第四屆國際漢學會議論文集

語言資訊和語言類型

鄭秋豫 主編

中央研究院

Papers from the Fourth International Conference on Sinology

Human Language Resources and Linguistic Typology

Edited by Chiu-yu Tseng

Academia Sinica

目　次

Contents

開幕致詞

大會主席　余英時院士

中央研究院第一次召開國際漢學會議是在一九八〇年，原本希望以後每十年召開一次。今年是第四屆，但距第三屆已十二年，算是遲了兩年。

事實上，早在一九五九年已有人提議本院應該召開漢學會議。當時的院長是胡適之先生，他認爲台灣的漢學研究無論在質還是在量的方面都還沒有達到他所期待的水平，因此力主緩議。這一提議卻引起了他對於所謂「漢學中心」的記憶和感慨。他說：

二十年前在北平和沈兼士、陳援菴兩位談起將來漢學中心的地方，究竟是在中國的北平，還是在日本的京都，還是在法國的巴黎？現在法國的伯希和等老輩都去世了，而日本一班漢學家現在連唐、宋沒有標點的文章，往往句讀也被他們讀破了。所以希望漢學中心現在是在台灣，將來仍在大陸。

但六十三年後的今天，我們對於胡先生的「漢學中心」說已有完全不同的理解。試一回顧過去五、六十年間漢學的發展，我們可以得到以下兩點認識：

第一，漢學已加速度地擴散到一切專門學科之中，不但人文和社會科學的每一部門中都包涵著越來越多的漢學研究，而且在中國科技史的廣大領域中，自然科學的各部門也和漢學日益緊密地結合在一起了。於是出現了一個奇詭的景象：漢學一望無際，觸處皆是，但是漢學作爲一個專門學

科（academic discipline）卻並不獨立存在，因爲漢學研究基本是寄托在其他學科之中的，如語言、文學、歷史、哲學、藝術、宗教之類。

第二，二戰以後各國漢學研究都取得了重要的成就，可是「漢學中心」卻未在任何地方出現。不用說，漢學研究在各國活躍的情形頗不一致，但即使是最活躍的國家也未曾取得公認的「中心」地位。事實上，如果我們分別考察各國漢學研究的大體趨向，便不難發現：主要由於研究的傳統和關注的問題彼此不同，每一地區的漢學都或多或少地展現出一種獨特的歷史和文化風貌。世界文化是多元的，漢學研究的傳統也不能不是多元的，這是我們今天共同承認的基本事實。

基於以上兩點新認識，我們可以十分肯定地說，胡適和他的朋友們當年最所縈心的「漢學中心」何在的問題，今天已自然而然地消逝了。如果有人堅持要在這個問題上討一個明確的答案，我祇好說：漢學猶如十六世紀布魯諾（Giordano Bruno）構想中的宇宙，其中心無所不在，其邊緣則無所在（“Its center is everywhere, its periphery nowhere.”）。

自本院一九八〇年召開第一次會議起，我個人每一屆都曾參與準備工作。因此我可以很負責地說，我們的唯一目的便是給世界各地漢學研究者提供一個充分交流的學術平台，所謂「漢學中心」問題從來不在我們的考慮之內。我們承認並且尊重每一地區漢學傳統的獨特風格，但是我卻不願看到任何漢學研究社群走上自我封閉的道路。因此不同傳統之間的互相溝通、互相認識和互相影響是極其必要的。我們相信，過去本院主持的三次會議多少曾發揮了這樣的功能。我們希望本屆會議也能作出同樣的貢獻。

我以最誠摯的心情感謝各地漢學同道們前來參加會議，特別是遠道冒暑而至的朋友們。我預祝大家有一次成功而愉快的學術聚會。

我個人因事不能到會，謹致最深的歉意！

二〇一二年六月二十日

序

中央研究院副院長　王汎森

中央研究院在 1980 年舉辦了第一屆國際漢學會議，其後以約 10 年的間隔舉辦一屆，至今已是第四屆。2010 年初我奉翁啓惠院長的任命，擔任副院長一職，當時思考著幾件應該進行的大事，其中之一就是漢學會議。在翁院長的大力支持下，第四屆國際漢學會議順利召開，上距第三屆已有 12 年，而距離第一屆更超過 30 年。30 多年來，與會的學者在改變，討論和關注的議題也隨著時空的推移，不斷的變遷。

當代社會資訊發達，空間距離縮小，連帶促進了地球村的形成，在這樣的時代背景下，使得「漢學」已不再是專屬華人的學問，而儼然成爲國際性的議題。學者關心的議題，也不再限於一時一地，而更多跨領域甚至跨國界的觀照。

從這一屆所發表的論文觀察，我感到較諸以往有幾點不同。在這裡僅以史學爲例，從「主題」、「材料」與「工具」三方面略作討論。

首先是「主題」的擴展，以歷史研究爲例，早已跨出傳統「帝王家譜」的框架，進而更多關注基層社會的歷史，關心人民的生活。討論的議題幾乎可說已經沒有局限而充分展現多元。

從「史料」方面講，傅斯年所揭示的「上窮碧落下黃泉，動手動腳找材料」的精神，繼續得到重視與發揚。尤其近年來大量新材料的出土，大幅改變了我們對問題的看法和認識。這裏所說的新材料，不只限於考古出土的史料，還應該包括過去被列爲禁燬或長期處於邊緣的史料，這些以往不能看或不易看到的東西，經過大量的刊印後，如今都觸手可及，無異於「新出土」，提供給學者們一個研究的新寶庫。此外，許多數位化資料庫

的建立，使得過去數十年皓首窮經都做不到的事情，現在也許一個按鍵彈指間就可獲得，爲學者帶來無窮的「近用性」。當然，我也必須強調，「檢索」不能代替閱讀。

從「工具」而言，科技帶來許多研究的新工具，數位工具的興起，使研究者與材料之間的關係得到全面的改變。另外如 GIS 的應用，方便我們用空間來思考事情，也使得研究的面向從單點連結到多點、甚至擴展爲全面。這些新工具的產生及應用，是過去幾屆漢學會議所無法想像的事。當然，以現代科技的日新月異，我們可以想像到第五屆時必然會有更新的工具出現，其新異的程度也可能是我們現在無法想像的。

第四屆漢學會議的成功召開，首先要感謝翁啓惠院長的全力支持，余英時院士慨允擔任大會主席，黃進興院士和黃樹民院士力任艱巨。世界各地學者的熱烈響應，更是我們必須表示敬意與謝意的。

在會議召開一年後，全帙 22 冊的會議論文集得以出版，這要歸功於出版委員會的不懈努力以及李宗焜博士、葉光輝博士與各冊主編和學者專家的配合。不論會議期間或出版過程，還有無數的助理同仁都付出了大量的心力。在全書即將付梓之際，我謹代表中央研究院對他們的貢獻表示衷心的感謝。

序

大會祕書處

二〇一〇年四月，翁啓惠院長由於關切人文社會科學的發展，特別指示，籌辦「第四屆國際漢學會議」。歷史語言研究所會同民族學研究所、近代史研究所、中國文哲研究所、臺灣史研究所、語言學研究所、人文社會科學研究中心等七個研究單位，共同配合辦理。經過整整兩年的努力，遂於二〇一二年六月二十日至二十二日如期召開大會。各國應邀發表論文的學者共計有二九二位，主持與評論者計四十三人，全體與會人員更達一千六百之眾，堪稱學術界的一大盛事。

本次會議具有四大特色：首先，進行深度討論，會議論文發表者，兼做主持人與評論，角色彼此替換；其二，跨學科、跨領域、跨地區的探討；其三，新領域與新議題的開發與探索；最後，迴異於往昔只有資深學人參加，此次廣邀富有潛力的年輕學者與會，副教授以下達總人數四分之一以上。

會議期間，雖值颱風侵襲，與會人員一本初衷，風雨無阻，踴躍出席，討論熱烈，成果甚爲豐碩。會後，復承王汎森副院長交待，成立編輯委員會，審議提交論文，俾便結集成冊。茲將最後定稿梓行，以供學界分享，並就教於諸方大家。

最後，本會議論文集得以刊行，必須感謝各分冊主編，尤其是出版委員會的編輯祕書李宗焜教授。

謹列本屆國際漢學會議攸關工作人員如下，並致謝忱：

主席兼召集人　余英時（中央研究院院士）
祕書長兼副召集人　黃進興（中央研究院院士、中央研究院歷史語言研究所）
副祕書長　黃樹民（中央研究院院士、中央研究院民族學研究所）
執行祕書　林富士（中央研究院歷史語言研究所）

籌備委員（依姓名筆畫順序）

胡曉真（中央研究院中國文哲研究所）
陳恭平（中央研究院人文社會科學研究中心）
黃克武（中央研究院近代史研究所）
鄭秋豫（中央研究院語言學研究所）
謝國興（中央研究院臺灣史研究所）

顧問（依姓名筆畫順序）

丁邦新（中央研究院院士）
王賡武（中央研究院院士、新加坡國立大學）
李亦園（中央研究院院士）
李歐梵（中央研究院院士、香港中文大學）
杜正勝（中央研究院院士）
杜維明（美國國家藝術與科學院院士、北京大學高等人文研究院）
郝延平（中央研究院院士）
張玉法（中央研究院院士）
梁其姿（中央研究院院士、香港大學香港人文社會研究所）
莊炳湟（中央研究院院士、Georgia Institute of Technology）
許文雄（東海大學歷史研究所）
許倬雲（中央研究院院士）
楊國樞（中央研究院院士）
葉文心（University of California, Berkeley）

序

中央研究院語言學研究所所長　鄭秋豫

2012 年 6 月 20 日至 22 日，中央研究院舉辦第四屆國際漢學會議，本屆會議語言學組有兩個主題，分別是「語言類型」和「語言資源」，會上一共發表了 26 篇論文。會議之後，經過收稿、送審、修訂、編輯等嚴謹的出版程序，終於以論文集的形式付梓。

本書共收錄了 10 篇文章，涵蓋本次會議語言學組的主題演講、院士演講和分場論文，其中與「語言類型」相關的論文有 6 篇，與「語言資源」相關的論文有 4 篇。

在「語言類型」方面，Walter Bisang〈Chinese from a Typological Perspective〉（〈從語言類型學視角來看中文〉）指出語言不僅有顯著性的語法標注手段，同時還有透過語用推理所造成的隱性面。本文比較了中文、西非與克里奧爾語言，具體呈現中文內高度複雜性的隱性面，進一步指出隱性面之形成不能單從語言接觸的觀點來解釋，而應當考慮到語用學在其中所起的重要作用。Randy J. LaPolla〈Arguments for a Construction-based Approach to the Analysis of Chinese〉（〈從構式語法角度分析現代漢語的結構〉）利用構式語法分析解決現代漢語長期存在的語法問題。本文認爲構式是最基本的分析單位，所以我們在分析的時候，只要看構式和組成構式成分的命題功能（如指示、修飾、敘述等），而不需要界定詞類和句法關係這種普遍性的範疇。James H-Y. Tai〈Reflections on Typological Characterization of Chinese Grammar〉（〈漢語語法類型學特徵的反思〉）指出，漢語被描述爲具有構詞獨立性與分析性，語法擬象性，同時是言談取向與語用取向。本文從語言相對論觀點，在功能主義與認知語法的架構

下，重新檢驗漢語語法類型學的特徵。由於漢語與手語、克里奧爾語言具有類似的特徵，可藉以擴大語言類型學的分析角度。連金發〈明清時期荔鏡／荔枝記閩南方言指示詞的演變：從指示詞到程度加強副詞或篇章標記〉一文深入分析了明清閩南語戲文中「遠／近指示詞＋樣式類別詞」所經歷的語法化現象，說明這些閩南語材料反映了所謂葉氏循環（Jespersen’s cycle）的新陳代謝現象。本文同時也探討了「指示詞＋數量詞」構式的語法化現象。劉丹青〈古今漢語的句法類型演變：跨方言的庫藏類型學視角〉首先介紹了庫藏類型學的基本觀念，然後以跨方言的庫藏類型學視角檢視古今漢語的若干重要的句法類型演變，提出 7 項主要的演變及其主要的方言差異。這些演變分別與「句法主導演變爲語用優先」及「名詞句法功能的萎縮和受限」兩大趨勢有密切關係。梅祖麟〈漢語方言裡的三個指代詞：「汝」、「渠他（佢）」、「許那」——再論魚虞有別與現代方言〉認爲現代漢語方言可以分爲兩組：官話、粵語、客家話導源於南北朝的北方通語；這三個方言魚虞相混；吳語、贛語、閩語各有一個導源於南朝通語的層次，在這個層次裡魚虞有別。本文深入地觀察魚韻三個代名詞成分（第二人稱「汝」、第三人稱「渠（佢）」和遠指代詞「許那」）進一步申論六朝時期南北通語的假設，並提出用指代詞和否定詞來區分現代漢語方言的具體方案。

在「語言資源」方面，Dafydd Gibbon〈Human Language Resources: their role in research, development and application〉（〈自然語言資源之於學術研究與技術研發應用的重要性〉）從自然語言資源發展的緣起與動機談起，並以音節語法、聲調分析、與詞彙學等實例說明自然語言資源須透過計算模型來呈現，才能提供實證而具體的語言學研究。詹衛東〈基於大規模中文樹庫的漢語句法知識獲取研究〉則利用已有標注的百萬字級漢語樹庫，在其中抽取不同層面上的句法知識，深入而具體地觀察幾個個案，例如：名詞性短語在不同句法位置上的差異，漢語中違反中心擴展規約和並列條件的短語結構的表現和短語結構歧義的情況。李愛軍〈Successive

Addition Boundary Tone of Chinese Emotional Intonation: Production and Perception〉（〈漢語情感語調的連續疊加邊界調：產出與感知〉）對連續疊加 (successive addition) 邊界調的聲學表現進行分析，並給出音系的表達。同時設計感知實驗，來考察連續疊加邊界調下降成分的聲學變化與整個語調情感表達的關係。通過對感知結果的線性回歸分析，發現連續疊加成分的聲學特性、邊界調的聲調類型以及句子的長短等因素對情感語氣的感知都有顯著的影響，但是對感知不同的情感影響不同。顧曰國〈A Conceptual Model of Chinese Illocution, Emotion and Prosody〉（〈論漢語施事行爲、情感與節律研究的概念模型〉）從「漢語現場即席話語語料庫」（Spoken Chinese Corpus of Situated Discourse，SCCSD）中提取相關語料實例，對漢語中施事行爲、情感與節律三者之間的互動關係提出詳盡的觀察。此外，本文還同時論證了幼兒話語裡「言、思、情整一」和「節律顯真情」這兩個假設。

綜觀全書，語言學組的論文集一方面充分顯示近十幾年來漢學領域中語言學在研究素材、方法及方向三方面都有新的開展，也引領漢語語言學未來的研究方向。屬於「語言資源」的論文，在研究上呈現出「跨領域」的特色，特別是資訊科學和語言學的結合，量化統計和質性分析的並用。屬於「語言類型」的論文，透過漢語及其方言的語言事實豐富了語言類型學的內涵，也透過語言比較來揭露不同語言之間在類型上的共性，呈現出「跨語言」的特色，提供了新的研究平台與格局，從漢語出發的語言學研究因此大步跨出區域性研究的侷限，在以後的十年、二十年裡扮演更關鍵的角色。

Human Language Resources and Linguistic Typology
Papers from the Fourth International Conference on Sinology
Taipei: Academia Sinica, 2013

Chinese from a Typological Perspective

Walter Bisang
University of Mainz

> Languages differ essentially in what they must convey and not in what they may convey.
> (Jakobson 1992[1959]:149)

The present paper starts out from the observation that complexity has two sides—an overt side that is determined by overt grammatical markers and a hidden side that depends on pragmatic inference. Most discussions on complexity only address overt complexity as it is motivated by explicitness. Hidden complexity as a result of economy is neglected. The present paper shows that Chinese is typologically characterized by its high degree of hidden complexity through a comparison with West African and Creole languages. It will also show from data on Creoles that a typologically high degree of hidden complexity cannot be explained exclusively by language contact. It is rather the result of a process of pragmatics-oriented maturation—a type of maturation that is complementary to Dahl's (2004) explicitness-oriented maturation.

Keywords: linguistic typology, complexity, pragmatics, pro-drop, language contact

1. Introduction

Chinese with its structural properties is a challenge to linguistic typology and to linguistic theory in various ways. An important challenge concerns the relevance of syntax in its interaction with semantics and pragmatics. The range of assumptions concerning the power of syntax in Chinese reaches from its total irrelevance to its equal importance across all languages. The former claim sees no necessity of assuming a specific level of grammar, since everything can be accounted for in terms of semantics and pragmatics (LaPolla 1990, 1993). The latter assumption is the implicit conclusion of formal linguistics, which sees syntax as a part of Universal Grammar that is the same across all human languages (cf. e.g. Huang et al. 2009).

The present paper will start out from Huang Yan's (1994:xiv) approach that "the extent to which syntax and pragmatics interact varies typologically" and that the grammar of Chinese and East and mainland Southeast Asian (EMSEA) languages[1] in general is characterized by the cross-linguistically very high degree with which it allows the use of pragmatic inference. To account for this fact, it will introduce a new perspective on complexity which integrates pragmatic inference. While current theoretical approaches to complexity are based on concrete linguistic form as it is overtly accessible in linguistic structures, it will be argued that complexity also has its hidden side which is not represented in linguistic form but operates through pragmatic inference (Bisang 2009). This "hidden complexity", which is complementary to "overt complexity", manifests itself in two different forms. One of them is the omission of information concerning grammatical categories, the other one is multifunctionality and

[1] The EMSEA languages belong to the following families/subfamilies: Sinitic, Mon-Khmer, Tai, Miao-Yao/Hmong-Mien and Chamic (Malayo-Polynesian).

categorical flexibility of individual markers.

The present paper will only address the omission side of hidden complexity (for examples on multifunctionality, cf. Bisang 2009). With this focus, it takes up a statement of Jakobson (1992[1959]) in which the question of what categories must be expressed is crucial for understanding cross-linguistic variation. To make this idea more concrete and quantifiable, omission is defined in the sense that a language does not have to express grammatical categories if they can be pragmatically inferred even though it has grammatical forms for expressing them if necessary (Bisang 2004, 2009, 2011). Thus, a language shows more hidden complexity with regard to a certain grammatical category than another language having the same category to the extent that it can omit its marking. Based on this definition, the following hypothesis will be tested for Chinese:

Hypothesis 1:
Chinese is characterized by a cross-linguistically high degree of hidden complexity.

To test this hypothesis, the paper will discuss the omission of grammatical information in the three domains of (i) zero arguments (pro-drop), (ii) the function of the head noun in the relative clause and (iii) the juxtaposition of clauses with no overt marking of the semantic and syntactic relation between them. It will start with the presentation of Chinese and it will continue with the analysis of data from Creole languages and from West African Kwa languages. Creole languages are selected because they are prominently claimed to be less complex than other languages (McWhorter 2001, 2005, Parkvall 2008). If it turns out that Creoles tend to show less hidden complexity than Chinese as predicted by *hypothesis 1*, the question is what determines the degree of hidden complexity in

Creoles. This question leads to *hypothesis 2*:

Hypothesis 2:
The degree of hidden complexity in Creoles is largely determined by the extent to which it is possible in the grammars of the substrate and superstrate languages involved.

To test this hypothesis, it is necessary to look at potential substrate languages (the properties of superstrate languages like English, French or Spanish will be presupposed). For that reason, it is necessary to discuss data from West African languages, since they are the substrate languages of most Atlantic Creoles, an important part of the population of Creole languages. Thus, Fongbe (Niger-Congo: Kwa) is an important substrate language of Haitian Creole. In addition, West African languages like Yoruba (Niger-Congo: Benue-Congo: Defoid) are sometimes treated as isolating languages that are supposed to share similarities with EMSEA languages.

If Chinese shows a higher degree of hidden complexity than at least a large number of Creoles, this property may not only be founded on the effects of a long history of contact with many other languages spoken in East and mainland Southeast Asia. As will be suggested in the conclusion, the answer should be sought in what Dahl (2004) called maturation, but not in a process of adding overt complexity through time but in a pragmatics-based process of maturation that supports hidden complexity.

The paper will be structured as follows: Section 2 will briefly introduce the notion of hidden complexity in comparison to complexity as an overt phenomenon as it is commonly dealt with in the literature. Sections 3 to 5 will present the data on omission. Chinese will be discussed in §3. The Kwa languages will be the topic

of §4, while §5 will be on Creoles. The paper will end with a short conclusion in §6, which will address the issue of pragmatics-based maturation.

2. Explicitness vs. economy and overt vs. hidden complexity

The idea that linguistic structures are motivated by the two competing forces of explicitness and economy is not new. Von der Gabelentz (1891:251) was probably the first who pointed out the impact of these two forces on the diachronic development of individual languages when he talked about the "instinct of clarity" (*Deutlichkeitstrieb* in German) and the "instinct of laziness" (*Bequemlichkeitstrieb*). Later on, the two forces were discussed in a number of different theoretical contexts, among them in Haiman's (1983) paper on iconicity vs. economy and in the two constraints of faithfulness vs. markedness in Optimality Theory (Prince & Smolensky 1993, Prince & Smolensky 1993).

The competition between explicitness and economy is motivated by what Levinson (2000:6, 27-30) calls the "articulatory bottleneck". What constitutes this bottleneck is "the remarkably slow transmission rate of human speech (conceived of as the rate at which phonetic representations can be encoded as discriminable acoustic signals), with a limit in the range of seven syllables or 18 segments per second" (Levinson 2000:28). Other aspects of speech production and comprehension are much less time-consuming. One of them is pragmatic inference. Since articulation and overt encoding of information are much more expensive than pragmatic inference, language structure has to opt for a certain degree of economy to the detriment of explicitness. The degree to which grammars of individual languages allow economy varies to a certain extent cross-linguistically.

Current approaches to complexity have a strong tendency to concentrate on

the explicitness side of grammar. Since they look at what is overtly expressed in individual languages, they deal with what I call "overt complexity" (Bisang 2009). McWhorter's (2001, 2005) surface-based approach is a straightforward case in point. Degrees of complexity are measured on the basis of markedness patterns and the number of rules that are needed for the production of grammatically acceptable utterances. In the case of recursion-based approaches (Hauser, Chomsky & Fitch 2002, Givón & Shibatani 2009), what matters is the overt linguistic manifestation of a small number of syntactic rules as they are used recursively. Research on overt complexity has produced a lot of interesting insights but it has neglected the economy side of language. If due to economy the surface structure of an utterance does not provide enough grammatical information its adequate interpretation depends on covert or hidden information which needs to be pragmatically inferred. Thus, economy is associated with a second type of inference-based complexity which I call "hidden complexity" in Bisang (2009).

If one understands grammatical structures as at least partially motivated by the competing forces of economy and explicitness overt complexity and hidden complexity are both to be expected. There are phenomena in individual languages that tend more to explicitness and others that give preference to economy but it is extremely unlikely if not impossible for a language to follow only one of the two principles. For that reason, the present paper does not claim that overt complexity and hidden complexity are mutually exclusive. What it claims is that Chinese (like most other East and mainland Southeast Asian languages) is characterized by a cross-linguistically high degree of hidden complexity (*hypothesis 1*) that manifests itself in the omission of grammatical categories.

3. Omission and hidden complexity in Chinese

Chinese and EMSEA languages are well known for the ease with which arguments can be dropped without any overt option of retrievability (not even through agreement). In the following constructed example, the subject and the object of the second clause are zero-marked because they can be inferred from context:

(1) 你昨天看了電影$_i$嗎？ — **ø**看**ø**$_i$了。

*Ni zuotian kan le dianying$_i$ ma? **ø** kan **ø**$_i$ le*

2.SG yesterday see PFV film Q see PF

'Did you see a film$_i$ yesterday? — [I] saw [one$_i$].'

Example (2) is from Lu Xun and shows how zero-marking is used in a concrete text:

(2) Chinese: Lack of overt arguments (from Lu Xun, *Fēigōng* 'Opposing aggression'):

子夏的徒弟公孫高來找墨子，已經好幾回了，總是不在家，見不著。大約是第四或者第五回罷，這才恰巧在門口遇見，…

Analysis: 子夏的徒弟公孫高$_i$來找墨子$_j$，已經好幾回了，**ø**$_j$總是不在家，**ø**$_i$見不著**ø**$_j$。大約是第四或者第五回罷，**ø**$_i$這才恰巧在門口遇見**ø**$_j$，…

Zixia de tudi Gongsun Gao$_i$ lai zhao Mozi$_j$,

Zhixia REL disciple Gongsun Gao come look for master Mo

*yijing haoji hui le, **ø**$_j$ zongshi bu zaijia, **ø**$_i$ jianbuzhao **ø**$_j$,*

already several times PF always NEG at home unable to meet

dayue shi disi huozhe diwu hui ba,
about be fourth or fifth time EXCL
***ø**$_i$ zhe cai qiaqiao zai menkou yujian **ø**$_j$.*
this finally by chance LOC doorway meet
'Gongsun Gao$_i$, a disciple of Zixia, was looking for master Mo$_j$ for several times and [he$_j$] was never at home, so [he$_i$] was unable to meet [him$_j$]. It was at about the fourth or the fifth time that [he$_i$] met [him$_j$] in the doorway ...'

The text in (2) is the beginning of a short story. It starts with the introduction of the two protagonists (Gongsun Gao and Mozi), who remain unmarked in the subject and object positions of the subsequent events. In the case of the predicate 在家 *zai jia* 'be at home', both protagonists are potential arguments. Since the reader knows that Gongsun Gao wants to see Mozi, it is reasonable to infer that it is Mozi who stays at home. The subject of the next predicate 見不著 *jian bu zhao* 'be unable to see/meet', however, is Gongsun Gao, since it is known that he is the one who wants to meet Mozi. Thus, not even a change of the subject is a sufficient reason for overt argument marking as long as the situation is clear from context. Finally, the same constellation of Gongsun Gao meeting Mozi is reported in the next event, again with no overt marking of subject and object.

Zero-arguments are not a recent phenomenon in Chinese. They continually exist since ancient times. The following example from classical Chinese may illustrate this fact:

(3) Classical Chinese:
王就見孟子曰，前日願見而不可得，得侍同朝甚喜。今又棄寡人而歸，不識可以繼此而得見乎。（孟子，公孫丑章句下）

Analysis: 王$_i$就見孟子$_j$**ø**$_i$曰，前日**ø**$_i$願見**ø**$_j$而**ø**$_i$不可得，**ø**$_i$得侍**ø**$_j$同朝甚喜。今又**ø**$_j$棄寡人而**ø**$_j$歸，**ø**$_i$不識**ø**$_i$可以繼此而**ø**$_i$得見**ø**$_j$乎。

Wang$_i$ *jiu* *jian* *Mengzi*$_j$ ***ø***$_i$ *yue*,
king then see Mencius say
qianri ***ø***$_i$ *yuan* *jian* ***ø***$_j$ *er* ***ø***$_i$ *bu* *ke* *de*,
previously want see but NEG can succeed
ø$_i$ *de* *shi* ***ø***$_j$ *tongchao* *shenxi*.
succeed attend court delighted
Jin *you* ***ø***$_j$ *qi* *guaren* *er* ***ø***$_j$ *gui*,
now again abandon 1.SG.HON and go home
ø$_i$ *bu* *shi* ***ø***$_i$ *keyi* *jici* *er* ***ø***$_i$ *de* *jian* ***ø***$_j$ *hu*.
NEG know can hereafter and succeed see Q

'The king$_i$ went to see Mencius and ø$_i$ said: "Previously, [I] wanted to see [you] but [I] did not succeed and when [I] succeeded to attend [you] in the same court [I] was highly delighted. Now, [you] abandon me again and [you] go home. [I] don't know if it will be possible hereafter that [I] see [you]."' (Mencius 2.B.10)

Data of this type show that Chinese is a radical pro-drop language (Neeleman & Szendröi 2007) in which there is no overt indication of the argument at all. In languages of this type, the *pro* element cannot be accounted for in terms of Rizzi's (1986) classical conditions. It is not formally licensed by a head (e.g. C, I, V, P) nor is it licensed by content through rich agreement on the verb.

Zero-arguments are not only involved with the problem of locally retrieving the omitted argument, they have a wide range of consequences for reference tracking or coindexation. In the case of relative-clause formation, the argument position which is coreferent with the head noun must be empty. If both positions

are empty in a relative clause with a transitive verb this creates problems of coindexation as illustrated in (4). Is the head noun 人 *ren* ‘people’ coindexed with the object position of the relative clause as in interpretation (4a) or with the subject position as in (4b)?

(4) Chinese:
找的人還沒有回來。
Analysis: [$\boldsymbol{ø}_{i?}$找$\boldsymbol{ø}_{j?}$的]人$_{j/i}$還沒有回來。

[$\boldsymbol{ø}_{i?}$	*zhao*	$\boldsymbol{ø}_{j?}$	*de*]	*ren*	*hai*	*meiyou*	*huilai*
	look for		REL	person	still	NEG	come back

a. Object coreference: ‘The people [we] were looking for are still not back.’
b. Subject coreference: ‘The people who looked for [us] are still not back.’

The grammatical system of Chinese allows for such structures in which coreference can only be determined by pragmatic inference. This is not possible in English relative clauses, nor is it possible in Yoruba, Fongbe and a number of Creole languages (cf. §4 and §5). In some instances, zero-marking is not limited to arguments, it can even be found in adjunct coreference (for details, cf. Ning 1993). Example (5) illustrates instrument coreference:

(5) Chinese:
[我寫信的]毛筆

[*wo*	*xiexin*	*de*]	*maobi*
1.SG	write letter	REL	pencil

‘the pencil$_i$ [I write a letter with ø$_i$]’

Arguments are not the only elements that can be omitted. Adverbial subordinators and complementizers are other instances of non-obligatory marking. If there is no overt marker the two (or more) clauses are simply juxtaposed and the relation between them is subject to pragmatic inference. Example (6) with its four interpretations is a very well-known example from Li & Thompson (1973). Another example is (7), which will be further discussed in §4 under (13) and (14).

(6) Chinese (Li & Thompson 1973):
你跪下來求張三。
Ni *guixialai* *qiu* *Zhangsan*
2.SG kneel beg Zhangsan
a. Purpose: 'You knelt down in order to beg Zhangsan.'
b. Consecutive action: 'You knelt down and then begged Zhangsan.'
c. Simultaneous action: 'You knelt down begging Zhangsan.'
d. Alternating action: 'You knelt down and begged Zhangsan.'

(7) Chinese (Li & Thompson 1981:595):
我買票進去。
Wo *mai* *piao* *jinqu*
1.SG buy ticket go in
a. Purpose: 'I bought a ticket to go in.'
b. Consecutive action: 'I bought a ticket and went in.'

Li & Thompson (1973, 1981) treat examples like (6) and (7) as single constructions irrespective of their different interpretations. That this analysis is inadequate was clearly demonstrated by several linguists, recently and prominently by Paul (2008). Each of the individual interpretations is linked to a different syntactic analysis and thus stands for a different construction. This fact is

associated with another important effect of omission and hidden complexity. Grammatical markers often are indicators of certain constructions. If they are dropped this indicator is missing and a given linguistic structure may be associated with two or more constructions. Thus, a seemingly simple surface structure may stand for more than one construction. It is needless to say that the question of which analysis is the right one can only be decided by pragmatic inference and thus reflects another aspect of hidden complexity. On the whole, hidden complexity through omission is about pragmatic inference of missing information concerning grammatical categories, reference tracking and the assignment of linguistic structures to constructions.

4. Omission and hidden complexity in West African languages

This section discusses the extent to which hidden complexity is possible in the three domains of zero arguments, relative clause formation and clause combining in West African languages. The languages selected are Yoruba and Fongbe (cf. §1). In both languages, zero-arguments are ungrammatical. In Yoruba, the presence of an overt subject and object is obligatory in independent declarative clauses with a transitive verb. Thus, the sentences in example (8) with the verb *rà* 'buy' are grammatical as long as there is an NP or a pronoun in the positions of the subject and the object as in (8a-d). If one or both of these positions are empty as in (8e-g) the sentence is unacceptable.

(8) Yoruba: On pro-drop with the verb *rà* 'buy':[2]

a.	*Ayọ̀*	*´*	*raṣọ.*		b.	*ó*	*raṣọ.*	
	Ayọ̀	´	rà	**aṣọ**		**ó**	rà	**aṣọ**
	Ayo	HTS	buy	clothes		3.SG	buy	clothes

'Ayo bought clothes.' 'He buys clothes.' 'Ayo buys it.' 'He buys it.'

c.	*Ayọ̀*	*´*	*ràá.*	d.	*ó*	*ràá.*	
	Ayọ̀	´	rà-**á**.		**ó**	rà-**á**.	
	Ayo	HTS	buy-3.OBJ		3.SG	buy-3.OBJ	

e.	**raṣọ.*			f.	**Ayọ̀*	*´*	*rà.*		g.	**rà.*	
	ø rà	**aṣọ**			**Ayọ̀**	´	rà	**ø**.		**Ø** rà	**ø**.
	buy	clothes			Ayo	HTS	buy				buy

Fongbe is largely similar to Yoruba. Independent declarative clauses with a transitive verb have an obligatory subject and an obligatory object in most instances but there are a few verbs called "verbs licensing expletive subjects" by Lefebvre & Brousseau (2002:246, 276-277) with optional subjects. One of these verbs is the raising verb *cí* 'seem' illustrated in (9). With other verbs, there must be an overt subject (10):

(9) Fongbe (Lefebvre & Brousseau 2002:276):

(*É*)	*cí*	*ɖɔ̀*	*Kɔ̀kú*	*jɛ̀ àzɔ̀n.*
3.SG	seem	COMP	Koku	be.sick

'It seems that Koku is sick.'

[2] HTS in the glosses stands for "High Tone Syllable". This tone occurs on the last syllable of the subject noun phrase in certain semantic contexts (Bisang & Sonaiya 1999).

(10) Fongbe (Lefebvre & Brousseau 2002:277):

É	*vɛ̀-wú*	*nú*	*Kɔ̀kú*	*ní*	*yì.*
3.SG	be-difficult	COMP	Koku	SUBORD	leave

'It is difficult that Koku leaves.'

In the case of relative clauses, the grammars of Yoruba and Fongbe simply do not leave any options for pragmatically inferring the function of the head noun. Even though both languages use rather different structures there is always a clear and obligatory distinction between subject coreference, object coreference and adjunct coreference. In Yoruba, relative clauses follow their head nouns and are introduced by the relative marker *tí*. In the case of subject coreference, the subject position must be filled by a subject pronoun (11a). In the case of object coreference, the object marker, which is obligatory in independent declarative clauses, must be dropped (11b). Finally, locative coreference entails the presence of the locative marker *ti* (11c):

(11) Yoruba: Relative clauses:

a. Subject coreference:

obìnrin	[*t'*	*ó*	*maa*	*ràá*]
obìnrin	tí	ó	maa	rà-**á**
woman	REL	3.SG.SUBJ	TA	buy-3.OBJ

'the woman who bought it'

b. Object coreference:

iṣu$_i$	[*tí*	*mo*	*rà*	***ø***$_i$	*lánàá*]
yam	REL	1.SG	buy		yesterday

'the yam I bought yesterday'

c. Locative coreference:

ó *mọ̀* *ọjà* [*tí* *mo* ***ti*** *ràá*].

3.SG know market REL 1.SG LOC buy:3.OBJ

'He knows the market where I bought it.'

In Fongbe, the head noun "is linked to a position within the relative clause through the lexical nominal operator, *ɖé*" (Lefebvre & Brousseau 2002:161). This operator attracts lexical material. In the case of relative clauses, it attracts resumptive pronouns by moving them from their initial position to the position immediately after *ɖé*. This process generates surface structures which clearly differ for each type of coreference. In the case of subject coreference (12a), the resumptive subject pronoun *é* [3.SG.SUBJ] is attracted to the nominal operator. In the case of object coreference, the operator *ɖé* attracts the object pronoun *è* [3.SG.OBJ] while the object position itself is empty (12b). If the head noun has the function of an adjunct the resumptive pronoun is combined with a postposition and both markers together are attracted to the nominal operator *ɖé* (12c). The extraction site is empty as in the case of object coreference.

(12) a. Fongbe: Subject coreference (Lefebvre & Brousseau 2002:161):

súnû [*ɖé-é* *wá*] *ɔ́*

man OP-3.SG.SUBJ come DEF

'the man who came'

b. Fongbe: Object coreference (Lefebvre & Brousseau 2002:161):

àsɔ́n$_i$ [*ɖé-è*$_i$ *Kɔ̀kú* *ɖù* *ø*$_i$] *ɔ́*

crab OP-3.SG.OBJ Koku eat DEF

'the crab that Koku ate'

c. Fongbe: Locative coreference (Lefebvre & Brousseau 2002:162):

xàsùn$_i$ [*ɖé-é-mɛ̀*$_i$ *Kɔ̀kú sɔ́ àwíì ɖó ø*$_i$] *ɔ́*

basket OP-3.SG.OBJ-in Koku take cat put DEF

'the basket in which Koku put the cat'

As in the case of relative clauses, the grammar strongly reduces options for pragmatic inference in clause combining. This will be shown by some examples from Yoruba. In this language, the unmarked juxtaposition of two clauses is only possible if the two clauses jointly have the function of sentence focus and are marked by the focus particle *ni*. In this particular case, juxtaposition indicates conceptual closeness. Thus, the juxtaposition of 'to buy a ticket' and 'to enter' as in example (7) from Chinese needs a highly specific context. Let's assume there is a soccer game and there are various options to get access to it. Some people make it into the stadium because they know the governor, while others have to buy a ticket for entering. If somebody asks 'How did you get in', the utterance in (13) is one possible answer, in which the actions of 'buying ticket' and 'enter' have become closely related through context:[3]

(13) Yoruba:

Mo ra tíkẹ́ẹ̀tì wọlé ni.

1.SG buy ticket enter FOC

'I entered by buying a ticket.'

If the two clauses are not focalized as a conceptually close unit, the speaker has to use different constructions depending on the semantic relation between the two clauses. The purpose construction is illustrated in example (14a), the

[3] I owe this example to my colleague Remi Sonaiya.

consecutive action in (14b):

(14) Yoruba:

a. *Mo* *ra* *tíkẹ́ẹ̀tì* ***láti*** *wọlé.*
1.SG buy ticket PURP enter
'I bought a ticket to go in.'

b. *Mo* *ra* *tíkẹ́ẹ̀tì*, *mo* ***sì*** *wọlé.*
1.SG buy ticket 1.SG and enter
'I bought a ticket and went in.'

A comparison of the data on Yoruba and Fongbe with Chinese clearly shows that the two African languages show almost no hidden complexity in the three grammatical domains to be analyzed. In Fongbe, there is a small number of verbs that do not have an overt subject. Apart from that, no pro-drop is allowed, neither in Fongbe nor in Yoruba. In the case of relative clauses and clause combining, overt marking is necessary for different coreferential relations between the head noun and its function in the relative clause and for expressing the relation between clauses.

5. Omission and hidden complexity in Creole languages

A comparison of these data with Creole languages shows that Creoles tend to exhibit the degree of hidden complexity of the substrate and superstrate languages involved. If these languages are not radically pro-drop as in the case of West African languages and European languages such as English, French, Spanish, Portuguese or Dutch there is a strong tendency for these languages not to give up overt marking of at least the subject argument. Thus, Haitian Creole, a French-based Atlantic Creole, does not seem to have radical pro-drop (Déprez

1994) even though there are exceptional cases with no overt subject marking. Example (15) illustrates the raising verb *genle* 'seem', which cannot take an expletive subject. Other constructions without expletives are existential predicates and weather predicates. However, exceptions like these do not prove that Haitian Creole is pro-drop (DeGraff 1993). In other contexts, among them simple declarative sentences (16), overt argument marking is necessary.

(15) Haitian Creole (DeGraff 1993:72):

a.	*ø*	*genle*	*Jak*	*damou*	b.	**li*	*genle*	*Jak*	*damou*
		seem	Jack	be.in.love		3.SG	seem	Jack	be.in.love

'Jack seems to be in love.'

(16) Haitian Creole (DeGraff 1993:72):

a.	*Li*	*pati*	b.	**ø*	*pati*
	3.SG	leave			leave

'S/He left.'

Angolar, another Atlantic Creole based on Portuguese and spoken on the island of São Tomé (Gulf of Guinea, Africa), seems to be one of the languages in which the subject position is not necessarily filled by an overt element. Even though the subject position is filled very frequently at least by a pronoun, Maurer (1995:62) points out that subject pronouns can be omitted if inferrable from context.

In other Creole languages, first of all those with no West African substrate, pro-drop is more common. One of these languages is Tok Pisin, a variety of Melanesian Pidgin English spoken in Papua New Guinea. Even though there are grammars which state that the subject must be marked overtly in simple independent clauses (e.g. Verhaar 1995) a look at some texts reveals that there

are null referential subjects. Thus, Meyerhoff (2000:134) found null subjects for third person singular in 39% of the clauses analyzed in her test sample. Another example is Zamboangueño, a variety of Chabacano or Philippine Creole Spanish spoken in and around Zamboanga City on the southern tip of Mindanao island. In this language, no overt subject is required in simple declarative sentences (since Zamboangueño is a VSO language, the ø-sign is put after the verb):

(17) Zamboangueño (Lipski & Santoro 2007:376, from Forman 1972:168):
Andá ø alyì na réyno.
go there to kingdom
'He goes there to the kingdom.'

In relative-clause formation, Creole languages tend to leave less room for pragmatic inference and hidden complexity than Chinese. In Haitian, there is a clear asymmetry between subject and non-subject coreference. In the case of subject coreference, relative clauses are marked by the complementizer *ki* (18) (cf. the analysis of DeGraff 1992, 2007:110), while there is no marker for object coreference (19). Non-arguments are taken up by a preposition followed by a resumptive pronoun in the relative clause.

(18) Haitian: Relative clause, subject coreference (Muysken & Veenstra 1995:154):
poul [***ki*** *kouvri pitit li*]
chicken **REL** cover little 3.SG
'a hen covering her chickens'

(19) Haitian: Relative clause, object coreference (DeGraff 2007:111):
Annou vote pou kandida [*nou vle*] *a.*
let.1.PL vote for candidate 1.PL want DEF
'Let's vote for the candidate we want.'

Angolar also shows asymmetry between subject and object coreference. Subject coreference is marked by the relativizer *ki* (20), while object and adjunct coreference are marked by *ma* with additional morphosyntactic differences between object (21) and adjunct (22) coreference (Maurer 1995:55-58).

