫的分外飽）
(70) 這裡有五虎先生，最算的好有，咱每那裡算去來。（同上，39a8；算的最好）
(71) 咳，這官人好尋思、計量大。（《朴通事諺解》22b8；尋思、計量大）
(72) 徐五的徒弟李大，……那廝十分做的好。（同上 25b9；做的十分好）

3.9.2 “程度副詞＋禁止詞”，與漢語詞序相反

程度副詞與禁止詞共現時，漢語通常的詞序是禁止詞在前，而《老》《朴》與漢語詞序相反。

- (73) 古本：你底似的休早行，俺聽得前頭路澀有。（8a；底似：抵死。用如“很、十分”）
諺解：你十分休要早行，我聽得前頭路澀。（24a3）

在清代的《新釋》《重刊》中，上句話被改爲“你們不要十分早行，……。”調

¹² 據李泰洙、江藍生 (2000)。

整了程度副詞與禁止詞“休”的位置。

《朴通事諺解》中也有與《老乞大》相同的情況，這說明“漢兒言語”程度副詞與禁止詞的語序跟現代漢語相反：

(74) 你十分休小看人。(36b)

(75) 我再沒高的了，官人十分休駁彈。(37b8)

這種異常的詞序是因為：阿爾泰語這類句子的詞序是“程度副詞（＋賓語）＋動詞＋否定詞”，漢語則為“否定詞＋程度副詞＋動詞（＋賓語）”，在接觸中為兼顧兩種語言的要求被調整為“程度副詞＋否定詞＋動詞（＋賓語）”。

3.10 處所名詞前置，前面不加介詞

在表示行為動作方向或場所的句子中，如果處所名詞前置，漢語通常要前加介詞，而元時經常不用介詞。¹³

(76) 古本： 主人家，別處快鋤刀借一個去。

諺解： 主人家，別處快鋤刀借一個來。

《新釋》、《重刊》把這句話改為：“主人家，你可往別處借一把快鋤刀來。”在處所詞“別處”前面加上了介詞“往”。再如：

(77) 古本： 咱每鋪裡商量去來。(26b)

諺解： 咱們鋪裡商量去來。(21b)

《新釋》、《重刊》把這句話改為：“咱們且到鋪裡商量去。”(30a, 22b)，在“鋪裡”前面加上了介詞“到”。

在這類句子中，處所名詞前不加介詞，有時會產生歧義，分不清處所名詞是指目的地還是出發地。例如：

(78) 我先番大都來時，……。 (古本《老乞大》8a)

¹³ 據江藍生、李泰洙(2000)。

根據上下文，“大都”應是終點，但是《新釋》、《重刊》誤解為起點，把這句改為“我前番從北京來時……”。

(79) 你稅了契時，到明日，俺下處送來。（古本 25b、諺解 18a）

“俺下處送來”，漢語標準說法應該為“送到俺下處來”，“俺下處”作處所補語。到了《新釋》本裡正是改為：“你稅了契，明日送到我下處來罷。”（29a；《重刊》本同此，只是句末沒有“罷”字）

另外，《古本》“房子裡搬入去”（32a）《諺解》改為“搬到帳房”（37b）；《古本》“自己睡臥房子裡放者”（32a），《諺解》改為“搬到自己睡處放下”（37b）等，都是把前置的處所詞後移為處所補語的例子。

3.11 比較句中用狀中式替代述補式¹⁴

古本和諺解本《老乞大》中有一些狀中表達式在漢語裡通常用述補式表達，例如：

(80) 這橋便是我夜來說的橋，比在前畝好有。（古本 11a10；‘比從前好得很’或‘比從前好多了’）

《諺解》後一句作“比在前十分好”（35a），也是狀語表達式。這種現象的產生應該跟蒙古語裡沒有像漢語那麼發達的補語表達方式有關，所以漢語用補語來表達的，蒙古語大都用狀語來表達（參看道布 1983）。¹⁵