(20) Angolar: Relative clause, subject coreference (Maurer 1995:55):
ome si [***ki*** *ba tamba*]
man DEF REL go catch.fish
'the man who left to catch fish'

(21) Angolar: Relative clause, object coreference (Maurer 1995:55):
ome si [***ma*** *m bê*]
man DEF REL 1.SG see
'the man I saw'

(22) Angolar: Relative clause, adjunct coreference (Maurer 1995:56):
[*ome si*]$_i$ [***ma*** *n ga taba ra ø*$_i$]
man DEF REL 1.SG TA work give
'the man for whom I work'

In many Creole languages, there is at least a clear-cut distinction between argument and non-argument coreference. In Zamboangueño, relative clauses are usually (but not obligatorily) introduced by the relativizers *ke* or *kyén*. In the case of argument coreference, the resumed noun phrase is represented by a zero-pronoun.

This is illustrated by (23) on subject coreference and (24) on object coreference. For non-argument coreference, there are various superstrate-influenced constructions as, for instance, pied-piping in (25).

(23) Zamboangueño: Relative clause, subject coreference (Lipski & Santoro 2007:383):
el mana hente$_i$ [***kyén*** *ya man tunúk* *ø*$_i$ *na gargánta*]
DEF PL people **REL** PST DRV be.prick.by.thorn in throat
'people who have gotten fish spines caught in their throat'

(24) Zamboangueño: Relative clause, object coreference (Lipski & Santoro 2007:383):
El hombre$_i$, [***ke*** *ya man enkontrá tu ø*$_i$], *mi hermano.*
DEF man REL PST DRV meet 2.SG my brother
'The man whom you met is my brother.'

(25) Zamboangueño: Relative clause, adjunct coreference (Lipski & Santoro 2007:383):
El persona [***kon-kyen*** *ta kombersá tu*] *byen bwéno gayót.*
DEF person with-REL PROG talk 2.SG very good EMPH
'The person you are talking to is very nice indeed.'

Clause combining is somewhat harder to deal with because most Creole grammars do not address the question of whether overt marking of the relation between two clauses is compulsory and thus excludes juxtaposition. In most cases, however, the omission of markers of adverbial subordination seems to be restricted. As a consequence, the mere juxtaposition of two or more clauses does not simply create multiply analyzable surface structures which increase the degree of hidden complexity. This can be illustrated by example (26) from Berbice Dutch.

This example is about conditionals, a grammatical domain in which juxtaposition is relatively common cross-linguistically. In spite of this, the mere omission of the conditional marker (*afi*, *aso*, *as*, *if* 'if') in Berbice Dutch does not necessarily produce a surface structure which can still be interpreted as a conditional. In (26a), the combination of past plus perfect in the protasis generates hypothetical meaning because of the presence of the adverbial subordinator *aso* 'if'. Without *aso*, the same tense-aspect-modality marking can no longer express hypothetical meaning, it must be interpreted as anterior tense. Thus, the utterance in (26b) gets another meaning which is not related to conditionality:

(26) Berbice Dutch: Conditionals (Kouwenberg 1994:115):

aso	***wa***	*krik-**it**-o*	*hiri*,	*o*	*wa*	*bato*	*doto.*
if	PST	get-PF-3.SG	here	3.SG	PST	kill:PF:3:SG	dead

a. With *aso*: 'If he had gotten it here, it would have killed him.'

b. Without *aso*: 'He had gotten it here. It killed him.'

A comparison of the above data with Chinese reveals that Creoles show less hidden complexity than Chinese. Zamboangueño has pro-drop in simple declarative sentences and it distinguishes argument coreference from non-argument coreference in relative clauses without subject/object asymmetry. With these properties, it is the Creole language that comes closest to Chinese. Clause combining is not easy to assess clearly from the grammatical description but the data in Lipski & Santoro (2007) indicate that the markers in Zamboangueño are similar to the ones in languages like Spanish or English. As was shown in (26) on Berbice Dutch, the morphosyntactic properties of the clauses involved has reached a degree of overt complexity that the simple omission of the adverbial subordinator produces either a change in meaning or maybe even an ungrammatical clause.

6. Conclusion

In *hypothesis 1* it is claimed that Chinese is characterized by a cross-linguistically high degree of hidden complexity. Table 1 summarizes the results from §3 to §5 for checking this hypothesis. It addresses the question of the extent to which the grammars of the languages involved leave options for the pragmatic inference of omitted grammatical information.

Table 1: The degree of hidden complexity in three grammatical domains

	Chinese	**West Africa: Yoruba, Fongbe**	**Creoles**
Zero arguments	Radical pro-drop: A lot of inference is possible	Subject and object arguments must be overtly expresses: No room for inference	Less options for inference than in Chinese in most Creole languages
Relative clause	Inference of subject vs. object coreference and even of adjunct corefence with certain head nouns	Obligatory distinction between subject, object and adjunct coreference: No room for inference	Less options for inference than in Chinese
Clause combining	Non-marking provides options for inference	Non-marking has specific meaning if it is possible at all, no room for inference	Non-marking provides limited options for inference, grammars are often unclear

As can be seen from Table 1, Chinese shows the highest degree of hidden complexity in each of the three domains of grammar. On the other end, we find Yoruba and Fongbe with the lowest degree of hidden complexity in each domain.

Creoles are situated mostly in between. Thus, Creoles are not the most simple languages of the world if one takes hidden complexity into account. Even if one accepts McWhorter's (2001, 2005) assumption that Creoles are the result of a reduction to those categories that are absolutely necessary for communication at the stage of pidginization, Creole grammars tend to require a considerable degree of overt marking which systematically prohibits the surface simplicity that is possible in Chinese. This is due to *hypothesis 2*, which claims that the degree of hidden complexity in Creoles is largely determined by the extent to which it is possible in the grammars of the substrate and superstrate languages involved.

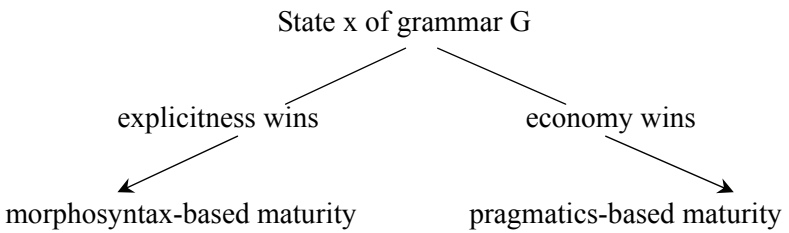
The Creole languages discussed in §5 confirm this hypothesis. Haitian Creole has no pro-drop except with a few verbs and thus comes very close to Fongbe (cf. the zero-arguments with the raising verbs in examples (9) and (15)). Similarly, its relative clauses show subject/object asymmetry as in French and Fongbe. In Angolar, pro-drop is possible but there is a distinction between subject, object and adjunct coreference in relative clauses. Its rigid coreferential properties in relative clauses seem to be supported by Portuguese as well as by the African substrate languages. The option of dropping the subject may be enhanced by Portuguese, which has pro-drop even though it is not a radical pro-drop language. Finally, the radical pro-drop properties of Zamboangueño are enhanced by its Philippine substrate languages. Its Spanish superstrate is also pro-drop, even though the subject can be retrieved from verb agreement. As for relative clauses, the coreference relation is clearly marked on the verb in the trigger system of Philippine languages (cf. Schachter 1993). In Spanish, the grammar does not allow the production of a relative clause that leaves the distinction of subject vs. object coreference to pragmatic inference. Thus, substrate and superstrate languages prevent a high degree of hidden complexity even though Zamboangueño relative clauses do not show rigid subject/object asymmetry.

The high degree of hidden complexity in Chinese (*hypothesis 1*) cannot simply be the result of extreme language contact as it is found in Creoles. Otherwise, all Creole languages would share a more or less equally high degree of hidden complexity, a scenario which has been proven inadequate in the context of *hypothesis 2*. Thus, the hidden complexity properties of Chinese cannot be due exclusively to language contact even though the impact of contact with other EMSEA languages can certainly not be denied. A look at arguments brought forward in the context of overt complexity will offer an idea of what may have driven Chinese grammar to its openness to pragmatic inference.

McWhorter (2001, 2005) argues that the relative simplicity of Creoles is due to their recent emergence. For that reason, they did not have enough time for the "contingent accumulation of "ornamental" elaboration" (McWhorter 2005:43) as it is found in other languages. Dahl (2004) discusses a similar phenomenon under the term of "maturation". He basically understands a mature phenomenon as the accumulation of lexical and/or morphosyntactic material in a grammar that was not present at an earlier stage of that grammar. Dahl (2004:114-115) mentions examples such as complex word structure (inflectional morphology, derivational morphology, incorporation), lexical idiosyncrasy (grammatical gender, inflectional classes), syntactic phenomena that are dependent on inflectional morphology (agreement, case marking) and specific marking of subordinate clauses. Definitions of maturation of this type are seen from the perspective of explicitness. In the course of time, grammars create tools for explicitly expressing more fine-grained grammatical distinctions. But linguistic structure is not only about explicitness. It is also about economy. Economy pushes a grammatical system to say only what is contextually needed and to leave the rest to pragmatics even if that system has a fine-grained inventory of grammatical distinctions. This is what triggers hidden complexity. In my view, there is an other type of maturation that leads to

hidden complexity. Thus, at any diachronic state x of the grammar of a language, there is a bifurcation that either leads to overt complexity or hidden complexity, if one takes the competing motivations of explicitness vs. economy seriously.

(27)



The development of a high degree of hidden complexity stands for the economy side of maturation. Chinese and mainland Southeast Asian languages are manifestations of a high degree of maturation motivated by economy.

List of glosses

1	1st person
2	2nd person
3	3rd person
COMP	Complementizer
DEF	Definiteness marker
DRV	Derivational affix
EMPH	Emphatic particle
EXCL	Exclamative
FOC	Focus particle
HON	Honorific
HTS	High Tone Syllable (cf. footnote 2)
LOC	Locative marker
NEG	Negation
OBJ	Object marker
OP	Nominal operator
PF	Perfect
PFV	Perfective
PL	Plural
PROG	Progressive aspect
PST	Past
PURP	Purpose marker
Q	Question marker
REL	Relativizer
SG	Singular
SUBJ	Subject
SUBORD	Subordination marker
TA	Tense-Aspect

References

Bisang, Walter. 2004. Grammaticalization without coevolution of form and meaning: the case of tense-aspect-modality in East and mainland Southeast Asia. *What Makes Grammaticalization? A Look from its Fringes and its Components*, ed. by Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer, 109-138. Berlin & New York: Mouton de Gruyter.

Bisang, Walter. 2009. On the evolution of complexity — sometimes less is more in East and mainland Southeast Asia. *Language Complexity as an Evolving Variable*, ed. by Geoffrey Sampson, David Gil & Peter Trudgill, 34-49. Oxford: Oxford University Press.

Bisang, Walter. 2011. Grammaticalization and typology. *The Handbook of Grammaticalization*, ed. by Heiko Narrog & Bernd Heine, 105-117. Oxford & New York: Oxford University Press.

Bisang, Walter, and Remi Sonaiya. 1999. The functional motivation of the High Tone Syllable in Yoruba. *Journal of African Languages and Linguistics* 20.1:1-19.

Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam & Philadelphia: John Benjamins.

DeGraff, Michel F. 1992. Haitian null subjects revisited. *Travaux de recherche sur le créole Haitien* 11:59-74.

DeGraff, Michel F. 1993. Is Haitian a Pro-Drop language? *Atlantic Meets Pacific: A Global View of Pidginization and Creolization*, ed. by Francis Byrne & John A. Holm, 71-90. Amsterdam & Philadelphia: John Benjamins.

DeGraff, Michel F. 2007. Kreyòl Ayisyen, or Haitian Creole. *Comparative Creole Syntax: Parallel Outline of 18 Creole Grammars*, ed. by John A. Holm & Peter L. Patrick, 101-126. London: Battlebridge.

Déprez, Viviane. 1994. Haitian Creole: a Pro-Drop language? *Journal of Pidgin and Creole Languages* 9.1:1-24.

Forman, Michael Lawrence. 1972. *Zamboangueño with Grammatical Analysis*. Ithaca: Cornell University dissertation.

Gabelentz, Georg von der. 1891. *Die Sprachwissenschaft*, *ihre Aufgaben*, *Methoden und bisherigen Ergebnisse*. Leipzig: T. O. Weigel Nachfolger.

Givón, Talmy, and Masayoshi Shibatani. 2009. *Syntactic Complexity: Diachrony*, *Acquisition*, *Neuro-Cognition*, *Evolution*. Amsterdam & Philadelphia: John Benjamins.

Haiman, John. 1983. Iconic and economic motivation. *Language* 59.4:781-819.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298:1569-1579.

Huang, C.-T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge & New York: Cambridge University Press.

Huang, Yan. 1994. *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Anaphora*. Cambridge & New York: Cambridge University Press.

Jakobson, Roman. 1992[1959]. On linguistic aspects of translation. *Theories of Translation: An Anthology of Essays from Dryden to Derrida*, ed. by Rainer Schulte & John Biguenet, 144-151. Chicago: University of Chicago Press. Originally in Reuben A. Brower (ed.), *On Translation*. Cambridge: Harvard University Press.

Kouwenberg, Silvia. 1994. *A Grammar of Berbice Dutch Creole*. Berlin & New York: Mouton de Gruyter.

LaPolla, Randy J. 1990. *Grammatical Relations in Chinese: Synchronic and Diachronic Considerations*. Berkeley: University of California dissertation.

LaPolla, Randy J. 1993. Arguments against ‘subject’ and ‘direct object’ as viable concepts in Chinese. *Bulletin of the Institute of History and Philology Academia Sinica* 63.4: 759-813.

Lefebvre, Claire, and Anne-Marie Brousseau. 2002. *A Grammar of Fongbe*. Berlin & New York: Mouton de Gruyter.

Levinson, Stephen C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge: MIT Press.

Li, Charles N., and Sandra A. Thompson. 1973. Serial verb constructions in Mandarin Chinese: subordination or coordination? *Chicago Linguistic Society* (*CLS*) 9:96-103. Chicago: Chicago Linguistic Society.

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional*

Reference Grammar. Berkeley: University of California Press.

Lipski, John M., and Salvatore Santoro. 2007. Zamboangueño Creole Spanish. *Comparative Creole Syntax: Parallel Outline of 18 Creole Grammars*, ed. by John A. Holm & Peter L. Patrick, 373-396. London: Battlebridge.

Maurer, Philippe. 1995. *L'angolar: Un créole afro-portugais parlé à São tomé*. Hamburg: Buske.

McCarthy, John, and Alan S. Prince. 1993. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Manuscript. Amherst: University of Massachusetts; New Brunswick: Rutgers University.

McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5.2-3:125-166.

McWhorter, John H. 2005. *Defining Creole*. Oxford & New York: Oxford University Press.

Meyerhoff, Miriam. 2000. *Constraints on Null Subjects in Bislama* (*Vanuatu*). Canberra: Pacific Linguistics.

Muysken, Pieter, and Tonjes Veenstra. 1995. Haitian. *Pidgins and Creoles: An Introduction*, ed. by Jacques Arends, Pieter Muysken & Norval Smith, 153-164. Amsterdam & Philadelphia: John Benjamins.

Neeleman, Ad, and Kriszta Szendröi. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38.4:671-714.

Ning, Chunyan. 1993. *The Overt Syntax of Relativization and Topicalization*. Irvine: University of California dissertation.

Parkvall, Mikael. 2008. The simplicity of creoles in a cross-linguistic perspective. *Language Complexity: Typology*, *Contact*, *Change*, ed. by Matti Miestamo, Kaius Sinnemäki & Fred Karlsson, 265-285. Amsterdam & Philadelphia: John Benjamins.

Paul, Waltraud. 2008. The serial verb construction in Chinese: a tenacious myth and a Gordian knot. *Linguistic Review* 25:367-411.

Prince, Alan S., and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. RuCCs Technical Report '2. Piscataway: Center for Cognitive Science, Rutgers University.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17.3:501-557.

Schachter, Paul. 1993. Tagalog. *Syntax: An International Handbook of Contemporary Research*, ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, 1418-1430. Berlin & New York: Mouton de Gruyter.

Verhaar, John W. M. 1995. *Toward a Reference Grammar of Tok Pisin: An Experiment in Corpus Linguistics*. Honolulu: University of Hawai‘i Press.

從語言類型學視角來看中文

Walter Bisang

美因茨大學

探究語言，其複雜性呈現於兩個方面，一面是由顯著性的語法手段標注，而另一面則需依靠語用推理獲知。迄今，對語言複雜性的討論大多數只停留在其顯性標注的一面，而往往忽略了其由力求經濟性而導致形成的隱性一面。筆者通過以中文和西非克里奧爾語言的比較，力現中文高度複雜性的隱性一面。來自克里奧爾語言的數據還將顯示：語言，其高度複雜性的隱性面之形成不能單靠語言接觸的觀點來進行解釋，而應當考慮到語用學的參與過程——一個以趨於成熟化爲導向的過程。由 Dahl (2004) 提出的「成熟化」這一概念只著眼於語言複雜性的顯性標注的一面。而以中文而言，一個建立在其複雜性的隱秘面上的成熟化過程同樣不容忽視。

關鍵詞：語言類型學，複雜性，語用學，代詞省略，語言接觸

Human Language Resources and Linguistic Typology
Papers from the Fourth International Conference on Sinology
Taipei: Academia Sinica, 2013

Arguments for a Construction-based Approach to the Analysis of Chinese

Randy J. LaPolla
Nanyang Technological University

This paper argues that long-standing problems in the analysis of Chinese, such as the question of word classes and grammatical relations, can be resolved, or actually done away with completely, if we take a constructionist approach in the analysis. This means the constructions are taken as basic, so we only need to look at the propositional functions of elements in the construction (referential, modifying, or predicative), and do not need to posit global categories such as word classes and grammatical relations.

Keywords: Mandarin Chinese, Construction Grammar, grammatical relations, word classes, clausal noun modifying structures

1. Constructions

The idea that language involves constructions rather than rules has been around a long time. Dwight Bolinger (1961, 1976) argued that much of language use involves recall of complete forms, including sentences, from memory rather than generation of totally new forms, as these remembered forms are what become fixed syntactic patterns (constructional schemata). Bolinger argued for something

like constructions, what he called 'idioms', and combinations of constructions, what he called 'syntactic blends' to form new syntactic structures, and pointed out 'the permeation of the entire grammatical structure by threads of idiom' (1961:366). (See also Pawley 1985, Grace 1987, Langacker 1987, Nunberg, Sag & Wasow 1994.) In Foley & Van Valin 1984 (see also Van Valin & LaPolla 1997: Ch. 6) we find the first treatment of grammatical relations as a construction-based phenomenon, that is, rather than seeing grammatical relations as a global phenomenon in a language, they saw it as a matter of individual constructions where one argument position in a construction is singled out for special prominence and treatment as the pivot of the construction, and there is a restricted neutralization of semantic roles in that pivot position. Different languages may grammaticalize different kinds of pivots, and in different constructions, if at all, and a single language may grammaticalize more than one type of pivot.[1] A key insight of that approach is that the meaning of the construction is more than just the sum of the parts. The construction as a whole forces a particular interpretation (i.e. the construction itself has a meaning). For example, consider the "Cross-clause co-reference constraint" in English. In this construction there are two conjoined clauses with one argument omitted from the second clause. This construction forces a particular interpretation of coreference between a particular argument of the first clause and the omitted argument of the second clause, as in (1), even when the interpretation doesn't make sense from a real-world perspective,

[1] It is important to distinguish simple topic-comment constructions from the type of grammaticalized pivot constructions that we are talking about here (as in (1) and (2) below). These pivot constructions constrain the interpretation of the role of referents in discourse, that is, they are a kind of referent tracking mechanism. Only they (and not the topic-comment structure) can be associated with syntactic relations (see LaPolla 2006a, 2006b for discussion). I have argued elsewhere that Chinese has not grammaticalized any such pivot constructions (e.g. LaPolla 1993, 2009).

as in (2). For an argument that isn't the actor of a transitive clause or the single argument of an intransitive clause to participate in this construction, the passive construction must be used, as in (1b).

(1) a. The man$_i$ saw the dog and Ø$_i$ went downhill.
b. The dog$_i$ was seen by the man and Ø$_i$ went downhill.
(2) The man$_i$ dropped the melon and Ø$_i$ burst. (Comrie 1988:191)

The modern constructionalist approach differs from earlier work in seeing all aspects of linguistic structure as constructions. In the constructionist approach, all grammatical knowledge is represented as constructions: pairings of form and meaning. Constructions can be simple or complex, schematic or substantive (or anything in between). No syntactic structure is needed aside from the constructions. There is also no separation of components as in the componential view of grammar that separates syntax, semantics, phonology, and pragmatics into different components that are separate from each other and from the lexicon, and must be linked together with linkage rules. In the constructionist approach there is no discussion of interfaces, as there are none. There is only the constructions, so there is also no separation of lexicon from the constructions; words are also constructions. In the Radical Construction Grammar of Croft (2001), the approach adopted here, there are no syntactic relations (which would be relations between parts of a construction other than the part-whole relation), only syntactic roles (part-whole relations). There are no syntactic categories other than the roles (propositional functions) the elements have in the constructions, and so there are no global syntactic categories in a language. As there are no universal syntactic construction types, there are no universal syntactic categories.[2]

[2] See the application of this view to transitivity in LaPolla (2011) and LaPolla, Kratochvíl & Coupe (2011).

2. Language structure as conventionalized constraints on inference

In several papers (LaPolla 1997, 2003, 2010, LaPolla & Poa 2002) we argued that there is no coding or decoding in communication, just ostension (performing an action in an ostensive way to show a desire to communicate) and inference (abductive inference, essentially guessing why the communicator made the ostensive act).[3] Language is not seen as a thing, but as behavior, governed and constituted by convention and habit. Our knowledge of language is our memory of the communicative behavior of ourselves and others. The role of language (as well as gestures and other ostensive acts) is to constrain the addressee's interpretation of the communicative intention of the communicator. Linguistic patterns emerge in discourse during the process of interaction (Bybee 2006, Hopper 2011, 2012, Thompson & Couper-Kuhlen 2005). Constructions then, as defined above, are conventionalized patterns of experience which constrain the interpretation in a particular way. This view entails that each language is unique, and each structure which conventionalizes in a language does so in a particular type of situation, so each construction is unique.

3. Chinese

3.1 Background

In discussions of Chinese grammar, two problems have loomed large: word classes and grammatical relations. In terms of the former, Y. R. Chao (1968) devoted 300+ pages of his *Grammar of Spoken Chinese* to defining word classes (pp.498-815), aside from another whole chapter trying to define "word" itself in

[3] This abductive inferential ability is a general cognitive ability, and is a survival instinct, used in understanding the natural world and the intentions of other humans. See LaPolla (2010) for more discussion.

Chinese, yet could not come up with any hard and fast definitions for classes, as there is a major problem of "overlapping classes":

> In Chinese 怪 *guay* is an adjective in 可是這很怪 *Keesh jeh heen guay* 'But this is odd', an adverb in 怪難看的 *guay nankann de* 'rather ugly', and a transitive verb in 別怪我! *Bye guay woo!* 'Don't find me odd—don't blame me!' … (1968:498)

His response was to try to define them anyway, with lots of exceptions and overlapping. He specifically mentioned not taking the line of Li Jinxi 黎錦熙 (1992[1924]:24), who had argued that "The class of a word depends on the sentence (it appears in), outside of a sentence it has no class" (凡詞, 依句辨品, 離句無品), and later, that "through (its) function (its) class becomes manifest" (由職顯類; 1953:10-11, my translations). Yet in the quote above it seems that Prof. Chao is in fact following Li Jinxi's view by saying the word *guay* (*guai*) belongs to a different class in each of the constructions it is used in. Recently two issues (40 and 41) of 語言學論叢 *Yuyanxue Luncong* (published by Beijing University) were devoted to defining word classes in Chinese; the controversies rage on. Most of the approaches taken take the distributional method for defining categories for granted; that is, they select constructions, and try to define global word classes based on particular constructions seen as criterial for the definition of those categories.

In terms of grammatical relations and the structure of the clause in Chinese, there are basically three opinions in the field: Topic-Comment (Y. R. Chao 1968, Lü Shuxiang 1979, LaPolla 2009); Topic-prominent (Li & Thompson 1976, 1981); and Subject-Predicate (Zhu Dexi 1982[4] and most formalist approaches).

[4] Although Prof. Zhu Dexi is generally seen as an exponent of the subject-predicate view, my own reading of his work is that he was closer to Prof. Chao in understanding "subject" to be topic.

Part of the difference is due to whether one assumes all languages must have the category of "subject", which usually also means one assumes this represents a global phenomenon in any one language. Similar to the case with lexical categories, constructions are often used for determining or defining grammatical relations. In the following we will look at several key constructions in Mandarin Chinese to see to what extent we find a consistent pattern of lexical class or grammatical relation when we look at the use of these constructions in natural language data.

3.2 The *ba* construction

The first construction we will look at is the *ba* construction. In the common view of this construction, for example as laid out by Thompson (1973), the construction is said to manifest the template in (3):[5]

(3) NP1 *ba* NP2 V1 (V2) (NP3)

In this view the particle *ba* is said to be preceded by the agent of the main verb (V1) and followed by a noun or noun phrase and is said to mark the "direct object" of V1 (e.g. Sun & Givón 1985). It isn't hard to find examples that seem to instantiate this template, for example in (4):

[5] I must emphasize that this is Thompson's view (and the common view) of this construction, which I am arguing against here; I do not accept the designations of "NP" and "V" used here.

(4) 他們計畫明年把共祭活動推廣到陵園和社區。

[Tamen]	jihua	mingnian[6]			
3PL	plan	next.year			
NP1					
ba	[gongji	huodong]	tuiguang	dao	[lingyuan
BA	public.obervance	activity	spread	arrive	cemetery
ba	**NP**		**V1**	**V2**	**NP3**
he	shequ]				
and	community				

http://news.sina.com.cn/c/2011-04-04/100722235984.shtml

Although this example seems to show what is uncontroversially a noun phrase that might be considered the direct object after the particle *ba*, as it is the theme or patient of *tuiguang* 'spread', there are many examples that show a different pattern, where the element that follows the particle *ba* is either not a theme or patient, or is not clearly a noun phrase. Consider the examples in (5):

[6] Abbreviations used include:

ADV	adverbial marker
AFF	the word 給 *gěi* when it is used as an affective marker
BA	the word 把 *bǎ* when it is used in the *bǎ* construction
CL	classifier
CSM	change of state marker
LOC	locative marker
MC	modifying clause
NEG	negation
NOM	nominalizer
PFV	perfective aspect marker

(5) a. 如果你把筆寫禿了，只要按一下藍色按鈕，筆芯就會馬上變尖。

Ruguo ni ba bi xie tu le
if 2SG BA pencil write blunt CSM
'If you make the pencil blunt from writing with it,'
http://www.ycxljy.com/jyky/ShowArticle.asp?ArticleID=16494

b. 這題太難了，會把頭髮愁白。

Zhe ti tai nan le, hui ba toufa chou bai
this question too difficult CSM will BA hair worry white
'This question is too difficult, it will make (one's) hair turn white.'
http://web.wenxuecity.com/BBSView.php?SubID=netiq_best&MsgID=582

c. 因爲這套衣服把我變的太可愛了，身材整個又矮腿又短全身看來就不成比例無言啦！

Zhe tao yifu ba wo bian de tai ke'ai le
this set clothing BA 1SG change ADV too cute CSM
'These clothes made me (look) too cute!'
http://www.wretch.cc/blog/gitbox/4658325

d. "我聽說，一般的女人不喝酒，女人不喝一般的酒，喝酒的女人不一般。" "我怕你了，那我就喝一小點兒，剩下的你全包了。" "好吧，千萬別強迫自己，真怕你喝醉了，做出什麼傻事兒來。" "才不會呢，我還怕這兩瓶酒把你給喝醉了。"

Wo hai pa zhe liang ping jiu ba ni gei he zui le
1SG still fear this two bottle liquor BA 2SG AFF drink drunk CSM
'I'm even afraid you'll get drunk from these two bottles.'
http://bbs.big5.voc.com.cn/topic-1964540-2-1.html

In (5a) the pencil is an instrument, not the patient, of 'write'; in (5b) *toufa* 'hair' is not a semantic argument of *chou* 'worry' at all; for (5c) there is no equivalent expression with the 1SG pronoun as the patient; and in (5d) the wine referred to by the referring expression before the particle *ba* is the patient of *he* 'drink', not the agent, and the agent of *he* 'drink' appears after the particle *ba.* So there is no consistency in terms of which roles appear in the post-*ba* slot, and in fact the referent referred to by the post-*ba* referring expression does not even have to be an argument of the verb.

Now let's turn to the type of phrase that can appear in the post-*ba* slot. Consider the examples in (6):

(6) a. 不要把吃飯變成一場戰爭。

Bu yao ba [chi fan] bian cheng yi-chang zhanzheng
NEG want BA eat rice change become one-CL war
'Don't make eating into a war.'
http://renyifei.172baby.com/posts/137278.html

b. 爲什麼有些人把吃飯睡覺當成最重要的？

Weishenme you xie ren ba [chi fan shuijiao]
why EXIST CL people BA eat rice sleep
dang cheng zui zhongyao de
take.as become most important NOM
'Why do some people take eating and sleeping as the most important (things)?'
http://zhidao.baidu.com/question/228560628.html

c. 爲什麼把吃飯各自付款稱爲AA？
Weishenme ba [chi fan gezi fukuan] cheng wei AA
why BA eat rice each pay call be AA
'Why is eating and each person paying for themselves called "AA"?'
http://iask.sina.com.cn/b/17752493.html

In (6a) a "verb phrase" appears in the post-*ba* slot; in (6b) two "verb phrases" appear in the post-*ba* slot; and in (6c) a whole clause appears in the post-*ba* slot. Trying to use the *ba* construction to define form classes or grammatical relations, then, will not work. What we can do is recognize that we have a construction that marks a secondary topic that is affected in some way by an action.

3.3 The *zhe ben shu de chuban* construction

A second construction that has been treated at length in discussions of form classes in Chinese is the 這本書的出版 "Zhe ben shu de chuban" construction (first discussed in Zhu, Lu & Ma 1961). This construction was considered problematic because we find what had been defined as a verb in other contexts acting as the head of a noun phrase. Natural examples highlighting the problem with this construction are given in (7):

(7) a. CNN循環播出中國國家形象宣傳片。
CNN xunhuan bochu Zhongguo guojia xingxiang
CNN circulate broadcast China country image
xuanchuanpian.
propaganda.film
'CNN repeatedly broadcast a propaganda film promoting China's national image.' *http://www.sina.com.cn/ 2011-01-19 07:49 新浪播客*

b. 有沒有看到那天的播出？

You	mei	you	kan	dao	nei	tian	de	bochu?
EXIST	NEG	EXIST	watch	arrive	that	day	NOM	broadcast

'Did (you) see the broadcast of that day?'

明日之星（電視節目）2011.06.11

Used in the construction in (7a), *bōchū* 'broadcast' has a predicative function, but used in the construction in (7b), *bōchū* 'broadcast' has a referring function. So the argument has been around whether such words are "really" nouns or verbs. Zhu et al.'s (1961) solution is to say that the word acting as a head is a verb, even though the phrase it is the head of is a noun phrase. This is problematic because it violates the sense of a phrase being a projection of the category of the head (see Shen 2007). This is completely driven by the assumed need for uniform global categories within a language. But no justification for the need for having uniform or global categories is ever given. In reality it is appearing in the particular position in the construction that gives the word its function; the word does not have (and does not need to have) any function outside of a particular construction.

3.4 The Topic-Comment construction

Another important construction is the Topic-Comment construction, the basic clause type in Chinese (Chao 1968, LaPolla & Poa 2005, 2006), and we can find the same phenomenon in this construction, that is, the same words and phrases can be used referentially or predicatively. For example, 吃 *chī* 'eat', when used in the comment of a Topic-Comment construction, as in (8a), has a predicative function, but when it appears in topic position, as in (8b), it has a referential function:

(8) a. 但不要吃那些加重近視的食物

dan bu yao chi naxie jiazhong jinshi de shiwu

but NEG want eat those increase nearsightedness NOM food

'but don't eat foods that will increase nearsightedness'

http://yk.bjybkf.com/news/ydbj/2007/9/079111727563384.html accessed 2012.09.02

b. 在中國及世界的許多國家，吃是一種文化。

chi shi yi zhong wenhua

eat COPULA one kind culture

'In China and many countries of the world, eating is a kind of culture.'

http://baike.baidu.com/view/13977.htm accessed 2011.07.10

In (9a) *tóngzhuōchīfàn* 'eat at the same table' appears in comment position, and so has a predicative function, but in (9b) the same expression appears in topic position, and so it has a referring function.

(9) a. 書記和我們同桌吃飯。

Shuji he women tong zhuo chifan

secretary and 1PL same table eat

'(The Party) secretary ate at the same table as us.'

http://d.wanfangdata.com.cn/periodical_ddkg201101023.aspx accessed 2011.07.10

b. 同桌吃飯也就具有了表演的性質。

Tong zhuo chifan ye juyou le biaoyan de xingzhi

same table eat also possess PFV perform NOM nature

'Eating at the same table also has the nature of a performance.'

http://baike.baidu.com/view/13977.htm accessed 2011.07.10

When used in the comment position of this construction, *shīrén* ‘poet’ and *dàxuéshēng* ‘student’ have a predicative function, as in (10a) and (10b) respectively:

(10) a. 這些人都很詩人。
Zhexie ren dou hen shiren
these person all very poet
‘These people are all like poets.’
http://hi.baidu.com/xmfine/blog/item/8c8b804404b6cb84b2b7dcbb.html

b. 都大學生了還這麼幼稚？
dou daxuesheng le hai zheme youzhi?
all university.student CSM still this.much naïve
‘(You) are already a university student, (but) still so naïve?’
http://video.baomihua.com/goodadv/12901470?P3P31

Example (11) is from a written advertisement for Baihuayou. If one looks in the dictionary, one will not find a predicative use for the word *zhái* 宅 ‘house’, but here it is used predicatively, and is easily understood.

(11) <u>除了整天宅在家</u>營養也是很重要的。
Chule zhengtian zhai zai jia…
aside.from all.day house LOC home
‘Aside from staying home all day…’
Ad for Baihuayou in MRT in Taipei, rec. 2012.03.12

3.5 Clausal referential phrase modifying construction with *de*

In Modern Mandarin, when a clause is used to modify a referring expression, it is nominalized by the particle *de* and then this nominalized clause precedes the

referring expression. This is actually the same structure as discussed in §3.3, but with a clause as the initial element. The clause with *de* is a complete referring expression and can be used alone (compare (12) and (13)).

(12) 坐在我後面的人

[[zuo zai wo houmian de]$_{MC}$ ren]$_{NP}$

sit be.at 1SG back NOM person

'the person sitting behind me'

http://tw.knowledge.yahoo.com/question/question?qid=1612031307463

(13) 站在我後面的是台灣人民

[zhan zai wo houmian de]$_{MC/NP}$ shi Taiwan renmin

stand be.at 1SG back NOM COPULA Taiwan people

'the (ones) who are standing behind me are the Taiwan people'

http://www.youtube.com/watch?v=iB2gZ5Zt4o0

In these particular examples it might be argued that the referring expression following the clause is an argument of the clause (although it would be hard to argue for a "gap", as there are no obligatory arguments), but the same structure in Chinese can be used when the head of the structure is clearly not an argument of the modifying clause, as in (14)-(18):

(14) 在每年地價稅開徵四十日前（即 9 月 22 日前），逾期申請者，<u>自申請的次年期開始適用</u>。

zi [[shenqing de]$_{MC}$ ci nianqi]$_{NP}$ kaishi shiyong

from apply NOM next year start use

'(it will be) effective starting in the year after applying'

http://www.kctax.gov.tw/tw/tax/LVT01.aspx

(15) 有人以前買菜的零錢，都會分給孩子，現在投竹筒，可以讓孩子一起來付出。

[[mai cai de]$_{MC}$ lingqian]$_{NP}$

buy vegetable NOM change

'change (left over after) buying groceries…'[7]

http://www.newdaai.tv/?id=49496&view=print

(16) 很快，室內響起了炒飯的聲音。

chao fan de shengyin

stir.fry rice NOM sound

'the sound of rice frying'

http://www.zwwx.com/book/10/10815/3146586.html

(17) 好吃又不會胖的甜點

haochi you bu hui pang de tiandian

tasty also NEG can fat NOM sweets

'sweets that (when you eat them they) won't make (you) fat'

http://yule.tv.tom.com/App_User_Video.php?video_id=21702

(18) 不用洗手的自動馬桶

[[bu yong xi shou de]$_{MC}$ zidong matong]$_{NP}$

NEG use wash hand NOM automatic toilet

'a toilet which (after having used it one) doesn't need to wash (one's) hands'

http://nimb.blogbus.com/logs/52825568.html

[7] In a different context this could mean 'the change to buy groceries with', but that is not what it was used to mean in the context in which it appeared. This is from a Buddhist web site where they are trying to get people to donate more money. The full translation is 'Some people, before they would take the money left over from buying groceries and give it to the children, now they put it into the collection box, this way they can have the children donate together'.

Compare the following examples, all with the expression 能寫的 *neng xie de* [able write NOM]:

(19) 後面能寫的紙都不浪費

houmian neng xie de zhi

back able write NOM paper

'paper on which you can write on the back…'

http://www.youtube.com/watch?v=V5DHRvnv_7M

(20) 想找枝能寫的筆

xiang zhao zhi neng xie de bi

think look.for CL able write NOM pen

'(I) want to find a pen (which) can write' (or 'which one can write with')

http://www.managertoday.com.tw/?p=2483

(21) 能寫的人，有福了！

neng xie de ren

able write NOM person

'person who can write…'

http://www.dk101.com/Discuz/viewthread.php?tid=93094

(22) 不能寫的人

bu neng xie de ren

NEG able write NOM person

'person that (one) cannot write about'

http://tw.knowledge.yahoo.com/question/question?qid=1510092303862

(23) 能說又能寫的能力

neng	shuo	you	neng	xie	de	nengli
able	speak	also	able	write	NOM	ability

'ability to speak and write (well)'

www.evis.com.tw/YOHAN_2012.pdf

(24) 學習英文，可以說是全民運動，可是真正學的好英文，能說又能寫的，實在少之又少。

neng	shuo	you	neng	xie	de
able	speak	also	able	write	NOM

'(those) who can speak and write (English well)'

http://save-coco.blogspot.com/2012/01/blog-post_05.html

(25) 將目前手中有的資料能寫的就寫吧

neng	xie	de	jiu	xie	ba
able	write	NOM	then	write	SUGGESTION

'… just write the (materials) that (you) can write'

http://zh.wikipedia.org/wiki/User:Irwin

The same structure in different contexts can be interpreted differently, as can be seen in comparing (21) with (22) and (24) with (25). When there is a head sometimes the interpretation of the nature of the referent of the head is a clue as to its relationship to the modifying clause, so if we compare, for example (19) and (20), it is only our understanding of the nature of 'paper' vs. 'pen' that tells us whether the referent of the modifying clause is what is written on or what is used to write. But even this is not fully deterministic, as the head is also open to various interpretations, as we can see from comparing (21) and (22).

These modifying constructions (including (12)) are structurally compound constructions made up of two referring expressions, similar to 學校的校長 *xuexiao de xiaozhang* [school NOM president] 'president of the school', though with a nominalized clause as the modifying element. There is no understood co-referential argument in the modifying clause, and while the first element modifies the second element, the relationship between the two can vary (cf. the discussions of English noun compounds in Downing 1977, Kay & Zimmer 1978, Levi 1978, Finin 1980). Semantically they can have a function similar to that of restrictive relative clauses in English, in that they can restrict the inference of the referent of the head. Once we start looking at the uses of this construction, we find that there are many possibilities, and in some cases the head might be said to be an argument of the clause, but in many cases it clearly isn't, and also whether there is a referring expression acting as head or not the interpretation of the referent of the modifying clause relies on inference from context, so the possibility of the head being omitted cannot be used as a criterion for distinguishing noun complements from relative clauses, and the construction cannot be used for determining grammatical relations. In the framework of LaPolla (2003), we would say that Chinese has not grammaticalized constraints on the identification of the relationship between the modifying clause and the head.

Given the many possible uses of this construction, rather than trying to artificially divide the possibilities into one or the other choice in the traditional dichotomy of relative clause and noun complement, and also trying to determine strict subcategorization frames or argument structures and relations, in Chinese we can simply follow a constructionist approach and recognize a single clausal modifying construction, which posits only a relationship between the modifying clause and the head. These two parts are both referring expressions, and so can be used alone or together. One of the core insights of the constructionist approach

is that the overall construction has meaning beyond the sum of the parts. It is the two elements (the modifying clause and the head) being together in the construction and in a particular context that allows the particular interpretation of the relation between the two and the interpretation of the referent.

4. Discussion

We have seen above that the constructions discussed are fully grammaticalized and have a consistent structural pattern, but the elements that enter into the constructions are not consistent in terms of type or semantic role. That is, these constructions do not constrain the interpretation of the identification of referents and their roles in discourse the way they do in some other languages (see also LaPolla 1993, LaPolla & Poa 2006). The most useful approach to Chinese grammar then is to take the constructions as basic, and not try to impose global categories on the language for which there is no empirical evidence, as taking the constructions as basic means there is no need for abstract global categories in individual languages or cross-linguistically. In the practice of the distributional method, one chooses constructions to define form classes, and then uses the form classes so defined to characterize the constructions. For example, a word is defined as a noun because it appears as the head of a certain type of construction, and then that construction is defined as a noun phrase because the head of the construction has been defined as a noun (on the basis of it appearing in that construction). As pointed out by Croft (2001), taking the constructionalist approach allows us to avoid this sort of circularity. Taking the constructions as basic also allows us to avoid the "methodological opportunism"[8] inherent in

[8] This term is from Croft (2001). It refers to the practice of picking and choosing which constructions one uses for defining form classes and grammatical relations to suit one's

using the distributional method for defining form classes and grammatical relations, and allows us to appreciate the diversity of structures in language, making for a much more empirically grounded linguistics.

For those of us involved in language documentation and linguistic typology, taking the constructionalist approach means that when we write grammars we do not need to have chapters on supposed global grammatical categories; we just need to present the constructions used for the propositional acts of referring, predicating, and attributing, and when we do cross-linguistic comparison we do not assume any global or universal categories and instead work inductively, looking to see what constructions are manifested in the languages, and whether the interpretation of a particular functional domain is constrained, and if so, to what extent, and also what form the construction used to constrain the interpretation in that way takes (see LaPolla 2003, 2006a-b for more discussion of this approach to typology, and particularly grammatical relations).

predilections. This problem, as well as the problem of overlapping categories, was recognized by the Structuralists (e.g. Bloomfield 1933, Harris 1946).

References

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.