通觀上面列舉的材料，我們不得不承認：“漢兒言語”在一定程度上偏離了漢語的正宗。這種偏離其實都是阿爾泰語的語法範疇和 SOV 型詞序對漢語的影

¹⁴ 據江藍生、李泰洙 (2000)。

¹⁵ 道布 (1983) 說蒙古語沒有像漢語那樣發達的補語表達式，往往用狀中式替代述補式。例如：

這毛太生長（這傢伙的毛長得太長了）

米太放入（米放得太多了）

啊，你好修理（啊，你修理得真好）

這花真美好縫，嘖嘖（這朵花繡得真漂亮）

好冷啊，厲害冷有啊（冷得厲害呀）

今天厲害熱有（熱得厲害）

嗨，你又錯做（你又做錯了）

泥稀和需要（泥要和得稀一點）

響與滲透所致。其中複數詞尾的泛用、第一人稱代詞複數區別包括式與排除式、方位詞充當格標記等是語法範疇的影響與滲透；而其他方面如賓語前置、係詞後置的判斷句、“像”義動詞後置的比擬式、處所名詞前置不用介詞以及副詞的異常位置等都是 SOV 型詞序的複製或結構再造。在漫長的歷史進程中，北方說阿爾泰語的民族在學說漢語的過程中把母語中的格範疇、詞序等帶到漢語口語中來；另一方面，長期生活在北方、跟北方各民族密切接觸的漢族也會自覺不自覺地向這種不太標準的漢語靠攏。這兩方面共同作用的結果，使得“漢兒言語”在語法上形成了上述各種特色。

“漢兒言語”在語法上的特點肯定不止以上諸點，這裡只是舉其大要者，隨著資料的開拓和研究的深入，相信對它的了解會更加全面，也更加細密。

4. 現代西北方言——“漢兒言語”的活化石

隨著元蒙王朝的滅亡，漢族在明朝重新執掌政權，“漢兒言語”走向衰落，迅速被共同語官話替代。有清一代，女真族的後裔滿人入主中原（清太宗諱言北宋末年由女真、金朝挑起的“靖康之變”，刻意模糊本民族的歷史，把國號“金”改為“清”，把“女真”改為“滿族”），但如前所說，早在金代女真人的漢化程度就很高，對本民族語言“或不通曉”者大有人在，歷明至清就更不用說了。滿族初入關時採用滿漢雙語制，到嘉慶時已通用漢語，“漢兒言語”失去了復辟的土壤。

“漢兒言語”雖然在東北華北一帶逐漸銷聲匿跡，但作為歷史上的真實存在，它還是留下了自己的遺跡。例如複數詞尾用於牲畜的特殊用法在河北藁城方言裡遺存（見楊耐思、沈士英 1958）；東北、華北、西北等廣大北方地區都保留了第一人稱複數代詞區分包括式與排除式的特點，跟南方（除極少數方言點外）無此區分形成鮮明的對照；¹⁶ 疊加式判斷句“俺是某某的便是”在戲劇、曲藝等文藝作品中時而使用；用動詞“有”表示“在”的用法保留在點名時的應答句中；“像”義動詞後置的比擬式沒有原樣保留，而是在“似”後添加助詞“的”，把動詞“似”改造為助詞“似的”後留存下來。如此等等，不一而足。

¹⁶ 萬波(1996)《方言》第2期介紹江西安義方言包括式：俺“我爾”的合音，指代說話者和聽話者雙方，只表雙數；表多數用“俺大家”。排除式：“我儕”不包括聽話一方，與北京話“我們”也可以是包括式不同。

北京話： 咱們（包括）

我們大家（多數，包括）

我們（排除/包括）

安義話： 俺（我和你，雙數，包括）

俺大家（多數，包括）

我儕（排除）

真正能集中反映“漢兒言語”遺風流緒的是甘肅、青海、寧夏、陝西、新疆等西北地區的漢語方言。把古代“漢兒言語”跟西北地區方言語法的描寫資料相比較，能夠清楚地看出二者之間的血緣關係，我們完全可以說：西北方言就是“漢兒言語”的活化石。下面擬用上文概括的特點作為參照點來觀察西北方言與“漢兒言語”之間的淵源關係。

4.1 關於複數詞尾的泛化用法

莫超(2008)介紹在甘肅省蘭州市區及所轄縣區、洮河下游、臨夏州等地區複數詞尾“們”可用在非指人名詞如牲口和無生命的東西後面，如“牲口們、樹們、衣裳們、米們、油們、酒們、這些磚們、那些新戲們”等。而且，“們”也能用在單數名詞後面，如“熱頭們（太陽）、月亮們、天氣們、這個老漢們、這本書們”等。可見，金元“漢兒言語”複數詞尾的用法至今仍較完整地保留在甘肅一些地區。過去學者們曾對元代文獻中的這類特殊用法是否反映當時的實際口語還保有疑惑，甘肅方言裡的上述用法對此作出了肯定的回答。

4.2 關於賓語前置

賓語前置的句式普遍存在於甘肅、青海、新疆等西北方言中。受事賓語後面分不帶格標記和帶格標記兩種類型。賓格標記多用“哈”（黃伯榮 1996），也有用“啦”或“啊”（讀 za；莫超 2008）的。

王森(1993)介紹甘肅臨夏方言的“賓+動”式：

- (81) 我箱子揭開者三塊錢拿出來了。（我打開箱子拿出了三塊錢）
- (82) 我我的親人想者。（我想我的親人）
- (83) 他他的成績知道了。（他知道了他的成績）
- (84) 我今個才他哈認下了。（我今天才認識了他）

動詞“有、沒有”的賓語也前置，如：

- (85) 你們學校幾個老師有喃？（你們學校有幾個老師？）
- (86) 解放前河州城裡醫院沒有，西醫沒有。
- (87) 南疆方言：我你這種女婿就沒有瞧上。（我就沒瞧上你這種女婿；引自石毓智 2008）

莫超 (2008) 指出，臨夏話裡如果語境中施事、受事明確，則不需要加賓語標記，如“警察小偷抓下了”也可以說“小偷警察抓下了”，不會產生誤會。如果可能產生歧義，就要在受事後面加上賓語標記“啦”或“啊”，例如：

- (88) 我他啦電話早打過了。（我給他早就打過電話了）
- (89) 我啦後人們錢給的多著哩。（兒女們給了我好多錢）

賓語前置的否定句，否定動詞位於句末，作“OV 的不要”、“OV 的不是”：

- (90) 新開荒地肥料上的不要。（新開荒的地不要上肥料）
- (91) 賬那麼個算的不是，致麼個算的是。（賬不是那麼算，是這麼算）

張成材 (1998) 介紹西寧方言有“賓+動”式：

- (92) 你茶喝，饅饅吃。（你喝茶，吃饅頭）
- (93) 爸爸一個洋糖給了。（叔叔給了一塊水果糖）

黃伯榮 (1996) 描寫青海話受事賓語帶格標記“哈”：

- (94) 我開水哈喝了（我喝了開水）
- (95) 我你哈沒見（我沒見你）

處所賓語也位於趨向動詞“來/去”之前：

- (96) 西寧話：家上海去過，我還沒去過。（他去過上海，我還沒去過；張成材 1998）
- (97) 陝西延川話：你學校去格來？（‘你去過學校了嗎？’“來”為語助詞；張崇 1990）

總的來看，受事賓語後帶不帶格標記並沒有嚴格的規定，但是可以看出有不帶格標記的趨勢，只有在可能產生歧義的情況下賓格標記才是必要的。這種趨勢的原因應該是普通話的影響。

另外，西北方言裡有些“把”字句的底層實際是 OV 式，“把”字相當於前置的賓格標記，這是“漢兒言語”在漢語詞序的影響下所作的結構調整。例如：¹⁷

- (98) 青海話： 我把你沒見。（也可以說：我你哈沒見。）
- (99) 西寧話： 到這會兒著你還把我打，把我罵。
- (100) 青海話： 王秘書把介紹信沒開。
- (101) 蘭州話： 我把他們的名字知道。
- (102) 西寧話： 我把你沒認得。
- (103) 南疆方言： 我把你還不知道？
- (104) 新疆話： 把吐魯番的葡萄任何人都喜歡。

在這些“把”字句中，謂語動詞有的是否定式（沒見、沒開、沒認得、不知道），有的是光杆動詞（打、罵、開），有的是低及物性的動詞（知道、認得、喜歡），標準的漢語在上述情況下一般都不會用“把”字句。