Bolinger, Dwight L. 1961. Syntactic blends and other matters. *Language* 37.3:366-381.

Bolinger, Dwight L. 1976. Meaning and memory. *Forum Linguisticum* 1.1:1-14.

Bybee, Joan L. 2006. From usage to grammar: the mind's response to repetition. *Language* 82.4:711-733.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese.* Berkeley: University of California Press.

Comrie, Bernard. 1988. Coreference and conjunction reduction in grammar and discourse. *Explaining Language Universals*, ed. by John A. Hawkins, 186-208. Oxford: Basil Blackwell.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective.* Oxford: Oxford University Press.

Croft, William. 2013. Radical Construction Grammar. *The Oxford Handbook of Construction Grammar*, ed. by Thomas Hoffmann & Graeme Trousdale, 211-232. Oxford & New York: Oxford University Press.

Downing, Pamela. 1977. On the creation and use of English compound nouns. *Language* 53.4:810-842.

Fillmore, Charles J. 1982. Frame semantics. *Linguistics in the Morning Calm*, ed. by the Linguistic Society of Korea, 111-137. Seoul: Hanshin.

Finin, Timothy Wilking. 1980. *The Semantic Interpretation of Compound Nominals*. University of Illinois at Urbana-Champaign dissertation.

Foley, William A., and Robert D. Van Valin, Jr. 1984. *Functional Syntax and Universal Grammar.* Cambridge & New York: Cambridge University Press.

Grace, George W. 1987. *The Linguistic Construction of Reality.* London & New York: Croom Helm.

Harris, Zellig S. 1946. From morpheme to utterance. *Language* 22.3:161-183.

Hopper, Paul J. 2011. Emergent grammar and temporality in interactional linguistics. *Constructions: Emerging and Emergent*, ed. by Peter Auer & Stefan Pfänder, 22-44.

Berlin & Boston: Walter de Gruyter.

Hopper, Paul J. 2012. Emergent Grammar. *The Routledge Handbook of Discourse Analysis*, ed. by James Paul Gee & Michael Handford, 301-314. London & New York: Routledge.

Kay, Paul, and Karl Zimmer. 1978. On the semantics of compounds and genitives in English. *Sixth California Linguistics Association Conference Proceedings*, ed. by R. Underhill, 29-35. San Diego, California.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, Vol. 1: *Theoretical Prerequisites*. Stanford: Stanford University Press.

LaPolla, Randy J. 1993. Arguments against 'subject' and 'direct object' as viable concepts in Chinese. *Bulletin of the Institute of History and Philology Academia Sinica* 63.4: 759-813.

LaPolla, Randy J. 1997. Grammaticalization as the fossilization of constraints on interpretation: towards a single theory of cognition, communication, and the development of language. Paper presented at the City University of Hong Kong Seminar in Linguistics, November 6, 1997. Available at http://tibeto-burman.net/rjlapolla/papers/LaPolla-1997-Grammaticalization_as_the_Fossilization_of_Constraints_on_Interpretation.pdf

LaPolla, Randy J. 2003. Why languages differ: variation in the conventionalization of constraints on inference. *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff*, ed. by David Bradley, Randy J. LaPolla, Boyd Michailovsky & Graham Thurgood, 113-144. Canberra: Pacific Linguistics.

LaPolla, Randy J. 2006a. On grammatical relations as constraints on referent identification. *Voice and Grammatical Relations: In Honor of Masayoshi Shibatani*, ed. by Tasaku Tsunoda & Taro Kageyama, 139-151. Amsterdam & Philadelphia: John Benjamins.

LaPolla, Randy J. 2006b. The how and why of syntactic relations. Invited plenary address and keynote of the Centre for Research on Language Change Workshop on Grammatical Change at the Annual Conference of the Australian Linguistics Society, University of Queensland, July 7-9, 2006. To appear in Christian Lehmann,

Stavros Skopeteas, Christian Marschke (eds.), *Evolution of Syntactic Relations.* Berlin: Mouton de Gruyter.

LaPolla, Randy J. 2009. Chinese as a Topic-Comment (not Topic-Prominent and not SVO) language. *Studies of Chinese Linguistics: Functional Approaches*, ed. by Janet Zhiqun Xing, 9-22. Hong Kong: Hong Kong University Press.

LaPolla, Randy J. 2010. On the logical necessity of a cultural connection for all aspects of linguistic structure. 10th RCLT International Workshop "The Shaping of Language", La Trobe University, July 14, 2010. Downloadable podcast and handout: http://itunes.apple.com/au/itunes-u/the-shaping-of-language/id391930814

LaPolla, Randy J. 2011. On transitivity in two Tibeto-Burman languages. *Studies in Language* 35.3:639-649.

LaPolla, Randy J., František Kratochvíl, and Alexander R. Coupe. 2011. On transitivity. *Studies in Language* 35.3:469-491.

LaPolla, Randy J., and Dory Poa. 2002. Xinxi chuanda de xingzhi yu yuyan de benzhi he yuyan de fazhan [The nature of communication and language, and their influence on language development]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 2002.3:203-209.

LaPolla, Randy J., and Dory Poa. 2005. Jiaodian jiegou de leixing ji qi dui Hanyu cixu de yingxiang [The typology of focus structures and their effect on word order in Chinese]. *Jiaodian Jiegou he Yuyi de Yanjiu* [*Studies on the Structure and Semantics of Focus*], ed. by Liejiong Xu & Haihua Pan, 57-78. Beijing: Beijing Foreign Studies University Press.

LaPolla, Randy J., and Dory Poa. 2006. On describing word order. *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K Ameka, Alan Charles Dench & Nicholas Evans, 269-295. Berlin & New York: Mouton de Gruyter.

Levi, Judith N. 1978. *The Syntax and Semantics of Complex Nominals.* New York: Academic Press.

Li, Charles N. (ed.) 1976. *Subject and Topic.* New York: Academic 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, 459-489. New York: Academic Press.

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar.* Berkeley: University of California Press.

Li, Jinxi. 1924[1992]. *Xin Zhu Guoyu Wenfa* [*Chinese Grammar on the National Language*]. Beijing: The Commercial Press.

Li, Jinxi. 1953. Zhongguo yufa de "cifa" yantao [Discussion of the grammar of words in Mandarin Chinese grammar]. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 1953.9:8-12.

Lü, Shuxiang. 1979. *Hanyu Yufa Fenxi Wenti* [*Problems in the Analysis of Chinese Grammar*]. Beijing: The Commercial Press.

Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow. 1994. Idioms. *Language* 70.3:491-538.

Pawley, Andrew. 1985. Lexicalization. *Language and Linguistics: The Interdependence of Theory*, *Data*, *and Application*, ed. by Deborah Tannen & James E. Alatis, 98-120. Washington, DC: Georgetown University Press.

Shen, Jiaxuan. 2007. Hanyu li de dongci he mingci [Verbs and nouns in Chinese]. *Journal of Sino-Tibetan Linguistics* 1:27-47. Beijing: The Commercial Press.

Sun, Chaofen, and Talmy Givón. 1985. On the so-called SOV word order in Mandarin Chinese: a quantified text study and its implication. *Language* 61.2:329-351.

Thompson, Sandra A. 1973. Transitivity and some problems with the *bǎ* construction in Mandarin Chinese. *Journal of Chinese Linguistics* 1.2:208-221.

Thompson, Sandra A., and Elizabeth Couper-Kuhlen. 2005. The clause as a locus of grammar and interaction. *Language and Linguistics* 6.4:807-837.

Van Valin, Robert D. Jr., and Randy J. LaPolla. 1997. *Syntax: Structure*, *Meaning*, *and Function.* Cambridge & New York: Cambridge University Press.

Zhu, Dexi. 1982. *Yufa Jiangyi* [*Lectures on Grammar*]. Beijing: The Commercial Press.

Zhu, Dexi, Jiawen Lu, and Zhen Ma. 1961. Guanyu dongci, xingrongci "mingwuhua" de wenti [On the problem of "nominalization" of verbs and adjectives]. *Journal of Peking University* 1961.4:51-64.

從構式語法角度分析現代漢語的結構

羅仁地

南洋理工大學

本文用構式語法分析法來解決現代漢語長期無法解決的語法問題。從構式語法分析法的角度來說，構式是最基本的分析單位，所以我們在分析構式的時候，只要看構式和組成構式成分的命題功能，如指示、修飾、敘述等。因爲以構式爲主，所以不需要定全語言性的範疇，因此詞類和句法關係這一類的問題就不存在。

關鍵詞：現代漢語，構式語法，句法關係，詞類，子句性修飾結構

Human Language Resources and Linguistic Typology
Papers from the Fourth International Conference on Sinology
Taipei: Academia Sinica, 2013

Reflections on Typological Characterization of Chinese Grammar*

James H-Y. Tai
National Chung Cheng University

Chinese has been characterized in different typological frameworks as isolating, analytic, discourse-oriented, pragmatics-oriented, topic-prominent, tense-lacking, pervasive in iconicity, and satellite-framed or equi-pollent-framed, and so on. This paper takes a relativist view and leans toward functionalism and cognitive grammar. It proposes that isolating and analytic features together with monosyllabicity cluster holistically as the 'gestalt' of Chinese. It also suggests the desirability of examining Chinese from the perspectives of sign languages and creoles.

Keywords: Chinese, typology, relativism, analyticity, iconicity, sign language, creoles

* This is a much revised version of the paper presented at the 4th International Conference on Sinology–Session on Linguistics, held at Academia Sinica, June 20-22, 2012. I am very grateful to the anonymous reviewers for their careful reading and very valuable comments and suggestions. Of course, I am solely responsible for any infelicities herein.

1. Introduction

The main purpose of this paper is to reflect on the characteristics of Chinese grammar within different frameworks of language typology. Since the characteristics of any language can only become apparent when they are compared to those of other languages within different typological frameworks, it is necessary to consider the goal of linguistic typology within different theoretical frameworks with different philosophical orientations.

The goal of linguistic typology is not only to provide a description of the range of variation among natural languages but also to search for general principles and language-specific rules. The general principles cannot properly be stated without reference to theoretical frameworks with different philosophical orientations toward nature of human language in both form and meaning (cf. Comrie 1981, Hammond et al. 1988, Shibatani & Bynon 1995).

There have been two different philosophical orientations that provide the backdrop for studies in linguistic typology since the historical and comparative study of languages in 18th and 19th centuries in Europe, namely, rationalism and empiricism. Linguistic universalism is based on rationalism, assuming that there are underlying mental invariants and variants exist only on the surface, and thus one can take a deductive approach to uncover invariants and surface variations by parameterization. In contrast, linguistic relativism is based on empiricism, taking the position that individual languages are historical entities developed by their speakers according to cultural needs for communication to construct their respective societies for survival. Each individual language thus represents a unique segmentation of the external world and the universe of human experience.

Linguistic universalists are rationalists. They adopt a deductive approach and focus on the underlying similarities across languages. They feel confident to

apply the metalanguage developed from one single language, say English, to those of other individual languages to uncover universality. They also adopt formal analyses for the purpose of constructing universals and parameterization principles. For example, in the works of practitioners of different versions of generative grammar, and their corresponding typological studies, categorial universalism is assumed, and cross-linguistic formal categories are adopted, and parameterization principles for variants are discrete (Newmeyer 2010).

Linguistic relativists are empiricists. They adopt an inductive approach and focus on the differences and variants rather than similarities and invariants. They take the position that individual languages are developed for communication in different cultures and societies, and they often opt for functional approaches and cognitive-functional approaches.[1] Their typological framework assumes categorial particularism and comparative concepts such as relative clause, passive voice, and adposition (Haspelmath 2010a, 2010b). Furthermore, instead of discrete parameterization, gradient divergence from prototype as stated in Sapir (1921) is the rule of thumb for language typology.

It should be noted that with respect to universalism vs. relativism, scholars in the past made a distinction between lexicon and syntax. For example, Humboldt supported universalism for grammar but he departed from the universalist position in claiming that concepts are language-specific.[2] Chomsky (1980) made a clear distinction between the computational and the conceptual system for human

[1] Language for communication is the basic tenet for functionalism. The caveat, as pointed out by one anonymous reviewer, is that "not all functional linguists reject the existence of linguistic universals, but they typically assume that any universals, should they exists, are likely to be grounded in certain functional or cognitive primitives." Also, in theory, formal analyses need not preclude an empirical approach as often in practice.

[2] See Losonsky (1999) on Humbolt's view of language diversity and Chomsky's misinterpretation of Humbolt's thoughts on human language.

languages. Computational system is universal, even innate, but conceptual system can be relative from culture to culture. One is tempted to assume that while computational system underlies the core syntax in minimalism, conceptual system involves conceptualization in lexicons and language-specific morphological and syntactic constructions which have construction meanings (Tai 2002a).[3] However, the distinction between lexicon and syntax along with innate universals has been rejected by cognitive grammarians (Langacker 1987, 1991, Goldberg 1995, 2013, Croft 2001). One important reason for me to adopt a relativist approach to Chinese grammar is that due to the lack of an indigenous metalanguage for analyzing Chinese grammar. Chinese grammarians have in general adopted metalanguage compatible with inflectional Indo-European languages since *Ma-shi Wentong* (馬氏文通) in 1898.[4]

2. Chinese in different typological frameworks

While it is commonly held that linguistic typology was developed by Greenberg in the 1960s, European scholars in the 19th century had already worked on language classification based on morphological characteristics (Greenberg 1974:10-34). Indeed, Shibatani & Bynon (1995:1-25) noted that August Schleicher (1821-68) observed that morphologically different types of languages have different ways of expressing grammatical relations of subject and object. He pointed out that in 'isolating' languages such as Chinese which do not have morphology, the grammatical relations of subject and object are expressed by word order; while in 'agglutinative' languages, they are expressed by affixes

[3] As pointed out by the anonymous reviewer that to mix minimalism and construction grammar would be to put two incompatible linguistic ideas together.

[4] See Tai (1989) for a brief discussion of metalanguage issues and his proposed cognition-based functional grammar of Chinese for a partial solution.

attached to the root, and in 'inflectional' languages, they are expressed fusionally with the unit expressing the root meaning. Chinese was characterized as an isolating language in which the grammatical relations are expressed by word order (Shibatani & Bynon 1995:1-25).[5] In this view, Chinese is an SVO language.[6]

Sapir (1921) also added a dimension of 'synthesis' in terms of morphological complexity in words. This dimension is also gradient, ranging from 'analytic' to 'synthetic' to 'polysynthetic', with 'mildly synthetic', and 'mildly polythetic'. Thus, Chinese is both 'isolating' and 'analytic', in contrast with 'fusional' and 'analytic' languages (e.g. English), 'agglutinative' and 'polysynthetic' languages (e.g. Nookta), and 'fusional' and 'polysynthetic' languages (e.g. Algonquin), and so on. In the total classificatory scheme arrived at by Sapir, Chinese is 'pure-relational, isolating, and analytic' (ibid.:138).

The analyticity and monosyllabicity of words in Chinese allows the language to use a small set of monosyllabic words to generate a larger set of compound words by subcategorization based on the modifier-head categorization principle. To wit,

(1) a. *che* 'vehicle': *huoche* 'train', *qiche* 'car', *kache* 'truck', etc.
b. *yu* 'fish': *guiyu* 'salmon', *xueyu* 'cod', *zunyu* 'trout', etc.
c. *hua* 'flower': *lanhua* 'orchid', *meiguihua* 'rose', *mudanhua* 'peony'
d. *cai* 'vegetable': *baoxincai* 'cabbage', *qincai* 'celery', *huayecai* 'cauliflower', etc.

[5] It appears that this important observation of correlative patterns in language anticipated the different versions of implicational universals proposed later by Sapir (1921), Jakobson (1941), and Greenberg (1966).

[6] However, Tai (1973) argued that Chinese can also be characterized as an SOV language based on Greenberg's implicational universals (1966).

Tai (1984) argues that Chinese in general doesn't have monosyllabic accomplishment verbs. Instead, resultative verb compounds which express action and result are used. Thus, there is no monosyllabic Chinese word for 'to kill' (Tai & Chou 1975). Instead, one finds,

(2) *shasi* 'to kill by using instruments', *dasi* 'to hit to death', *dusi* 'poison', etc.

I have taken the position that the resultative complements are semantic heads serving as the center of predication, with the action verbs serving as modifiers specifying the various methods to 'to cause to die' (Tai 2003). Here, *si* (死) 'to die, to cause to die' is subcategorized by different kinds of action verbs. The contrast between opaque English accomplishment verbs and transparent Chinese action-verb compounds again attest to the explanatory value of Sapir's characterization of Chinese as analytic. Chinese resultative verb compounds also present a problem for Talmy's (1985) well-known typological dichotomy between 'verb-framed' languages like French and Spanish on the one hand, and 'satellite-framed' languages like English and German. Talmy (1985) also treated Chinese as a 'satellite-framed' language since the cognitive component of Manner is incorporated in verbs in this language. However, Slobin (2000) argues that serial verb languages including Chinese should belong to the third-type, referred to as 'equi-pollent' language.[7]

Constructions based on modifier-head categorization principle are also pervasive in nominal phrases and relative clauses (both gapped and gapless)

[7] Slobin (2000) remarks that "serial-verb languages like Chinese may represent a third type of lexicalization pattern, lying between S-languages and V-languages." (ibid.:134) This third type is referred to by Slobin as 'equi-pollent' in contrast with 'satellite-framed' and 'verb-framed' languages.

constructed with the modifier marker *de* (的) in Chinese. An illustration,

(3) ta de shu
he DE book
'his book'

(4) hao de shu
good DE book
'good books'

(5) shujia-shang de shu
bookshelf-top DE book
'books on the bookshelf'

(6) ta xie Ø de shu (gapped relative clause)
he write GAP DE book
'the books he wrote'

(7) ta shuo hua de shengyin (gapless relative clause)
he speak word DE voice
'the voice of his talking'

The function of relative clauses is to restrict and identify referents, and categorization is a necessary first step for referential identification. Examples (3)-(7) also show that the *de* (的) construction does not differentiate between relative clauses from other types of modifying constructions. Comrie (1996) and Shuanfan Huang (2007) therefore argue that Chinese doesn't have relative clauses. On the other hand, the existence of relative clauses is supported by formal arguments made by Huang, Li & Li (2009) and the experimental findings made by Charles Lin (2008) that Chinese subject-extracted relative clauses are easier

than object-extracted relative clauses to process, no different from English and other languages. Nonetheless, in accordance with a series of works by Matsumoto (1988, 1997) on Japanese relative clauses, I would like to take a relativist view to the effect that both East Asian languages and European languages have structurally-operated relative clauses for the same function of restricting and identifying referents, but East Asian languages use the nominal categorization marker *de* (的) to group both gapped and gapless relative clauses under the general category of noun modification constructions.

In recent years, C.-T. James Huang (2005, 2012) has adopted the concept of analyticity to the analysis of light verb in Chinese to account for some systematic differences between Chinese and English syntax, including the absence of English denominal verbs (e.g. *to water*, *to shelf*, etc.). Thus, it appears that 'analyticity' as well as 'isolating' typological features string lexicon and syntax together, as been observed by earlier Chinese scholars, morphology and syntax in Chinese share the same combinatory rules.

Sapir's approach, though morphology-based, took two important departures from the 19th century classical morphological typology in Europe. First, it is gradient and relative, as opposed to absolute. Second, languages types are defined in terms of combination of properties, as opposed to single features. In essence, the classical theory is based on the classical approach to categorization, while Sapir's approach is based on the prototype theory of categorization. Greenberg was a student of Sapir. The well-known implicational universals proposed by Greenberg (1966:73) take the logic form of X → Y, that is, in a particular language, if we find X, we always find Y, but not conversely. The implicational universals are based on surface syntactic structures sampled from 30 languages with a wide range of genetic and areal coverage. Generative grammarians of different generations, however, have translated Greenberg's statistical tendencies

across languages into law-like implicational universals for the sake of formal formulation as well as for exhibiting the explanatory power of different versions of generative grammars. It appears that while most cognitive linguists subscribe to prototype theory of categorization, generativists subscribe to the classical theory of categorization. The notion of 'type' in Sapir's framework is 'holistic', arising from a 'cluster of properties' (Greenberg 1974). In essence, the 'analyticity' together with 'monosyllabicity' can be used to characterize 'holistically' the 'gestalt' of Chinese.

With this 'holistic' view, we can give a brief review of typological characteristics of Chinese which have been stated by various Chinese grammarians. First, Chinese has been characterized by various scholars as more 'discourse-oriented' or 'pragmatic-oriented' (Chao 1968, Tsao 1990, Y. Huang 1994, Chu 1998). Chao's (1968) proposal of double-subjects in Chinese was redefined by Li & Thompson's (1976) characterization of Chinese as a topic-prominent language. In a similar vein, Tsao (1990) proposed that discourse chunks linked to the same topic form a long but one single sentence rather finite sentences defined by tense in English. C.-T. James Huang's (1984, 1989) proposal of Chinese as a pro-drop language in GB, and researches on word order and definiteness/specificity of bare nouns, and occurrences of BA and BEI sentences in conversations and texts are all pointing to the discourse-oriented nature of the Chinese grammar. In accounting of tense/aspect interpretations, Li & Thompson (1981) use a functional approach to describe how Chinese uses aspect to derive tense interpretation, and in a series of work on tense and aspect in the framework of formal semantics, Jo-wang Lin (2006) also views Chinese as a pragmatic-oriented language lacking formal tense categories but relies on aspect marking to determine tense in a sentence.

Tai (1985, 2002b, 2011) has demonstrated that word order in Chinese can be explained in terms of the principle of temporal sequence to the effect that 'the relative word order between two syntactic units is determined by the temporal order of the states which they represent in the conceptual world' (Tai 1985:50). An illustration:

(8) a. Zhangsan [dao tushuguan] [na shu].
Zhangsan reach library take book
'John went to the library to get the book.'

b. Zhangsan [na shu] [dao tushuguan].
Zhangsan take book reach library
'John took the book to the library.'

While the principle of temporal sequence is characterized by Tai as a syntactic principle of linearization in Chinese, Newmeyer (1992, 1998) argues that temporal sequence in Chinese is not a syntactic principle, but rather a grammaticalized discourse principle. He argues that the meaning difference in the two sentences in (4) can be accounted for by the well-known conversational implicature in temporally-ordered conjoined sentences.

(9) a. Mary bought some motor oil and went to the supermarket.
b. Mary went to the supermarket and bought some motor oil.

Putting the debate between Newmeyer and Tai aside, it is clear that Chinese word order is more in line with temporal sequence than English word order as can be illustrated by (10).

(10) a. Ta cong gongyuan zou-dao tushuguan.
he from park walk-arrive library
1 2 3
'He walked from the library to the park.'

b. *Ta zou-dao tushuguan cong gongyuan.
he walk-arrive library from park
'He walked to the park from the library.'

Newmeyer cannot explain why temporal sequence is more pervasive in word order in Chinese. As a matter of fact, the so-called prepositions in Chinese were treated as 'co-verbs' (DeFrancis 1964, Li & Thompson 1974). Tai (2011) argues that 'co-verbs', like main verbs, are susceptible to temporal interpretation. Thus, 'cong' means 'starting from' and not just 'from', and 'dao' doesn't mean 'to' but means 'to arrive at'. Various types of sentences including serial verb constructions as illustrated below are subject to the principle of temporal sequence.

(11) Didi da-po-le beizi. (action-result)
younger brother hit-break-ASP cup
'The younger brother broken the cup.'

(12) Mama qu shichang mai cai. (action-purpose)
mother go market buy vegetable
'The mother went to the market to buy vegetable.'

(13) Meimei zhuanshen likai fangjian. (consecutive actions)
younger sister turn around leave room
'The younger sister turned around and left the room.'

(14) Chushi yong dao qie rou. (manner-action)
cook use knife cut meat
'The cook cut the meat with a knife.'

Chinese is 'pragmatic-oriented' can further be observed in the relatively free selection of surface arguments in Chinese (Tzong-Hong Lin 2001) as illustrated below,

(15) chi niuroumian (theme object)
eat beef noodles
'eat beef noodles'

(16) chi Sichuan guan (location)
eat Sichuan restaurant
'dine at a Sichuan restaurant'

(17) chi da wan (instrument)
eat big bowl
'eat the food in the big bowl'

(18) chi wanshang (time)
eat evening
'(The banquet) is in the evening.'

(19) chi touteng (reason)
eat headache
'(The medicine) is for headache.'

(20) Zaijia chi fumu, chuwai chi pengyou. (metaphorical)
at home eat parents go outside eat friend
'One lives on his parents when at home, but on friends when traveling.'

Examples (15)-(20) show that a transitive verb in Mandarin Chinese like *chi* ‘to eat’, besides its regular theme object argument, can take location, instrument, time, reason, and metaphorical expressions as its object argument. Tzong-Hong Lin (2001) adopted light verb analysis to account for this and other kinds of selection of subject and object arguments in Mandarin Chinese. Thus, the surface transitive verb *chi* embedded under the empty higher light verb phrase and verb phrases containing abstract verbs such as AT, USE, and FOR. However, the formal account would not be able to explain why the transitive verb *he* ‘to drink’ cannot have the same set of unselected object arguments as *chi* ‘to eat’. One is tempted to speculate here that eating activities occupy a central place in Chinese culture, and for communicative efficiency, its syntax is simplified with pragmatic inferences. In a frequency count by Tao (2000), the frequency of *chi* (吃) ‘to eat’ is much higher than *he* (喝) ‘to drink’ and other related verbs. Tzong-Hong Lin (2001) also observes similar phenomenon in subject selection for Chinese verbs. Thus, in addition to agent argument (21), locative and patient arguments can also be selected in the subject position as shown in (22) and (23).

(21) Laozhang kai-le yi-liang tanke-che. (agent)
Laozhang drive-ASP one-CL tank
‘Laozhang drove a tank.’

(22) Gaosu-gonglu kai-zhe yi-pai tanke-che. (location)
express-highway drive-ASP one-CL tank
‘There is a line of tanks on the expressway.’

(23) Zhe-liang poche kai-de wo xia-si le. (patient)
this-classifier broken car drive-DE I scare-die LE
‘Driving this broken car made me scared to death.’

The relative free selection of arguments on the surface structure abounds in Chinese. While the light verb theory accounts for the phenomena in elegant syntactic manners, I am tempted to think that on closer examination, the light verb theory would run into the same kind of complication as it is applied to the derivation of denominal verbs (to be discussed in §3). It is important to observe that while Chinese grammar tends to be more pragmatically oriented in argument selection than English, it is not so pragmatically oriented as English in the formation of denominal verbs (Tai 1997). Perhaps, following Sapir's idea of gradient typology, we can develop a framework in which we can measure different degrees of being 'pragmatically-oriented'.

As a matter of fact, pragmatics entrenches every language with various kinds of deictic categories as can be illustrated by (24) of which all three deictic categories of person, time, and place are employed.

(24) I'll be here tomorrow.

The famous Chinese ambiguous sentences like (25) are no different from English sentences like (26). Both are subject to different interpretations depending on contexts.

(25) Ji chi-le.
chicken eat-ASP
'The chicken was fed.' / '(We) have eaten the chicken.'

(26) Missionaries are ready to eat.

We also find some aspects of English syntax to be more pragmatically oriented than in Chinese. This has to do with abundance of denominal verbs, be they established or innovative.

3. Denominal verbs

In English and many other languages including French, German, Spanish, and Indonesian, certain words naming concrete objects, such as ‘bottle’, ‘skin’, ‘truck’ and ‘water’, can also be used as verbs. These verbs, ‘to bottle’, ‘to skin’, ‘to truck’ and ‘to water’, are used to name events associated with the corresponding concrete objects. In the literature on English grammar, these verbs have been referred to as ‘denominal verbs’ and has been treated by Jespersen (1942) as a shift in morphological category from noun to verb. This morphological process has been referred to as ‘conversion’ or ‘zero derivation’ (Lyons 1977:522ff). Generative semanticists such as McCawley (1971) and Green (1974) derived denominal verbs from a conflation of underlying universal semantic constants, such as ‘to cause something to be in the bottle’ and ‘to cause the skin to be removed’. In a more recent treatment by generative syntacticians such as Hale & Keyser (2002), the Larsonian VP-shells theory was applied to derive denominal verbs from moving a noun upward to position of ‘light’ verb in the VP-shell. Culicover & Jackendoff (2005) correctly point out that Hale & Keyser’s treatment of denominal verbs is inadequate for at least three reasons. First, one cannot predict which particular nouns can become verbs. For instance, ‘chair’ and ‘table’ can become ‘to chair’ and ‘to table’ but ‘desk’ and ‘sofa’ cannot. Second, the theory cannot predict idiosyncratic meanings associated with denominal verbs. For instance, ‘to father a child’ means ‘to bring about the child’s birth’, but ‘to mother his students’ means ‘to treat his students like a mother.’ Third, the light verb treatment cannot predict the thematic status of the associated nouns. For instance, ‘to carpet the van’ means ‘to cover the van with carpet’, but not ‘to put her van in the carpet’ as in ‘to garage the van’, which means to put the van in the garage. In fact, we need only to take a good look at the denominal verbs in

English as documented in Clark & Clark (1979) to realize that the creation and meaning of denominal verbs in English as well in other language are determined by pragmatic and not syntactic factors.

Clark & Clark (1979) argue that denominal verbs should be treated as contextual expressions rather than denotational or indexical expressions. Particularly with respect to innovative denominal verbs, such as 'to porch the newspaper' (meaning 'to put the newspaper on the porch' as by the newspaper carrier) and 'to Houdini one's way of the locked closet' (meaning 'to escape by trickery'), they propose that such contextual expressions shifted sense and denotation according to different contexts. These are distinguished from denotational expressions, such as 'man' and 'bachelor', which have fixed sense and denotation, and from indexical expression, such as 'he' and 'the bachelor'.

Clark & Clark propose a denominal verb convention to treat innovative denominal verbs in English. This convention, the Innovative Denominal Verb Convention (IDVC), patterned after Grice's cooperative principle, is stated as below:

The Innovative Denominal Verb Convention (IDVC)

> In using an innovative denominal verb sincerely, the speaker means to denote the kind of situation that he has good reason to believe that on this occasion the listener can readily compute **uniquely** on the basis of their mutual knowledge in such a way that the parent noun denotes one role in the situation, and the remaining surface arguments of the denominal verb denote other roles in the situation. (Clark & Clark 1979:787)

The main idea in Clark & Clark's theory is that, in using an innovative denominal verb, the speaker intends the listener to come to a unique inter-

pretation of what he has said, not only from the meanings of the words alone, but also from the context as well on the basis of what they mutually know. Thus, as contextual expressions, innovative denominal verbs can have, in theory, a very large, if not infinite, number of senses. Clark & Clark's theory appears to account for established denominal verbs as well as innovative ones. It explains the phenomenon in English that an established denominal verb can often have a number of conventionalized meanings. For example, 'to water' in English can mean 'to moisten, to sprinkle, to soak with water'; in addition, it has other meanings, including 'to supply with water for drink', 'to supply water to' and 'to dilute by the addition of water'.

It should be noted that the demarcation between innovative verbs and established innovative verbs cannot always be clearly made. Once an innovative denominal verb is introduced, it may become fully established. Alternatively, it may have become established for some speakers but not for others in a speech community; or, it may even fall into disuse completely. For example, 'to parent' is still not acceptable to many speakers ever though it is widely used. The denominal verb 'to money' now seems unacceptable in British English even though the *Compact Oxford English Dictionary* lists the following meanings: 'to mint money; to supply with money; to furnish money for an undertaking'. In short, every language is pragmatically oriented, but in different degrees and in different manners. Thus, with respect to some aspects of syntax, Chinese is more pragmatically-oriented than English, but with respects to some other aspects of syntax, English is more pragmatically-oriented than English. It is only after we have identified all or most of the principles of pragmatic entrenchment in syntax that we are able to construct a continuum for the degree of pragmatic entrenchment in different languages.

4. Iconicity

In a series of works arguing against the arbitrariness and autonomy thesis of syntax, Tai (1985, 1993, 1994) has demonstrated the pervasiveness of iconicity in Chinese grammar. Important iconic motivations identified by Haiman (1980, 1985) are (1) order motivation, (2) distance motivation, (3) separateness motivation, (4) juxtaposition motivation, and (5) reduplication motivation. They appear more clearly in Chinese than in English. For instance, the order motivation can be stated as "the order of linguistic expressions corresponds to their order in the conceptual order". The principle of temporal sequence is just an obvious application of this principle. Iconicity in signs in human languages originally proposed by Peirce (1932), expounded later mainly by Jakobson (1968[1941]), and Haiman (1980, 1983) presents a fundamental different view against Saussurian doctrine of arbitrariness of human languages. As a matter of fact, Saussure (1959[1916]) also proposes the notion of motivation to account for the transparency and translucency of internal structure of lexical items.

We have earlier noted the analytic nature of Chinese noun and verb compounds. It is also worth noting that a large number of nominal expressions such as names for the twelve months and names for the seven days in a week are constructed with numbers. For example, *shieryue* (十二月) 'December' and *xingqiliu* (星期六) 'Saturday'. The decimal system for accounting in Chinese also provides an epitome of analyticity and transparency of numerals. For example, *shiyi* (十一) 'eleven' and *shier* (十二) 'twelve'. Recently, Tsai (2011) has also argued that iconicity in Chinese syntax is also largely due to the analyticity of syntactic structures for the transparent mapping between syntax and semantics.

5. Chinese, Creole and sign language

There are two major modalities of natural human language: visual-gestural modality of signed languages and auditory-vocal modality of spoken languages. These two modalities of human language share several fundamental properties. However, there are also modality effects which contribute to the drastic differences between signed and spoken languages in lexicon, morphology and syntax. Two most important effects are iconic representation of objects and actions and indexic/ostensive identification of referents in signed languages (Meier 2002). These two modality effects result in relative uniformity in morphology and syntax across signed languages.

At the same time, signed languages share some similarities with creoles because of their similar ambience of language acquisition (Fischer 1978, Aronoff, Meir & Sandler 2005). As summarized in Aronoff, Meir & Sandler (2005), "These commonalities [between creoles and sign languages] include: no distinction between tensed and infinitival clauses, no tense marking but a rich aspectual system, no pleonastic subjects, no true passives, the occurrence of transitive verbs with agent subjects as intransitives with patient/theme subjects as well, pervasive topic-comment word order; both young creole languages and ASL make extensive use of content words as grammatical markers; neither young creole languages nor ASL use prepositions to introduce oblique cases; both use preverbal free morphemes to express completive aspect; and both rely heavily on prosodic cues for expressing certain syntactic relations (such as those encoded by relative clauses and conditionals in other languages)." (ibid.:307)

One cannot fail to notice that Chinese grammar also exhibits common structural features shared between sign languages and young creole languages. Yet, Chinese is not a young language. The Chinese puzzle is therefore created: if

the development of inflectional morphology in spoken language is a function of age as cogently argued by (Aronoff, Meir & Sandler 2005), why didn't Chinese, during the course of its long history, develop a rich inflectional morphology, as did European languages? The question can be tentatively answered, if pending further evidence, we hypothesize that Chinese was a creole language to start with, and that Chinese had opted to utilize functional mappings rather than inflections for making distinctions among different word classes to indicate different parts of speech. This strategy of functional mapping is compatible with Nisbett's (2003) theory that Chinese cognition focuses on relations between individuals rather than on the attributes of an individual. The introduction of Chinese characters for monosyllabic words in the early history of this language may also have contributed to the perpetuation of the monosyllabic structure.

6. Conclusion

In this paper, I have briefly examined different frameworks of language typology and how Chinese grammar has been characterized in these different frameworks. In summary, Chinese is isolating, analytic, verb-initial or verb-final, discourse-oriented, pragmatic oriented, topic-prominent, tense-lacking, pervasive in iconicity, creole-like, and satellite-framed or equi-pollent-framed, and so on, depending on different philosophical orientations, different theoretical frameworks and approaches to linguistic analysis and language universals. One thing that seems to be clear is the continuum with respect to any typological features employed for cross-linguistic comparison. On this continuum, lexicon is then largely relative, while syntax is limitedly relative. Relative clause constructions serve as a good case in point. While recursive and embedded structure along with communication function of identifying a referent in discourse is cross-linguistic

the same, other details can be very different. Thus, relative clauses in Chinese subsume both 'gapped' and 'gapless' under the modification construction marked with *de* (的). In contrast, English only allows 'gapped' relative clauses and uses *wh*-words as relative pronouns. Similarly, grammatical categories including traditional notions of parts of speech can be relative within and across languages (McCawley 1992, Bisang 2008).

Following Sapir (1921), I have explored language typology of Chinese in an attempt to construct a holistic view with which we can hope to find the 'gestalt' of Chinese grammar. Jakobson (1971) points out that in contrast to Saussure's claim that "signs that are wholly arbitrary realize better than the others the ideal of the semiological process" (1959[1916]:68), Peirce's thesis is that a system of signs blending as equally as possible all three types of signs—namely, icon, index and symbol—is 'the most perfect of signs' (Jakobson 1971). This means that each language blends the three types of signs differently, and the 'gestalt' of each language is to be in the particular way it blends all the grammatical rules of the three types.

In this paper, I have also suggested the desirability to use typological characteristics drawn from signed languages to provide new perspectives for language typology. Design features of human language and language universals as well as language typology cannot be constructed solely based on spoken languages.

Along with sign language, co-speech gestures have recently gained the attention of cognitive linguistics (McNeill 2000, Kendon 2004). From the view point of functionalism, we communicate not only with speech or sign but also with the accompanying gestures. If we do not want to limit ourselves to only idealized spoken language, then a comprehensive theory of language universals

and typology must encompass our deep understanding of both co-speech gestures and signed languages.[8]

References

Aronoff, Mark, Irit Meir, and Wendy Sandler. 2005. The paradox of sign language morphology. *Language* 81.2:301-344.

Bisang, Walter. 2008. Precategoriality and syntax-based parts-of-speech: the case of Late Archaic Chinese. *Studies in Language* 32.3:568-589.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese.* Berkeley: University of California Press.

Chomsky, Noam. 1980. *Rules and Representations*. New York: Columbia University Press.

Chu, Chauncey C. 1998. *A Discourse Grammar of Chinese*. New York: Peter Lang.

Chui, Kawai. 2002. Categorization of gestures in communication. *Form and Function: Linguistic Studies in Honor of Shuanfan Huang*, ed. by Lily I-wen Su, Chinfa Lien & Kawai Chui, 105-129. Taipei: Crane.

Clark, Eve V., and Herbert H. Clark. 1979. When nouns surface as verbs. *Language* 55.4:767-811.

Comrie, Bernard. 1981. *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: University of Chicago Press.

Comrie, Bernard. 1996. The unity of noun-modifying clauses in Asian languages. *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics*, Vol. 3, 1077-1088. Salaya: Institute of Language and Culture for Rural Development, Mahidol University.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological*

[8] As a case in point, while Spanish is verb-framed languages with 'manner' component of the main verbs expressed with a adjunctive clause expressing manner (e.g. 'He entered the room running' in English), Spanish speakers can simply substitute the 'running' by gesture (McNeill 2000). In fact, it has been reported by Chui (2002) that gesture can complement a Chinese verb such as *zai* (載) 'to carry' with gestures indicating to carry with a motorcycle or car.

Perspective. Oxford & New York: Oxford University Press.

Culicover, Peter W., and Ray Jackendoff. 2005. *Simpler Syntax*. Oxford & New York: Oxford University Press.

DeFrancis, John. 1964. *Intermediate Chinese.* New Haven: Yale University Press.

Fillmore, Charles J. 1997. *Lectures on Deixis.* Stanford: CSLI.

Fischer, Susan D. 1978. Sign language and creoles. *Understanding Language through Sign Language Research*, ed. by Patricia A. Siple, 309-331. New York: Academic Press.

Goldberg, Adele E. 1995. *A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago Press.

Goldberg, Adele E. 2013. Explanation and constructions: response to Adger. *Mind and Language* 28.4:479-491.

Green, Georgia M. 1974. *Semantics and Syntactic Regularity*. Bloomington: Indiana University Press.

Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Language* (2nd edition), ed. by Joseph H. Greenberg, 73-113. Cambridge: MIT Press.

Greenberg, Joseph H. 1974. *Language Typology: A Historical and Analytic Overview:* The Hague: Mouton.

Haiman, John. 1980. The iconicity of grammar: isomorphism and motivation. *Language* 56.3:515-540.

Haiman, John. 1983. Iconic and economic motivation. *Language* 59.4:781-819.

Haiman, John. (ed.) 1985. *Iconicity in Syntax.* Amsterdam & Philadelphia: John Benjamins.

Hale, Kenneth L., and Samuel Jay Keyser. 2002. *Prolegomena to a Theory of Argument Structure.* Cambridge: MIT Press.

Hammond, Michael, Edith A. Moravcsik, and Jessica R. Wirth. 1988. *Studies in Syntactic Typology.* Amsterdam & Philadelphia: John Benjamins.

Haspelmath, Martin. 2010a. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86.3:663-687.

Haspelmath, Martin. 2010b. The interplay between comparative concepts and descriptive categories (Reply to Newmeyer). *Language* 86.3:696-699.

Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15.4:531-574.

Huang, C.-T. James. 1989. Pro-drop in Chinese: a generalized control theory. *The Null Subject Parameter*, ed. by Osvaldo Jaeggli & Kenneth J. Safir, 185-214. Dordrecht & Boston: Kluwer.

Huang, C.-T. James. 2005. On syntactic analyticity and the other end of the parameter. Lecture notes from 2005 LSA Linguistic Institute course. Manuscript. Cambridge: Harvard University.

Huang, C.-T. James. 2012. On macrovariations and microvariations in parametric theory. *Proceedings of the 13th International Symposium on Chinese Languages and Linguistics* (*IsCLL 13*), ed. by Yung-O Biq & Lindsey Chen, 1-18. Taipei: National Taiwan Normal University.

Huang, C.-T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese.* Cambridge & New York: Cambridge University Press.

Huang, Shuanfan. 2007. Is there a relative construction in spoken Chinese? Paper presented at the First Conference on Language, Discourse and Cognition, May 18-19, 2007. Taipei: National Taiwan University.

Huang, Yan. 1994. *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Chinese.* Cambridge & New York: Cambridge University Press.

Jackendoff, Ray. 1990. *Semantic Structures.* Cambridge*:* MIT Press.

Jakobson, Roman. 1968[1941]. *Child Language Aphasia and Phonological Universals*. The Hague: Mouton.

Jakobson, Roman. 1971. Quest for the essence of language. *Selected Writings*, Vol. 2: *Word and Language*, 343-359. The Hague: Mouton.

Jespersen, Otto. 1942. *A Modern English Grammar on Historical Principles*, Vol. 6: *Morphology.* Copenhagen: Munksgaard.