4.3 關於特殊判斷句

目前尚未掌握西北方言中是否有“S+N 是”式判斷句，但甘肅臨夏方言中可見與其相對的否定式“S+N 的不是”（據謝曉安、張淑敏 1990，王森 1993）：

- (105) 我謙虛的不是，也保守的不是。（我不是謙虛，也不是保守）
- (106) 我新衣裳沒有的，有了沒穿的不是。（我是沒有新衣裳，不是有而不穿）

這跟古本《老乞大》中“尋常的不是，有玲瓏花樣的”（33b）、“俺買呵，買一兩個，自穿的不是，一發買將去要覓些利錢”（36a）完全一樣（《諺解》本改爲：“我買時，不是買自穿的”（54a））。

疊加式判斷句“S+是+N+就是”存留在臨夏方言的疑問句中：

- (107) 你是學生就是啦？（你是學生吧？）

¹⁷ 引自石毓智(2008)。

- (108) 這個車子是你的就是啦？（這輛車子是你的吧？）
(109) 兀個年輕人是皮革廠的廠長就是啦？（那個年輕人是皮革廠的廠長吧？）

這說明“漢兒言語”的疊加式判斷句的使用已有了限制，僅存留在疑問句中；把元明時期的“便是”換成了口語詞“就是”，也符合時代的特色。

與肯定式不同，疊加式判斷句的否定式使用比較自由，不限於疑問句：

- (110) 阿哥是怕人的人不是。（阿哥不是怕人的人）
(111) 這個水是冰的不是。（這水不是涼的）
(112) 你是工人不是□？（你不是工人吧）

4.4 關於副詞的異常位置

4.4.1 副詞遠離被飾成分

莫超 (2008) 介紹在甘肅各地方言中以下二例有歧義：

- | | | |
|-------|------------|-------------------------------------|
| (113) | 人都還沒來 | 北京話：一個也沒來
甘肅話：A 一個也沒來；B 人還沒都來 |
| (114) | 這次考試都不要給及格 | 北京話：全都不給及格
甘肅話：A 全都不給及格；B 不要都給及格 |

歧義句的造成是由於表示範圍的副詞“都”沒有緊挨在它所修飾的動詞前面，如果把第一句改爲“人還沒都來”、“這次考試不要都給及格”就不會產生歧解了。

臨夏話“這個還不算，還比這麻達的事情有呢”（比這麻煩的事情還有呢）這句話符合阿爾泰語“狀語＋賓語＋動詞”的詞序。

4.4.2 “程度副詞＋否定詞＋VP”表示程度輕

當程度副詞與否定詞連用表示程度不重時，北京話用“不＋程度副詞＋VP”，如“不很疼”表示不太疼。莫超 (2008) 指出甘肅話與此相反，“程度副詞＋不＋VP”表示“不大VP”、“不太VP”：

(115) 這個人很不說話。(不大說話)

(116) 他很不喝酒。(不大喝酒)

(117) 我的槍法很不準。(不太準)

(iii) “狀語+禁止副詞/疑問副詞+VP”

甘肅中部、銀川、西寧等地在否定句中普遍採用這一詞序，以甘肅為例：

(118) 你胡不要拉。(不要胡拉)

(119) 你亂別說。(不要亂說)

(120) 你話怎麼好好不說。(怎麼不好好說話)

(121) 這幾年我們家這麼著沒團圓過。(沒像這樣團圓過)

有時狀語為單音節形容詞：

(122) 你深不要挖。(不要深挖)

(123) 多不買煞？(不多買點嗎)

(124) 東西亂不要擺。(不要亂擺)

當謂語動詞帶賓語時，賓語前置，如臨夏方言：

(125) 有的小學生書好好地不念者。

(126) 房子各處胡不要蓋。(“不要”連讀)

張成材(1998)也指出了西寧方言中否定詞在其所修飾的狀語後面：

(127) 他常常按時不上班。

(128) 你阿蒙好好不學習？(阿蒙：怎麼)

(129) 之個小說我仔細沒看過。

李樹儼、張安生(1996)指出：銀川方言否定句中程度副詞“甚、太”等可以放在禁止詞“罷(不要)”之前，也可以放在其後。

(130) 辣子甚罷種得稠了。(辣椒別種太稠了)

- (131) 辣子罷甚種得稠了。
(132) 飯太罷舀得滿了。（飯別舀得太滿了）
(133) 飯罷太舀得滿了。

前一種詞序是“漢兒言語”式的，後一種正在向普通話靠攏，但程度副詞還是遠離被飾成分。

4.5 在表示行為動作方向或場所的句子中，處所名詞前置，不用介詞

- (134) 西寧話： 昨晚夕，我你家裡去了。（昨晚我到你家去了；張成材 1998）
(135) 臨夏話： 娃娃病下了，我還醫院裡沒看去。（我還沒到醫院看去；王森 1993）

這兩例裡的方位詞“裡”並非必須，把它看作處所補語標記的遺存也不無可能。

- (136) 臨夏話： 別人家背後你的脊樑不要叫戳著。（別叫別人在背後戳你的脊樑）

張崇(1990)描寫陝西延川話“主語＋處所名詞＋趨向動詞”的詞序，由於處所名詞前不用介詞，因此往往造成歧義，分不清處所詞是目的地還是出發地：

- (137) 你山裡上去（你上山去/你從山裡上去）
(138) 他延川回來了（他回到延川了/他從延川回來了）

這跟前面介紹的“漢兒言語”的情況相同。

5. 結語

迄今為止漢語學界對西北方言的描寫和研究是很不夠的，但即使從這些一鱗半爪的現象中，也可以看出它跟“漢兒言語”血脈相連的關係，有些特點簡直如出一轍地驚人相似。西北方言的現狀讓我們更加確信：“漢兒言語”是北朝以來

特別是遼金元各代在北方地區口語中普遍通行的語言，它保留了漢語的基本特點，但也一定程度地吸收了阿爾泰語的語法範疇、複製重組了阿爾泰語的某些句型。今天的西北方言是“漢兒言語”的流，它能夠歷經歷史的滄桑之變而較好地保留“漢兒言語”的面貌，是因為這一大片地方自古以來就是民族雜居混處的地區，語言接觸密切，而且，這種環境很少因為改朝換代而改變，只要人文語言環境不變，它就可以繼續存在下去。過去因為西北方言既不是少數民族語言，又不像東南方言那樣跟普通話有很大的差別，因此一直未能得到應有的重視，這是十分可惜的。今天，隨著社會的開放、經濟的發展，那種相對封閉的環境日漸被打破，西北方言中的“漢兒言語”元素勢必會受到普通話越來越強有力的衝擊，語言研究者應該懷著緊迫感加強對它的調查研究，讓這塊寶貴的語言化石在漢語史、中華民族交融史研究中發揮應有的作用。