Kendon, Adam. 2004. *Gesture: Visible Action as Utterance*. 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. 1991. *Foundations of Cognitive Grammar*, Vol. 2: *Descriptive Application.* Stanford: Stanford University Press.

Li, Charles N., and Sandra A. Thompson. 1974. Co-verbs in Mandarin Chinese: verbs or prepositions? *Journal of Chinese Linguistics* 2.3:257-278.

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.

Lin, Chien-Jer Charles. 2008. The processing foundation of head-final relative clauses. *Language and Linguistics* 9.4:813-838.

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.

Losonsky, Michael. (ed.) 1999. *On Language: On the Diversity of Human Language Construction and its Influence on the Mental Development of the Human Species.* Cambridge & New York: Cambridge University Press.

Lyons, John. 1977. *Semantics*, Vol. 2. Cambridge & New York: Cambridge University Press.

Matsumoto, Yoshiko. 1988. *Grammar and Semantics of Adnominal Clauses in Japanese*. Berkeley: University of California dissertation.

Matsumoto, Yoshiko. 1997. *Noun-modifying Constructions in Japanese: A Frame Semantic Approach.* Amsterdam & Philadelphia: John Benjamins.

McCawley, James D. 1971. Prelexical syntax. *Linguistic Development of the Sixties: Viewpoints for the Seventies*, ed. by Richard J. O'Brien, 19-33. Washington, DC: Georgetown University Press.

McCawley, James D. 1992. Justifying part-of-speech assignments in Mandarin Chinese.

Journal of Chinese Linguistics 20.2:211-245.

McNeill, David. (ed.) 2000. *Language and Gesture.* Cambridge & New York: Cambridge University Press.

Meier, Richard P. 2002. Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. *Modality and Structure in Signed and Spoken Languages*, ed. by Richard P. Meier, Kearsy Cormier & David Quinto-Pozos, 1-25. Cambridge & New York: Cambridge University Press.

Newmeyer, Frederick J. 1992. Iconicity and generative grammar. *Language* 68.4:756-796.

Newmeyer, Frederick J. 1998. *Language Form and Language Function.* Cambridge: MIT Press.

Newmeyer, Frederick J. 2010. On comparative concepts and descriptive categories: a reply to Haspelmath. *Language* 86.3:688-695.

Nisbett, Richard E. 2003. *The Geography of Thought: How Asians and Westerners Think Differently and Why.* New York: Free Press.

Peirce, Charles S. 1932. *Elements of Logic.* Cambridge: Belknap Press of Harvard University Press.

Pinker, Steven. 1989. *Learnability and Cognition.* Cambridge: MIT Press.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech.* New York: Harcourt, Brace and Company.

Saussure, Ferdinand de. 1959[1916]. *Course in General Linguistics*, ed. by Charles Bally & Albert Sechehaye, translated by Wade Baskin. New York: McGraw-Hill.

Shibatani, Masayoshi, and Theodora Bynon. 1995. Approaches to language typology: a conspectus. *Approaches to Language Typology*, ed. by Masayoshi Shibatani & Theodora Bynon, 1-25. Oxford: Clarendon Press.

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.

Tai, James H-Y. 1973. Chinese as a SOV language. *Chicago Linguistic Society* (*CLS*) 9:659-671. Chicago: Chicago Linguistic Society.

Tai, James H-Y. 1984. Verbs and times in Chinese: Vendler's four categories. *Papers*

from the Parasession on Lexical Semantics, 289-296. Chicago: Chicago Linguistic Society.

Tai, James H-Y. 1985. Temporal sequence and Chinese word order. *Iconicity in Syntax*, ed. by John Haiman, 49-72. Amsterdam & Philadelphia: John Benjamins.

Tai, James H-Y. 1989. Towards a cognition-based functional grammar of Chinese. *Functionalism and Chinese Grammar*, ed. by James H-Y. Tai & Frank F. S. Hsueh, 187-226. South Orange: Chinese Language Teachers Association.

Tai, James H-Y. 1993. Iconicity: motivations in Chinese grammar. *Principles and Prediction: The Analysis of Natural Language*, ed. by Mushira Eid & Gregory K. Iverson, 153-174. Amsterdam & Philadelphia: John Benjamins.

Tai, James H-Y. 1994. Chinese classifier systems and human categorization. *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*, ed. by Matthew Y. Chen & Ovid J. L. Tzeng, 479-494. Taipei: Pyramid Press.

Tai, James H-Y. 1997. Category shifts and word-formation redundancy rules in Chinese. *Chinese Languages and Linguistics*, Vol. 3: *Morphology and Lexicon*, ed. by Feng-fu Tsao & H. Samuel Wang, 435-468. Taipei: Institute of History and Philology, Academia Sinica.

Tai, James H-Y. 2002a. Gainian jiegou yu feizizhuxing yufa: Hanyu yufa gainian xitong chutan [Conceptual structures and non-autonomous syntax: some conceptualization principles in Chinese grammar]. *Dangdai Yuyanxue* [*Contemporary Linguistics*] 2002.1:1-12.

Tai, James H-Y. 2002b. Temporal sequence in Chinese: a rejoinder. *Form and Function: Linguistic Studies in Honor of Shuanfan Huang*, ed. by Lily I-wen Su, Chinfa Lien & Kawai Chui, 331-351. Taipei: Crane.

Tai, James H-Y. 2003. Cognitive relativism: resultative construction in Chinese. *Language and Linguistics* 4.2:301-316.

Tai, James H-Y. 2005. Conceptual structure and conceptualizations in Chinese. *Language and Linguistics* 6.4:539-574.

Tai, James H-Y. 2006. On modality effects and relative uniformity of sign languages. Pre-Conference Proceedings of 14th Annual Conference of the International

Association of Chinese Linguistics (IACL-14) & 10th International Symposium on Chinese Languages and Linguistics (IsCLL-10), 222-240. Taipei: Academia Sinica.

Tai, James H-Y. 2007. Zhongwen gouci yu jufa de gainian jiegou [Conceptualization in Chinese grammar]. *Journal of Chinese Language Teaching* 4.1:1-30.

Tai, James H-Y. 2008a. The nature of Chinese grammar: perspectives from sign language. *Proceedings of the 20th North American Conference on Chinese Linguistics* (*NACCL-20*), Vol. 1, ed. by Marjorie K. M. Chan & Hana Kang, 21-40. Columbus: The Ohio State University.

Tai, James H-Y. 2008b. Entrenchment of pragmatics in syntax. Paper presented at the Second Conference on Language, Discourse, and Cognition, May 17-18, 2008. Taipei: National Taiwan University.

Tai, James H-Y. 2011. Zailun shijian shunxu yuanze [On principle of temporal consequence revisited]. *Renzhi Yufa yu Duiwai Hanyu Jiaoxue Lunji* [*Monograph on Cognitive Grammar and Teaching Chinese as a Second Language*], ed. by Xiliang Cui, 66-85. Beijing: Peking University Press.

Tai, James H-Y., and Jane Chou. 1975. On the equivalent of 'kill' in Mandarin Chinese. *Journal of the Chinese Language Teachers Association* 10.2:48-52.

Talmy, Leonard. 1983. How language structures space. *Spatial Orientation: Theory*, *Research*, *and Application*, ed. by Herbert L. Pick & Linda P. Acredolo, 225-282. New York: Plenum Press.

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. 2000. *Toward a Cognitive Semantics*, Vol. 1 & 2. Cambridge: MIT Press.

Tao, Hongyin. 2000. Cong "chi" kan dongci lunyuan jiegou de dongtai tezheng ['Eating' and emergent argument structure]. *Yuyan Yanjiu* [*Studies in Language and Linguistics*] 2000.3:21-38.

Tsai, Wei-tien Dylan. 2011. Cong yufa-renzhi jiemian tan Hanyu yufa jiaoxue: yi yiwen zhuangyu he fanshen zhuangyu wei li [On pedagogy of teaching Chinese as a second language from syntax-cognition interface: a case study of *wh*- and reflexive adverbials]. *Renzhi Yufa yu Duiwai Hanyu Jiaoxue Lunji* [*Monograph on Cognitive Grammar and Teaching Chinese as a Second Language*], ed. by Xiliang Cui, 1-17. Beijing: Peking University Press.

Tsao, Feng-fu. 1990. *Sentence and Clause Structure in Chinese: A Functional Perspective.* Taipei: Student Book.

漢語語法類型學特徵的反思

戴浩一

國立中正大學

在不同的類型學架構下，漢語被描述爲具有構詞獨立性與分析性，語法擬象性，同時是言談取向與語用取向，主題比主語更重要的語言。本文從語言相對論觀點，在功能主義與認知語法的架構下，重新檢驗漢語語法類型學的特徵。由於漢語與手語及克里歐語言具有類似的特徵，可藉以擴增語言類型學的分析角度。

關鍵詞：漢語，語言類型學，語言相對論，分析性，擬象性，手語，克里歐語

《語言資訊和語言類型》
第四屆國際漢學會議論文集
臺北：中央研究院，2013 年

古今漢語的句法類型演變：跨方言的庫藏類型學視角*

劉丹青
中國社會科學院

本文首先簡要介紹了庫藏類型學的基本觀念，然後以跨方言的庫藏類型學視角檢視古今漢語的若干重要的句法類型演變，提出七項主要的演變及其主要的方言差異。並將其中的六項概括為兩個大勢。前三項涉及話題化、語序類型和連動式的發展，總結為在基本語序等方面由句法主導演變為語用優先。後三項涉及名詞句法語義功能的消長，量詞的使用和方位後置詞的使用，概括為名詞句法功能的萎縮和受限，名動由均衡對立演變為名弱動強。中間一項是動補式的發展，本文認為這一變化並不像某些學者認為的那樣純粹體現了漢語由綜合性演變為分析性的過程。

關鍵詞：庫藏類型學，古今漢語，句法類型演變，方言差異

* 本文寫作獲中國社會科學院重點課題「語言庫藏類型學」資助，除於第四屆國際漢學會議上（臺北，2012.6）報告外，還在中國語言學會年會（昆明，2012.8）報告。在兩會上獲梅祖麟、黃正德、羅仁地、蔡維天、張洪明、張誼生等眾多與會先生的討論意見。另唐正大、陳玉潔、王芳諸博士和博士生高再蘭、曹瑞炯、白鴿、嚴豔群、盛益民等均對初稿提出過意見建議，匿名審稿人也提出了很有益的意見。在此向以上各方一併致謝。尚存問題均由作者負責。

1. 引言：語言庫藏類型學：從共時的「在庫」到歷時的「進庫」

本文以筆者所倡導的語言庫藏類型學（Linguistic Inventory Typology，見劉丹青 2011a）爲工作框架，參考既往的研究成果，包括筆者的個人研究（如劉丹青 2001a, 2003, 2011b），對漢語的重要句法類型特點進行古今和跨方言的比較，試圖揭示古今漢語在類型上影響深遠的若干演變以及這些演變在現代方言中的共性和差異。其中古代漢語以先秦典籍所反映的語言爲依據，暫不考慮甲骨文等更早的特殊文體語言。

語言庫藏類型學關注語種的形式庫藏和語義（含語用）範疇的雙向互動，尤其強調語法形式庫藏及其顯赫程度對語義範疇的影響。其主要視角是由低到高的兩大庫藏現象：

(1) 在庫 (available in the inventory) 及進庫 (entering the inventory)：研究哪些語義範疇有機會在一種語言中成爲語法庫藏，即由專用的語法形式手段表示。在庫是共時視角，觀察語義範疇在語法庫藏中的存在；進庫或稱庫藏化 (inventorization)，是歷時視角，觀察語義範疇形成語法庫藏要素的過程。在庫和進庫統稱入庫 (being in the inventory)。進庫的反面是離庫 (out of the inventory) 或稱去庫藏化 (deinventorization)，即從語法庫藏中消失。

(2) 顯赫度 (mightiness)：研究哪些入庫的手段造就庫藏中的顯赫範疇，即表達該範疇的手段常用、能產及可類推，富有範疇擴張力，可以表達由近及遠的其他語義範疇。

本文從庫藏類型的角度出發，認爲古今漢語的下列語法演變具有顯著的庫藏類型學意義，影響遍及漢語的主要方言，但演變的結果又因方言的不同而略有差異。

1) 話題結構由古漢語中的普通入庫範疇晉階爲現代漢語中的顯赫範疇。

2) 在整體語序類型上，漢語由古代的句法主導語言演變爲現代的語用優先語言。

3) 基本句法關係方面由古代的並列顯赫語言演變爲後代的連動顯赫語言。

4) 動結式、動趨式、能性動補式等動補結構發展成獨立於連動式的新的顯赫範疇。

5) 由名詞動詞均衡對立的類型經過功能上的名消動長形成名弱動強的動詞型語言。

6) 從非量詞型語言轉型爲量詞型語言，削弱了名詞表示個體的功能。

7) 隨著方位名詞不同程度的語法化和使用強制化，普通名詞失去了指處所的功能。

下文將逐一探討這些類型演變及其方言差異，最後對類型演變的總趨勢進行討論。多數現象是以往文獻討論過的（如太田辰夫 2003[1958]，王力 1989，魏培泉 2003），本文則從跨方言的庫藏類型學角度進行新的解讀。文獻有涉的演變細節不再一一詳列，請參看有關文獻。

2. 話題結構由普通入庫範疇晉階為顯赫範疇

馬建忠 (1983[1898]:390-392)《馬氏文通》已有專節討論古漢語中非施事性成分充當「起詞」的情況（卷之十，彖一，系七），其所引例句大都同時存在常規的主語，因此充當全句「起詞」的成分多具有話題性質，如下面諸例（長例或有刪節。帶下標同指標記的「之$_i$/其$_i$」是複指話題的代詞，若話題化只留下空位而沒有代詞複指，則以帶括弧的語跡標記 [t_i] 表示）：

(1) 夫顓頊者$_i$，昔者先王以 [t_i] 為東蒙主，且在邦域之中矣，(《論語・季氏》)

(2) 子路$_i$，人告之$_i$以有過則喜，(《孟子・公孫丑上》)

(3) 甯武子$_i$，邦有道則 [t_i] **知**，邦無道則 [t_i] 愚。(《論語・公冶長》)

(4) 之人$_i$也，物莫**之**$_i$傷，大浸稽天而 [t_i] 不溺，大旱金石流土山焦而 [t_i] 不熱。(《莊子・逍遙遊》)

(5) 鳥$_i$吾知**其**$_i$能飛，魚$_j$吾知**其**$_j$能遊，獸$_k$吾知**其**$_k$能走。(《史記・老莊申韓列傳》)

(6) 馬，蹄可以踐霜雪，毛可以禦風寒，齕草飲水，翹足而陸，此馬之真性也。(《莊子・馬蹄》)

馬建忠所舉的以上例子，已經代表了古代漢語話題結構的諸多主要類型。例 (1) 話題提取了「以 X 爲 Y」句式中的 X 做話題，空位所在小句的主語則爲「先王」。「以」可以視爲動詞或介詞，話題化造成的結果是兼語的空位或介詞的懸空。兩種空語類都爲現代漢語所不容。[1] 例 (2) 的話題提取了動詞「告」的賓語，賓語位置由「之」複指，「之」所在小句的主語則爲「人」。例 (3) 的話題隔了條件從句與兩個結果句的空語類主語同指，因而不再另有主語。其謂語分別是「知」和「愚」。例 (4) 的話題帶了助詞「也」做話題標記，與第一個分句的否定句前置代詞賓語「之」同指，所在小句的主語同時與後面兩個轉折複句中的空主語同指。例 (5) 的三個主句都以「吾」爲主語，被話題化的論元原位嵌入較深，是主句動詞「知」的賓語從句內的主語，作爲內嵌句主語按古漢語的規則由領格代詞「其」(＝NP＋之）充當。例 (6) 的話題與兩個述題小句的主語有語義上的領屬關係，與現代漢語的「這棵樹葉子大，枝條多」這一整體–部分類的話題結構一致，沒有複指成分，述題部分沒有明顯的句法空位。

以上例子顯示，在古代漢語中，可以話題化的成分和所構成的話題結構種類都相當多樣，而且這些結構也屬普通常見而非特異的形式。與非話題優先的 SVO 語言如英語相比，古漢語可以話題化的成分更多，話題化

[1] 這種懸空是古漢語固有的，不限於話題結構，因此不能據此認為古漢語話題化比現代更發達。

的樣式更多。(1)-(6) 這些話題結構大多無法直譯爲英語的話題結構，而只能譯爲普通的主謂句。即使勉強說成話題結構，也是屬於非常特異、難得一用的句式，不是常見自然的表達（參看 Steele 1978）。由此可見，話題結構在古代漢語語法庫藏中確實存在，是入庫的句法範疇。

但是，跟現代漢語的話題優先相比，古代漢語話題優先的程度有相當的距離。程麗霞 (2005) 對古今幾種主要的話題結構做了統計分析，在這方面提供了較好的證據。

程文將述題內有代詞複指受事話題的結構稱爲左偏置結構 TC_1，將述題內只有空位的稱爲 TC_2，後者是更加句法化的話題結構，因爲涉及賓語的位移，而前者不改變原有句法結構，只是加了一個句首話題，更接近語用操作。程文統計了從先秦的《論語》到 1990 年的現代小說《血玲瓏》共 8 種歷代文獻的話題結構，其 TC_1 和 TC_2 的出現次數之比如下：

《論語》	TC_1: 32 (91%)	～	TC_2: 3 (8.6%)
《世說新語》	TC_1: 9 (75%)	～	TC_2: 3 (25%)
《祖堂集》（節選）[2]	TC_1: 5 (35.7%)	～	TC_2: 9 (64%)
《朱子語類》（節選）	TC_1: 4 (30.8%)	～	TC_2: 9 (69.2%)
《老乞大》（節選）	TC_1: 7 (28%)	～	TC_2: 18 (72%)
《水滸傳》（會話部分）	TC_1: 16 (26.2%)	～	TC_2: 45 (73.8%)
《紅樓夢》（會話部分）	TC_1: 21 (10.8%)	～	TC_2: 174 (89.2%)
《血玲瓏》（會話部分）	TC_1: 0 (0%)	～	TC_2: 46 (100%)

以上數字顯示，語用屬性較強的 TC_1 由盛而衰，句法屬性較強的 TC_2 則從弱到強，後者逐步取代了前者。TC_2 是從 TC_1 發展而來的。此外，後代還逐步發展出了 TC_3（受事位於施事主語和動詞之間，如：我昨日冷酒吃多

[2] 三部著作的節選都是收入劉堅、蔣紹愚《近代漢語語法資料彙編》（唐五代卷 1990、宋代卷 1992、元代明代卷 1995）的部分，北京：商務印書館。

了）、TC_4（沒有施事主語，只有受事話題，如：那話怎敢說？）[3]、TC_5（整體–部分型話題結構，如：這三個火伴，兩個是買馬的，一個是牙子）。這些都是英語之類非話題優先語言更爲缺乏的話題結構。據此，通過與英語話題句演化的比較，程麗霞認爲，「漢語的話題結構在兩千多年時間裡已由原來的附接性成分演化爲具有獨立語法地位和豐富語義內涵的句法成分」。

現代漢語話題結構的重要性還不僅於此。除了以上這些話題結構外，現代漢語還出現了一些在世界語言中更加少見的話題結構（詳見劉丹青 2001b, 2009，Liu 2004），如：

(7) 我黃魚買了三條。
(8) 他主任倒是一個主任。
(9) 小張學習認真，工作積極，處理事務很有效率，關心同事仔細入微。

例 (7) 是分裂式話題結構，單一的受事論元「三條黃魚」被分裂成次話題「黃魚」和賓語「三條」兩個句法成分，分置於動詞兩端。(8) 是同一性話題結構，話題和述題中的一個成分一致，話題沒有爲動詞增加論元，是語義冗餘成分。(9) 是行爲表達式的話題化，由語義上表行爲的 VP 在主語後充當次話題，讓 VP 去焦點化，而由本應充當行爲動詞狀語性成分的詞語升格爲謂語並成爲自然焦點。這幾類的共性是，在保留 SVO 語序的前提下，借助於漢語句法庫藏中的話題位置，放置與話題屬性相容（但語篇話題性不一定高）的成分，該成分既不依賴代詞複指、也不留下空位，同時有別於主語，以此來表達一些在其他語言中不採用話題化的內容。對於句法庫藏中沒有專用話題位置的語言來說，甚至話題僅僅入庫但不夠顯赫的

[3] 引者按：TC_4，即施事和被動標記都不出現的受事主語句，是古代漢語就有的，不應列在此處，如《莊子・胠篋》:「昔者龍逢斬，比干剖」。

語種（如古漢語）來說，這些句法結構都是無法接受的。

因此，在現代漢語中，話題結構不僅存在於句法庫藏，而且非常顯赫，具有超強的擴展能力。不過，現代漢語不同方言在話題結構的顯赫程度上仍有一定差異。

初步觀察顯示，吳語的話題顯赫特徵最明顯，粵語較接近典型的 SVO 語言，話題的顯赫度偏低。以普通話爲代表的官話則介於兩者之間。下面只舉兩項表現來說。

表現之一，吳語（上海話爲例）強烈排斥有定成分做賓語，有定受事成分優先充當話題（主語前的主話題或主語後的次話題）；粵語（廣州話爲例，參考了單韻鳴博士的個人通信意見，謹致謝意）對有定賓語較寬容，受事前置及處置句都很受制約；普通話對有定賓語的容忍度介於兩者之間，有定受事如前置，優先採用帶介詞「把」的處置句，其次才是話題句（主要限於主話題）。如（參看劉丹青 2000, 2001a）：

(10) <普> 你快把那把鑰匙還給我＞那把鑰匙你快還給我＞?你快還我那把鑰匙

(11) <滬> 儂搿把鑰匙快點還撥我/搿把鑰匙儂快點還撥我＞儂快點拿搿把鑰匙還撥我＞??儂快點還撥我搿把鑰匙

(12) <穗> 你快啲畀返條鎖匙我＞?你將條鎖匙快啲畀返我/?條鎖匙你快啲畀返我

假如受事成分是受到全量算子約束的有定成分，則吳語和普通話更強烈排斥受事後置做賓語，分別傾向於使用受事話題句和處置句，粵語則仍然自由地使用 VO 句。如：

(13) <普> 他把那些水果全吃完了＞那些水果他全吃完了＞*他全吃完了那些水果

(14) <滬> 伊搿眼水果儕吃脫了/搿眼水果伊儕吃脫了＞伊拿搿眼水果儕吃脫了＞*伊儕吃脫了搿眼水果

(15) <穗> 佢食晒啲生果喇＞?佢將啲生果食晒喇/?啲生果佢食晒喇

根據李英哲 (2001:168)，臺灣閩語使用主題–評論（即話題–述題）的傾向比普通話更強，尤其排斥某些動結式後面的有定賓語，如：

(16) a. 彼領衫穿無破。

～ b. *穿無破彼領衫。

（比較：那件衣服沒穿破。～沒穿破那件衣服。）

表現之二，分裂式話題結構是強烈體現話題優先特徵的結構，很難存在於非話題優先語言。[4] 在吳語和閩語中（參看劉丹青 2001b），(7) 那樣的分裂式話題結構是很常用的結構，在溫州等南部吳語中甚至成爲基本結構，「渠飯吃一碗爻了」是表達該意義的最常用的敘述句式。在普通話中，分裂式話題結構屬於合語法而並不常用的句式（入庫而不顯赫），只見於對比性語境，如「米飯你吃了幾碗？」（需要米飯與其他食物對比的語境），作爲問話遠不如「你吃了幾碗米飯？」自然常用。而在粵語中，「佢恤衫買咗三件」比普通話相應句子還不自然，除非在一些很特殊的語境下才會說。

[4] 在通常被認為是主語話題並重語言的韓國語中，據曹瑞炯（個人交流）介紹，存在 SOV 型的分裂式話題結構，如：

(i) kay-nun nay-ka sey-mali-lul khiwu-n-ta.／개는 내가 세 마리를 키운다.
狗-話題 我-主格 三-隻-賓格 養-現在時-直陳（狗我養著三隻）

(ii) nay-ka syechu-nun phalan-sayk-ul cohaha-n-ta.／내가 셔츠는 파란 색을 좋아한다.
我-主格 襯衫-話題 藍-色-賓格 喜歡-現在時-直陳（我襯衫喜歡藍的）

韓語有格標記，這些例句清楚地體現出被一分為二的論元前部帶話題標記，後部帶賓格標記，並且這兩種標記不能互換。

粵語的話題結構不如其他方言顯赫，不等於說粵語的話題結構沒有入庫。與非話題優先的 SVO 語言相比，粵語的話題結構仍是一個顯著的存在（參看馬詩帆、楊月英 2003）。

3. 整體語序類型由古代的句法主導走向現代的語用優先

在語序類型學最重要的兩個參項——及物小句基本語序和介詞類型方面，古代漢語相對接近典型的 SVO-前置詞語言。雖然它也兼有 SOV 部分特點，並且偶有後置詞現象，但這些複雜現象大多可由句法規則控制，也有部分與語用因素有關。比較起來，古代漢語至少在這些語序方面更接近句法主導的語言；而現代漢語則離 SVO 語言的常規特徵較遠，比古漢語更多地體現劉丹青 (1995) 所說的漢語作爲語用優先語言的特點。下面具體分析。

在甲骨文語言中，(S)VO 具有絕對主導地位。根據 Djamouri (2001) 統計，甲骨文中 VO 和 OV 之比爲 93.8%：6.2%。先秦文獻語言也以 (S)VO 爲基本語序；此外，存在幾種大家熟知的 (S)OV 結構，其出現條件分析如下。

1) 疑問代詞賓語前置。如《論語・子罕》「吾誰欺，欺天乎」。疑問代詞是個高度封閉的小詞類，它們的統一前置，屬於明確的句法規則。而且，類型學上，代詞賓語比名詞賓語更容易前置（參看 Greenberg 1966 共性 15 條），在 SVO 語言中代詞賓語前置並不罕見，如法語作爲 SVO 語言代詞賓語就是前置的。

2) 否定句中的代詞賓語前置。如《左傳・襄公九年》「今楚師至，晉不我救，則楚強矣。」周法高 (1959:39) 認爲雖然有一些否定句代詞賓語後置，但這些例外「在漢以前文獻中，僅佔極少數比例，並不能影響此規則的正確性。」並且他認爲這些例外是有一些原因的，如肯定否定對舉、傳抄失誤、加「矣」強調等。這條規則涉及的否定詞和代詞都是高度封閉的

語法類，因此其賓語前置規則有一定的句法強制性。而且，謙稱尊稱「臣、僕」、「子、君」等在真值語義上分別等於第一、二人稱代詞，但它們在語法上不屬於代詞，就不受這條前置規則的制約，可見，這一代詞賓語前置規則受制於語法屬性而不是語義語用屬性。不過，否定句代詞賓語前置的強制性不像疑問代詞賓語前置那麼強，存在不少例外，在實際語境中，賓語的前後取捨與語用也有一定關係。

3)「唯 O 是 V」型焦點句，如《左傳・宣公十二年》:「率師以來，唯敵是求」。這種賓語前置的動因是焦點化，因而是一種語用操作。由於這種焦點句已經高度構式化，因而成爲先秦句法庫藏中的一種構式。在典型的 SVO 語言中，用句法手段將賓語焦點化並導致前置，是常見的現象，如英語的分裂句：It is the enemy that I have been seeking。

古代漢語的介詞（前置詞）都由動詞虛化而來，因此也保留了一些 OV 語序帶來的後置詞用法，主要見於代詞尤其是疑問代詞賓語，與上述現象 1) 和 2) 密切相關，如「是以、何以、何爲」等。而先秦時的方位詞還有很強的名詞性，尚構不成詞類庫藏中的後置詞。

綜上所述，古代漢語中的 OV 現象，基本上都是 SVO 語言內部受句法規則控制的一些變體，類似現象也常見於其他 SVO 語言，不影響古代漢語 SVO 語序的句法主導地位。

現代漢語也以 SVO 爲基本語序，而且在中古以後就不再有以上三項賓語前置規則，但在許多方面卻離 SVO 語言的類型特點更遠。突出表現在：

1) 中古以後興起並盛行至今的「把/將」處置式，成爲近現代漢語及物結構的顯赫句型，這種受事前置的句式顯著減少了受事成分在動詞後的比例，而處置式與 VO 句的選用，只有一部分是句法制約，如現代漢語表現出來的動後限制（張敏 2010），導致賓語與多種其他成分（包括介詞短語）在動詞後的位置相互排斥，這時使用受事前置的句式，確屬句法制約；而其他很多情況下，無法用句法規則說清，而受制於指稱、信息結構、語篇等複雜的語義語用因素。

2) 受事前置的話題結構通常無須再在賓語位置加複指成分，構成上引程麗霞文所說的 TC_2，取代了古代漢語中佔優勢的代詞複指型的 TC_1。TC_1 不改變 SVO 語言的基本結構，只是在句首加了一個左出位 (left-dislocated) 的成分，該成分語用屬性強，沒有進入小句的核心句法成分。而 TC_2 改變了句子的 SVO 性質，其受事話題也作爲受事的唯一表徵而成爲小句更內在的句法論元成分。

現代漢語的以上兩點表現，罕見於其他 SVO 語言，也不見於古漢語，從而使現代漢語離 SVO 的句法常態很遠。重要的是，以上賓語前置現象跟古代漢語的賓語前置規則，語序表現不一，有不同的性質。古代漢語的賓語前置，主要是強制性句法規則，是由論元結構、語類、句法結構等因素造成。而現代漢語偏離 SVO 語序的結構，都不是強制性規則。從語法庫藏的角度講，現代漢語的處置式是一種進入構式庫藏的句法結構，現代漢語的話題結構則擁有比古代漢語的話題更顯赫的句法位置，可以用來放置比古代漢語更加多樣的成分。但是，用 VO 句，還是處置句，還是話題結構？是讓賓語後置還是讓介詞短語等其他成分後置？其間的選擇在現代漢語中主要不是由語類、論元結構這些句法條件決定的，而是取決於語境中的有定無定、信息新舊、話題還是焦點等語用因素。當然，古漢語的話題化也受信息結構的制約，但古代的話題化對基本語序影響不大，被話題化的受事類成分仍然由複指代詞佔據著賓語的位置。正是在這一意義上，我們認爲古代漢語小句內的基本語序主要是句法導向的，而現代漢語的基本語序更多體現了語用對現代漢語句法的強烈制約。劉丹青 (1995) 曾從包括語序在內的很多方面，論證漢語是一種句法限制較小的語用優先語言。LaPolla (1995) 認爲現代漢語的主語賓語等成分主要是話題焦點一類語用成分而不是真正的句法成分。張伯江 (2011) 認爲「漢語的句法結構很大程度上反映的是語用結構，而不是像英語那樣主要地反映語義結構。」這些論述對話題的句法化程度的認知不盡相同，但程度不同地強調了在現代漢語語法中語用的作用較大，而這並不完全反映古代漢語的情況。總體

上，從古到今，至少在基本語序方面，漢語語法庫藏中強制性的句法規則在弱化，而由語用決定的規則或傾向在增加。

另一方面，在總體上語用優先的前提下，漢語不同方言在句法和語用的相對份量上仍有一定差異，語用優先的程度一般也與話題優先的程度相稱。

吳語受指稱、量化和信息結構等語用因素的影響最大。不但有定、類指的受事成分強烈傾向前置（分裂式話題結構也是將受事論元分爲光杆名詞和限定成分兩部分，將表類指的光杆名詞前置，將含限定成分的焦點信息後置），而且全量受事成分，不管用什麼限定詞，都幾乎強制性地前置，比較以下普通話的常見說法和相應的蘇州吳語的說法：

(17) <普>

a. 見每一個客人

b. 不理睬所有的朋友

c. 沒看一本書

d. 不帶走一絲雲彩

<蘇>

a. 每個客人儕見

b. 所有朋友儕弗見

c. 一本書也勿看

d. 一點雲也弗帶走

這些蘇州話的句子很難將全量的受事移到動詞之後，如：*見每個客人｜*勿看一本書。可見吳語的語序比普通話對語用因素更敏感。

而粵語不但有定成分後置比普通話和吳語自由（見 (10)-(12)），並可以照搬 (17a-d) 這些普通話全量成分後置的語序，而且還存在普通話也難以接受的有悖於現代漢語信息結構大勢的語序現象。試舉兩例。

其一，普通話（及北部吳語）對動詞的狀語性成分有個基本的語序分工，屬於行爲事件的附加性信息的在動詞前（稱爲狀語，如「快走」，重點是「走」,「快」是附帶的方式信息），屬於焦點信息的在動詞後（稱爲補語。如「走得快」,「走」是已知信息，重點在走的方式「快」)。這也是現代漢語信息決定語序的表現之一。但是，粵語卻有一批屬於附帶信息的副詞狀語用在動詞之後，與作爲焦點的補語有相同的語序和很不同的信息地位。如「走**先**（先走）、食碗**添**（再吃一碗）、食**多**啲嘢（多吃些東西）」等。這些後置狀語在其他 SVO 語言中也常做動詞後狀語，顯示粵語更多採納 SVO 語言的常規句法規則而更少受語用因素的影響。

其二，粵語更多偏離句末焦點的規則。漢語有很強的自然焦點居末的傾向。漢語特色話題句類型之一是所謂主謂謂語句，如「象鼻子長」。這一特色句型將自然焦點（上句中是「長」）置於句末，高度切合漢語信息結構的語序規則。同類的句子，在非話題優先語言中的表現形式之一是用「有」字領有句將焦點置於領有賓語的定語位置，並不一定位於句末，而位於句末的可以是沒有信息量的名詞，如：An elephant has a long nose. He has ugly teeth。「象有鼻子」,「他有牙齒」，都缺乏信息量，只有「長」、「醜陋」才是焦點所在。它們的普通話直譯句「象有一根長鼻子」、「他有一口醜陋的牙齒」等，遠不如焦點在末的「象鼻子長」和「他牙齒醜陋」自然（參看劉丹青 2011c）。粵語同樣使用焦點居末的話題句，但是它有一種「形＋名/量」結構可以用作謂語和定語，在做謂語時可以表達普通話上述話題結構的語義。麥耘 (1995) 列舉了大量這類組合及其用例，如：

(18) a. 呢個女人好**利口**。（這個女人嘴巴很厲害。）
　　b. 佢做嘢好**快手**。（他做事情動作很快。）

(19) a. 我嘅生意好**細單**嘛。（我的這宗生意很小的。或：我的每宗生意都很小的。）
　　b. 呢啲藥丸好**細粒**。（這些藥丸很小。）

這類組合的最大語用特點是，自然焦點不在句末，而在句末成分的定語上，與上述英語例句有共同之處，而句末的名詞或量詞則沒有信息量。以 (18) 爲例。每個人都有「口」（借指說話能力），句末的「口」信息量很低，焦點在「利」（厲害）上。餘例同理。信息結構以如此語序呈現，很不合普通話、吳語的語感，而在粵語中大量存在。這說明粵語對句末焦點的敏感性不如普通話。這反映了語用優先在粵語中不如在官話和吳語中那麼強烈。

4. 基本句法關係由並列顯赫型演變為連動顯赫型

梅廣 (2003) 基於古代漢語 VV 並列式的常用和並列連詞「而」的常用及多樣化，認爲「上古漢語是一種以並列爲結構主體的語言；中古以降，漢語變成一種以主從爲結構主體的語言」。我們贊同梅先生關於上古漢語的判斷，但是不盡同意關於中古以降的看法。中古以降漢語很難說變成了以主從爲結構主體的語言，而更像是變成了一個以連動結構爲顯赫範疇的語言。我們同意某些連動式進一步變成了有主從結構性質的動補式（參看第四節），但是連動式家族在兩千年來一直是漢語中有活力的結構，並非只是朝向主從結構演化的中間站。我們同意張敏、李予湘 (2009) 的看法：上古漢語連動式非常不發達，中古以後連動式逐漸發達，現代漢語是一種連動發達的語言。因此，漢語從古到今經歷了由並列顯赫到連動顯赫的類型演變（有關連動句的描寫可參看呂冀平 1985[1958]，高增霞 2006）。

主從（即偏正，包括動補）是人類語言的普遍句法範疇，是語皆有。並列結構的有無也不是類型參項，關鍵要看它是否在一種語言中表現得比在其他語言更強勢、更具有擴展能力——可以用來表達其他語言中屬於其他結構關係的語義。誠如梅廣先生所見，中古以後主從式比上古漢語更加強勢和活躍，但是，放眼其他語言，中古至現代漢語用主從結構表達的內容，在其他語言中用並列類結構來表示的情況並不多見。反之，確有一些

在其他語言中多用主從結構表達的內容，在現代漢語中傾向於用主從以外的結構來表示。例如，我們用「象（，）鼻子長」這類話題結構，代替了其他語言更傾向用的「象的鼻子長」、「象有很長的鼻子」之類含主從結構的句式。下面將要談的連動，更是表達其他語言主從結構內容的常見手段。

連動是僅見於一小部分語言的句法結構。[5] 世界上很多語言的語法庫藏中不存在連動式，它們以主從或並列等結構表達連動類的語義關係。因此，任何語言只要存在連動式，大致就劃歸連動型語言。假如連動式的功能發達，則可以進一步確定爲連動顯赫型語言。上古漢語連動式處於不發達狀態，連用的動詞多以並列連詞「而」來連結，連動式處於句法庫藏中的邊緣位置。中古以降，連動式愈益發達，可以表達的語義關係越來越多。近現代漢語連動式的極其豐富多樣，已有諸多文獻論及。這裡需要指出兩點：

1) 古代漢語很多用並列連詞「而」連接的多動詞結構，「而」在其中不能省略，它們在英語之類非連動語言中也可以用並列連詞 and 之類連接，但在現代漢語中最自然的表達法是連動式，而且在很多情況下不容加進合適的並列連詞。[6] 比較：

(20) a. 竭澤**而**漁
~ b. to drain the pool ***and*** catch fish
~ c. 抽乾池塘的水[ ]捉魚

(21) a. 棄甲曳兵**而**走
~ b. to put down the weapons ***and*** flee
~ c. 丟了鎧甲拖著武器[ ]逃跑了

---

[5] Aikhenvald (2006:1)：「連動式廣泛存在於克里奧爾語，西非、大亞洲、東南亞、亞馬遜和新幾內亞的語言。這也就意味著這些有限區域以外的很多語言是不存在連動式的」。

[6] 現代漢語沒有中性的連接謂語動詞的並列連詞，「和」主要連接名詞及論元位置的謂詞，「並」帶有輕度遞進，「而」帶有輕度轉折，謂語的直接並列和帶連詞並列都受到很大限制。參看劉丹青 (2008)。

(20) 例中，a 句按先秦的規則不能省略「而」，b 句也不能省略 and，而 c 句不但不出現並列連詞，事實上也不能自然地加進連詞。這就意味著，現代漢語連動句可以表達在其他時代或語言中帶連詞的並列結構所表達的內容，是連動範疇向並列範疇的擴展。

2) 現代漢語的連動結構可以表達其他語言常用主從結構表達的語義關係。例如：

(22) 笑著說｜站著說｜炒著吃｜開著窗戶睡覺｜唱著歌走路｜睜著眼睛說瞎話

例 (22) 中前面的「V 著 (O)」，語義上是後面動詞的方式修飾成分，在英語中要用副詞狀語、介詞短語或分詞短語來表示，都屬於從屬成分。漢語中 (22) 這類結構有時也被分析爲狀中結構。但是，從句法形式上看，「V_1 著 (O) V_2」只能是連動結構，這個「著」可以加到顯然並非從屬成分的謂語上，如「他們喝著茶談著生意」。「著」是體助詞而不是狀語標記（狀語標記是「地」de 或介詞，如「笑嘻嘻地說」、「他以微笑迎接客人」）。這裡，連動式擴展到了主從範疇，表達在其他語言中包裝爲主從結構的內容。[7]

[7] 有審稿先生認為，(22) 中的「著」不可省略，可以看做從屬語標記，正如英語 -ing 既是體標記，又是從屬語標記。我們覺得這一看法值得重視，但還難以完全成立。「V 著 (O)」在語義上可以屬於表方式的背景信息，但是「著」本身仍是進行持續體標記，可以用在主句謂語上，如上舉的「談著生意」，甚至還有方式背景信息不帶「著」，主要謂語帶「著」的情況，如「他舉起雙手托著箱子」。而體標記用在兩個謂詞之間不能省略也不限於「著」，「了」也有類似的不能省略的用法，如「豬就吃了睡，睡了吃」，不能據此就說「了」是從屬語標記。至於「著」和英語 -ing，兩者有重要區別。V-ing 本身不能做謂語，它表示進行體時，V 前必須有係動詞 to be，它才是承擔時態、一致關係的謂語句法核心，-ing 確實有從屬語標記的性質。而帶「著」的動詞包括動態動詞有做謂語的功能，如：人們唱著、跳著｜雪正下著呢｜他們正開著會呢（引自《現代漢語八百詞》「著」條）。即使句中出現副詞「正」和語氣詞「呢」，它們也都不是謂語性的。上舉「談著生意」的「談」也是有動態性的行為動詞。因此，只能說「著」有幫助表達語義上的輔助性或背景性信息的作用，但是句法上它不是從屬語標記，不改變 (22) 的連動句性質。

根據這些情況，我們認定中古以後漢語並列結構有所衰落，主從結構佔了古漢語並列結構的一些地盤，但是並沒有得到特別顯著的擴展。反而是連動結構，可以左右開弓：既擴展到並列的語義域，又擴展到主從的語義域，從而真正發展爲顯赫範疇。

另一方面，漢語連動式的顯赫度，存在跨方言的差異，下面舉例說明。

北部吳語普遍存在一個可以用於名詞、動詞、形容詞和各種句法成分的泛用並列後置連詞，即放在並列結構前肢之後的虛詞，若要插入停頓，要在該詞之後，如上海話的「咾」、蘇州話和無錫話的「勒」等。例如（括弧中允許的停頓顯示這個連詞是後置詞性的）：

(23) a. 張三咾（，）李四
張三和李四
b. 我辣屋裡吃茶咾（，）看報。
我在家喝茶、看報

在 20 世紀早期留下來的上海方言作品中，「咾」的使用十分廣泛，其中有不少用例自然對應普通話和其他方言的連動式。如：

(24) 伊是靠鬼王咾趕脫鬼個。
他是靠鬼王[ ]趕走鬼的
（《方言聖人行實・摘錄聖女亞加大・童貞致命》，上海：土山灣印書館，1913 年）

(25) 今朝既然巧極大家儕到拉此地，格末多坐一歇歇咾談談。
今天既然巧極了大家都到了這裡，那就多坐一會兒[ ]聊聊
（Albert Bourgeois 編 LeÇons sur le Dialect de Changhai（上海話課本）301 頁，上海：Cours Moyen，1939 年）