引用文獻

- Chen, Yinke (陳寅恪). 1936. Dongjin Nanchao zhi Wuyu 東晉南朝之吳語 [The Wu language in the Eastern Jin and Southern Dynasties]. *Bulletin of the Institute of History and Philology Academia Sinica* 7.1:1-4.
- Daobu (道布). 1983. *Mengguyu Jianzhi* 蒙古語簡志 [An Introduction to Mongolian]. Beijing: Nationalities Press.
- Fu, Lehuan (傅樂煥). 1984. Songren shi Liao yulu xingcheng kao 宋人使遼語錄行程考 [A study on the reports of Song diplomats]. *Liaoshi Congkao* 遼史叢考 [A History of Liao Dynasty: From Political and Ethnic Aspects], 1-28. Beijing: Zhonghua Book Company.
- Hashimoto, Mantaro J. (橋本萬太郎). 1983. Beifang Hanyu de jiegou fazhan 北方漢語的結構發展 [The structural development of northern Chinese]. *Yuyan Yanjiu* 語言研究 [Studies in Language and Linguistics] 1983.1:88-99.
- Hu, Shuangbao (胡雙寶). 1986. Du Qiaoben Wantailang Yuyan Dili Leixingxue 讀橋本萬太郎《語言地理類型學》 [A review on Hashimoto Mantaro's *Typogeographical Linguistics*]. *Yuwen Yanjiu* 語文研究 [Linguistic Researches] 1986.2: 47-51.
- Huang, Borong (黃伯榮). (ed.) 1996. *Hanyu Fangyan Yufa Leibian* 漢語方言語法類編 [Grammar Index to Chinese Dialects]. Qingdao: Qingdao Press.
- Huang, Zheng (黃征). 1996. Dunhuang suyufa yanjiu zhi yi 敦煌俗語法研究之一 [Syntactic study of colloquial Dunhuang language]. *Dunhuang Tulufan Yanjiu* 敦煌吐魯番研究 [Journal of the Dunhuang and Turfan Studies] 1:65-77. Beijing: Peking University Press.
- Jia, Jingyan (賈敬顏). 1985. Hanren kao 漢人考 [An enquiry into the term "Han people"]. *Zhongguo Shehui Kexue* 中國社會科學 [Social Sciences in China] 1985.6:99-109.
- Jiang, Lansheng (江藍生). 1992. Zhuci 'si de' de yufa yiyi ji qi lai yuan 助詞“似的”的語法意義及其來源 [The grammatical meaning of the particle *si de* and its sources]. *Zhongguo Yuwen* 中國語文 [Studies of the Chinese Language] 1992.6: 445-452.
- Jiang, Lansheng (江藍生). 1998. Houzhici 'xing' kaobian 後置詞“行”考辨 [Enquiry into postposition *xing*]. *Yuwen Yanjiu* 語文研究 [Linguistic Researches] 1998.1:1-10.
- Jiang, Lansheng (江藍生). 1999a. Cong yuyan shentou kan Hanyu binishi de fazhan 從語言滲透看漢語比擬式的發展 [The development of analogy in the Chinese language as seen from language infiltration]. *Zhongguo Shehui Kexue* 中國社會科學 [Social Sciences in China] 1999.4:169-179.

- Jiang, Lansheng (江藍生). 1999b. Chongdu Liu Zhiyuan Zhugongdiao 重讀《劉知遠諸宮調》[Rereading of *Liu Zhiyuan Zhugongdiao* (the various musical poems)]. *Wenshi* 文史 [Literature and History] 1999.3:263-276.
- Jiang, Lansheng (江藍生). 2003. Yuyan jiechu yu Yuan Ming shiqi de teshu panduanju 語言接觸與元明時期的特殊判斷句 [Language contact and the special copula sentence in Yuan and Ming Dynasties]. *Yuyanxue Luncong* 語言學論叢 [Essays on Linguistics] 28:43-60. Beijing: The Commercial Press.
- Kang, Baocheng (康保成). 1997. Yuan zaju hu qi wei dasao yu xiongdi gong qi gusu 元雜劇呼妻爲大嫂與兄弟共妻古俗 [The wife being called “sister-in-law” in the Yuan poetic drama and the ancient custom of brothers sharing a wife]. *Journal of Yangzhou University (Humanity and Social Sciences Edition)* 1997.6:23-30.
- Lee, Tae-Soo (李泰洙). 2003. *Laoqida Si Zhong Banben Yuyan Yanjiu* 《老乞大》四種版本語言研究 [Analysis on Language Usage Trends from Four Editions of *Laoqida*]. Beijing: Language and Culture Press.
- Lee, Tae-Soo (李泰洙), and Lansheng Jiang (江藍生). 2000. *Laoqida yuxu yanjiu* 《老乞大》語序研究 [On word order in *Laoqida*]. *Yuyan Yanjiu* 語言研究 [Studies in Language and Linguistics] 2000.3:71-82.
- Li, Chongxing (李崇興), Shuxian Huang (黃樹先), and Zesui Shao (邵則遂). 1998. *Yuan Yuyan Cidian* 元語言詞典 [The Lexicon of Language in the Yuan Dynasty]. Shanghai: Shanghai Education Press.
- Li, Shuyan (李樹儼), and Ansheng Zhang (張安生). 1996. *Yinchuan Fangyan Cidian* 銀川方言詞典 [Yinchuan Dialect Lexicon]. Nanjing: Jiangsu Education Press.
- Ma, Changshou (馬長壽). 1985. *Beiming Suo Jian Qianqin zhi Suichu de Guanzhong Buzu* 碑銘所見前秦至隋初的關中部族 [Information about the Tribes in Guanzhong from Inscription Tablets]. Beijing: Zhonghua Book Company.
- Mo, Chao (莫超). 2008. Gansu Hanyu fangyan yufa tedian zonglun 甘肅漢語方言語法特點綜論 [Généralités sur les caractéristiques grammaticales des dialectes chinois du Gansu]. Paper presented at Contacts de langues et changements linguistiques dans le cas du chinois et des langues altaïques, November 23, 2008. Paris: EHESS, CRLAO.
- Ota, Tatsuo (太田辰夫). 1953. *Laoqida de yuyan* 《老乞大》的語言 [The language of *Laoqida*]. *Chugokugogaku Kenkyukai Ronshu* 中國語學研究會論集, Vol. 1. Tanba: Chugokugogaku Kenkyukai.
- Ota, Tatsuo (太田辰夫). 1954. Guanyu Han'er yanyu: shilun baihua fazhanshi 關於漢兒言語：試論白話發展史 [About the Han'er language]. *Kobe Gaidai Ronso* 神戶外大論叢 [Kobe City University Journal] 5-3.
- Ota, Tatsuo (太田辰夫). 1988. *Chugokugoshi Tsūkō* 中國語史通考 [Comprehensive Studies of the History of the Chinese Language]. Tokyo: Hakuteisha.