注釋句中的空括弧顯示加連詞「咾」的地方在普通話中不能加連詞、只表現爲連動式。

此外，也有些「咾」用在語義上的主從關係雙方的中間，猶如古漢語的「而」用在「十一而稅」這類主從關係中的例子：

(26) 告伊拉趕緊咾頂真調查。
告訴他們趕緊地認真調查
（同 (25) 例書 137 頁）

(27) 准許伊照聖母個命咾做。
准許他按聖母的命令做
《方言聖人行實・聖達尼老各斯加・耶穌會》

由此可見，至少在老上海話中，並列結構仍是顯赫範疇，並列連詞可以從並列擴展到其他方言中用連動和主從來表達的關係，與古漢語的並列強勢類型有共同之處。

5. 來自連動結構的多種動補式成為新的顯赫範疇

現代漢語包括動結式、動趨式和能性動補式在內的動補結構非常常用，是構成漢語語法類型特點的重要方面。雖然對這些結構的具體出現時間和形成過程還有一些不同意見，但基本共識是它們並非自古就有，而是在上古晚期到中古漢語中形成並逐步強勢（參看太田辰夫 2003[1958]:192-207, 213-219，王力 1989: Ch.17, 19）。動補結構的形成與連動式有較密切關係，某種程度上是連動式顯赫的衍生物，由連動式凝固而來，但是它們今天已經與連動式割斷了臍帶，由動詞和補語謂詞按照動補規則大量形成，不再經過連動的中介階段。

動補式脫離連動自身成爲顯赫範疇，有以下幾點表現。

1) 上古漢語由單個動詞表示的行爲，在中古以後要用動結式來表示。這是很多漢語史文獻都提到的。[8] 例如：破之＞打破/敲破/摔破它；傷其臂＞打傷/擊傷/壓傷了他的胳膊，活之＞救活他。再如上古漢語和其他語言用情態助動詞表示的可能情態，在近代漢語以來常常用能性動補式表示，尤其是在否定時，如：吃得/不進、舉得/不高，想得/不到。

2) 動補式的詞彙搭配面大大擴大。能構成動結式的組合數量大大超過一些存在動結式卻很有限的語言。英語有 paint the wall white, shot him dead 等動結式，但是能用的動詞和補語都非常少。現代漢語中能用於此式的動詞和補語都非常多，甚至還有一些動–補之間、動–賓之間都很難組合，但借助補語和賓語的聯繫將三者組合起來的動結式，如「笑痛了肚子、看瞎了眼，跑丟了錢包，走濕了鞋」等。而能用於動結式的組合，大都能插入「得/不」，變成能性動補式。動趨式的組合也很自由，所有單音和複合趨向動詞能都做補語。

3) 動結式形成之初，補語主要用來表示動作的預期結果或自然導向的結果，如「射死」（射的目的就是讓目標死），「吹開」（開是吹的自然結果）。隨著動結式的強勢擴展，其搭配面越來越廣，表達的語義關係也越來越多樣，並造成動、補兩個謂詞的論元結構整合過程的複雜性。譬如，補語可以不是行爲的預期結果或自然結果，而是行爲的某種意外或附帶的後果，涉及的賓語也可以不是行爲的受事或施事（這也是動結式論元整合規則複雜化的原因之一），而是後果牽連的對象。如「（洗衣服）洗濕了鞋」，洗的受事是衣服，洗的目的是讓衣服乾淨，而這個賓語「鞋」不是洗的受事，補語「濕」不是洗的預期結果或自然結果，而是洗衣時意外發生在旁涉對象上的後果。類似的或稍有不同的還有「（蓋房子）蓋窮了一家人」、「（逛商場）逛丟了孩子」、「這段山路跑累了他」，「吃飽了肚子」、

[8] 如太田辰夫 (2003[1958]:193-194)：「在古代漢語中自動、他動兩用動詞很多。但是時代往後，產生了兩用動詞固定為自動用法的傾向。因此，作為古代他動功能的繼承，使成複合動詞就成為必要的了」。這裡說的「使成複合動詞」就是動結式。

「跑崴了腳」、「笑痛了肚子」等等，這些結構對原型動結關係的偏離，反映動結式已經發展出強大的語義擴展力。

4) 動結式和動趨式的補語，以唯補詞（劉丹青 1994）中介，成爲體標記和動相 (phase) 標記等虛詞的主要來源。「了、著、過、起來、下去、上、下、來、去、完」等普通話或方言的體標記都來自補語。只有在語言中常用和重要的要素才能成爲語法化的來源。

漢語方言全都具備以上三種動補式，但是發展的程度和表現形式仍有差異。

1) 動結式的鬆緊度不同。眾多重要文獻將現代漢語動結式視爲複合詞（周遲明 1957，太田辰夫 2003[1958]:193，呂叔湘 1980:17, 34，梅廣 2003，董秀芳 2007），因其結構緊密，不帶連接成分。在動結式發展過程中，曾有過可合可分的階段，動結之間插入賓語的例子如「風吹羅裳開」、「當打汝頭破」、「斫右手斷」等（轉引自蔣紹愚、曹廣順 2005: Ch.10）。這種隔開型動結式，顯然不是複合詞，而是句法組合。現代漢語方言中，仍然存在這種鬆散的動結式。泉州閩語在以 VCO 動結式爲主的格局下，也殘存了少量 VOC 結構，如「入三矸滇」（字面：裝三瓶滿）、「食日晝飽」（字面：吃午飯飽）（例見李如龍 1996）。而在桂南平話賓陽方言中（覃東生手稿，2010，基於三材村、委栗村本地話），只有 VOR 才是能產而自然的語序：

(28) a. 張三昨日吃酒醉啦。～*張三昨日吃醉酒啦。
b. 我吃飯飽啦。～*我吃飽飯啦。

(29) a. 阿三敲水缸壞啊。＞??阿三敲壞水缸啊。（很不自然）
b. 你再偷東西，我就拗你手斷去。～*你再偷東西，我就拗斷你手去。

另有些方言的動結式中間要加連結性助詞，狀似普通話帶「得」的組

合式補語。這樣的動結式也不是複合詞，甚至不是黏合性補語。如泰如片江淮官話普遍有一個動詞後的虛詞「啊」（[a]，在入聲韻和陽聲韻字後分別讀 [ka] 和 [ŋa]），它是個功能極多的虛詞（參看李人鑒 1957，汪如東 2006, 2008），其中包括用在動結式及動趨式中間，如海安話：

(30) 吃啊傷啊，跌啊青啊，學啊會啊，煮啊爛啊，走啊慢啊（汪如東 2006:385）

這些例子表示「吃傷了、跌青了……走慢了」之意。其中後一個「啊」是體標記，相當於「了」，前一個「啊」則是動詞和結果補語之間的聯繫項。泰如方言沒有不帶「啊」的「吃傷、跌青」等說法。同時，這個「啊」也可以用在普通話狀態補語前「得」的位置，如如皋話「好吃啊沒得命」、「重啊連我總拿不動」（鮑明煒、王均 2002:476-477）。

2) 動趨式中間有聯繫項。普通話本身存在「舉起一塊牌子來」、「扔了一支煙過去」等部分或全部趨向補語與動詞隔開的情況，這是比動結式更鬆散的表徵。此外，在不少方言中，動趨式中間可以或要求用連接性助詞。蘇州方言動詞帶單音節趨向補語時中間要用「得」連接，如「走得來個（走來的）、明朝帶得去（明天帶去）」。山西的很多晉語和中原官話都要像某些近代漢語語料那樣在動詞和趨向補語間用「將」（或念「張」）或功能相同的「得」等（喬全生 2000: Ch.10），如中陽「荷將來啦」、平遙「送將去 [tiʌʔ] 啦」、大同「送張來/去」，臨汾「荷得來/去」等。喬全生認爲「將」、「得」等是連接動詞和趨向補語的結構助詞。

3) 能性動補式存在相反方向的差異。很多晉方言的肯定式可能補語不加「得」，而與前面動詞形成黏合式動補式，同時在補語後加「了」（音 [lɔ] 或 [liɔ]）（否定可能式則從普通話到方言普遍不帶「得」，否定詞直接用在核心動詞後）。如：平遙「搬轉了」（能搬開）、「窗子糊住了糊不住？」（糊得住糊不住？）（侯精一 1999:408）。這些能性動補式顯然比普通話的

能性動補式更加緊密。反過來，也有些方言的可能補語比普通話鬆散，可以後置於賓語（一般限於簡短的代詞賓語），如上海話「舉伊勿起」、「推儂勿動」等。

各種隔開型補語的存在，對後文總結漢語類型演變大勢有重要意義。

6. 動詞和名詞的詞類地位：從名動兩立到名弱動強

劉丹青 (2010) 曾從多個角度論證漢語是一種動詞型語言，區別於英語的名詞型語言。其中突出表現在兩點：

1) 從獨立成句能力、句子內部構造到詞義引申和兒童語言習得，漢語動詞都比英語動詞強勢，英語名詞及名詞化成分比漢語名詞性成分強勢。

2) 動詞可以自由地充當名詞的典型句法成分，即論元（主賓語等），而名詞不能自由地充當動詞的典型成分，即謂語核心。

但是，回顧歷史，現代漢語這樣的動強名弱格局，不是自古而然的。雖然沒有理由說上古漢語像英語那樣是名詞型語言，但是上古漢語中名詞的句法地位確實比現代漢語名詞的句法地位重要得多。突出表現在以下方面。

1) 名詞做謂語。上古漢語中的典型判斷句是不用係動詞的，名詞在語氣詞「也」幫助下獨立做判斷謂語。如：「石碏，純臣也」(《左傳・隱公四年》)，「南冥者，天池也（《莊子・逍遙遊》)。自從「是」由指示代詞演變爲係詞進入庫藏後，名詞謂語判斷句就逐漸淡出庫藏。現代漢語普通話中名詞做謂語的功能已局限於個別小類的名詞（某地人、日期等）。如果排除短語、只考慮單詞，名詞做謂語更受限制。漢語整體上從非係詞語言轉化爲係詞語言。這一轉化顯著壓縮了名詞性成分在語言中的活動空間，強化了動詞和名詞的不對等：動詞可以自由地佔據本屬名詞的論元位置，名詞則不能自由佔據本屬動詞的謂語位置。

2) 名詞做狀語。上古漢語名詞可以比較自由地充當狀語，例如（轉引

自楊伯峻、何樂士 2001:58-62）比況式的狀語：「豕人立而啼（《左傳・莊公八年》）」、「庶民子來」（《詩經・大雅・靈台》、「嫂蛇行匍伏」（《戰國策・秦策一》；工具狀語：「晉楚不務德而兵爭」（《左傳・宣公十一年》；空間狀語：「若野賜之……」（《左傳・昭公元年》、「猶欲之楚而北求之也」（《荀子・樂論》）。後代名詞的這種功能基本消失，只有小部分以詞彙化的化石形式遺留下來，如「蜂擁、蠶食、魚貫」等。方位詞做狀語現代限於對舉句式，如「上有老下有小」。名詞的狀語功能總體上離庫了。

3) 名詞動詞互轉機制的變化。漢語動詞自古至今基本上都能自由地用於論元位置，雖然在類型學上主賓語絕不是動詞的基本功能。上古漢語名詞可以比較自由地用作動詞，即除了充當判斷句謂語外，還可以用作普通的動詞（包括帶賓語），非常像英語中的 denominal verbs (Clark & Clark 1979)。幾乎每一本古漢語語法書都會設立專節討論名詞做（判斷謂語以外的）謂語的用法。周法高 (1961:65-81) 有「名謂式」一節，討論了普通名謂式、名詞的使謂式和意謂式三類做謂語的情況。楊伯峻、何樂士 (2001: 177-120) 專門討論了名詞（包括代名詞）用作動詞的情況。僅兩書涉及的名詞就非常廣泛，從專有名詞、具體器物名詞到身分職銜名詞、抽象名詞和時空方位名詞等等，很難用語義標準來排除不能做謂語的小類。兩書所涉的大部分名詞（及本文在括弧中添加的現代對應詞）在現代漢語中沒有謂語用法，如：君（君主）、臣（臣子、臣民）、父（父親）、子（兒子）、聲、味、物（東西）、轅（車轅、轅木）、路、肘、手、口（嘴）、牙、吏（官吏）、天、家、風、門、角、屍（屍體）、耳（耳朵）、目（眼睛）、溝、公、墓、墳、井、王、天、虎、堂、蟲、疆（邊疆）、兵、夜、性、身、蠶、孺子、封豕（大豬）、茅、縣、吳王、國、夷狄、夫人（以上選自周著）、雨、器、棺、妻、女、夏商、心、卿、爾、汝、吾（以上補自楊何著）。以上例子中，少數有變調構詞模式（可能來自上古韻尾交替），如「王」做動詞念去聲，屬於構詞形態，大部分讀音不變，是名詞本身用作謂語。

4) 重要的名詞性量化詞語被動詞短語取代，新生的量化成分多由動詞性成分詞彙化而成。全量否定詞「莫」被「沒有人、沒有什麼」等短語取代，存在量化詞「或」被「有人」等短語取代。後起的「有的」、「所有」、「一切」、「舉（國）」、「滿（地）」、「遍（地）」等量化詞也由動詞語義引申組合而來。這是重要詞語上的動進名退的局面。

由於以上演變，動詞強勢依舊，而名詞功能萎縮。在某種意義上，強勢就是具有接近其他詞類的功能，萎縮就是失去類似其他詞類的功能。而從庫藏類型學的角度看，只有句法功能強大、在語言心理中容易激活的詞類才有擴展到其他詞類的能力，猶如隱喻借喻也總是以顯著的詞語或物象來代表較不顯著的詞語或物象。詞類的功能多樣性本質上還是該詞類作爲顯赫範疇的功能擴展力。動強名弱的配置，使現代漢語成爲更加典型的動詞型語言。

動詞型特徵在漢語方言中表現大體一致，只有輕微的方言差異，主要表現在有些方言能接受比普通話更多的名詞謂語判斷句。

如「張明是中學生」一句，在普通話中很難省略「是」，但是在劉丹青、唐正大 (2005) 的 19 個點的方言語法語料庫中，有 8 個點可以不用「是」，說成「張明中學生」，它們是：長沙、南昌、溫州、湘鄉、揚州、都昌、南寧、西寧。在對比句語境中，如「他是醫生，我是護士」，「是」的省略更加容易，在語料庫中增加到 13 個點，除以上 8 處外，增加了福州、成都、蘇州、廈門、衡陽。不難看出，除了西部的西寧外，更能接受名詞謂語判斷句的方言都是南方的官話（西南官話、江淮官話）和非官話方言，包括吳、湘、贛、閩等語。語料庫中南方只有廣州粵語和梅州客家話不能接受名詞謂語判斷句，這似乎與它們較強的 SVO 類型特徵有關。與之形成對比，西寧方言由於藏緬和阿爾泰等民族語言的影響，本身帶有 SOV 特徵，這可能也是它較能接受無係詞句的原因，因爲 SOV 語言不在主語和表語之間用係詞。

7. 從非量詞型語言到量詞型語言的轉型

類型學上具有重要指標意義的個體量詞，在古代漢語中尚處在萌芽狀態。數詞可以直接限制名詞，這一特徵使上古漢語屬於非量詞型語言。隨著量詞庫藏由小變大和量詞在數名組合中出現的強制性，漢語逐步過渡到量詞型語言。現代漢語中量詞已經是數詞和名詞組合中不可缺少的要素。比較（古漢語例引自周法高 1961:274）：

(31) 生丈夫：二壺酒，一犬；生女子，二壺酒，一豚。(《國語・越語上》

(32) 魯君與之一乘車，兩馬，一豎子，(《史記・孔子世家》)

(33) 他有一個孩子，家裡還養著一條狗，一頭豬，兩匹馬。

(31)-(32) 中「壺」是臨時容器量詞，「乘」在古代被作爲集體量詞，因爲車被視爲一車和數馬的集合。數詞和其他名詞即「犬、豚、馬、豎子」相接，都不加量詞。而相應的現代表達，數名之間的個體量詞「個、條、頭、匹」等都是不可或缺的。量詞已經成爲漢語語法庫藏中名詞計數時須強制添加的手段。

雖然所有漢語方言均屬量詞型語言，但量詞的入庫，並不意味著顯赫。事實上漢語量詞在不同方言間的顯赫程度存在著較大差異。根據已有研究的豐富成果，吳語粵語是量詞最顯赫的方言，而北方，量詞的顯赫度存在顯著落差。差距主要表現在以下方面：

1) 北方方言基本指示詞（這、那）和名詞之間量詞可免，如「這（本）書、那（只）包」。分別以蘇州話、廣州話爲代表的吳語、粵語的指示詞和名詞不能直接組合，中間必須用量詞。如蘇州話「搿*（本）書」，廣州話「嗰*（只）包」。實際上江淮官話及其以南的方言都極少使用指示詞直接限制名詞的結構，只是有些方言不像吳語粵語那樣絕不允許。

2) 北京話等部分北方官話口語中，基數詞「一、兩、三」和名詞之間的量詞也可以脫落，由讀固定陽平調的「一」和「倆、仨」顯示其後有一個隱性的「個」，如「來了一朋友」、「買倆蘋果」。某些官話中這種脫落量詞的數詞可以由「一」到「十」，如陝西商州話（張成材 2007）。南方官話和非官話都不允許這種量詞得不到顯現的計量結構。

3) 很多吳語、粵語方言和部分湘語、徽語、閩語（潮州話）及個別江淮官話（漣水南祿，王健、顧勁松 2006）可以由「量詞＋名詞」不依賴數詞和指示詞直接用在主語、主語的定語等位置上，主要表示有定的指稱義（個別方言也用於無定）。如蘇州話、廣州話都可以由「本書」、「只包」等做主語。其他官話方言（除煙台方言，見劉探宙、石定栩 2012）都未見這種用法。

4) 在蘇州吳語、廣州粵語等方言中，量詞可以用在領屬語等定語和核心名詞之間，不需要相當於「的」的專用定語標記，由量詞兼做定語標記，同時表示有定的語義。如蘇州話「小張買本書」（小張買的那本書），廣州話「你只包」（你的這只包）。很多漢語方言，包括幾乎全部官話，不能有這種用法。

5) 吳語、粵語量詞在 4) 中的用法也可以省略核心名詞，直接由量詞代替被飾名詞充當名詞語的核心，如蘇州話「小張買本」（小張買的那本），廣州話「我只」（我的那只（包））。官話和其他很多方言沒有這種用法。

以上幾點顯示，在吳語、粵語等南方方言中，必用量詞和可用量詞的句法位置都比官話多，量詞的表義作用也更加多樣，量詞在南方漢語中更爲顯赫，其顯赫性相應降低了名詞的重要性，因爲涉及個體的位置都要有量詞出現，從而使名詞失去或弱化了表個體的功能，成爲更加典型的類別名詞，區別於英語等語言的單數物體名詞 (sort noun vs. singular object noun, Rijkhoff 2002:54)。

8. 方位名詞的語法化和普通名詞的去處所化

漢語中的方位詞從古到今在語法性質上發生過重要的變化。

先秦時期的方位詞是名詞的一個次類，在語句中主要發揮其詞彙義功能而非語法標記功能。請看下例（見劉丹青 2003:130；另參看魏培泉 2003: §2.16）

(34) 豕人立而啼。公懼，隊=墜**於車**。傷足，喪屨。反，誅屨於徒人費。弗得，鞭之，見血。走出，遇賊**於門**。劫而束之。費曰：「我奚禦哉？」袒而示之背。信之。費請先入。伏公而出，鬥，死**於門中**。石之紛如死**於階下**。遂入，殺孟陽**於床**。曰：「非君也，不類。」見公之足**於戶下**，遂弑之，而立無知。(《左傳・莊公八年》)

由 (34) 可見，在介詞短語中，方位名詞的使用完全是由語義決定的。「於」是上古漢語介引處所題元的最重要的前置介詞。「於」支配的表處所的短語句法上並不要求加方位詞，影響選擇的是語義。所以，完全相同的句法條件下，「於車、於門、於床」不出現方位詞，「於門中、於階下、於戶下」則帶方位詞。

在現代漢語中，介引方所的前置詞在帶普通名詞作處所題元時，必須要帶方位詞，如以上例子要說成：

(35) 在車*（上）、在門*（邊）、在床*（上）、在：門*（裡面）、在臺階*（下）、在門*（下）

再比較：

(36) a. 手上/手中/手裡有錢～*手有錢
b. 坐在地上/地下/地板上～*坐在地/*坐在地板

例 (36) 說明，這裡的「上、中、裡、下」等語義基本中和，互相換用不改變語義，但是不能省略，可見主要是因句法需要而用的。這也與雙音節要求無關。(36b) 如果「地」換成「地板」，只要不加方位詞，仍然不合格。

此外，作爲名詞小類，古漢語方位詞可以自由地受帶「之」的定語修飾，如「疾不可爲也，在肓之上、膏之下」(《左傳・成公十年》)，「公入而賦：『大隧之中，其樂也融融。』姜姓出而賦：『大隧之外，其樂也泄泄。』」(《左傳・隱公元年》) 現代至少單音節方位詞已失去名詞性，不能受帶定語標記的定語修飾，如「在桌子（*的）上、在舞台（*的）下、在地道（*的）裡、在車廂（*的）外」。當然，現代也有「之上、之下」等組合，但那是古漢語以化石方式留下的詞彙化成分。它們不能和現代的定語標記組合，就表明它們不是活的名詞了。

因此，現代漢語中的方位詞已經程度不同地朝向介引處所題元的後置介詞方向語法化了，其中部分主要用於後置的單音節方位詞已經離名詞詞類很遠，接近典型的後置詞。

處所類前置詞後面的名詞強制性要加方位後置詞，這使得現代漢語名詞本身失去了充當處所名詞的功能，其功能在又一個方面比古代漢語名詞萎縮。

這個方面也存在一定的方言差異。吳語比普通話和粵語更徹底地丟失普通名詞用作處所名詞的功能。例如蘇州話「在」義存在動詞兼介詞「勒」之後不能直接帶下列名詞，而這些名詞在普通話「在」後加方位詞是兩可的，粵語「喺」後則不加更自然：

(37) 蘇州：勒學堂*（裡）/郵電局*（裡）/操場*（浪上）/圖書館*（裡）/公司*（裡）/車站*（裡）

(38) 普通話：在學校（裡）/郵局（裡）/操場（上）/圖書館（裡）/公司（裡）/車站（裡）

(39) 廣州：喺學校/郵局/操場/圖書館/公司/車站

以上六、七兩點都進一步顯示了漢語名詞功能的弱化。

9. 討論與小結：漢語語法類型演變的大勢

以上討論了古今漢語的七點類型演變。其中的六點可以概括爲兩大趨勢：

1、第一、二、三點，都反映了漢語語法的一些重要方面從句法主導到語用優先的趨勢。古代漢語語法至少在基本語序等方面更多體現句法導向，語法庫藏中有更多句法性規則，而現代漢語語法更多體現語用優先，語用在語法中的作用變得更大。就第一點來說，話題比主語有更多語用屬性，其出現與否也受制於語用因素。話題在語法體系中變得更加顯赫，意味著語用在語法中的作用得到加強。第二點是第一點的泛化，即除了話題結構，在其他的語序方面，現代漢語也更多遵循語用規則而較少受句法規則的制約。第三點，連動式顯赫。連動比之用連詞「而」之類連接的並列和用「之/的」介引的主從關係，更少句法制約，更多靠象似性等認知原則來組織結構。認知原則也是一種語用因素。

在方言差異方面，前兩點，涉及語用因素中的信息結構。吳語受信息結構影響最大，粵語最小。普通話爲代表的官話則介於兩者之間。在連動顯赫所體現的認知因素方面，正好相反，老上海話等吳語由於更多使用帶連詞的並列結構，連動式顯赫度不如其他方言，從而顯得在這點上比其他方言更接近於古代漢語的句法主導型。這也說明不同的方言有不同方面的存古度，不宜單純根據一個方面來判斷哪個方言存古度更高。

2、第五、六、七這三點，都反映了古代漢語名詞在語法庫藏中比後代強勢和活躍，在功能方面與動詞形成均衡對立的格局。第五點是名詞整體活躍程度和活動能力變得不如動詞。第六點表明它失去了充當個體名詞的能力，要靠量詞來獲得個體性。第七點說明它失去了充當處所名詞的能力，要靠後置詞來獲得處所性。這些情況相加顯著減弱了漢語名詞的語法

活力，強化了動詞和名詞在漢語中顯赫度的不對稱。這些情況在方言中大致相同，有一些細微的差別：有些方言允許較多的判斷句不使用係詞，有些方言量詞在某些條件下可以只以隱性方式存在，有些方言一部分名詞還保留用作處所名詞的功能，有些方言則有與之相反的表現，名詞的功能比普通話更弱。[9]

以上沒有討論到的是第四點動補結構的發達和顯赫。關於動補式發達所體現的漢語演變大勢，有一些學者提出了一些值得重視的看法。

梅廣 (2003) 指出，「動詞合併在上古漢語用得很多。這表示上古漢語的動詞結構具有較多綜合性，而後來的結構則朝向分析性 (analytic) 方向發展，終於完全以使成式取代了使動動詞」。黃正德 (2008) 也基於其他一些語言現象持現代漢語比古漢語有更多解析性（＝分析性）的觀點，認爲「相對於現代漢語，英語與古代漢語有相當豐富的 L-Syntax，許多動詞可以在詞彙部門派生而成（如 phone, fish, joke 和『電、漁、戲』等都可以做動詞用，其來源是詞法派生），但現代漢語缺少相應的派生詞法，所以用『打電話、捕魚、開玩笑』這類複式述語來表達同樣的概念。所謂複式述語 (complex predicate, complex verb) 就是以短語的手段來表達相當於一個簡單動詞的概念。（Massam (2001) 把這種結構稱爲 pseudo-incorporation『准併入結構』）」。他也將古代漢語「孤之事君」這一類 VP 名詞化用法的

[9] 有審稿人提出，第六點（量詞的使用）和第七點（方位詞的使用）其實反映了現代漢語表達更加有分析性，古漢語用一個名詞可以表達現代漢語量詞＋名詞或名詞＋方位詞的組合的內容，古漢語更有綜合性。我們覺得這一解釋有一定道理，但不能解釋問題的全部。量詞的顯赫性不僅表現為在數詞和名詞之間要加量詞，而且更多表現為量詞的多樣化功能，特別是在吳語粵語中表定指、表定中關係、甚至單獨做論元的功能，這些都與分析性無關。而方位詞用於名詞後也並不都增加分析性。例 (36) 各句顯示，方位詞的語義在其中基本不起作用，其使用完全是一種句法位置的需求，因此並不增加單位表義的分析性。方位詞也發展出了一些擴展功能，與空間位置無關，如「原則上、根本上、理論上、暗地裡」等，這些組合與空間表達的分析性與否也沒有直接關係。再如，有些名詞＋方位詞的組合反而不表空間，而表組織機構或領導群體，如「組織上、縣上、省裡、領導上（＝領導）」，這些方位詞也談不上增加空間表達分析性的問題，而是方位後置詞作為顯赫範疇的功能擴展。

消退與屈折形態的衰落聯繫起來，進一步闡述這一觀點，認爲「漢語屈折形態的消失導致『主之謂』結構的衰退，而派生形態的衰弱則引起輕動詞結構與准併入結構的興起，並輾轉導致僞定語結構的滋長。」

關於古今漢語從綜合到解析的觀點，可以解釋一系列現象，如上述學者提到的由單個謂詞的使動用法到兩個謂詞構成的動結式，由單個名詞做謂語動詞到用動詞短語來表達同樣意思。但是，綜合＞解析論也面對一些很難解釋的現象，包括本文所總結的一些演變趨勢。

中古以後顯赫起來的連動式，在上古漢語中常以帶「而」的並列結構出現。連動結構是不借助任何連接手段緊密結成一個整體的單位，而帶「而」的並列式是由連詞明確分割開的兩個單位的組合。如「爬樹找魚」和「緣木而求魚」，後者顯然比前者有更強的解析性。連動式的顯赫度增加是由解析性較強到綜合性較強的演變。

再看動結式。梅廣先生將動結式與更早時的使動用法相比較，由此例證漢語綜合性向解析性的演變。這一比較本身是可信的。但是，古漢語表達致使範疇，不僅靠綜合性的使動用法（其中有些可能有形態變化，也屬於綜合性手段），同時也用致使性兼語句，如：

(40) 五色令人目盲，五音令人耳聾，五味令人口爽，馳騁田獵令人心發狂，難得之貨令人行妨。(《老子十二》)

這些兼語句的意思，也頗能用後代的動結式來表達，如：

(41) 五色炫瞎眼睛，五音震聾耳朵，五味吃壞人嘴，馳騁打獵打狂了人心，珍奇好貨誘壞了行為。

在致使手段類型學中，用形態手段表達屬於綜合性手段（使動用法可視爲零形態的綜合手段），用另造詞語表達致使屬於詞彙性手段（如吃–喂，英語 die-kill），兩者可合併爲綜合性手段；用兼語句一類句法手段屬

於分析性手段（參看牛順心 2004:§1.2 及所引 Comrie、Shibatani 等的論述）。而漢語表致使的動結式，被眾多文獻視爲複合詞，其性質介於綜合性及詞彙性致使式和分析性句法致使式的中間，而更靠攏詞彙式一些（牛順心 2004:§2.6），亦即更偏向綜合手段一側。動結式的顯赫，既部分取代了古代漢語的使動用法（沒有全部取代，現代也有「溫暖人心、嚴肅紀律、壞了事兒」等使動用法），也大大分割了分析性致使結構的領地。像「跑累了馬」、「笑痛了肚子」、「唱壞了嗓子」、「洗濕了鞋」這類常用的現代漢語動結式所表示的語義，在其他語言中常要用更加分析性的句法結構才能表達。因此，即使是動結式的顯赫，也不單純是分析性的加強。所以，我們認爲根據目前提出的證據，從綜合性到解析性還不足以構成古今演變的大趨勢。動結式的顯赫與漢語庫藏調整大勢的關係還有待進一步考察。

總體上，本文提到的這些類型演變及其概括出來的大勢，是遍及漢語方言的根本性演變。本文的跨方言比較，一方面讓我們看到在演變速度和範圍上的差異，另一方面也讓我們更加確信這些演變的大勢是漢語各方言的大勢。

引用文獻

Aikhenvald, Alexandra Y. 2006. Serial verb constructions in typological perspective. *Serial Verb Constructions: A Cross-linguistic Typology*, ed. by Alexandra Y. Aikhenvald & R. M. W. Dixon, 1-68. Oxford & New York: Oxford University Press.

Bao, Mingwei (鮑明煒), and Jun Wang (王均). (eds.) 2002. *Nantong Diqu Fangyan Yanjiu* 南通地區方言研究 [*A Study on the Dialects in Nantong Area*]. Nanjing: Jiangsu Education Press.

Cheng, Lixia (程麗霞). 2005. Zuo pianzhi jiegou pinlü tongji yu huati jiegou de xianxian 左偏置結構頻率統計與話題結構的顯現 [Frequency of left-dislocation and the emergence of the topic construction]. *Yufahua yu Yufa Yanjiu* 語法化與語法研究 [*Studies on Grammaticalization and Grammar*], Vol. 2, ed. by Jiaxuan Shen (沈家煊), Fuxiang Wu (吳福祥) & Beijia Ma (馬貝加), 1-17. Beijing: The Commercial Press.

Chou, Fa-kao (周法高). 1959. *Zhongguo Gudai Yufa: Chengdai Pian* 中國古代語法：稱代編 [*A Historical Grammar of Ancient Chinese: Substitution*]. Taipei: Institute of History and Philology, Academia Sinica.

Chou, Fa-kao (周法高). 1961. *Zhongguo Gudai Yufa: Zaoju Pian* 中國古代語法：造句編 [*A Historical Grammar of Ancient Chinese: Syntax*], Vol. 1. Taipei: Institute of History and Philology, Academia Sinica.

Clark, Eve V., and Herbert H. Clark. 1979. When nouns surface as verbs. *Language* 55.4:767-811.

Djamouri, Redouane. 2001. Markers of predication in Shang bone inscriptions. *Sinitic Grammar: Synchronic and Diachronic Perspectives*, ed. by Hilary Chappell, 143-171. Oxford & New York: Oxford University Press.

Dong, Xiufang (董秀芳). 2007. Cong cihuihua de jiaodu kan nianheshi dongbu jiegou de xingzhi 從詞彙化的角度看粘合式動補結構的性質 [On the property of resultative construction: from the perspective of lexicalization]. *Yuyan Kexue* 語言科學 [*Linguistic Sciences*] 2007.1:40-47.

Gao, Zengxia (高增霞). 2006. *Xiandai Hanyu Liandongshi de Yufahua Shijiao* 現代漢語

連動式的語法化視角 [*Serial Verb Constructions in Mandarin: A Perspective of Grammaticalization*]. Beijing: China Archives Information Press.

Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Language* (2nd edition), ed. by Joseph H. Greenberg, 73-113. Cambridge: MIT Press.

Hou, Jingyi (侯精一). 1999. *Xiandai Jinyu de Yanjiu* 現代晉語的研究 [*Studies in Modern Jin Dialects*]. Beijing: The Commercial Press.

Huang, C.-T. James (黃正德). 2008. Cong 'ta de laoshi dangdehao' tanqi 從"他的老師當得好"談起 [On *ta de laoshi dangdehao* 'he teaches well' and related problems]. *Yuyan Kexue* 語言科學 [*Linguistic Sciences*] 2008.3:225-241.

Jiang, Shaoyu (蔣紹愚), and Guangshun Cao (曹廣順). (eds.) 2005. *Jindai Hanyu Yufashi Yanjiu Zongshu* 近代漢語語法史研究綜述 [*A Sketch of the Research of the History of Pre-modern Chinese Grammar*]. Beijing: The Commercial Press.

LaPolla, Randy J. 1995. Pragmatic relations and Chinese word order. *Word Order in Discourse*, ed. by Pamela Downing & Michael Noonan, 297-330. Amsterdam & Philadelphia: John Benjamins.

Li, Renjian (李人鑒). 1957. Taixing fangyan zhong dongci de houfu chengfen 泰興方言中動詞的後附成分 [Enclitics to verbs in the Taixing dialect]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1957.5:16-20.

Li, Rulong (李如龍). 1996. Quanzhou fangyan de ti 泉州方言的體 [Aspect system of the Quanzhou dialect]. *Dongci de ti* 動詞的體 [*Aspect of Verbs*], ed. by Song-Hing Chang (張雙慶), 195-224. Hong Kong: T. T. Ng Chinese Language Research Centre, Institute of Chinese Studies, The Chinese University of Hong Kong.

Li, Ying-che (李英哲). 2001. Guanhua he Taiwan Minyu de yixie dongcizu jiegou bijiao yanjiu 官話和臺灣閩語的一些動詞組結構比較研究 [A comparative study of certain verb phrases in Mandarin and Taiwanese Minnan]. *Hanyu Lishi Gongshi Yufa Lunji* 漢語歷時共時語法論集 [*Chinese Grammar Then and Now: Writings on Chinese Diachronic and Synchronic Syntax*], 163-173. Beijing: Beijing Language and Culture University Press.

Liu, Danqing (劉丹青). 1994. 'Weibuci' chutan "唯補詞"初探 [A preliminary study

of complement-only words]. *Hanyu Xuexi* 漢語學習 [*Chinese Language Learning*] 1994.3:23-27.

Liu, Danqing (劉丹青). 1995. Yuyi youxian haishi yuyong youxian: Hanyu yufaxue tixi jianshe duanxiang 語義優先還是語用優先？——漢語語法學體系建設斷想 [Semantic prominence or pragmatic prominence: some segmental thoughts on the reconstruction of the theoretical system of the Chinese grammar]. *Yuwen Yanjiu* 語文研究 [*Linguistic Researches*] 1995.2:10-15.

Liu, Danqing (劉丹青). 2000. Yueyu jufa de leixingxue tedian 粵語句法的類型學特點 [Typological characteristics of Cantonese syntax]. *Asia Pacific Journal of Language in Education* 3.2:1-29.

Liu, Danqing (劉丹青). 2001a. Hanyu fangyan yuxu leixing de bijiao 漢語方言語序類型的比較 [A typological comparison of word order among Chinese dialects]. *Contemporary Research in Modern Chinese* 2:24-38.

Liu, Danqing (劉丹青). 2001b. Lunyuan fenlieshi huati jiegou chutan 論元分裂式話題結構初探 [A preliminary study on argument-split topic structure]. *Yuyan Yanjiu Zai Renshi: Qingzhu Zhang Bin Xiansheng Congjiao Wushi Zhounian ji Bashi Huadan* 語言研究再認識：慶祝張斌先生從教五十周年暨八十華誕 [*New Approaches to Linguistic Research: A Festschrift in Honor of Professor Zhang Bin's Fiftieth Teaching Year and Eightieth Birthday*], 141-154. Shanghai: Shanghai Education Press. Also in *Huati yu Jiaodian Xinlun* 話題與焦點新論 [*New Ideas about Topic and Focus*], ed. by Liejiong Xu (徐烈炯) & Danqing Liu (劉丹青). Shanghai: Shanghai Education Press, 2003.

Liu, Danqing (劉丹青). 2003. *Yuxu Leixingxue yu Jieci Lilun* 語序類型學與介詞理論 [*Word Order Typology and a Theory of Adpositions*]. Beijing: The Commercial Press.

Liu, Danqing. 2004. Identical topics: a more characteristic property of topic prominent languages. *Journal of Chinese Linguistics* 32.1:20-64.

Liu, Danqing (劉丹青). 2008. Binglie jiegou de jufa xianzhi ji qi chubu jieshi 並列結構的句法限制及其初步解釋 [Syntactic constraints on coordination and some preliminary explanations]. *Yufa Yanjiu he Tansuo* 語法研究和探索 [*Researches*

and Explorations in Chinese Grammar] 14:1-21. Beijing: The Commercial Press.

Liu, Danqing (劉丹青). 2009. Huati youxian de jufa houguo 話題優先的句法後果 [Syntactic consequences of topic prominence]. *Hanyu de Xingshi yu Gongneng Yanjiu* 漢語的形式與功能研究 [*Formal and Functional Studies in Chinese*], ed. by Gong Cheng (程工) & Danqing Liu (劉丹青), 50-71. Beijing: The Commercial Press.

Liu, Danqing (劉丹青). 2010. Hanyu shi yizhong dongcixing yuyan: shishuo dongcixing yuyan he mingcixing yuyan de leixing chayi 漢語是一種動詞型語言：試說動詞型語言和名詞型語言的類型差異 [Chinese as a verby language: on typological differences between verby languages and nouny languages]. *Shijie Hanyu Jiaoxue* 世界漢語教學 [*Chinese Teaching in the World*] 2010.1:3-17.

Liu, Danqing (劉丹青). 2011a. Yuyan kucang leixingxue gouxiang 語言庫藏類型學構想 [Linguistic Inventory Typology: a proposal for a new approach of linguistic typology]. *Dangdai Yuyanxue* 當代語言學 [*Contemporary Linguistics*] 2011.4: 289-303.

Liu, Danqing (劉丹青). 2011b. Hanyushi yufa leixing tedian zai xiandai fangyan zhong de cunfei 漢語史語法類型特點在現代方言中的存廢 [The reservation and disappearance of typological features in the history of Chinese: a cross-dialectical examination]. *Yuyan Jiaoxue yu Yanjiu* 語言教學與研究 [*Language Teaching and Linguistic Studies*] 2011.3:28-38.

Liu, Danqing (劉丹青). 2011c. 'You' zi lingyouju de yuyi qingxiang he xinxi jiegou "有" 字領有句的語義傾向和信息結構 [The semantic preference and information structure of the possessive construction with the verb 'you (have)']. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2011.2:99-109.

Liu, Danqing (劉丹青), and Zhengda Tang (唐正大). (eds.) 2005. *Xiandai Hanyu Fangyan Yufa Yuliaoku* 現代漢語方言語法語料庫 [*Corpus of Contemporary Chinese Dialectal Grammar*]. Beijing: Chinese Academy of Social Sciences.

Liu, Jian (劉堅), and Shaoyu Jiang (蔣紹愚). (eds.) 1990-95. *Jindai Hanyu Yufa Ziliao Huibian* 近代漢語語法資料彙編 [*Comprehensive Source Materials for the Study of Grammar of Early Modern Chinese*], 3 Vols. Beijing: The Commercial Press.

Liu, Tanzhou (劉探宙), and Dingxu Shi (石定栩). 2012. Yantaihua zhong budai zhishici huo shuci de liangci jiegou 煙台話中不帶指示詞或數詞的量詞結構 [Bare classifier phrases in the Yantai dialect]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2012.1:38-49.

Lü, Jiping (吕冀平). 1985[1958]. *Fuza Weiyu* 複雜謂語 [*Complex Predicates*]. Shanghai: Shanghai Education Press.

Lü, Shuxiang (吕叔湘). (ed.) 1980. *Xiandai Hanyu Babaici* 現代漢語八百詞 [*Eight Hundred Words in Modern Chinese*]. Beijing: The Commercial Press.

Ma, Jianzhong (馬建忠). 1983[1898]. *Ma Shi Wen Tong* 馬氏文通. Beijing: The Commercial Press.

Mai, Yun (麥耘). 1995. Guangzhouhua de houbushi xingrongci 廣州話的後補式形容詞 [Adjectives formed as 'adjective-noun/classifier' in Cantonese]. *Guangzhouhua Yanjiu yu Jiaoxue* 廣州話研究與教學 [*Research and Teaching of Cantonese*] 2:11-21. Guangzhou: Sun Yat-sen University Press.

Matthews, Stephen (馬詩帆), and Louisa Y. Y. Yeung (楊月英). 2003. Guangdonghua huatihua de chuli dongji 廣東話話題化的處理動機 [A processing advantage for topicalization in Cantonese]. *Huati yu Jiaodian Xinlun* 話題與焦點新論 [*New Ideas about Topic and Focus*], ed. by Liejiong Xu (徐烈炯) & Danqing Liu (劉丹青), 145-163. Shanghai: Shanghai Education Press.

Mei, Kuang (梅廣). 2003. Yingjie yige kaozhengxue he yuyanxue jiehe de Hanyu yufashi yanjiu xin jumian 迎接一個考證學和語言學結合的漢語語法史研究新局面 [Anticipating a new horizon of the historical syntax of Chinese incorporating philology and linguistics]. *Historical Development of Chinese Language*, ed. by Dah-an Ho (何大安), 23-47. Taipei: Institute of Linguistics, Academia Sinica.

Niu, Shunxin (牛順心). 2004. *Hanyu zhong Zhishi Fanchou de Jiegou Leixing Yanjiu* 漢語中致使範疇的結構類型研究 [*A Typological Study of Causatives in Chinese*]. Shanghai: Shanghai Normal University dissertation.

Ota, Tatsuo (太田辰夫). 2003[1958]. *Zhongguoyu Lishi Wenfa* 中國語歷史文法 [*A Historical Grammar of Modern Chinese*] (2nd edition), translated by Shaoyu Jiang (蔣紹愚) & Changhua Xu (徐昌華). Beijing: Peking University Press.

Qiao, Quansheng (喬全生). 2000. *Jin Fangyan Yufa Yanjiu* 晉方言語法研究 [*A Study on the Grammar of Shanxi Dialects*]. Beijing: The Commercial Press.

Rijkhoff, Jan. 2002. *The Noun Phrase*. Oxford: Oxford University Press.

Steele, Susan. 1978. Word order variation: a typological study. *Universals of Human Language*, Vol. 4: *Syntax*, ed. by Joseph H. Greenberg, 585-623. Stanford: Stanford University Press.

Wang, Jian (王健), and Jinsong Gu (顧勁松). 2006. Lianshui (Nanlu) hua liangci de teshu yongfa 漣水（南祿）話量詞的特殊用法 [Special usage of classifiers in the Lianshui dialect]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2006.3:237-241.

Wang, Li (王力). 1989. *Hanyu Yufashi* 漢語語法史 [*A History of Chinese Grammar*]. Beijing: The Commercial Press.

Wang, Rudong (汪如東). 2006. *Haian Fangyan Yanjiu* 海安方言研究 [*Studies in the Haian Dialect*]. Beijing: Xinhua Press.

Wang, Rudong (汪如東). 2008. Tairupian fangyan zhong dongci houfu ‘a (啊)’ de liangzhong jiegou 泰如片方言中動詞後附“啊”的兩種結構 [A study on two structures with the suffix ‘a (啊)’ in Tairu dialects]. *Yuyan Kexue* 語言科學 [*Linguistic Sciences*] 2008.3:329-336.

Wei, Pei-chuan (魏培泉). 2003. Shanggu Hanyu dao Zhonggu Hanyu yufa de zhongyao fazhan 上古漢語到中古漢語語法的重要發展 [The important developments of grammar from Old Chinese to Middle Chinese]. *Historical Development of Chinese Language*, ed. by Dah-an Ho (何大安), 75-106. Taipei: Institute of Linguistics, Academia Sinica.

Yang, Bojun (楊伯峻), and Leshi He (何樂士). 2001. *Gudai Hanyu Yufa ji qi Fazhan* 古代漢語語法及其發展 [*The Grammar of Classic Chinese and its Development*] (revised edition). Beijing: Language and Culture Press.

Zhang, Bojiang (張伯江). 2011. Hanyu de jufa jiegou he yuyong jiegou 漢語的句法結構和語用結構 [Mapping from pragmatic structures to syntactic structures in Chinese]. *Hanyu Xuexi* 漢語學習 [*Chinese Language Learning*] 2011.2:3-12.

Zhang, Chengcai (張成材). 2007. Shangzhou fangyan kouyu zhong de shuliang lianyin

bianhua 商州方言口語中的數量連音變化 [Sandhis between a numeral and a classifier in the Shangzhou dialect]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2007.2:159.

Zhang, Min (張敏). 2010. 'Donghou xianzhi' de quyu tuiyi ji qi shizhi "動後限制" 的區域推移及其實質 [Areal transition and the property of the Postverbal Constraint phenomenon]. Paper presented at the International Symposium for Comparative and Typological Research on Languages of China, May 8-9, 2010. Hong Kong: The Hong Kong Polytechnic University.

Zhang, Min (張敏), and Yuxiang Li (李予湘). 2009. Xianqin Lianghan Hanyu quxiang dongci jiegou de leixingxue diwei ji qi bianqian 先秦兩漢漢語趨向動詞結構的類型學地位及其變遷 [Typological status and its change of verb-directional constructions in Chinese during Pre-Qin and Han dynasties]. Paper presented at the Workshop on Chinese Directionals: History and Dialectal Variation, in conjunction with the 6th Cross-Strait Conference on Chinese Historical Grammar, August 26-27, 2009. Taipei: Institute of Linguistics, Academia Sinica.

Zhou, Chiming (周遲明). 1957. Hanyu de shidongxing fushi dongci 漢語的使動性複式動詞 [Causative compound verbs in Chinese]. *Shandong Daxue Xuebao* 山東大學學報 [*Journal of Shandong University*] 1957.1:187-227.

Typological Change of Syntax from Classic to Modern Chinese: A Cross-dialectal Perspective of Inventory Typology

Danqing Liu

Chinese Academy of Social Sciences

Beginning with a brief introduction of some basic concepts of linguistic inventory typology, this article examines several important typological changes of syntax from Classic Chinese to Modern Chinese from a cross-dialectal perspective of inventory typology. Seven major changes and their dialectal diversity are discussed, six of which are grouped into two general diachronic tendencies. Changes 1-3 concern the evolution of topicalization, basic word orders and serial verb constructions, illustrating a change from syntax-dominant type into a pragmatic-prominent type. Changes 5-7 concern a general decline of syntactic functions of nouns, the usage of classifiers and postpositional localizers, demonstrating a change from a situation of verb-noun balance to a model of verbs dominant over nouns. The middle change, item 4, involves the emergence and development of verb-complement (resultative, directive and potential complements) constructions. It is claimed that the robust roles of verb-complement constructions cannot be summarized as to be an indicator of a general evolution of Chinese from a more synthetic language in old times to a more analytic one in modern times.

Keywords: inventory typology, classic and modern Chinese, typological changes of syntax, dialectal diversity

《語言資訊和語言類型》
第四屆國際漢學會議論文集
臺北：中央研究院，2013 年

明清時期荔鏡／荔枝記閩南方言指示詞的演變：從指示詞到程度加強副詞或篇章標記*

連金發
國立清華大學

本文旨在檢視明清閩南語戲文中「遠/近指示詞＋樣式類別詞」所經歷的語法化現象，從戲文中可以看出「遠/近指示詞＋樣式」構式，即「許樣」和「只樣」，有些例子跳脫定語詞組，成為修飾程度形容詞的加強語，連帶促成合音詞「向」和「障」，合音詞中的樣式義淡化後所留下的空缺還可以加上樣式類義詞「般」、「生」、「年」。從「指示＋樣式」到程度加強，又回歸到「指示＋樣式」，反映葉氏循環的新陳代謝現象。「向」和「障」又分別和「生」、「年」結合，形成篇章標記。此外，本文也探討「指示詞＋數量詞」構式的語法化現象，「只」＋「夥」形成合音詞「拙」，「拙」可以指「這些」或「這麼」，前意較能反映原來的組合意，後意是進一步演化的結果。「喝」或「歇」為遠指詞「許」和「夥」的合音，不過數量極少。

關鍵詞：指示詞，加強副詞，篇章標記，閩南語，明清戲文，葉氏循環

* 本文的研究承蒙國科會 TYOMIN (NSC 101-2923-H-007-001) 的資助，謹此致謝。語料爬梳，論文整理輸入王小梅居功厥偉，在此一併感謝。原稿修改期間，兩位匿名評審多方指正，受益良多，銘感於心。本文或有疏漏，責在筆者。

1. 前言

本篇論文從歷時和類型的角度探索閩南語指示詞的演變，本文據以討論的文本是明清時期的荔鏡/荔枝記戲文。[1] 近年來有些學者認爲限定詞組 (determiner phrase) (DP) 中限定詞佔據中心語 (head) 的位置。[2] 閩南語限定詞組由限定詞、數量詞、類別詞、名詞組構成，其中類別詞可分爲個體類別詞和種類類別詞，如 tsi^{2} tsit^{8} ki^{1} pit^{4}「只一枝筆」(這一枝筆)、tsi^{2} tsit^{8} khuan^{2} pit^{4}「只一款筆」(這一種筆)。[3] 限定詞組中限定詞的位置可出現近指和遠指指示詞，如 tsi^{2}「只」和 hur^{2}「許」。限定詞組的組成分子並不需要每個都出現，可以將其中一部分抽離出來使用，如「只一款」和「許一款」。指示詞常和數詞「一」形成合音詞，如「只一款」和「許一款」合音後「一」成爲「只」和「許」的韻尾。「只一款」和「許一款」的合音式一般寫爲「只款」和「許款」。明清時期的戲文中指示詞和樣式語詞（爲種類類別詞的一種），如 tsi^{2} iunn^{7}「只樣」和 hir^{2} iunn^{7}「許樣」，形成合音詞，用同音詞表示，分別以假借字 tsiunn^{3}「障」、hiunn^{3}「向」書寫。[4] 合音現象反映充當定語變爲程度形容詞的加強副詞，即原論元變作謂語的

---

[1] 明清閩南荔鏡/荔枝記戲文有五個版本：嘉靖本 (1566)、萬曆本 (1581)、順治本 (1651)、光緒本 (1884)（參見吳守禮 2001a-d）、道光本 (1831)（泉州市文化局泉州地方戲曲研究社 2010）。本文標音以泉腔為主，但考慮早期泉州方言的特點，去聲還是保留陰陽調。標示法詳參註 3。

[2] 有關定語詞組的內部結構深入分析請參閱 Valois (1991) 及 Giusti (1997)。

[3] 本文的閩南語的音標主要根據 (Douglas 1873)，為了輸入方便起見，音標做了若干的調整。/ko/「高」和 /kɔ/「孤」的元音區別以 /o/ 和 /oo/ 表示。元音鼻化以 -nn 表示、如「三」/sã/ 轉寫為 /sann/。陰平、陰上、陰去、陰入、陽平、陽去、陽入調分別以上標符號阿拉伯數字 1、2、3、4、5、7、8 表示。0 表示輕讀，如 au^{7}-jit^{0} 後日（後天）。後高不圓唇元音 /ɯ/ 以 /ir/ 表示。聲母 ch/chh 和 ts/tsh 無音位對立一律做 ts/tsh。後中不圓唇元音 /ɤ/ 以 /er/ 表示。本文的標示法與教育部公告之「臺灣閩南語羅馬字音標方案」基本上吻合，細節說明可參看。

[4] 專就明清閩南戲文的指示詞做理論的探索有黃漢君、連金發 (2007)、陳麗雪 (2009)。兩篇都只根據一個文本：第一篇針對萬曆本，第二篇針對嘉靖本。本文擴大範圍，涵蓋五個文本，橫跨三個世紀多。

附加語。現代臺灣閩南語指示詞和樣式和種類語詞所形成的「只樣」和「許樣」、「只款」和「許款」、「只種」和「許種」並沒有轉化爲加強副詞，而是由指示詞和另一種來歷未明的語詞合音後再加上 ni[1]「爾」所形成的 tsiah[4] ni[1]「即爾」（這麼）和 hiah[4] ni[1]「郝爾」（那麼）表示，如「即爾厚」（這麼濃/多）、「郝爾熱」（這麼熱）。[5] 共通語「這麼」和「那麼」除做加強詞外還可充當方式副詞，這個功能臺灣現代閩南語是用 an[2] ni[1]「按爾」表示，如 An[2] ni[1] tsoh[4] tsiah[4] ue[7] sai[2] tit[4]「按爾做即會使得」（這麼做才可以）。此外「按爾」還有照應詞的功能，如 A[1]-ing[1] tsin[1] khiau[2], a[1]-lan[5] ma[7] an[2] ni[1]「阿英真巧，阿蘭嘛按爾」（阿英真聰明，阿蘭也是這樣）。[6]

本文預計以明清閩南戲文荔鏡/荔枝記中遠/近指示詞＋樣式/數量詞的構式爲研究焦點，探討這類結構所演化出的各種語法功能。特別著重討論原始遠/近指示樣式構式及合音詞之間在功能上的異同，並連帶討論語言演變中葉氏循環現象。

2. 遠/近指示詞＋樣式詞

近指指示詞（以下簡稱「近指詞」）tsi[2]「只」和遠指指示詞（以下簡稱「遠指詞」）hir[2]「許」和樣式類別詞 iunn[7]「樣」結合成 tsi[2] iunn[7]「只樣」和 hir[2] iunn[7]「許樣」，「只樣」和「許樣」可縮成合音詞 tsiunn[3]「障」

---

[5] 「爾」有三種變體：ni[1]、nih[8]、ne[1]。舒聲韻促化，即母音喉塞化，是語法化常伴隨的現象（鄭張尚芳 1995，戴昭銘 2004）。「即」和「赫」都可以充當副詞性程度加強詞，「爾」可有可無，這表示「即」和「赫」是由近遠指示詞和未明的種類詞合音而成，也許主流臺灣閩南語比較反映漳腔的現象。「即爾」和「赫爾」可以分別縮約為合音詞 tsian[3] 和 hian[3]，意指這麼、那麼（張相 2009:100, 166）。這樣的現象說明主流臺灣閩南語也反映了葉氏循環現象。

[6] 現代閩南語指示代詞的研究請參閱黃丁華 (1961)、Lien (1999)、劉秀雪 (2005)、李如龍 (2007)。明清時代閩南語和現代閩南語的指示＋樣式構式的比較研究會很有意義。本文先建立歷時的基礎，比較研究將另文發表。

和 hiunn[3]「向」。[7]

明清戲文中只有「只樣」，沒有「許樣」，但是都有「障」和「向」，可見「只樣」變入合音詞「障」和「許樣」變入合音詞「向」的速度不是一致的。前者還殘存原始形式，後者原始形式已經銷聲匿跡。

3. 韻律和語法演變

近代漢語，閩南語也不例外，雙音詞和其他語詞合用時有其韻律上的限制，一般以對稱爲原則。雙音詞接雙音節通常不接單音詞，即 2＋2，*2＋1，比如「只樣」只能接雙音節，「障」可接單音節，形成新雙音節詞。「只樣」可以獨用，但「障」不可以。這種音韻上的限制只能看成一種傾向，並非絕對的，單雙音節或音節對稱不對稱可能會有其他因素涉及其中（比較黃漢君、連金發 2007）。

4. 原始構式、合音詞、合音詞再加後綴

由實詞 (lexeme) 所組合成的複合詞融合成單音節詞，原始詞中的各個實詞的語義可能磨損掉。爲了表現原來的詞義，就須加上新的語詞，比如「只樣」融合成「障」，其中表示樣式的「樣」的語義磨損掉，爲了彌補這個義素，就另外加上「般」這個詞，形成新起詞「障般」。

4.1「只樣」和「障」、「只樣」和「向」

「只樣」本屬定語詞組，即「只近指詞」＋「樣樣式類別詞」＋名詞組。但是「只樣」後頭接的不是名詞組，而是動詞組或形容詞組，這樣一來就會跳

[7] 現在潮州音（包括汕頭話）「障」做 tsionn[3] 和「向」做 hionn[3]（曾憲通 1991，施其生 1996）。

脫名詞詞組的格局成爲副詞。

與「只樣」合用的形容詞都是程度形容詞 (gradable adjectives)，「只樣」充當加強副詞，意指「這麼」。明清戲文中與「只樣」合用的形容詞只有「無狀」這個類型，「無狀」[8] 是由「形狀」派生而來的否定形容詞。與動詞組合用的「只樣」是指稱樣式或方式的副詞，意指「這樣」，帶有鮮明的當場的指示示證 (evidentiality) 語意。

「只樣」可以縮約成合音詞「障」，「許樣」也可縮約爲「向」，例見如下：

(1) liah^{8} gua^{2} tsit^{8} sin^{1} phuah^{4} tit^{4} tsiunn^{3} tam^{5}
力我一身潑得障濕（22.119 嘉靖）
把我潑得全身這麼濕

(2) li^{2} hiunn^{3} ai^{3} tsinn^{5}
你向愛錢（14.311 嘉靖）
你那麼愛錢

合音詞「障」「向」中「樣」的語意逐漸淡化，因此可以再加上類義詞「般」、「年」、「生」。這些都是意指「樣式」的類義詞，分屬不同的時間層次。「樣」、「般」是現代樣式類別詞。「年」（＜爾/若）上古漢語即存在。[9]「生」最早出現於唐朝中葉（志村良治 1984:323-335, 1995:303-

[8] 「名詞」前加 u^{7}「有」和 bo^{5}「無」（＜否定＋「有」）有些可以轉為程度形容詞，如 u^{7} tsinn^{5}「有錢」可以用加強詞「真」tsin^{1} 修飾，如「真有錢」。Bo^{5}「無」的作用也如此，如 bo^{5} le^{2}-mau^{7}「無禮貌」可以用「足」tsiok^{4} 來修飾。

[9] 「年」可能來自上古漢語鼻音泥聲母系列的「爾/若/然」等指代詞，和喉牙音聲母系列的指代詞都是遠指詞，但是前者是較近的遠指詞，而後者是較遠的遠指詞（藤堂明保 1952，志村良治 1984:131-152, 1995:103-141）。六朝時期「爾」多用作遠指指示詞。「爾」常與「馨」連用，形成「爾馨」（其變體為「寧馨」），「爾」帶有「這樣」「那樣」之意。後來這整個雙音節的語意由「寧」（或「能」）所承續（志村良治 1984:131-152, 1995:103-143）。明清時期的「年」可能就是傳承於此。

315)。這些不同時代層次的語詞並存於明清戲文的文本中，各有相似而略不同的語意、句法甚至篇章功能。以下依次討論「障般」/「向般」、「障年」/「向年」、「障生」/「向生」的用法。接著連帶討論相關的語式。

4.2「障般」和「向般」

新起詞 tsiunn^{3} puann^{1}「障般」和原始詞「只樣」用法有重疊之處，但不盡相同，比如同樣的名詞組和動詞組都可以與 hiunn^{3} puann^{1}「障般」或「只樣」合用，如 tsi^{2} iunn^{7}/tsiunn^{3} puann^{1} jin^{5}-but^{8}「只樣/障般人物」、tsi^{2} iunn^{7}/tsiunn^{3} puann^{1} kiann^{5} lai^{0}「只樣/障般行來」，兩者基本上以出現於定語詞組爲常，但後起詞「障般」較有活力，能產性較高。

「障般」和「只樣」都反映說話者主觀的情感，憎惡和讚賞都有，所接名詞組有部分是詈詞，如 ok^{4} tshat^{8}「惡賊」、si^{2} tshat^{8} kann^{2}「死賊仔」、kau^{2} sing^{3}「狗性」、tshat^{8} loo^{5}「賊奴」等，但也有主觀讚嘆口吻，如 ho^{2} king^{2}「好景」、tshin^{1} tshinn^{2} tsir^{1} niu^{5}「親淺孜娘」（貌美的女人）等。帶主觀情緒的名詞組容易轉變爲程度詞。[10]

4.3「障年」和「向年」

十七世紀以前現存的閩南語戲文 tsiunn^{3} ni^{1}「障年」只出現於萬曆本中，都爲回指的篇章功能，前頭常和「做」或「是」合用，意指「這樣」。「向年」有承上啓下的功能，意指「那樣的話」。[11]「向年」指稱前述事，引出隨後的行爲，相當於共通語「那/這樣的話，就接續下面的行爲」。例見如下：

---

[10] Tshin^{1}-tsinn^{2}，Douglas (1873:81) 釋意為 pretty。明清戲文書寫為「親淺」或「親醒」，南音文獻記為「親醒」（李麗敏 2006:121）。

[11] 「障年」tsiunn^{3}-ni^{1} 和「向年」hiunn^{3}-ni^{1} 可能是潮州話的特徵詞，除萬曆本外還出現於有潮州方言特色的《金花女》和《蘇六娘》。「障年」除帶有本文所論的「這樣」之意外，還可以表示「怎樣」之意（曾憲通 1991）。

(3) (大) gua^{2} tir^{7} ke^{1} kau^{3}-hun^{3} sio^{2} tong^{5}
我在家教訓小童（35.081 萬曆）
我在家教導小孩
(公) tsiann^{3} si^{7} tsiunn^{3} ni^{1}
正是障年（35.082 萬曆）
正是這樣
gua^{2} tshut^{4} tsit^{8} tue^{5}-bok^{8} tshiann^{2} ko^{2}-tshi^{3} tsit^{8} e^{0}
我出一題目，請考試一下（35.083 萬曆）
我出一個題目，來考一考
(4) (旦) lir^{2} tann^{1} iu^{7} berh^{4} gun^{2} sin^{1} ge^{0}
你今又要阮新个（24.396 萬曆）
你現在又要我新的
hiunn^{3}-ni^{1} tshiu^{2}-phe^{3} tir^{7} tsia^{1} khit^{4} lin^{2}
向年手帕在只，乞恁（24.397 萬曆）
那麼手帕在這兒，給你

4.4「障生」和「向生」

Sinn^{1}「生」是唐中葉到宋代能產力高的後綴，常用於「太…生」的固定語式中。[12] Tsiunn^{3} sinn^{1}「障生」和 hiunn^{3} sinn^{1}「向生」的「生」多少承襲了這個傳統。明清戲文中「障生」是「這樣」之意，具有篇章指稱的功能，有前指 (cataphora) 的用法。

(5) (生) sio^{2} ber^{7} lir^{2} u^{7} mih^{8} ke^{3} ue^{7} sing^{5}-tsiu^{7} tsi^{2} in^{1}-ien^{5}
小妹你有乜計會成就只姻緣？（14.289 順治）
小妹，你有什麼計策可以成就這段姻緣呢？

[12] 參見志村良治 (1984, 1995:303-315)、張相 (2009:133-136)。「生」潮州音做 senn^{1}。

(貼) tann^{1} tsiunn^{3}-sinn^{1}
今障生，(14.290 順治)
現在這樣吧，
lir^{2} tann^{1} tsiong^{1} sim^{1}-pak^{4} sia^{2} tsit^{8} hong^{1} sir^{1}
你今將心腹寫一封書，(14.291 順治)
你就將內心的話寫成一封信，
gua^{2} kang^{7} lir^{2} sang^{3} khir^{3} thoo^{7} gun^{2} a^{1}-niu^{5}
我共你送去度阮亞娘。(14.292 順治)
我替你把它送給我們小姐。

「障生」有時用於言外 (exophora) 當場的情境，如下所示：

(6) (旦白) tshut^{4} lai^{0} tshu^{3} to^{1} m^{7} sau^{3} si^{7} mih^{8} ue^{7} hiunn^{3} tsue^{7}
出來，厝都不掃，是乜話向多。(24.313 萬曆)
出來，屋內都不掃，什麼話那麼多。
(生唱) lir^{2} kinn^{3} gua^{2} tsiunn^{3}-sinn^{1} khi^{2} ue^{7} sau^{3} mih^{8} tshu^{3}
你見我障生，豈會掃乜厝？(24.314 萬曆)
你看我這個樣子，難道能掃什麼屋子嗎？

「障生」也有回指 (anaphora) 的指稱功能，如例 (7)。

(7) (貼) thai^{7} sio^{2}-ber^{7} tsiong^{1} sit^{8}-tsing^{5} kang^{7} gun^{2} a^{1}-niu^{5} serh^{4}
待小妹將實情共阮啞娘說 (25.216 嘉靖)
等小妹將實情跟我們小姐說
kio^{3} lin^{2} ia^{7} si^{7} ho^{2} lang^{5} kann^{2} ji^{5}
叫：恁也是好人仔兒，(25.217 嘉靖)
說：您也是體面人家的兒子，

m^{7} ui^{3} in^{1}-ien^{5} put^{4} tsiu^{7}
不畏姻緣不就。(25.218 嘉靖)
不怕姻緣不成。
(生) sio^{2}-ber^{7} tsiann^{3} si^{7} tsiunn^{3}-sinn^{1}
小妹，正是障生。(25.219 嘉靖)
小妹，正是這樣。

「向生」是「那樣」之意，具有篇章指稱的功能，通常用於回指 (anaphora)。

4.5「障說」和「向說」

Tsiunn^{3} serh^{4}「障說」是固定用語，意指「這麼說」，前頭常帶「見是」或「既然/雖是」，如下例所示，「既是障說」意指「既然這麼說」。

(8) (貼) a^{1}-niu^{5} ki^{3}-jien^{5} bo^{5} sim^{1}
啞娘既然無心，(24.016 嘉靖)
小姐既然沒有心思，
tang^{5} khir^{3} au^{7} hue^{1}-hng^{5} lai^{7} siunn^{2} hue^{1} kue^{2} bun^{5} tsit^{7} huan^{1}
同去後花園內，賞花解悶一番。(24.017 嘉靖)
一同去後花園裡，賞花解解悶。
(旦) ki^{3} si^{7} tsiunn^{3} serh^{4}
既是障說，(24.018 嘉靖)
既然這樣說，
siann^{1} kang^{7} lir^{2} khir^{3}
相共你去。(24.019 嘉靖)
和你一塊兒去。

「障說」回指前頭所敘述的情況，此固定用語也有引介以下隨即行爲的作

用，因此有承前啓後的功能。Hiunn^{3} serh^{4}「向說」也是固定用語，意指「那麼說」，前面不帶其他語詞。下例中「向說」回指淨所說的情況，然後引出貼（即益春）所建議的後續行爲，因此也有承前啓後的作用。

(9) (淨) gua^{2} tsuah^{4} si^{5} io^{1} thiann^{3} to^{1} bue^{7} tshut^{4} lai^{0}
我拙時腰痛都袂出來，（18.082 嘉靖）
我這些時候腰痛都沒法子出來，
bing^{5}-jit^{8} kah^{4} gua^{2} sai^{1}-kann^{2} lai^{5} bua^{5}
明日甲我師仔來磨。（18.083 嘉靖）
明天叫我師父來磨。
(貼) hiunn^{3} serh^{4}
向說，（18.084 嘉靖）
那麼說，
bin^{5} tuann^{3} tsa^{2} lin^{2} sai^{1}-kann^{2} pang^{3} tsa^{2} lai^{5} bua^{5}
明旦早恁師仔放早來磨。（18.085 嘉靖）
明天早上您師父提早來磨。

我們可以看出，凝結特列詞，不論「障說」或「向說」，都具有篇章的回指功能，有承先啓後的作用。說話者以「這麼說」或「那麼說」來回指對方所說的話，然後針對對方的話帶出自己因應的行爲或對策。

4.6「只夥」和「拙」、「許夥」和「喝」、「歇」

Tsuah^{4}「拙」是 tsi^{2}「只」和 ua^{7}「夥」，[13] 但是明清文獻並沒有「夥」

[13] 據梅祖麟 (2002)，閩南語「偌」dzua^{7}（即本文標示的 jua^{7}）是「若夥」的合音，如「偌久」（多麼久？）。「夥」ua^{7}（多）見於 Douglas (1873:346)，用於 bo^{5} ua^{7}「無夥」（不多）。就音義而言，「拙」意指「這些」，可以看成是「只」和「夥」的合音。現在泉州話「拙」有舒聲促化的現象，其讀音為 tsuah^{4}（李麗敏 2006:106，陳燕玲 2008）。「夥」原出現於揚雄的《方言》（見 Norman 1983）。

的痕跡，「只夥」是近指詞＋數量詞。「拙」做爲合音詞意指「這些」，但是合音之後語義逐漸轉變爲程度的加強詞，帶「這麼」之意。帶「這些」之意時，還是存在於限定詞組中，用來修飾名詞，如 ni^{5}「年」、her^{3}「歲」、tsiu^{2}「酒」、loo^{7}「路」、lan^{5} gun^{5}「零銀」、ue^{7} kann^{2}「話（仔）」、jit^{8} thau^{5}「日頭」、lai^{5} in^{1}「來因」等。

「拙」做爲副詞性程度加強詞，用於修飾程度形容詞，這類形容詞可以指狀態或屬性，如：單音節：tsue^{7} 多、uann^{3} 晏（晚）、tua^{7} 大、tim^{5} 沉、hong^{1} 慌、ku^{2} 久、oh^{4} 惡（難、慢）、tsa^{2} 早、ho^{2} 好、tng^{5} 長、ter^{2} 短、lau^{7} 老、hing^{3} 興、mua^{2} 滿；雙音節：ho^{2} thiann^{1} 好聽、ho^{2} kua^{1} 好歌、tshinn^{1} ling^{2} 青冷、khoo^{2} khui^{3} 苦氣、pinn^{3} bin^{7} 變面（翻臉）、tua^{7} han^{3} 大漢（體形大）、tshin^{1} tshinn^{2} 親淺、song^{2} li^{7} 爽利（賞心悅目）。

「拙」不出現於萬曆本荔枝記，此戲文一般判定爲反映潮州方言，可見是泉州方言的特定功能詞。與「拙」相對的是 huah^{4}「喝」，「喝」也是合音詞，由遠指詞「許」＋「夥」縮約而成。與「拙」平行，「喝」可以指數量，及「那些」之意，如 (13)，或充當程度加強詞，即「那麼」之意，如 (14)、(15)。萬曆本是潮州戲文，此文本沒有合音詞「拙」和「喝」可以提供佐證。[14]

與「喝」功能類似的是 hiah^{4}「歇」，但「歇」只有「那麼」的語意，不帶數量之意。「歇」hiah^{4} 意指「那麼」，用例最早出現於順治本，用來修飾程度形容詞或「無」＋名詞組所形成的程度詞。[15]

---

[14] 現代泉州話有「拙」tsuah^{4}，也有「喝」huah^{4}（陳燕玲 2008）也和這個現象一脈相承。至今還可在二十世紀五十年出版歌仔冊中找到「喝」的足跡，如「阿司總無化年戇」A^{1}-sai^{1} tsong^{2}-bo^{5} huah^{4}-ni^{5} gong^{7}（阿司總不會那麼笨）（見林春榮 1955）。

[15] 這個詞是「歇」hioh^{4}（棲息，歇息）的借音字。程度加強詞的 hiah^{4} 和動詞的 hioh^{4} 兩者沒有詞源上的關係。「歇」hiah^{4} 實際上就是前言所說的「赫」，兩者只是書寫上的不同。一般歌仔冊也寫作借音字「靴」。程度加強詞的 hiah^{4} 最早出現於順治本，後頭不帶相當於「爾」的「年」。帶有遠指指示成分的「喝」huah^{4} 和「歇」hiah^{4} 可能是分別代表泉腔和漳腔的特徵詞。

(10) tsai^{7} ni^{1} hiah^{4} bo^{5} hok^{4}

偆年歇無福。（17.175 順治）

怎麼那麼沒有福氣。

(11) hiah^{4} tsue^{7} khit^{4} i^{1} mih^{8} tai^{7}

歇多乞伊乜事？（28.057 光緒）

為什麼那麼多給他？

(12) lir^{2} hiah^{4} bo^{5} ki^{3} sing^{3}

恁歇無記性。（4.025 光緒）

你那麼沒有記性。

(13) gua^{2} ku^{3} khiam^{3} i^{1} huah^{4} ge^{5} tsinn^{5} niu5

我句欠伊喝個錢糧。（3.374 順治）

我卻欠他那些個錢糧。

(14) iah^{4} tat^{8} tsit^{8} ter^{3} gun^{5} huah^{4} tua^{7}

亦值一塊銀喝大（26.228 嘉靖）

也值得一塊銀元那麼大

(15) lin^{2} huah^{4} bo^{5} ki^{3} sing^{3}

恁喝無記性（5.016 道光）

您那麼沒有記性

(16) tsi^{2} nng^{7} ge^{5} a^{1}-kann^{2} hiah^{4} sui^{2} tsiann^{3} si^{7} mih^{8} lang^{5}

只二個阿仔歇水，正是乜人？（4.070 光緒）

這兩個年輕小夥子，正是什麼人？

4.7「只乜」和「乜」

Tsi^{2}「只」＋mih^{4}「乜」的構式是由近指詞和變項「乜」（＜「物」）所組成的。[16] 其轉化爲程度加強詞和上述的指示＋樣式的演變相類似，可資

[16] 「乜」充當非疑問詞時，聲調為陰入 (Douglas 1873:330)。有關「物」如何由指物變為疑

比較。「只乜」只出現於光緒本荔枝記，如以下例句所示：

(17) sann^{1} tia^{1} tsi^{2} mih^{4} tshau^{3}-kha^{2}
三爹只乜湊巧（15.035 光緒）
三爹這麼湊巧

(18) jin^{4}-but^{8} sinn^{1} tit^{4} u^{7} tsi^{2} mih^{4} tsue^{5}-tsing^{2}
人物生得有只乜齊整（16.005 光緒）
人物生得這麼整齊

(19) lir^{2} tsi^{2} mih^{4} khui^{1}-sim^{1} mih^{4} kiann^{5} phai^{2}
你只乜虧心乜行歹（23.207 光緒）
你這麼虧心，品行這麼壞

「只乜」由來已久，在變文中有「只沒」、「只麼」（比較「熠沒」）等例證（以下以「只沒」爲代表）。「只沒」意指「這麼、如此」，[17] 那時期「沒」也可以單獨出現，做「這麼」解釋。[18] Mih^{4}「乜」相當於變文中的「沒」或「麼」，「乜」做「這麼」解，可以理解爲省略了近指詞「只」，例子如下：[19]

(20) tsit^{8} ui^{7} niu^{5}-kann^{2} tsi^{2} mih^{4} tshin^{1}-tshinn^{2}
一位娘仔乜親淺。（8.010 嘉靖）
一位姑娘這麼漂亮。

問變項，可參見周法高 (1972:191-192)、志村良治 (1970, 1984, 1985)、太田辰夫 (1988: 124-141)。

17 參見蔣禮鴻 (1997:515-516)。

18 參見梅祖麟 (1983)、張相 (2009:313-315)。

19 近代漢語「麼」（源自「物」，相當於閩南語的「乜」）單獨可充當程度加強詞，帶「這麼」、「那麼」的語義（張相 2009:310-311）。有關明清戲文中「乜」充當程度加強詞的深入分析參見趙靜雅、連金發 (2009)。

(21) lim^{5} tshu^{0} kuann^{1}-lang^{5} sinn^{1} tit^{4} mih^{4} tshin^{1}-tshinn^{2}
林厝官人生得乜親淺。（5.272 順治）
林家少爺長得這麼英俊。

5. 葉氏循環現象

「只乜」中「只」的語意融入「乜」而自身略去，這種現象很像法語的「ne…pas」，否定義由 ne 表示，然後把否定傳染給 pas，使 pas 帶否定義，ne 成爲可以省略的成分，終究略去，這也反映葉氏循環 (Jespersen's Cycle) 現象。

從「遠/近指＋樣式詞」的構式形成合音詞，合音詞再加上另外的樣式詞，可以看出語言演化新陳代謝的現象，這樣的演化過程具體複述如下：

「只樣」合音爲「障」，合音詞「障」可以加上另外的樣式詞，如「般」、「生」、「年」，「許樣」合音爲「向」，合音詞「向」也可以加上上列的樣式詞，因此有「障般」、「障生」、「障年」，也有「向般」、「向生」、「向年」這類新起複合詞。複合詞融合成合音詞，合音詞因其中樣式詞的樣式義減弱，再加上別的樣式詞，構成一個循環，[20] 這種循環與葉氏循環類似。

[20] 葉氏循環現象 Jespersen's Cycle 是指英語否定詞的新陳代謝語法演變現象。葉氏循環論也可以用來說明法語否定詞的演變。法語否定詞的演變可以分成三個階段：第一階段否定義由動詞前的否定詞表示，第二階段否定義由動前和動後的兩個不相鄰的成分表示，第三階段動前成分消失，單由動後成分表示否定。有關葉氏循環現象參見 Jespersen (1917)，語法演變的葉氏循環現象自然不只限於否定詞的演變。近年來語法演變中葉氏循環現象逐漸受到學者的矚目。參看 Dahl (1979)，Lucas (2007)，及 Gelderen (2009)。第三項是論文集，其中一系列論文以葉氏循環現象為本展開各項理論探索。就如 Hopper & Traugott (2003:20-21) 所指陳的，十九世紀末葉德國語言學者暨漢學家 Gabelentz（甲柏連孜）就注意到語言中的新陳代謝現象，但是他傾向認為是漩渦式 (spiral) 的演變 (Gabelentz 1891:241-242)。漢語方言詞法也有雷同的葉氏循環現象，參見曹逢甫 (2006)。

6. 遠/近指示＋樣式構式演變總趨勢

從下表可以看出明清時期五個荔鏡/荔枝記的版本中，遠指和近指＋樣式構式的不對稱性。[21]「只樣」有例證，「許樣」沒有例證，近指合音詞「障」比遠指合音詞「向」數量多，[22]「障般」也比「向般」多，唯一例外的是萬曆本：遠指的「向生」比近指的「障生」多出很多。[23]「障年」只出現於萬曆本，「向年」亦然。這可能是萬曆本是記錄潮州話使然。

	只樣	障	障般	障般樣	障生	障年
嘉靖	8(100%)	80(61%)	22(92%)	1(50%)	4(40%)	0(0%)
萬曆	3(100%)	35(49%)	22(100%)	0(0%)	2(6%)	6(20%)
順治	14(100%)	99(61%)	12(100%)	4(80%)	6(55%)	0(0%)
道光	5(100%)	103(69%)	9(100%)	3(100%)	10(83%)	0(0%)
光緒	8(100%)	102(69%)	10(100%)	2(100%)	9(90%)	0(0%)
總計	38(100%)	419(63%)	75(97%)	10(83%)	31(41%)	6(17%)
	許樣	向	向般	向般樣	向生	向年
嘉靖	0(0%)	51(39%)	2(8%)	1(50%)	6(60%)	0(0%)
萬曆	0(0%)	36(51%)	0(0%)	0(0%)	30(94%)	24(80%)
順治	0(0%)	64(39%)	0(0%)	1(20%)	5(45%)	0(0%)
道光	0(0%)	47(31%)	0(0%)	0(0%)	2(17%)	2(100%)
光緒	0(0%)	45(31%)	0(0%)	0(0%)	1(10%)	3(100%)
總計	0(0%)	243(37%)	2(3%)	2(17%)	44(59%)	29(83%)

[21] 本表列出嘉靖、萬曆、順治、道光、光緒本中遠/近指＋樣式語式的出現次數。上半部是近指＋樣式的出現頻率，下半部是遠指＋樣式的出現頻率。括弧中的百分比表示每個文本中每對遠/近指示樣式構式彼此所佔的比例。

[22] 「障」、「向」的數據實際上包含不帶「-般」、「-樣」、「-生」、「-年」的所有例子，因此也含有像「障說」、「向說」這樣高頻率的雙音節詞。

[23] 「障生」、「向生」在現代的潮州話次方言汕頭方言還保存著（施其生 1996:247-249）。

如下表所示，「拙」可能是「只」和「夥」的合音詞，而「喝」可能是「許」和「夥」的合音詞。雖然合音詞「歇」很可能是由遠指詞「許」和另一個語詞組合而成的，但是那個語詞究竟是甚麼有待確認。「拙」不見於萬曆本，這表示「拙」、「喝」/「歇」都是泉州方言或非潮州系統閩南語的特有語詞。

	拙	拙年	喝	歇
嘉靖	43	5	1	0
萬曆	0	0	0	0
順治	64	1	0	6
道光	55	1	6	1
光緒	53	1	0	4
總	215	8	7	11

7. 結語

本文主旨在探討明清閩南語戲文荔鏡記/荔枝記中「遠/近指示詞＋樣式類別詞」語法化現象。「只近指」＋「樣樣式」、「許遠指」＋「樣樣式」分別形成合音詞「障」和「向」，這兩個構式原來都帶指示樣式的語意，屬於定語詞組的內部成分，在某些語境中產生與程度詞連用的現象，促使它由樣式詞轉爲程度加強詞。語意的改變也導致形式的改變，由原始構式變爲合音詞，合音詞雖由原始詞縮約而來，但兩者並不完全等同。樣式意在合音詞中可能弱化，甚至耗損掉，留下空缺由其他類義詞彌補。文本中有「障般」、「障生」、「障年」及「向般」、「向生」、「向年」的例證，「般」、「生」、「年」都是「樣」的類義詞。從指示樣式到程度加強，再回歸到指示樣式，這就形成了葉氏循環，反映語言生生不息的新陳代謝現象。「障生」、「障年」及「向生」、「向年」另發展出篇章標記的功能。

「指示＋樣式」構式外，本文還討論了「指示」＋「數量」構式，即「只」＋「夥」,「夥」是不定數量詞,「只夥」是「這些」之意,「只夥」縮約之後形成合音詞「拙」,「拙」可以和指稱「時」、「地」、「人」、「事」名詞詞組連用，也可以和程度形容詞連用。第二種用法「拙」已不帶原有的數量義。與「拙」對稱的是「喝」，具有「那些」或「那麼」之意，前意例證較少。

語言是變動不居的事物，其變化有條理可循，但是因各種動因起作用，總有參差不齊的現象。閩南語指示 (deixis) 的語法範疇，其價值 (value) 和屬性的選取是二元的，即遠指和近指，詞彙化爲「許」和「只」，和「樣式詞」結合成指示樣式構式，但是在明清戲文中只見「只樣」，不見「許樣」，而合音詞「障」、「向」都出現，因此遠/近指的原始構式演化速度並不同步。近指的「只樣」、「障」並存，而遠指只有「向」，表示遠指樣式構式都已轉化成程度加強詞。

「指示＋數量」的構式可能形成得更早，明清戲文中找不到「指示＋數量」的原始構式「只夥」，這可能表示「只夥」到「拙」的變化已經完成。此外，有「只夥」的合音詞，也有「許夥」的合音詞[24]「喝」或「歇」，不過數量極少。「喝」和「歇」可能是分別代表泉腔和漳腔的特徵詞。

[24] 現代泉州話倒有「許夥」的合音詞「喝」hua̍ʔ^{4}（陳燕玲 2008）。

引用文獻

Chao, Chingya Anne (趙靜雅), and Chinfa Lien (連金發). 2009. Zai 'Li Jing Ji' de gantan yu yiwen zhijian 在《荔鏡記》的感嘆與疑問之間 [Between exclamatives and interrogatives in *Li Jing Ji*]. *Bulletin of Chinese Linguistics* 3.2:45-68.

Chen, Li-Hsueh (陳麗雪). 2009. Shiliu shiji Minnanyu zhishici de yufahua xianxiang 十六世紀閩南語指示詞的語法化現象 [Grammaticalization of demonstratives in sixteenth century Southern Min]. *Chinese Studies* 27.4:179-195.

Chen, Yanling (陳燕玲). 2008. Quanzhou fangyan jibenxing zhishi daici fenlei yu shiyong chutan 泉州方言基本形指示代詞分類與使用初探 [Classification and usuage of demonstrative pronouns with basic form of Quanzhou dialect]. *Quanzhou Shifan Xueyuan Xuebao* 泉州師範學院學報 [*Journal of Quanzhou Normal University*] 2008.3:19-23.

Chou, Fa-kao (周法高). 1972. *Zhongguo Gudai Yufa: Chengdai Pian* 中國古代語法：稱代編 [*A Historical Grammar of Ancient Chinese: Substitution*]. Taipei: Tailian Guofeng Press.

Dahl, Östen. 1979. Typology of sentence negation. *Linguistics* 17.1-2:79-106.