- 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.
- Shi, Yuzhi (石毓智). 2008. Hanyu fangyan yuxu bianhua de liang zhong dongyin ji qi xingzhi chayi 漢語方言語序變化的兩種動因及其性質差異 [The motivation of the word-order change in Chinese dialect and their differences]. *Minzu Yuwen* 民族語文 [*Minority Languages of China*] 2008.6:19-27.
- Wang, Sen (王森). 1993. Gansu Linxia fangyan de liang zhong yuxu 甘肅臨夏方言的兩種語序 [Special word orders in the Linxia dialect of Gansu Province]. *Fangyan* 方言 [*Dialect*] 1993.3:191-194.
- Xie, Xiaolan (謝曉安), and Shumin Zhang (張淑敏). 1990. Gansu Linxia fangyan de yiwenju 甘肅臨夏方言的疑問句 [Interrogative sentences of the Linxia dialect of Gansu Province]. *Journal of Lanzhou University (Social Sciences)* 1990.3:141-146.
- Xue, Ruizhao (薛瑞兆). 2004. *Jindai Keju* 金代科舉 [*Imperial Examination in Jin Dynasty*]. Beijing: China Social Sciences Press.
- Yamakawa, Hidehiko (山川英彦). 1967. 元朝秘史總訳語法札記 [Notes on the grammar in the Chinese version of *The Secret History of Mongols*]. *The Journal of the Faculty of Letters, Nagoya University* 67:63-79.
- Yang, Naisi (楊耐思), and Shiyang Shen (沈士英). 1958. Gaocheng fangyan li de 'men' 藁城方言裡的“們” [*Men in Gaocheng dialect*]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1958.6:278.
- Yu, Zhihong (余志鴻). 1983. Yuandai Hanyu de houzhici 'xing' 元代漢語的後置詞“行” [Postposition *xing* in Yuan Chinese]. *Yuwen Yanjiu* 語文研究 [*Linguistic Researches*] 1983.3:48-50.
- Yu, Zhihong (余志鴻). 1987. Yuandai Hanyu '-xing' de yufa yiyi 元代漢語“～行”的語法意義 [Semantic meaning of *-xing* in Yuan Chinese]. *Yuwen Yanjiu* 語文研究 [*Linguistic Researches*] 1987.2:16-20.
- Yu, Zhihong (余志鴻). 1988. Menggu Mishi de teshu yufa: lun Yuandai Hanyu de shitizhi 《蒙古秘史》的特殊語法：論元代漢語的時體制 [The special grammar of *The Secret History of the Mongols*]. *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 1988.1:29-41.
- Yu, Zhihong (余志鴻). 1992. Yuandai Hanyu de houzhici xitong 元代漢語的後置詞系統 [Postposition system of Chinese in Yuan Dynasty]. *Minzu Yuwen* 民族語文 [*Minority Languages of China*] 1992.3:1-10.
- Zhang, Chengcai (張成材). 1998. *Xining Fangyan Cidian* 西寧方言詞典 [*Xining Dialect Lexicon*]. Nanjing: Jiangsu Education Press.
- Zhang, Chong (張崇). 1990. *Yanchuan Fangyanzhi* 延川方言志 [*A Record of Yanchuan Dialect*]. Beijing: Language and Culture Press.