Dai, Zhaoming (戴昭銘). 2004. Ruohua, cuhua, xuhua he yufahua: Wu fangyan zhong yizhong zhongyao de yanbian xianxiang 弱化、促化、虛化和語法化：吳方言中一種重要的演變現象 [On reduction, glottalization, semantic shift, and grammaticalization: an important evolution in Wu dialect]. *Hanyu Xuebao* 漢語學報 [*Chinese Linguistics*] 8:26-34.

Douglas, Rev. Carstairs. 1873. *Chinese-English Dictionary of the Vernacular or Spoken Language of Amoy with the Principal Variations of the Chang-chew and Chin-chew Dialects*. London: Trübner and Co.

Gabelentz, George von der (甲柏連孜). 1891. *Die Sprachwissenschaft: ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel.

Gelderen, Elly van. (ed.) 2009. *Cyclical Change*. Amsterdam & Philadelphia: John Benjamins.

Giusti, Giuliana. 1997. The categorical status of determiners. *The New Comparative*

Syntax, ed. by Liliane M. V. Haegeman, 95-123. London: Longman.

Hopper, Paul J., and Elizabeth Closs Traugott. 2003. *Grammaticalization* (2nd edition). Cambridge & New York: Cambridge University Press.

Huang, Dinghua (黃丁華). 1961. Minnan fangyan li de zhishi daici 閩南方言裡的指示代詞 [The demonstrative pronouns in Southern Min dialects]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1961.12:23-29.

Huang, Han-Chun (黃漢君), and Chinfa Lien (連金發). 2007. Wanliben Li Zhi Ji zhishici yanjiu 萬曆本荔枝記指示詞研究 [Demonstratives in the Wanli version of *Li Zhi Ji*]. *Tsing Hua Journal of Chinese Studies*, New Series 37.2:561-577.

Jespersen, Otto. 1917. *Negation in English and Other Languages*. København: A. F. Høst & Søn.

Jiang, Lihong (蔣禮鴻). 1997. *Dunhuang Bianwen Ziyi Tongshi* 敦煌變文字義通釋 [*The Translation and Explanation of Dunhuang Bianwen*] (Revised edition). Shanghai: Shanghai Ancient Books Publishing House.

Li, Limin (李麗敏). 2006. *Quanzhou Fangyin Jiaocheng* 泉州方音教程 [*Course on the Quanzhou Dialect*]. Xiamen: Xiamen University Press.

Li, Rulong (李如龍). 2007. Minnan fangyan de daici 閩南方言的代詞 [The pronouns in Southern Min dialects]. *Daici* 代詞 [*Pronoun*], ed. by Rulong Li (李如龍) & Song-Hing Chang (張雙慶). 265-288. Guangzhou: Jinan University Press.

Lien, Chinfa. 1999. Place deixis in Taiwan Southern Min. *Studies on Chinese Historical Syntax and Morphology: Linguistic Essays in Honor of Mei Tsu-Lin*, ed. by Alain Peyraube & Chaofen Sun, 73-88. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale, Ecole des Hautes Etudes en Sciences Sociales.

Lin, Chunrong (林春榮). 1955. *Xinge Lintou Jie* 新歌林投姊 [*New Song: Lin Tou Jie*]. Taichung: Xingxin Press.

Liu, Hsiu-Hsueh (劉秀雪). 2005. Qionglin fangyan de zhishi daici: gouci yuyi he yuyong tanxi 瓊林方言的指示代詞：構詞語意和語用探析 [Place-deixis in the Qionglin dialect]. *Language and Linguistics* 6.1:133-152.

Lucas, Christopher. 2007. Jespersen's cycle in Arabic and Berber. *Transactions of the Philological Society* 105.3:398-431.

Mei, Tsu-Lin (梅祖麟). 1983. Dunhuang Bianwen li de 'yimo' he 'ju' zi 敦煌變文裡的

“熠沒”和“舉”字 [*Yimo* and *ju* in *Dunhuang Bianwen*]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1983.1:44-50.

Mei, Tsu-Lin (梅祖麟). 2002. Jige Minyu xuci zai wenxian shang he fangyan zhong chuxian de niandai 幾個閩語虛詞在文獻上和方言中出現的年代 [The occurrences in the literatures of some function words in the Min languages]. *Dialect Variations in Chinese*, ed. by Dah-an Ho (何大安), 1-21. Taipei: Institute of Linguistics, Academia Sinica.

Norman, Jerry. 1983. Some ancient Chinese dialect words in the Min dialects. *Fangyan* 方言 [*Dialect*] 1983.3:202-211.

Ota, Tatsuo (太田辰夫). 1988. *Chūgokugoshi Tsūkō* 中国語史通考 [*Comprehensive Studies of the History of the Chinese Language*]. Tokyo: Hakuteisha.

Quanzhou Department of Culture Affairs (泉州市文化局), and Quanzhou Local Opera Study Club (泉州地方戲曲研究社). 2010. *Li Jing Ji*, *Li Zhi Ji Sizhong: Disanzhong Qingdai Daoguang Kanben Li Zhi Ji Shuying ji Jiaodingben* 荔鏡記荔枝記四種：第三種清代道光刊本《荔枝記》書影及校訂本 [*Four Playscripts of Li Jing Ji*/*Li Zhi Ji: The Third Playscript*, *the Daoguang Edition of Li Zhi Ji*]. Beijing: Chinese Theatre Book Press.

Shi, Qisheng (施其生). 1996. *Fangyan Lungao* 方言論稿 [*Papers on Dialects*]. Guangzhou: Guangdong People’s Publishing House.

Shimura, Ryoji (志村良治). 1970. ‘Renme’ kō「恁麼」考 [The investigation of the phrase ‘renme’]. *Tohogaku* 東方学 [*Eastern Studies*] 39:126-141.

Shimura, Ryoji (志村良治). 1984. *Chūgoku Chūsei Gohōshi Kenkyū* 中国中世語法史研究 [*A Study of the History of Middle Chinese Grammar*]. Tokyo: Santoushya.

Shimura, Ryoji (志村良治). 1995. *Zhongguo Zhongshi Yufashi Yanjiu* 中國中世語法史研究 [*A Study of the History of Middle Chinese Grammar*], translated by Lansheng Jiang (江藍生) & Weiguo Bai (白維國). Beijing: Zhonghua Book Company.

Todo, Akiyasu (藤堂明保). 1952. Jūko kango ni okeru shijishi no kinō 上古漢語に於ける指示詞の機能 [The function of demonstratives in Old Chinese]. *Bulletin of the Sinological Society* 4:1-18.

Tsao, Fengfu (曹逢甫). 2006. Yufahua lunhui de yanjiu: yi Hanyu biyinwei/biyin

xiaochengci wei li 語法化輪迴的研究：以漢語鼻音尾/鼻音小稱詞爲例 [On grammaticalization cycle]. *Hanyu Xuebao* 漢語學報 [*Chinese Linguistics*] 14:2-15.

Valois, Daniel. 1991. *The Internal Syntax of DP*. Los Angeles: University of California dissertation.

Wu, Shouli (吳守禮). 2001a. *Ming Jiajing Kan Li Jing Ji Xiwen Jiaoli* 明嘉靖刊荔鏡記戲文校理 [*Annotated Texts of the Romance of Li Jing Ji of Ming Jiajing Edition*]. Taipei: Ts'ung-I Workshop.

Wu, Shouli (吳守禮). 2001b. *Ming Wanli Kan Li Zhi Ji Xiwen Jiaoli* 明萬曆刊荔枝記戲文校理 [*Annotated Texts of the Romance of Li Zhi Ji of Ming Wanli Edition*]. Taipei: Ts'ung-I Workshop.

Wu, Shouli (吳守禮). 2001c. *Qing Shunzhi Kan Li Zhi Ji Xiwen Jiaoli* 清順治刊荔枝記戲文校理 [*Annotated Texts of the Romance of Li Zhi Ji of Qing Shunzhi Edition*]. Taipei: Ts'ung-I Workshop.

Wu, Shouli (吳守禮). 2001d. *Qing Guangxu Kan Li Zhi Ji Xiwen Jiaoli* 清光緒刊荔枝記戲文校理 [*Annotated Texts of the Romance of Li Zhi Ji of Qing Guangxu Edition*]. Taipei: Ts'ung-I Workshop.

Zeng, Xiantong (曾憲通). 1991. Mingben Chaozhou xiwen suo jian Chaozhou fangyan shulue 明本潮州戲文所見潮州方言述略 [A brief description of the Chaozhou dialectal expressions in Chaozhou playscripts of the Ming]. *Fangyan* 方言 [*Dialect*] 1991.1:10-29.

Zhang, Xiang (張相). 2009. *Shi*, *Ci*, *Qu Yuci Huishi* 詩詞曲語辭匯釋 [*The Collection and Analysis of Words and Phrases in Classical Chinese Poetry*]. Shanghai: Shanghai Ancient Books Publishing House.

Zheng Zhang, Shangfang (鄭張尚芳). 1995. Fangyan zhong shusheng cuhua xianxiang 方言中舒聲促化現象 [On glottalization of syllables with non-stop ending in dialects]. *Zhongguo Yuyan Xuebao* 中國語言學報 [*Journal of Chinese Linguistics*] 5:172-183.

Grammaticalization of Demonstratives in Early Southern Min Texts: From Demonstratives to Adverbs of Intensification or Discourse Markers

Chinfa Lien

National Tsing Hua University

The present paper aims at exploring the grammaticalization of distal/proximal demonstrative + kind classifier/quantifier in early Southern Min texts. Some instances of the construction featuring distal/proximal demonstrative + kind classifier, exemplified by *hur*^{2}許 / *tsi*^{2}只 + *iunn*^{7}樣 decoupled from the determiner phrases lead to the adverbial intensifiers as modifiers of gradable adjectives yielding the fusional words *hiunn*^{3}向 and *tsiunn*^{3}障. The gap left by the bleaching of kind sense in the fusional words is filled by other kind classifiers such as *puann*^{1}般, *sinn*^{1}生 and *ni*^{1}年. The change of demonstrative + kind classifier to intensifier and further the refurbishment of new kind classifier is a process of rejuvenation called Jespersen's cycle. *Hiunn*^{3}向 and *tsiunn*^{3}障 respectively combined with *sinn*^{1}生 and *ni*^{1}年 further yield discourse markers. I also tackle the formation of fusional words from the demonstrative + quantifier construction. The fusional word *tsuah*^{4}拙 stems from contracting *tsi*^{2}只 and *ua*^{7}夥. It denotes 'these' or 'so'. The first sense derives compositionally from its source, whereas the second sense is a result of grammaticalization. Examples of *huah*^{4}喝 or *hiah*^{4}歇 as a fusional form of *hur*^{2}許 + *ua*^{7}夥 are fairly scanty.

Keywords: demonstrative, intensifier, discourse marker, Southern Min, Ming/Qing playscripts, Jespersen's cycle

《語言資訊和語言類型》
第四屆國際漢學會議論文集
臺北：中央研究院，2013 年

漢語方言裡的三個指代詞：「汝」、「渠他（佢）」、「許那」——再論魚虞有別與現代方言

梅祖麟
康乃爾大學

我以前提出一個假設：現代漢語方言可以分為兩組。官話、粵語、客家話導源於南北朝的北方通語；這三個方言魚虞相混。吳語、贛語、閩語各有一個導源於南朝通語的層次，在這個層次裡魚虞有別。本文希望能證實這個假設。

第二節提出一個四個吳語（蘇州、桐廬、遂昌、慶元）兩個贛語（南城、黎川）魚虞有別層次中魚韻字的比較詞表，目的在於說明「汝」、「渠他（佢）」、「許那」這三個魚韻的指代詞連袂出現於吳語和贛語，由此可見吳語和贛語的魚虞有別層次同根生，都是導源於六世紀的南朝通語。

第三節說明廣東的客家話（梅縣、河源）魚虞相混；贛南的客家話（于都、寧都）也是魚虞相混。贛南客家話魚虞有別的層次借自贛語。

第四節討論漢語方言的兩個「佢」字，一個開口，中古音 gie2，一個合口，中古音 giu2。這節還會討論「佢」字上古音 *gjag 怎樣會分裂為中古的 gie、giu 兩音。

第五節介紹 6×6 的方陣，可以顯示現代六大方言的形成過程。

關鍵詞：魚虞有別，時間層次，比較方法，歷史方言學

1. 前言

我在〈方言本字研究的兩種方法〉(1995) 曾經說：「吳語中魚虞有別的層次，可以跟《切韻》序的幾句話聯繫起來：『又支脂魚虞，共爲不韻；先仙尤侯，俱論是切』。顏之推 (531-581)《顏氏家訓》中又曾舉例說明北人魚虞相混，南人魚虞有別。據此，我們暫且把吳語中魚虞有別的層次年代訂在南北朝」。又說：「舉例而言，上面提到的吳語的現象可以說明三個虛詞的本字。(1) 蘇州話『倷 [nɛ6]』的本字是『汝』。(2)『辣海』的『海』本字是『許 [hɛ3]』。(3) 吳語『吃仔飯』的『仔』本字是『著 [tsɿ5]』」。

寫那篇文章時，我深受李榮先生〈吳語本字舉例〉(1980) 的影響。李文 (1980:140) 說：

古聲母	見	溪	群	疑	曉
例字	鋸	去	渠	魚	虛
常熟文	tɕi5	tɕhi5	dʑi2 渠道	i2	ɕi1
常熟白	kɛ5	khɛ5	gɛ2 他	ŋɛ2	hɛ1 浮腫
溫嶺文	ky5	khy5	gy2 渠道	ny2	xy1
溫嶺白	kie5	khie5	gie2 他	ŋ̍2	he1
南昌文			tɕhy2 渠道		許 ɕy3
南昌白	kie5	tɕhie3	tɕhie2 他	ȵie5	許 he3 那、那麼

九十年代初期贛語的資料不易見到。我憑羅常培《臨川音系》(1940) 和李文提供的南昌話的資料推測北部贛語也有個魚虞有別的層次。以後在三篇文章（梅祖麟 1993, 2000, 2001）裡，我提出一個漢語方言形成過程的假設：現代漢語方言可以分爲兩組。官話方言、粵語、客家話導源於南北朝的北方通語的後身（所以它們魚虞相混），吳語、北部贛語、閩語各有一個導源於南朝通語的層次（所以在這個層次裡魚虞有別）。

北方通語、南朝通語丁邦新先生 (1995) 稱爲「鄴下切韻」、「金陵切韻」，我 (1993) 稱爲「河北方言」、「江東方言」。用的名詞雖然不同，我們都認爲 (1) 漢語在南北朝已經分裂爲南北兩大方言，《切韻》是南北通語的疊合音系（請參看周祖謨〈《切韻》的性質和它的音系基礎〉(1966)）。(2) 現代漢語六大方言可以按照中古來源分爲兩組。

現在舊話重提。第一，李榮先生 (1980) 比較兩個吳方言（溫嶺、常熟）一個贛方言（南昌）。現在吳語和贛語的資料大有增加。第二節要做一個四個吳方言（蘇州、桐廬、遂昌、慶元）兩個贛語（南城、黎川）魚虞有別層次中魚韻字的比較詞表。目的在於說明 (1) 吳語和贛語都以「汝」爲第二人稱代詞。(2)「汝」、「渠他」、「許那」這三個魚韻的指代詞連袂出現於吳語和贛語，由此可見吳語和贛語的魚虞有別層次同根生，都是導源於六世紀的南朝通語。

第二，「佢」字出現於吳、贛、客、粵四個方言。按照我們的看法，前兩個方言導源於南朝通語，後兩個方言導源於唐代的北方通語。我們希望用「佢」字在這四個方言裡的音韻演變來檢驗我們「六個方言分成兩組」的假設。

上面引的資料可以說明贛語的「佢」字和吳語的「佢」字同出一源。

佢　吳語：常熟 gɛ2　　溫嶺 gie2
　　贛語：南昌 tɕhie1

重要的是，吳語的「佢」和贛語的「佢」都是開口（不圓唇）的韻母：-ie、-ɛ。宋代山陰人今浙江紹興陸游 (1125-1210)《老學庵筆記》卷六：「吳人訛魚字，則一韻皆開口」。吳語和贛語的「佢」字讀開口，與陸游描寫的宋代吳語符合。

再去看粵語和客家話的「佢」字。

「佢」字粵語廣州話說 [khøy4]：

廣州

魚韻	渠他	渠水渠	鋸	舉	居
	khøy4	khøy2	køy5 文 kæ5 白	køy1	køy3
虞韻	瞿	句	具	區	
	khøy2	køy6	køy6	khøy1	

很清楚，廣州話魚虞相混，而且魚韻讀合口。

碰到客家話，事情就沒有這麼簡單。第一，《客贛方言調查報告》(1992:420) 記錄贛語「佢」字的讀音，如新餘 kiɛ6、宜豐 tɕiɛ6、餘干 tɕiɛ2、南城 kiɛ3。同時又記錄寧都的客家話「佢」字讀 tsie1（第二人稱代詞說 nie1）。客家話「佢」字韻母讀 -ie 僅有寧都話「佢 tsie1」字一例。這事有點蹊蹺。

第二，客家話的代表方言梅縣話「佢」字讀 [ki2]，翁源、連南、揭西等客方言「佢」字也讀 [ki2]，永定讀 [ki2]，清溪、河源讀 [khi2]。按照我們的推測，客家話「佢」字應該讀合口。這麼多的客家話讀 [ki2]、[khi2] 等音實在難解。

由於偶然的機遇，我去讀謝留文先生的《客家方言語音研究》(2003)，多年困擾我的兩個問題都得以解決。

第一，謝書 54-56 頁列舉寧都、于都、石城、南康等贛南客家話裡魚虞有別的證據，其中就有寧都話的：汝 nie1、nie3，鋸 kie5，渠他 tɕie1，豬 tɕie1 等字。

我去查顏森《黎川方言詞典》，發現贛語黎川話魚韻韻母讀 [ɛ, iɛ] 的字包括：汝 nɛ3、鋸 kɛ5、渠他 tɕiɛ3、豬 tɕiɛ1、苧 tɕhiɛ1、去 khɛ5、tɕhiɛ5。拿贛語黎川話和贛南的客家話比較可見寧都話的「汝鋸渠他豬箸苧煮薯去」等魚韻字全部借自贛語。

第二，謝書 282 頁指出于都的客家話「佢」字讀 [ku6]。受了謝書的

啓發，我去查書，發現瑞金的客家話「佢」字讀 [ku1]，而閩西南的秀篆客家話「佢」字讀 [ky2]。客家話也有合口的「佢」字！下一步可以推測梅縣「佢 ki2」字的 -i 韻母是後起的，來自更早的 *ky。

下面第三節「客家話的啓示」就是我讀謝書的心得。

第四節給開口的「佢」擬構 *gie，給合口的「佢」擬構 *giu。*gie、*giu 分別爲唐代南方通語「佢」字，唐代北方通語「佢」字的語音。這節還討論「佢」上古音 *gjag 怎樣會分裂爲中古的 gie、giu 兩音。

第五節介紹一個 6×6 的方陣，縱欄是現代六大方言：官話、粵、客、吳、贛、閩。橫行是四個指代詞和兩個否定詞。

2. 贛語以及吳語裡的魚虞有別

這節提出四個吳方言和兩個贛方言魚虞有別層次中魚韻字的比較詞表。這四個吳方言蘇州在蘇南、太湖的北岸，桐廬在浙中，遂昌、慶元在浙南處衢地區。兩個贛語方言南城、黎川在贛語的臨川片，再往西南走就到達寧都、于都、石城、瑞金等贛南的客語區。資料來源是：(1) 汪平《蘇州方言語音研究》(1996)、葉祥苓《蘇州方言志》(1988)。(2)《桐廬方言志》(1992)。(3) 南城：李如龍、張雙慶《客贛方言調查報告》(1992)，陳昌儀《贛方言概要》(1991)。(4) 顏森《黎川方言詞典》(1995)。(5)、(6) 遂昌、慶元：曹志耘、秋谷裕幸等《吳語處衢方言研究》(2000)。

〈表 1〉吳語和贛語魚虞有別比較字表

	蘇州	桐廬	黎川	南城	遂昌	慶元
豬	tsɿ1	tsɿ1	tɕiɛ1	tɕiɛ1	[ta]	[ʔdo]
苧	zɿ6	dzɿ6	tɕhiɛ1	tɕhiɛ1	dʑie4	tɕye4
煮	tsɿ3	tsɿ3	tɕiɛ3	tɕiɛ3	ie3	ie3
鼠	sɿ3	[tɕhy]	ɕiɛ3	ɕiɛ3	tɕhie3	tɕhie3

箸	zɿ6[1]	dzɿ6	tɕhiɛ6	tɕiɛ6	dʑie6	tɕye6
鋤	zɿ2	[dzu]	thɛ2	[thu]	[za]	[so]
梳	sɿ1	[su]	ɕiɛ1	ɕiɛ1	[sa]	sɤ
虛	hɛ1	[ɕy]	hɛ1	[ɕy]	—	—
許允許	hɛ3	[ɕy]	hɛ3	hɛ3 姓	xɤ3	xɤ3
許那	hɛ3	—	hɛ3	—	—	—
鋸	kɛ1	ke5	kɛ5	kiɛ5	kɤɯ5	kɤ5
去	tɕhi5	khi5	khɛ5	khiɛ5	khɤ5	khɤ5
渠他	ge2(常熟)	ji2	kɛ3	kiɛ3	gɤ2	kɤ4
汝你	nɛ6 [倷]	ne3 [你]	nɛ3 [汝]	nɛ3 [你]	ȵie4 [你]	ȵie4 [你]

以上最值得注意的是第二人稱代詞「汝」字。方括弧裡寫的是資料來源給的漢字。其中只有顏森《黎川方言詞典》26 頁給的本字完全正確：「[汝] nɛ˥你」。(1) 在 1995 年的文章裡我說：蘇州話「倷 [nɛ6]」的本字是「汝」。「汝」和「鋸、虛、許」等字都是魚韻三等字。當 -j- 介音失落時，「汝」的日母變爲泥母 [n-]，韻母的演變和「虛、鋸」等一樣，變成 [ɛ]。(2) 桐廬第二人稱代詞 ne3，韻母和「鋸 [ke5]」一樣，所以本字也是「汝」。(3) 南城第二人稱代詞 nɛ3 和黎川「汝 [nɛ3]」一樣，所以本字也是「汝」。[2] (4) 遂昌、慶元第二人稱代詞 ȵie4 的韻母和本方言魚韻字「鼠 tɕhie」、「煮 ie3」的韻母一樣，都是 [-ie]，所以 ȵie4 的本字也是「汝」。

此外，當「你」講的嘉興 ne3、嘉善 nə6、海鹽 ne6、桐鄉 nəɤ6，[3] 本字也是「汝」。

---

[1] 乾隆《寶山縣志》載：「有諱惡字而呼為美字者。如傘諱散，呼為豎笠，箸諱帶，呼為筷子之類。」今蘇州呼箸為「筷兒 khuɛ5 ṅ」。如果蘇州話用「箸」的話，此字應讀為 zɿ6。請參看許寶華，游汝杰〈方志所見上海方言初探〉，《吳語論叢》(1988)，189 頁。

[2] 李榮〈漢語方言裡當「你」講的「爾」（上）〉(1997) 已經說明浙江桐廬 ne3、江西黎川 nɛ3、南城 ne3 都用「汝」字。

[3] 徐越《浙北杭嘉湖方言語音研究》(2007)，126 頁。

另一方面，「汝」字在某些吳、贛方言會變成聲化韻 [n̩] 或 [ŋ̩]。

趙元任《現代吳語的研究》(1928:96) 記蘇州話第二人稱代詞複數（你們）的兩個形式：nh doq「唔篤」和 neh doq「倷篤」。我在〈蘇州話的「唔篤」（你們）和漢代的「若屬」〉(2004:249) 寫道：「上面說明『唔篤』[n̩6 toʔ7] 來自『倷篤』[nɛ6 toʔ7]。因此可以說，『唔篤』、『倷篤』的本字都是『汝屬』」。

現在重抄桐廬、蘇州魚韻字的比較字表，再把第二人稱代詞複數加在後面。

	豬	煮	苧	徐	去	鋸	汝	你們
桐廬	tsɿ1	tsɿ3	dzɿ6	ʑi2	khi5	ke5	ne3	ŋ3 təʔ「五得」
蘇州	tsɿ1	tsɿ3	zɿ6	zi2	tɕhi5	kɛ1	nɛ6	n̩6 toʔ「唔篤」～nɛ6 toʔ「倷」

《桐廬方言志》62 頁 [təʔ7] 音下列「得德篤督」等字，可見桐廬的「五得」就是蘇州的「唔篤」；也就是昆山、寶山的「唔得你們」（趙元任 1928:96）。那麼「五得」、「唔篤」、「唔得」下字的本字是「屬」，上字的本字是「汝」；「汝」[nɛ] 弱化了就變成聲化韻 [n̩]～[ŋ̩]。

按照李榮先生〈漢語方言裡當「你」講的「爾」（上）〉(1997:83) 的說法，桐廬話第二人稱代詞單數用「汝」[ne3]，複數用「爾得音」[ŋ3 təʔ7]。我們覺得這種分析法不太合適。

李榮先生 (1997:83) 指出新建望城第二人稱代詞 nɛ3 的本字是「汝」。查地圖，新建就在南昌市旁邊。李榮先生 (1980) 首次說明南昌話也有個魚虞有別的層次。下面把陳昌儀《贛方言概要》提供的望城的資料寫在李文 (1980) 的資料的下面。

	鋸	去	渠他	魚	許那	對稱
常熟白	kɛ5	khɛ5	gɛ2	ŋɛ2	—	neeng 能上聲
南昌白	kie5	tɕhie3	tɕhie2	ŋ̍ie5	he3	n̩3
望城	kɛ5	—	tɕhiɛ2	ŋ̍iɛ5	hɛ3	nɛ3 汝

《漢語方言詞匯》（第二版）548 頁「你」字下：「南昌　你 ṇ˙✓˥」。本文認爲南昌 [ṇ3] 的本字是「汝」。換句話說，南昌的昨天就是望城的今天。

吳語和贛語有些方言三身代詞說「我儂」、「汝儂」、「渠儂」，例如（曹志耘等 2000:416，游汝杰 1995:46）：

	我	你	他
遂昌	①ŋa　我 ②ŋa nəŋ　我儂	①ṇie　汝 ②ṇie nəŋ　汝儂	①gɤ　渠 ②gɤ nəŋ　渠儂
開化	①ŋɔ ②ŋɔ nəŋ	①n ②n nəŋ	①ge ②ge nəŋ
湯溪	①ɦa ②ɦa noŋ	①ŋ ②ŋ noŋ	①gɯ ②gɯ noŋ
蘭溪	uːe noŋ	noŋ	gi noŋ

遂昌有兩套三身代詞，一套是「我、汝、渠」，另一套是「我儂、汝儂、渠儂」。「我儂」、「渠儂」在開化、湯溪、蘭溪維持不變，「汝儂」卻有所改變。(1) 在開化、湯溪，帶著 v 元音的「汝 *nv」字音節化而變爲 n～ŋ。(2) 在蘭溪 n-nong 變爲 nong「儂你」。

上海、浦東、寶山第二人稱單數說「儂你」nong。本文認爲「儂你」的來源是「汝儂」。

呂叔湘先生《近代漢語指代詞》51 頁引過馮夢龍《古今譚概・雜誌》第三十六：「嘉定近海處，鄉人自稱曰吾『吾儂』、『我儂』，稱他人曰『渠儂』、『你儂』，問人曰『誰儂』。夜聞有叩門者，主人問曰：『誰儂？』外客曰：『我儂。』主人不知何人，開門才識，乃曰：『卻是你儂』。後人因名其處爲三儂之地。」

嘉定就在上海、寶山附近。但是我們不要把馮夢龍的記載看得太死。馮氏記載的三身代詞是「我儂、你儂、渠儂」。清代的《寶山縣志》、《青浦縣志》記載寶山、青浦第二人稱說「你儂」，《上海縣志》、《松江縣志》

記載上海、松江說「爾儂」。[4] 本文認爲「你儂」的「你」,「爾儂」的「爾」都是訓讀字，本字是「汝」*ne。

爲什麼要如此說？因爲上海、松江、寶山附近有明確的「汝」，如蘇州 nɛ6、嘉興、海鹽 ne6、嘉善 nə6 等等。但是沒有明確的「你」，也沒有明確的「爾」。

據上所述，上海–蘇州地區的第二人稱代詞有①「汝」、②「汝儂」兩式。①式變爲蘇州 nɛ6、嘉興、海鹽 ne6。②式的變化是：

「汝儂人」*ne nong > ṅ nong > nong「儂你」

上海、寶山、嘉定、上虞、嵊縣、奉化、金華的「儂你」來自「汝儂人」。常熟、崑山當「你」講的「能 [neng]」也來自「汝儂」。

贛語都昌話人稱代詞也以「儂」爲附加成分。第一人稱單數可以說「我」，也可以說「我儂」。李如龍、張雙慶主編的《客贛方言調查報告》、盧繼芳《都昌陽峰研究報告》都報導都昌話。當「你」講的 [ṅ3 (nuŋ^0)] 前者 419 頁寫作「你（儂）」，本文照抄，但把「你」放在方括弧內。後者 208 頁寫作「□儂」，方框框表示來源不明。本文認爲 [ṅ3] 的本字是「汝」。理由前面已經說過了，這裡不贅。

	我	你	他
都昌	我(儂) ŋɔ3 (nuŋ^0)	[你] (儂) ṅ (nuŋ^0)	渠(儂) iɛ2 (nuŋ)
本字	我(儂)	汝(儂)	渠(儂)

「我們」、「你們」、「他們」在吳語慶元的說法如下：

[4] 請參看許寶華、游汝杰〈方志所見上海方言初探〉(1988)，189 頁。

	我們	你們	他們
慶元	ŋo noŋ 我儂	ŋ̍ie noŋ 汝儂	kɤ noŋ 渠儂

「我們」、「你們」、「他們」在贛語南昌、土塘的說法如下。陳昌儀《贛方言概要》324 頁把南昌 [ṇ^{213}]、土塘 [ṇ^{315}] 寫作「你」。本文認爲本字是「汝」。下面照抄陳書的寫法，但把「你」放在括弧內。

	我們	你們	他們
南昌	我東 ŋo3 tuŋ	[你] 東 ṇ3 tuŋ	渠東 tɕiɛ2 tuŋ
土塘	俺儂 nan^{35} noŋ	[你] 儂 ṇ noŋ	渠儂 ie2 noŋ

南昌「東 tuŋ」是「多儂人」的合音詞。這三個人稱複數代詞的本字在南昌是：「我多儂」、「汝多儂」、「渠多儂」；在土塘是「俺儂」、「汝儂」、「渠儂」。

做個小結：(1) 吳語和贛語都以「汝」爲第二人稱代詞。(2)「汝」、「渠他」、「許那」這三個魚韻的指代詞連袂出現於吳語和贛語，「儂人」字也出現於吳語和贛語，由此可見吳語和贛語魚虞有別的層次同根生，都是導源於六世紀的南朝通語。

還有兩個問題需要討論。

(3) 爲什麼吳語和贛語南朝的層次第二人稱代詞用「汝」而不用「爾」、「你」？周法高《中國古代語法：稱代編》(1959:70-72)、何大安〈語詞的脈絡、階級與體式：中古代詞爾、汝、卿的用法與異同〉(1993) 都曾經說明對稱在南朝文獻（《世說新語》、王羲之的尺牘）用「汝」，北朝文獻（《北齊書》、《隋書》）用「爾」、「你」。

(4) 贛語魚虞有別的層次是從什麼地方來的？什麼時候來的？贛語是怎樣形成的？

答曰：贛語有兩個時間層次。早的魚虞有別，而且人稱代詞以「儂人」爲後綴。這個層次在南北朝從蘇南浙北地區（江東方言的老家）傳入江西北部的鄱陽平原。

晚的層次魚虞相混，是從北方傳來的。至於傳入的年代，我們猜想主體是在宋代，理由有二。一、贛語否定詞用 *pət（如南昌、奉新、餘干 pət、臨川 put），官話方言也用 *pət（如太原、揚州 pəʔ）。它們的來源是北宋末年在北方興起的奔物切的「不」puət 字。(「弗」字失落 -j- 介音就會變成 puət（< pjuət「弗」))。王觀國（1119 年進士）《學林》說：「不字舉世讀爲奔物切 (puət)」。換句話說，漢語否定詞在北宋末年以前沒有 puət 字。贛語既然以 pət (< puət) 爲否定詞，可見這個語詞是南宋或更晚傳入的。

二、《南昌方言詞典》16 頁說，南昌話的語法特點之一在於：普通話的「V 得 CO」「V 不 CO」，南昌話習慣的說法是「V 得 OC」和「VO 不 C」。例如：喫得飯進，喫飯不進｜睏得覺著，睏覺不著｜叫得門開，叫門不開｜打得渠他贏，打渠不贏。

呂叔湘 (1984:14) 指出，「V 得 OC」、「VO 不 C」這類語法結構在宋代常見，如 VO 不 C：「無中，做恕不出來。（上蔡，中 5）」，「軾亦自知相公門下用軾不著。（曲洧舊聞 5.1）」。「人當放肆怠惰時，方敬，便扶策得此心起（朱語 70）」。

當然，北方話可能從中唐開始已經陸續遷入贛地，不過就可以斷代的特徵來說，宋代留下的痕跡比較顯著。

3. 客家話的啟示

謝留文《客家方言語音研究》(2003:56) 認爲「魚虞有別見於贛南和閩西的客家方言，不見於廣東的客家方言」，又在 54-56 頁列舉寧都、石城、于都、南康、長汀、寧化等贛南、閩西的客家話的資料來說明這些方言中的魚虞之別。下面轉錄寧都、于都的資料，也把瑞金的資料附在後面。

寧都	魚韻	女 nie3	蛆 tɕhie1	箸 tɕhie6	豬 tɕie1	苧 tɕhie1
		煮 tɕie3	書 ɕie1	鼠 sa3	薯 sa2 蕃薯 ɕie2 腳板薯	
		鋸 kie5	魚 nie2	漁 nie2	去 ɕie5	墟 ɕie1
		汝 nie1 nie3		渠他 tɕie1		
	魚韻	徐 ɕiu2	除 tshu2	初 tshu1	舒 su1	居 tsu1
		許 su3				
	虞韻	娶 tɕhiu3	住 tshu6	數名詞 su5	樹 su6	句 tsu5
		芋 iu6				
于都	魚韻	女 nie3	蛆 tshe1	豬 tʂe1	苧 tʂe1	煮 tʂe3
		鼠 ʂe3	薯 ʂe1	鋸 kie5		
	魚韻	徐 siu2	除 tʂhu2	初 tshiu1	舒 ʂu1	居 tɕiu1
		許 ɕiu3	渠他 ku6			
	虞韻	娶 tɕhiu3	住 tʂhu6	數名詞 su5	樹 ʂu6	句 tɕiu5
		芋 iu6				
瑞金	魚韻	女 ȵie3	蛆 sɿ1	豬 tɕie1	苧 tɕhie1	煮 tɕie3
		鼠 ɕie3	薯 ɕie2	濾 lie5		
	魚韻	魚 ȵiu2	除 tshu2	初 tshu1	書 su1	處 tshu3
		去 ɕiu5	許 ɕiu3	佢 ku1		
	虞韻	娶 tɕhiu3	遇 ȵiu5	樹 su5	雨 iu2	芋 iu5

魚虞有別也見於贛語。顏森《黎川方言詞典》17 頁「黎川方言的特點」:「魚韻有許多字韻母讀 [ɛ iɛ]，區別於虞韻，如蛆 thɛ1≠趨 thy1｜書 ɕiɛ1≠輸 ɕy1」。又「文白異讀」5 頁:「遇攝三等魚韻字，文讀韻母 [y]（莊組讀 [u]，[ɵ]），白讀韻母 [ɛ] 或 [iɛ]，如虛 hy1~心，hɛ1 蘿蔔~了｜鼠 ɕy3 子~丑牛，ɕiɛ3 老~」。下面列舉黎川和南城的魚虞韻字。

黎川	魚韻	女 niɛ3	魚 niɛ1	豬 tɕiɛ1	煮 tɕiɛ3	鼠 ɕiɛ3
		苧 tɕhiɛ1	去 tɕhiɛ5	箸 tɕhiɛ6	書 ɕiɛ1	梳 ɕiɛ1
		薯 ɕiɛ2	絮 ɕiɛ5	余 iɛ2	渠他 tɕiɛ3 (或 kɛ3)	
		蛆 thɛ1	鋤 thɛ2	鋸 kɛ5	去 khɛ5	虛 hɛ1
		許姓 hɛ3	許那 ɛ3	汝 nɛ3		
	魚韻	居 ky1	舒 ɕy1	虛 hy1	書 ɕy1	鼠 ɕy3
		初 thu1				
	虞韻	樹 ɕy6	句 ky5	取 thy3	數動詞 su3	芋 y6
南城	魚韻	女 ȵiɛ3	徐 thiɛ2	豬 tɕiɛ1	苧 tɕhiɛ1	梳 ɕiɛ1
		煮 tɕiɛ3	書 ɕiɛ1	鼠 ɕiɛ3	薯 ɕiɛ6	鋸 kiɛ5
		去 khiɛ5	魚 ȵiɛ2	余 iɛ2	汝 nɛ3	渠他 kiɛ3
	魚韻	初 thu1	鋤 thu6	除 tɕhy2	虛 ɕy1	許 ɕy3
	虞韻	數名詞 su5	樹 ɕy6	句 tɕy5	娶 tɕhy3	芋 y6
		住 tɕhy6				

比較兩個贛方言（黎川、南城）和兩個贛南的客家方言（寧都、于都），可見寧都、于都魚虞分韻的魚韻字全部來自贛語。寧都連第二人稱代詞「汝 nie1」、第三人稱「渠 tɕie1」也是來自贛語。

顏森《江西方言的分區（稿）》(1986:23) 說「贛南客家話可以分成東西兩片。東片包括興國、寧都、石城、瑞金、會昌、尋烏、安遠、定南、全南、信豐共十一縣。西片包括大余、崇義、上猶、南康、贛縣、于都共六個縣。……」然後說：「東片有人說他們自己的方言是地道的客家話，而西片受贛語影響較多，客家話的影響『好像一陣大水一樣，沖到西邊』，已成強弩之末。」

另外，根據羅香林《客家研究導論》(1992[1933]:94) 所說，江西一省，計有純客住縣：尋鄔（尋烏）、安遠、定南、龍南、虔南、信豐、南康、大庾（大余）、崇義、上猶等十縣；非純客住縣，已知的，則有贛縣、興國、于都、會昌、寧都、石城、瑞金、廣昌、永豐、萬安、遂川、

吉安、萬載、萍鄉、修水、吉水、泰和等十七縣。並說遷入的客家人「與湘贛系人們，雜錯居住，交涉繁夥」。換句話說，客家人是後來遷入的。[5] 早遷入的贛人，他們的語言和黎川、南城的贛語相仿。

這些晚遷入的客家人，在于都、石城、寧都、瑞金等地和贛系人們雜居，年深日久，于都、寧都、瑞金等客家方言就有個贛語的層次。據上所述，這個層次魚虞有別。

贛南的客家方言，剝掉贛語的底層，就可以看到客家話自己的層次。以于都話爲例，它的第三人稱代詞說「渠 ku6」（瑞金也說「渠 ku1」）。上面說過，寧都「渠他 tɕie1」是從贛語借來的，所以我們猜想于都話「渠 ku6」是客家話自己的第三人稱代詞。于都話「渠他」字怎麼會讀 [ku6] 音？爲了回答這個問題，下面比較兩個廣東的客家話（梅縣、河源）、一個閩西南的客家話（秀篆）、兩個贛南的客家話（于都、瑞金）（李如龍、張雙慶 1991，謝留文 2003）。

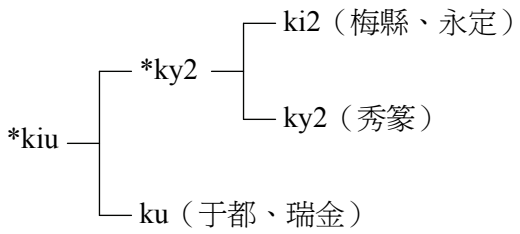
	梅縣	秀篆	河源	于都	瑞金
豬	tsu1	ʧy1	tsy1	[tɕie1]	[tɕie1]
徐	tshi2	ʃy2	tshy2	ɕiu2	—
除	tshu2	ʧy2	tshy2	tshu2	tshu2

[5] 請參看吳瑞文 (2004:171-172)：「綜合上述的兩條線索，我們可以有把握地說，于都方言事實上存在一個贛方言的底層，這個底層覃談有別。之後傳入了客家話，帶入了一個覃談不分層次，成為表層。……另外，根據羅香林《客家研究導論》(1992:94) 所說，雩都（也就是于都）非純客住縣，並說遷入的客家人『與湘贛系人們，雜錯居住，交涉繁夥』。換句話說，客家人是後來遷入的。」
還有，羅肇錦《瑞金方言》(1977:160-161)：「瑞金地處贛客之間，自然會受兩個不同方言的影響而變成一種適合於他自己環境的方言。……現在，把瑞金話在描寫中所發現的特點羅列出來，再和客語、贛語做個總比較，才能得到客觀的歸屬。……」
「從這十項統計中，瑞金和客語一致的有六項，和贛語一致的有四項，加上瑞金人自己說他們是客家人的證據，本文的結論是『瑞金話本來是客語，受了贛語的影響而發生改變，再加上借入一些官話的讀書音，形成今天的瑞金話。』」
羅肇錦先生自己說的是梅縣客語。「在一個偶然的機會裡，聽到這種像客家話又不像客語的瑞金話時，立刻引起探求到底的興趣」（頁 6）。羅肇錦先生的意見值得重視。

苧	tshu1	ʧhy6	tsy1	[tɕhie1]	[tɕie1]
初	tshɿ1	tshu1	tshuɔ1	tshu1	tshu1
煮	tsu3	ʧy3	ʧy3	[tɕie3]	[tɕie3]
箸	—	ʧhy6	tshy6	[tɕhie6]	—
薯	su1	ʃy2	sy2	[ɕie2]	[ɕie2]
住	tshu6	ʧhy6	tshy6	tʂhu6	—
樹	su6	ʃy6	sy6	ʂu6	su6
數	sɿ5	su5	su5	su5	—
娶	tshɿ3	tshy3	tshy3	tɕhiu3	tɕhiu3
魚	n2	m2	ŋy2	[nie2]	ŋiu2
鋸	ki5	ky5	ky5	[kie5]	—
許	hi3	ʃy3	hy3	ɕiu3	ɕiu3
句	ki5	ky5	ky5	tɕiu5	—
芋	vu6	vu6	jy6	iu6	iu6
雨	i3	vu3	jy6	iu3	iu2
虛	hi1	hy1	hy1	ɕiu1 墟	ɕiu1 墟
去	khi5　hi5	khy5	khy5	ɕiɤ1	ɕiu1
佢	ki2	ky2	—	ku6	ku1