- Zhang, Qing (張慶). 1997. Tangdai funü de liuxing fuzhuang 唐代婦女的流行服裝 [Costume in trend: women style in Tang Dynasty]. *Wenshi Zhishi* 文史知識 [Chinese Literature and History] 1997.3:61-64.
- Zhou, Yiliang (=Chou, Yi-Liang) (周一良). 1938. Nanchao jingnei zhi ge zhong ren ji zhengfu duidai zhi zhengce 南朝境內之各種人及政府對待之政策 [Native tribes in the territory of the Southern Dynasties]. *Bulletin of the Institute of History and Philology Academia Sinica* 7.4:449-504.
- Zhu, Dexi (朱德熙). 1958. Laoqida Yanjie, Piaotongshi Yanjie shu hou 《老乞大諺解》《朴通事諺解》書後 [Comments on two Korean books of proverbs]. *Journal of Peking University* 1958.2:69-75.
- Zu, Shengli (祖生利). 2000. *Yuandai Baihua Beiwen Yanjiu* 元代白話碑文研究 [A Study of the Inscription Tablets in the Yuan Dynasty]. Beijing: Chinese Academy of Social Sciences dissertation.
- Zu, Shengli (祖生利). 2001. Yuandai baihua beiwen zhong daici de teshu yongfa 元代白話碑文中代詞的特殊用法 [Special pronoun usages in Chinese vernacular on inscription tablets in the Yuan Dynasty]. *Minzu Yuwen* 民族語文 [Minority Languages of China] 2001.5:48-62.
- 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.

主要引用書目

- 《三朝北盟會編》，宋·徐夢莘編著，光緒三十四年許涵度校刊本。
- 《攻媿集》卷一一一《北行日錄》上，宋·樓鑰，四部叢刊。
- 《元代漢語本《老乞大》》，（韓）慶北大學出版部古典叢書 9，慶北大學出版部影印，2000。
- 《老乞大諺解》，（韓）奎章閣叢書第九，京城大學法文學部影印，1844。
- 《朴通事諺解》，（韓）奎章閣叢書第八，京城大學法文學部影印，1844。
- 《老乞大新釋》，（韓）奎章閣藏書 4871 號。
- 《重刊老乞大諺解》，（韓）弘文閣，1984。
- 《大元聖政國朝典章·刑部》，祖生利、李崇興點校。太原：山西古籍出版社，2004。
- 《《蒙古秘史》校勘本》，額爾登泰、烏雲達賚校勘。呼和浩特：內蒙古人民出版社，1980。
- 《孝經直解》（原題《新刊全相成齋孝經直解》），來薰閣影元刊本，1938。
- 《元代白話碑集錄》，蔡美彪。北京：科學出版社，1955。
- 《新校元刊雜劇三十種》，徐沁君點校。北京：中華書局，1980。
- 《元曲選》，臧晉叔編。北京：中華書局，1979。
- 《劉知遠諸宮調》，北京：文物出版社影印金刻本，1958。

The “Han’er Language” Revisited

Lansheng Jiang

Chinese Academy of Social Sciences

This paper traces the rise of the “Han’er language” and reviews main materials relevant to the study of the language. Eleven syntactic characteristics of the “Han’er language” are summarized from the materials and it is argued that these characteristics are the results of Altaic influence. Furthermore, based on a comparison of the “Han’er language” and some varieties of Modern Northwestern Mandarin dialect we propose that Modern Northwestern Mandarin dialect is a “living fossil” descended from the “Han’er language” and it is worthy of further research.

Key words: “Han’er” language, Altaic language, Northwestern Mandarin dialect, language contact

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