原始客家話「佢」的聲母是個陽平調的 k- (< *g-)，韻母是 [-iu]。*kiu 失落 -i- 介音就變爲于都、瑞金 [ku]。*kiu2 > ky2 就變爲秀篆 [ky2]。在梅縣 ky > ki，*ky2 就變爲梅縣 [ki2]。換句話說，梅縣的 ki2 的 -i 是後起的。

客家話的「佢」



謝留文先生 (2003) 在《客家方言語音研究》說：「梅祖麟先生認爲客家話是沒有江東方言層次的方言，這個結論恐怕值得商榷。從上文第二節的討論，我們知道，『魚虞有別』和『支與脂之有別』普遍見於客家方言。」

這裡討論的是兩種客家方言。謝書 56 頁說「從分布上看，魚、虞有別見於江西和閩西的客家方言，不見於廣東的客家方言」。上面五個客家話（梅縣、秀篆、河源、于都、瑞金）的比較詞彙說明魚、虞有別不見於廣東的客家方言（如梅縣、河源），也不見於贛南客家話（如于都、瑞金）的客語層次。

另一種就是謝留文在第三節討論的客家方言，如于都、寧都、石城、瑞金等。謝先生詳細的指出魚虞有別見於這些贛南、閩西的客家話。本文認爲它們的魚虞有別應該歸給這些客方言的贛語層次，不能算作魚虞有別見於客家方言的證據。

上面的論證可以幫助我們解答若干方言史以及移民史的問題。

(1) 贛語有個魚虞有別的層次。客家話沒有，客家話是個魚虞相混的方言。這樣，就可以從中古（6-10 世紀）的層次來分辨客、贛兩大方言。

(2) 上面看到客家話魚虞相混，而且在保守的客家話（如秀篆、河源）魚虞兩韻都讀合口。這是客家人來自北方最好的證據。

(3) 以前追蹤江東方言在江西的傳播，看到魚虞有別在臨川片留下很多痕跡；臨川、崇仁、南城、黎川都有十幾個魚韻和虞韻分立的魚韻字（孫宜志 2007:179）。到了贛南，這陣大水好像完全乾涸。當時得到的解釋是：贛南是客家人的地盤，客家話魚虞相混，所以在贛南找不到魚虞有別的蹤跡。

謝留文 (2003) 的書糾正了我的錯覺。說江東方言的先民在南朝從蘇南浙北進入江西，然後逐步散播到贛中、贛南。贛南客家話裡的魚虞有別就是江東方言留下的足跡。

4. 漢語方言裡的兩個「佢」字

4.1 楔子

上面看到南朝通語魚虞有別，而且現代方言（如吳語、贛語）顯示，在南朝通語魚韻是個開口韻。又看到北方通語魚虞相混，而且現代方言（如官話、粵語、客語）顯示，在北方通語魚韻是個合口韻。

高本漢的《切韻》音把魚韻擬爲合口韻 -jwo，同時又給《切韻》擬構了一個不分南北的中古音系。這兩點都需要修正。至於高本漢《切韻》音系的其他部分，本文認爲暫且可以維持不變。

根據上面說過的線索，這節打算給唐代南朝通語的「佢」擬構 gie2，給唐代北方通語的「佢」擬構 giu2。

4.2 中古的兩個魚韻：南 -ie／北 -iu

第三人稱代詞在漢語方言一共有三個。官話方言用「他」。閩語除了閩西以外用「伊」。其他四個方言（吳、贛、粵、客）用「佢」。下面把吳語、贛語的「佢」字放在一堆，把粵語、客語（以及閩西閩語）的「佢」放在另一堆。

「佢$_1$」

吳語：常熟 ge2、黃巖 giɛ2、溫嶺 gie2、平陽 gi2、溫州 gei2、
遂昌 gɤ2、開化 ge6、常山 ŋə2。
嚴州　淳安 khɯ2、遂安 khɯ2、壽昌 kəɯ2。
徽州　休寧 khɤ3、祁門 tɕi2～tɕhi2、婺源 tɕhiɛ2。

贛語：南昌 tɕiɛ3；黎川 tɕiɛ3、kɛ3；高安 tɕiɛ1、kiɛ1；餘干 tɕhiɛ2。

「佢$_2$」

粵語：廣州 khøy4、陽江 khei2、珠海（前山）khy4、台山 khui1。

客語：梅縣 ki2、永定 ki2、秀篆 ky2、于都 ku6、瑞金 ku1。

閩西閩語：建甌 ky4、永安、將樂 ky3。

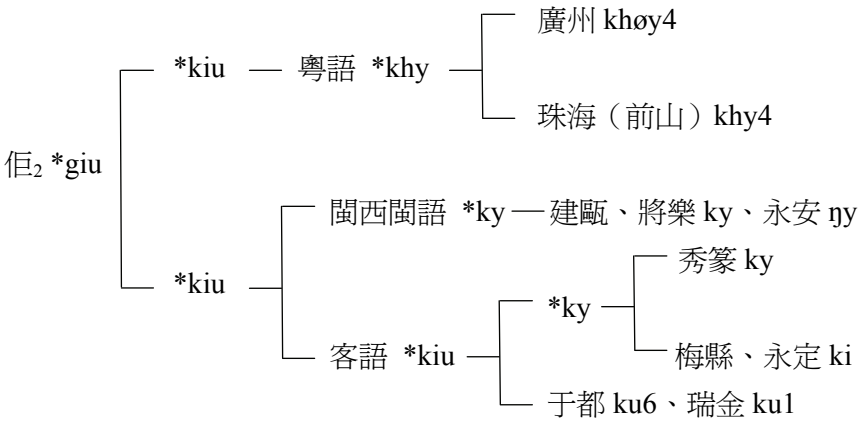
先討論合口的「佢$_2$」。

上一節看到客家話「佢」*kiu2 失落 -i- 介音就變爲于都、瑞金 [ku]，保留 -i- 介音則 *kiu 變爲秀篆 [ky2]。*ky 在梅縣、永定變爲 [ki2]。

同樣的，「佢$_2$」*kiu 變爲 *ky，就可以解釋閩西閩語建甌 ky4、永安 ŋy2、將樂 ky3的來源。

在粵話，*khy > 珠海（前山）khy3、中山（石岐）khy2、香港（新界錦田）ky1。ky > køi [køy] 是常見的演變，例如 Jerry Norman “The Proto-Min finals” (1981:37) 引的：「鋸」福州、福安 køi5；建甌、建陽、永安、將樂 ky5；「箸」福州、福安 tøi6；建甌、建陽、永安 ty5。由此可見粵語「佢」*khy > 廣州 khøy，再開口化就變爲陽江 khei2。另外，*khøy > 台山 khui1。

總之，給「佢$_2$」自擬構北方通語 *giu > *gy 的讀音，就可以解釋廣州 khøy4、秀篆 ky2、于都 ku6、梅縣 ki2、建甌 ky4 的來源。



上面的擬構假設魚虞兩韻在唐代的北方通語讀 *-iu，證據之一是客家話于都、瑞金魚虞兩韻都有讀 -iu 的字。

于都	魚韻	徐 siu2	居 tɕiu1	許 ɕiu3
	虞韻	娶 tɕhiu3	句 tɕiu5	芋 iu6
瑞金	魚韻	魚 ȵiu2	去 ɕiu5	許 ɕiu5
	虞韻	娶 tɕhiu3	遇 ȵiu5	芋 iu6　雨 iu2

虞韻中古音讀 -ju，例如：取 *tshjugx > tshju，主 *tjugx > tśju，樹 *djugh > źju，儒 *njug > ńźju，瞿 *kwjagh > kju，虞 *ngwjag > ngju，于 *gwjag > ju（李方桂《上古音研究》1980:56-60, 71-72）。

北方通語魚虞相混。既然虞韻讀 -ju，跟虞韻相混的魚韻也讀 -ju。這正是周祖謨〈唐五代的北方語音〉(1988:219) 得到的結論。「魚韻，包括魚虞模三韻和尤侯韻的唇音字。iu、u。魚虞音近 iu，變文中有與之部押韻的例子」。換句話說，于都、瑞金魚虞韻音 -iu，可以算是唐代北方通語的活標本。

現在討論開口的「佢$_1$」。

下面是六個處衢地區吳方言中魚虞有別層次中的魚韻字。

	開化	常山	玉山	遂昌	雲和	慶元
魚韻	ie / e	ie / ə	ie / ə	ie / ɤ	ɿ / i	ie / ɤ
苧	die6	də4	dəˀ$^{-22}$	dʑie4	dʒy4	tɕye4
箸	dʑie6	dʑie6	dʑie6	dʑie6	dzɿ6	tɕye4
煮	ie3	ie3	ie3	ie3	i3	ie3
書	ɕie1	ɕie1	ɕie1	ɕyɤ1	ʃy1	ɕye1
鼠	tɕhie3	tɕhie3	tɕhie3	tɕhie3	tshɿ3	tɕhie3
鋸	kie5	gə6	kə5	kɤɯ5	ʧy5	kɤ5
去	khie5	khə5	khə5	khɤ5	khi5	khɤ5
魚	ŋe2	ŋə2	ŋə2	ŋɤ2	ɲy2	ŋɤ5
許	xie3	xə3	xə3	ŋɤ2	ʃy3	xɤ3

渠他	ge6	ŋə3	ŋə1	gɤ2	gi2	kɤ4
對稱	n6 汝	n4 汝	ȵi 汝(?)	ȵie4 汝	ɲi4 汝	ȵie4 汝

開化、常山、玉山、遂昌、慶元這五個方言的魚韻可以擬爲 *ie。(i) 在開化舌根音聲母 ŋ-、g- 後面，介音 -i- 失落，因而「魚」*ŋie > ŋe、「佢」*gie > ge。(ii) 在常山、玉山、遂昌、慶元的舌根音聲母 K- 後面，介音 -i- 失落，*Kie > *Ke，然後：

e → ə / K _ （常山、玉山）
e → ɤ / K _ （遂昌、慶元）

雲和的情形不同。魚韻 *-ie 變爲 *-i。逢見系聲母，i 元音維持不變（如「佢 gi2」、「去 khi5」）。逢知章系聲母，i 元音變爲 ɿ（如「箸 dzɿ6」、「鼠 tshɿ3」）。

下面是浙中（龍游、金華）、嚴州（建德、壽昌、淳安、遂安）、徽州（歙縣、休寧）魚虞有別裡的魚韻字。

	龍游	金華	建德	壽昌	淳安	遂安	歙縣	休寧
魚韻	ɿ / əɯ	ɿ / ɤ, əʔ	ɿ / i	ɿ / əɯ	ɿ / ɯ	ɿ / ɯ	/ i	/ ɤ
苧	dzɿ2	—	tshɿ2	tshɿ4	[tɕhya]	[tɕhy]	—	[tɕy]
箸	dzɿ6	[dʑy]	—	tshɿ5/6	—	—	—	—
煮	[tɕy5]	tsɿ3	[tɕy]	tsɿ3	[tɕhya]	[tɕy]	[tɕy]	[tɕy]
薯	zɿ^{0}	—	[ɕy]	sɿ2	—	sɿ2	[ɕy]	—
鼠	tshɿ3	tshɿ3	tshɿ3	[tɕhy]	[tɕhya]	[tɕhy]	[tɕhy]	[tɕhy]
鋸	kəɯ5	kɤ5	khi5	kəɯ5	khɯ5	kɯ5	[tɕy]	[tɕy]
去	khəʔ7	khɯ^{0}	khi5	khəɯ5	khɯ5	khəɯ5	tɕhi5	khɤ5
魚	ŋəɯ2	[ȵy]	[y]	[ȵy]	[ya]	[y]	[ny]	[ȵy]
許	xəɯ3	hɤ3	[ɕy]	[ɕy]	[ɕya]	[ɕy]	[ɕy]	[ɕy]
渠他	gəɯ4	ŋəʔ	ki2	kəɯ2	khɯ2	khɯ2	ti2	khɤ2
對稱	ṅ 汝	noŋ 儂	ṇ (上)汝	ṇ 汝	ṇ 汝	i 汝	n 汝	n 汝

這八個方言的魚韻我們還是擬構爲 *-ie，在舌根音聲母 K- 後面，-ie（失落 -i- ）而變爲 -e，然後：

金華、休寧 e → ɤ / K-（金華「佢」*gɤ > gəʔ）
龍游、壽昌 e → əɯ / K-
淳安、遂安 e → ɯ / K-

在其他（非見曉系）聲母後面，*-ie > -i。逢知章系聲母，i 元音變爲 ɿ，如：「苧」龍游 dzɿ2，壽昌 tshɿ2；「鼠」tshɿ3；「薯」龍游 zɿ^{0}，壽昌、遂安 sɿ2。

建德、歙縣的情形不同。在建德、歙縣魚韻 *-ie 在所有的聲母之後都變爲 *-i，以致在建德「苧」音 tshɿ (< *tɕhi)、「佢」音 ki2、「去」音 khi5。在歙縣「去」khi5 > tɕhi5，「佢」ki2 > tɕi2 > ti2。

回去看李榮先生 (1980) 指出的吳、贛兩種方言魚虞有別層次中的對應關係。

	鋸	去	渠他	魚	虛	許那
溫嶺	kie5	khie5	gie2	ŋ̍2	he1	—
南昌	kie5	tɕhie3	tɕhie2	ȵie5	—	he3

上面說過，最值得注意的是「佢」字溫嶺音 gie2，南昌音 tɕhie2。現在把溫嶺和贛語的資料放在更大的框架中。

		蘇州	溫嶺	南城	遂昌
豬	ɿ / ie	tsɿ1	tsɿ1	tɕiɛ1	—
苧	ɿ / ie	zɿ6	dzɿ6	tɕhiɛ1	dʑie4
煮	ɿ / ie	tsɿ3	tsɿ3	tɕiɛ3	ie3
鼠	ɿ / ie	sɿ3	—	tɕiɛ3	ie3
鋤	ɿ /	zɿ2	—	[thu]	—

鋸	/ ie, e	kɛ5	kie5	kiɛ5	kɤɯ5
虛	/ ie, e	hɛ1	he1	[ɕy]	—
去	i / ie, ɤ	tɕhi5	khie5	khiɛ5	khɤ5
許	/ e, ɤ	hɛ3	—	hɛ3	xɤ3
渠他	/ ie, e	ge2 常熟	gie2	kiɛ3	gɤ2
汝	/ ie, e	nɛ6	ʔn3	nɛ3	ȵie4

(1) 南城、遂昌的資料說明魚韻字逢見曉系聲母和非見曉系聲母都讀 *-ie。

(2)「佢」字可以擬構爲 *gie2。逢見曉系聲母，*-ie（在有些方言裡）會失落 -i- 介音，而「佢」*gie > ge（常熟）；*gie > *ge > gɤ（遂昌）、khɯ（淳安、遂安）、gəɯ（龍游）、kəɯ（壽昌）等等。

(3)「汝」字可以擬構爲 *ńie。(i) *ńie > 遂昌、慶元 ȵie。(ii) 失落 -i- 介音，*ńie 變爲 *ne。*ne > ne（嘉興、海鹽、桐廬）、nɛ6（蘇州）、nɛ3（望城、南城、黎川）。

還有，李榮先生〈漢語方言裡當「你」講的「爾」（上）〉(1997:81-82) 認爲溫嶺第二人稱代詞 [ʔni] 陰上的本字是「爾」。本文認爲本字是「汝」，理由上面已經說過了，這裡不贅。

做個小結。這節給唐代北方通語的魚韻擬構 *-iu，「佢$_2$」自擬構 *giu。同時給唐代南方通語的魚韻擬構 *-ie，「佢$_1$」字擬構 *gie。

最後要用呂叔湘《近代漢語指代詞》15-16 頁的資料來說明「渠他」字中古南北兩地的文獻裡都有用例。

(一) 南方文獻裡的「渠他」字

(1) 女婿昨來，必是渠所竊。（吳志 18，趙達）

(2) 無奈人心復有憶，今暝將渠共不眠。（庾子山集 6.4）

(二) 北方文獻裡的「渠」字

(3) 即今無自在，高下任渠攀。（游仙窟 12）

(4) 今朝忽見渠姿首，不絕慇勤着心口。令人頻作許叮嚀，渠家太劇難求守。（游仙窟 15）

(5) 回頭指大男，渠是弓弩手。（杜甫，遭田父泥飲美嚴中丞）

(6) 莫掩夜窗扉，共渠相伴宿。（白居易，北窗竹石）

例 (1)-(2) 是南方文獻裡的「渠」字。它是開口的「佢」字的來源。例 (3)-(6) 是北方文獻裡的「渠」字，也就是合口的「佢$_2$」字的來源。易言之，初唐時代南北兩地指代詞的差異是：

	對稱	他稱	遠指
北方	你 < 爾	佢$_2$	那
南方	汝	佢$_1$	許

上面給南方通語的魚韻擬構 -ie，北方通語的魚韻擬構 -iu。偶讀羅杰瑞〈漢語方言通音〉(2011)，得知羅杰瑞給漢語方言通音的魚韻擬構 -ie，虞韻擬構 -iu，而且他非常重視顏森《黎川方言詞典》裡的資料。羅氏此文可謂先得吾心。

4.3 中古兩個魚韻的來源：南 -ie < -jaï／北 -ju < -jo < -jaï

下面簡單地說一下本文擬構的兩個魚韻（北 -iu／南 -ie）怎樣和漢語音韻史配合。丁邦新《魏晉音韻研究》(1975) 第 238-247 頁是「上古到中古韻部演變大勢表」，下面節錄 239 頁的魚虞支三韻。

	上古	西漢	東漢	魏晉	南北朝	中古	例字
侯	-jug	-juag	-juag	(-juaï)…-juo	虞 -juo	虞 -ju	朱府裕
魚	-jag	-jag	-jag	(-jaï)…-jo	魚 -jo 虞模 -juo	魚 -juo 虞 -ju	書許 無雨傅
歌	-jar	-jar	-jei	-jei	支 -jæi	支 -jě	倚義移
佳	-jig	支 -jieï	-jiei	-jiei	支 -jæi	支 -jě, -jiě	枝易

經過本文的修改，上面的表可以重寫如下：

	上古	兩漢	魏晉	南北朝	中古	例字
侯	-jug	-juag	(-juaï)…-juo	虞 -juo	虞 -ju	朱府裕
魚	-jag	-jag	北 (-jaï)…-jo 南 -jai	魚 -jo	魚 -ju 魚 -je	書　許 ʂu1 ɕiu3（于都） ɕie1 xie（開化）
歌	-jal	-jai	-jei	支 -jæi	支 -jé	倚義移離

應該做兩點解釋。

(1) -g > -ï > -i。李方桂〈〔貴州〕獨山話的塞音尾〉(1971:198) 說：

台語 -k 韻尾變成後高不圓唇的元音或滑音 [-ï]，跟上古漢語 *-g 尾的演變相似。上古 *-g 尾變爲 -ï，然後變爲 -i 尾。以之部的「亥 hâi < *gəg」爲例，借入台語，在仲家話 (Dioi) 讀 kaeu (=kai)，在怒語 (Lü) 讀 kai。

假設上古的 *-g 尾和 *-d 尾元音化後分別變爲 -ï 和 -i，我們不但可以說明台語 -aï < -əï < *-əg 的來源，還可以解釋爲什麼東漢詩文押韻中之 (-əï < *-əg) 微 (-əi < *-əd) 有別（請參看羅常培、周祖謨《漢魏晉南北朝韻部演變研究》（第一分冊）(1958)）。到了六世紀或稍早 -əï 和 -əi 才合流爲咍韻 -âi。

李方桂先生的說法非常重要。丁邦新 (1975:239) 已經用來說明魚部三等在北方通語的演變：-jag > 魏晉 -jaï > 南北朝魚 -jo。[6] 本文用李文的 -g > -ï > -i 來說明，在南方通語魚部 -jag > 魏晉 -jaï > -jai > 中古 -ĵe (=-ie)。

(2) 遂昌、慶元支魚兩韻都讀 -ie。

遂昌、慶元都是魚虞有別，支與脂之有別，而且支魚兩韻都讀 -ie。例如：

---

[6] 龔煌城 (2002:219)〈從原始漢藏語到上古漢語以及原始藏緬語的韻母演變〉指出漢藏比較中漢語 -k 尾對應藏緬語 -u 尾。
上古漢語「角」*kruk：PTB *kruw ‘horn’ < PST *kruk ~ *krug。龔煌城的「*-g > *-w」和李方桂的「*-g > -ï」都可以幫助我們了解為什麼魚部 *-jag > -jaï 會在北方變為合口的魚韻 -jo。

遂昌	魚 [ie]	齒盛 tie1	苧 dʑie2	箸 dʑie6	藷 dʑie2	鼠 tɕhie3
		煮 ie3	汝 ȵie4			
	支 [ie]	脾 bie2	荔 lie6	支枝梔 tɕie1	紙 tɕie3	刺 tɕhie3
		匙 dʑie2	兒 ȵie	施 ɕie1	氏 ʑie4	豉 ʑie6
		移 ie2	易 ie6			
慶元	魚 [ie]	齒 ʔdie5	藷 tɕie2	鼠 tɕhie3	煮 ie3	汝 ȵie4
	支 [ie]	籬 lie2	支枝 tɕie1	紙 tɕie3	舐 tɕie4	刺 tɕhie5
		匙 ɕie2	施 ɕie1	氏 ɕie4	豉 ɕie4	移 ie2
		易 ie6				

參看丁邦新的表，歌部三等 -ĵar（上古）> -ĵai（兩漢）> 中古支 -jě (= -ie)。本文假設魚部三等 -jag（上古）> jaï > jai（魏晉）。按照歌部兩漢音 -jai 的演變規律，魚部魏晉音也會變爲 -je (=ie)。

-ie 是支部中古的音值 (-jě)。我們擬構的魚韻 -ie，跟大家擬構的支韻的中古音 -ie 同音。由此可見遂昌、慶元等處衢方言是保存了唐代南方通語支魚兩韻的 -ie 韻母。

5. 用指代詞和否定詞來區分漢語六大方言

	自稱	對稱	他稱	遠指近指	否定（不）	沒有（無）	魚虞有別
官話	我	你	他	那	弗$_1$ (pəʔ < pət)	沒	－
粵語	我	你	佢合 *giu2	箇	m̩6（無）	毛（< 無有）	－
客語	偓	爾	佢合	箇	m̩6	毛	－
吳語	我	汝	佢開 *gie2	許　箇	弗 (fəʔ < pjuət)	m-, 沒	＋
贛語	我	汝	佢開	許　箇	弗$_1$ (pət)	毛	＋
閩語	我	汝	伊	許	m̩～ŋ̍	毛	＋

現在簡單地介紹這個方案。

(1) 第一人稱代詞

客家話第一人稱代詞用 ŋai2，本字是「我」。用「偓」這個方言字來寫，更可以顯示客家話的特徵。

(2) 第二人稱代詞

南北朝末葉，第二人稱代詞已經一分爲二。南朝用「汝」，北朝用「爾、你」。官話用「你」、粵語用「你」、客語用「爾」（請參看項夢冰《連城客家話語法研究》(1997:149-150)），都可以說明這三個方言導源於北方通語。同樣的，贛語、吳語、閩語第二人稱代詞用「汝」，都是南朝通語留下的痕跡。

(3) 第三人稱代詞

據上所述，吳語、贛語用開口的「佢 *gie2」，客家話、粵語用合口的「佢 *giu2」。所以表裡把前者寫作「佢開」，後者寫作「佢合」。

(4) 指示詞

(A)「許那」

在〈幾個閩語語法成分的時間層次〉(1995)，我曾經說明 1)「許那」字單用作爲指示詞的最早用例見於南朝樂府，如「風吹冬帘起，許時寒薄飛」（子夜歌，《清商曲辭一》），「督護初征時，儂亦惡聞許」（宋武帝〈丁督護歌〉，《清商曲辭》）。2) 閩南話的遠指詞 hit7 是「許那一」的合音詞。3) 本文曾經引徵「許那」字黎川音 hɛ3，南昌音 he3。還有，上海話、蘇州話「辣海」（在那兒）的「海 hɛ3」本字是「許」。

(B)「箇」

呂叔湘《近代漢語指代詞》(1985:243)、江藍生、曹廣順《唐五代語言詞典》(1997:137) 說明指示詞「箇」字在中古南北兩地的文獻都有用例。

(5) 否定詞

羅杰瑞 (1988:181-182) 指出漢語方言的否定詞可以分成 p-/m- 兩系。北京 pu3、蘇州 fəʔ、南昌 pət 屬於 p- 系，它們都來自「弗」pjuət 或「弗」的同源詞。南方的三大方言——粵語、客家話、閩語——的否定詞

屬於 m- 系，如廣州 m2，梅縣 m2；廈門 m6、揭陽 m4，福州 ŋ6。羅杰瑞〈建陽方言否定詞探源〉(1995) 更進一步說明，廣州「唔 m2」、廈門「伓 m6」、福州「伓 ŋ6」等來自 *m̩，而 *m̩ 來自上古漢語的「無」字 (*m(j)a)。

另一方面，《切韻》時代以及更早的漢語否定詞用方久切「不」pjə̂u3 < *pjə̂gx 字。至晚在晚唐五代，「弗」pjuət 字在北方官話替代「不」pjə̂u3 字而成爲新的否定詞，但字形還是寫作「不」。然後「弗」字隨著安史之亂以後的移民潮進入吳語區，終於產生蘇州等吳方言的 fəʔ 字。表裡寫作「弗」，就是吳語 fəʔ 的本字。

在北宋末年，「弗」pjuət 失落 -j- 介音而在北方話變爲奔物切的 puət。否定詞在官話方言說 *pət（如太原、揚州 pəʔ），在贛語說 *pət（如南昌、奉新、餘干 pət），這幾個否定詞的來源都是北宋末年興起的 puət 字。這個字是失落 -j- 介音的「弗」字，所以本文寫作「弗$_1$」。

(6) 沒有（無）

羅杰瑞 (1995:32) 說明當「沒有」講的廣州 mou2（冇）、梅縣 mo2、福州 mo2、廈門 bo2 都可以看作「無有」的合音：*ma（無）＋*wu（有）→ maw。又引錢大昕《十駕齋養新錄》卷五：「『無』又轉如『毛』。後漢書馮衍傳：『飢者毛食』注云：按衍集『毛』字作『無』。漢書功臣表序『靡有孑遺秏矣』注：孟康曰：『秏』音『毛』。師古曰：『今俗語猶謂「無」爲「秏」』，大昕按：今江西湖南方音讀『無』如『冒』即毛之去聲」。

《漢語方言詞典》475 頁「沒有（無）」條下記錄南昌「冒*有 mau6 iu3」，梅縣「無 mɔ2」、廣州「冇 mou4」、廈門「無 bo2」、福州「毛 mɔ2」。據上所述，「冒*」、「無」、「冇」的「本字」都是「毛」，「毛」是「無有」的合音詞。

「沒 muət」字在唐代北方崛起，如白居易〈浪淘沙〉「誰道小郎拋小婦，船頭一去沒回期」。周法高 (1953) 指出，muət 來自「勿 mjuət」，而否定詞「沒」在寒山拾得的詩中寫作「勿」，如「寒山深，稱我心。純白

石，勿黃金。」（寒山詩頁四八），「誰來幽谷餐仙食？獨向雲泉更勿人。」（拾得詩頁五六）。

官話方言一般用「沒」，如太原話「沒 məʔ」、揚州話「沒得 məʔ təʔ」。吳語的情形複雜。上海話用「沒 məʔ」。溫州話用「冇 nau3」，本文認爲「冇 [nau3]」是「無 m̩～n̩」和「有 jiau4」的合音詞。蘇州話 [m̩pɤʔ] 可能是「無 m̩」和「沒 mɤʔ」合成的複詞。

上面討論的「特字」可以分成三個時間層次：《切韻》以前（＝秦漢），《切韻》(601) 時代，《切韻》以後（＝唐宋）。

(1)《切韻》以前

*m「無」、*mau「毛 < 無有」。這兩個否定詞跨越閩、粵、客三大方言。按照羅杰瑞 (1988:210-212) 的說法，秦漢時代來到嶺南、閩地的漢族移民，帶來了古江南方言 (Old Southern)，其中就有 *m「無」、*mau「毛」、*ŋai「我」。閩人、客人、粵人的祖先，先後來到中國最南端的地區，就吸收了古江南方言中的 *m「無」（不）、*mau「毛」（沒有）、「偓」（我）等語詞。

(2)《切韻》時代

按照顏之推 (531-581)《顏氏家訓》的說法，漢語在《切韻》時代已經分裂爲南北兩大方言，而且他還告訴我們，南朝通語魚虞有別，支亦不與脂之相混。相反的，北方通語魚虞相混，支脂無別。

據上所述，南北兩大方言在《切韻》時代的情況是：

	自稱	對稱	他稱	遠指	近指	否定	沒有
北方通語	我	爾、你	佢合	那	箇	不（方久切）	無
南朝通語	我	汝	佢開	許	箇	不（方久切）	無（無有）

贛語、吳語承繼了南朝通語的「我、汝、佢開、許那」。北方通語指代詞的流傳比較複雜。官話方言承繼了「你、那」，粵語承繼了「你、佢合」，客

家話承繼了「爾、佢合」。

(3)《切韻》以後

否定詞吳語用「弗」(fəʔ)，是晚唐五代或更早從北方傳入江浙地區的否定詞。贛語用「弗$_1$」(pət)，是北宋末年以後從北方傳入江西的否定詞。從音韻的觀點去看「弗」pjuət 比「弗$_1$」puət 早。

上面看到在南朝通語的層次，吳語和贛語完全一樣。那麼，贛語和吳語豈不是變成雙胞胎了嗎？不是的，贛語和吳語除了共有的南朝層次以外，各有一個較晚的，從北方傳來的層次（其中魚虞相混，都讀合口；否定詞用「弗」或「弗$_1$」）。就否定詞來看，吳語的 *fət 比贛語的 *pət 早，而且要早三百年。另外，我們猜想贛語受官話的影響較深，吳語受官話的影響較淺。

粵語和客家話都是中古 (601-) 以後從北方遷來的。它們不同之點在於 (i) 客語的「爾」比粵語的「你」早，(ii) 客語的「佢 *giu」比粵語「佢 *gy」早。粵語一般認爲是晚唐五代從北方遷來的。那麼，客家話南遷的時間大概可以訂在唐代。

引用文獻

Cao, Zhiyun (曹志耘). 1996a. *Jinhua Fangyan Cidian* 金華方言詞典 [*A Dictionary of the Jinhua Dialect*]. Nanjing: Jiangsu Education Press.

Cao, Zhiyun (曹志耘). 1996b. *Yanzhou Fangyan Yanjiu* 嚴州方言研究 [*Studies in the Yanzhou Dialects*]. Tokyo: Kohbun.

Cao, Zhiyun (曹志耘), Hiroyuki Akitani (秋谷裕幸), Itsuku Ota (太田齋), and Rixin Zhao (趙日新). 2000. *Wuyu Chuqu Fangyan Yanjiu* 吳語處衢方言研究 [*A Study of the Wu Dialects of the Chuqu Region in Southern Zhejiang*]. Tokyo: Kohbun.

Chang, Kun (張琨). 1985. Lun Wuyu fangyan 論吳語方言 [The Wu dialects]. *Bulletin of the Institute of History and Philology Academia Sinica* 56.2:215-260.

Chao, Yuen Ren (趙元任). 1928. *Xiandai Wuyu de Yanjiu* 現代吳語的研究 [*Studies in the Modern Wu Dialects*]. Beijing: Tsing Hua College Research Institute.

Chen, Changyi (陳昌儀). 1991. *Gan Fangyan Gaiyao* 贛方言概要 [*A Sketch of the Gan Dialects*]. Nanchang: Jiangxi Education Press.

Chou, Fa-kao (周法高). 1953. Zhongguo yufa zhaji 中國語法札記 [Notes on the Chinese grammar]. *Bulletin of the Institute of History and Philology Academia Sinica* 24:191-281.

Chou, Fa-kao (周法高). 1959. *Zhongguo Gudai Yufa: Chengdai Pian* 中國古代語法：稱代編 [*A Historical Grammar of Ancient Chinese: Substitution*]. Taipei: Institute of History and Philology, Academia Sinica.

Du, Wenlan (杜文瀾). (ed.) 1958. *Gu Yao Yan* 古謠諺 [*Traditional Chinese Ballads and Proverbs*]. Beijing: Zhonghua Book Company.

Gong, Hwang-cherng (龔煌城). 2002. *Collected Papers on Sino-Tibetan Linguistics*. Taipei: Institute of Linguistics, Academia Sinica.

Hirata, Shoji (平田昌司). 1998. *Huizhou Fangyan Yanjiu* 徽州方言研究 [*Studies in the Huizhou Dialects*]. Tokyo: Kohbun.

Ho, Dah-an (何大安). 1993. Yuci de mailuo, jieji yu tishi: zhonggu daici ‘er’, ‘ru’, ‘qing’ de yongfa yu yitong 語詞的脈絡、階級與體勢：中古代詞爾、汝、卿的用法與異同. *Wang Shumin Xiansheng Bashi Shouqing Lunwenji* 王叔岷先生八十壽慶

論文集 [*Festschrift in Honor of Professor Wang Shumin on his 80th Birthday*], 901-909. Taipei: Da'an Press.

Huang, Xuezhen (黃雪貞). 1983. Yongding (Xiayang) fangyan cihui 永定（下洋）方言詞彙 [A glossary of the Yongding dialect]. *Fangyan* 方言 [*Dialect*] 1983.2:148-160, 1983.3:220-240, 1983.4:297-304.

Huang, Xuezhen (黃雪貞). 1995. *Meixian Fangyan Cidian* 梅縣方言詞典 [*A Dictionary of Meixian Dialect*]. Nanjing: Jiangsu Education Press.

Jiang, Lansheng (江藍生), and Guangshun Cao (曹廣順). 1997. *Tang Wudai Yuyan Cidian* 唐五代語言詞典 [*A Dictionary of Words and Expressions of the Tang Dynasty and Five Dynasties Period*]. Shanghai: Shanghai Education Press.

Karlgren, Bernhard (高本漢). 1915-1926. *Études sur la phonologie chinoise*. Upsala: K. W. Appelberg.

Karlgren, Bernhard (高本漢). 1948. *Zhongguo Yinyunxue Yanjiu* 中國音韻學研究 [*The Study of Chinese Phonology*], translated by Yuen Ren Chao (趙元任), Changpei Luo (羅常培) & Fang-Kuei Li (李方桂). Shanghai: The Commercial Press.

Li, Fang-Kuei (李方桂). 1971. The final stops in Tushan. *Bulletin of the Institute of History and Philology Academia Sinica* 43.2:195-200.

Li, Fang-Kuei (李方桂). 1980. *Shangguyin Yanjiu* 上古音研究 [*Studies on Archaic Chinese Phonology*]. Beijing: The Commercial Press.

Li, Rong (李榮). 1978. Wenling fangyan de bianyin 溫嶺方言的變音 [Tonal suppletion in the Wenling dialect]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1978.2:96-103.

Li, Rong (李榮). 1979. Wenling fangyan de liandu biandiao 溫嶺方言的連讀變調 [Tone sandhi in the Wenling dialect]. *Fangyan* 方言 [*Dialect*] 1979.1:1-29.

Li, Rong (李榮). 1980. Wuyu benzi juli 吳語本字舉例 [Notes on Wu etymology]. *Fangyan* 方言 [*Dialect*] 1980.2:137-140.

Li, Rong (李榮). 1997. Hanyu fangyan li dang 'ni' jiang de 'er' 漢語方言裡當「你」講的「爾」[The second person pronoun 'er 爾' in Chinese dialects], Part I. *Fangyan* 方言 [*Dialect*] 1997.2:81-87.

Li, Rulong (李如龍), and Song-Hing Chang (張雙慶). 1992. *Ke Gan Fangyan Diaocha*

Baogao 客贛方言調查報告 [*A Report on a Survey of the Kejia and Gan Dialects*]. Xiamen: Xiamen University Press.

Li, Xiaofan (李小凡). 1998. *Suzhou Fangyan Yufa Yanjiu* 蘇州方言語法研究 [*Studies in Suzhou Grammar*]. Beijing: Peking University Press.

Liu, Zemin (劉澤民). 2005. *Ke Gan Fangyan Lishi Cengci Yanjiu* 客贛方言歷史層次研究 [*A Study of Historical Strata in Hakka and Gan Dialects*]. Lanzhou: Gansu Nationalities Publishing House.

Lo, Hsiang-lin (羅香林). 1933. *Kejia Yanjiu Daolun* 客家研究導論 [*An Introduction to the Study of Hakkas in its Ethnic*, *Historical*, *and Cultural Aspects*]. Shanghai: Shanghai Literature and Art Publishing Group.

Lo, Seo-Gim (羅肇錦). 1977. *Ruijin Fangyan* 瑞金方言 [*The Ruijin Dialect*]. Taipei: National Taiwan Normal University MA thesis.

Luo, Changpei (羅常培). 1931. '*Qieyun*' yu yu de yinzhi ji qi suo jiu fangyan kao《切韻》魚虞的音值及其所據方言考. *Bulletin of the Institute of History and Philology Academia Sinica* 2:358-385.

Luo, Changpei (羅常培). 1940. *Linchuan Yinxi* 臨川音系 [*The Phonetics and Phonology of Lin-Ch'uan Dialect*]. Taipei: Institute of History and Philology, Academia Sinica.

Luo, Changpei (羅常培), and Zumo Zhou (周祖謨). 1958. *Han Wei Jin Nanbeichao Yunbu Yanbian Yanjiu* 漢魏晉南北朝韻部演變研究 [*A Study of the Evolution of Rhyme Categories during Han*, *Wei-Jin*, *and Nanbeichao Period*], Vol. 1. Beijing: Science Press.

Lu, Jifang (盧繼芳). 2007. *Duchang Yangfeng Fangyan Yanjiu* 都昌陽峰方言研究 [*Studies in the Duchang Yangfeng Dialect*]. Beijing: Culture and Art Publishing House.

Lü, Shuxiang (呂叔湘). 1984. *Hanyu Yufa Lunwenji* 漢語語法論文集 [*Studies on Chinese Grammar*] (revised edition). Beijing: The Commercial Press.

Lü, Shuxiang (呂叔湘). 1985. *Jindai Hanyu Zhidaici* 近代漢語指代詞 [*Pronouns and Demonstratives in Early Modern Chinese*]. Shanghai: Xuelin Press.

Mei, Tsu-Lin (梅祖麟). 1993. Nanbeichao de Jiangdong fangyan he xiandai fangyan 南北朝的江東方言和現代方言 [The Jiangdong dialect of the 6th century and

modern Chinese dialects]. Manuscript.

Mei, Tsu-Lin (梅祖麟). 1995. Fangyan benzi yanjiu de liangzhong fangfa 方言本字研究的兩種方法 [Two ways of studying etymology]. *Wuyu he Minyu de Bijiao Yanjiu* 吳語和閩語的比較研究 [*Comparative Studies among Wu and Min Dialects*], 1-14. Shanghai: Shanghai Education Press.

Mei, Tsu-Lin (梅祖麟). 2000. Zhongguo yuyanxue de chuantong he chuangxin 中國語言學的傳統和創新 [Tradition and innovation in Chinese linguistics]. *Historical and Methodological Reflections on the Development of Modern Chinese Humanities*, 475-500. Taipei: Institute of History and Philology, Academia Sinica.

Mei, Tsu-Lin (梅祖麟). 2001. Xiandai Wuyu he 'zhi, zhi, yu, yu, gong wei bu yun' 現代吳語和「支脂魚虞，共爲不韻」[The survival of two pairs of *Qieyun* distinctions in Southern Wu dialects]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 2001.1:3-15.

Mei, Tsu-Lin (梅祖麟). 2004. Suzhouhua de 'wudu' (nimen) he Handai de 'ruoshu' 蘇州話的「唔篤」(你們）和漢代的「若屬」[An etymology for [n̩˨ toʔ˩˧] 'you (plural)' in the Suzhou dialect]. *Fangyan* 方言 [*Dialect*] 2004.3:247-255.

Mei, Tsu-Lin (梅祖麟). 2011. Jiangdong fangyan de 'ru' zi (> suzhou 'ne') ji qi xiangguan wenti 江東方言的「汝」字（> 蘇州 nɛ「倷」）及其相關問題 [The second person pronoun 'ru 汝' in Jiangdong dialect and other related issues]. *Dongfang Yuyanxue* 東方語言學 [*East Linguistics*] 9:1-12.

Mei, Tsu-Lin (梅祖麟), and Hsiu-fang Yang (楊秀芳). 1995. Jige Minyu yufa chengfen de shijian cengci 幾個閩語語法成分的時間層次 [Chronological strata in the grammar of the Min dialects]. *Bulletin of the Institute of History and Philology Academia Sinica* 66.1:1-21.

Morino, Shigeo (森野繁夫), and Toshiyuki Sato (佐藤利行). 1985. *O Gishi no Shokan* 王羲之の書翰 [*Wang Xizhi's Correspondence*]. Tokyo: Daiichi Gakushusha.

Norman, Jerry (羅杰瑞). 1981. The Proto-Min finals. *Proceedings of the First International Conference on Sinology: Linguistics*, 35-74. Taipei: Academia Sinica.

Norman, Jerry (羅杰瑞). 1988. *Chinese*. Cambridge & New York: Cambridge University Press.

Norman, Jerry (羅杰瑞). 1995. *Hanyu Gaishuo* 漢語概說 [*Chinese*], translated by Huiying Zhang (張惠英). Beijing: Language and Culture Press.

Norman, Jerry (羅杰瑞). 2006. Common dialectal Chinese. *The Chinese Rime-Tables: Linguistic Philosophy and Historical-Comparative Phonology*, ed. by David Prager Branner, 233-254. Amsterdam & Philadelphia: John Benjamins.

Norman, Jerry (羅杰瑞). 2011. Hanyu fangyan tongyin 漢語方言通音 [Common dialectal Chinese], translated by R. VanNess Simmons (史皓元) & Yanhong Zhang (張艷紅). *Fangyan* 方言 [*Dialect*] 2011.2:97-116.

Ota, Tatsuo (太田辰夫). 1991. Guanyu Han'er yanyu 關於漢兒言語 [On Han'er language]. *Hanyushi Tongkao* 漢語史通考 [*Comprehensive Studies of the History of the Chinese Language*], translated by Lansheng Jiang (江藍生) & Weiguo Bai (白維國), 181-211. Chongqing: Chongqing Press.

Peking University (北京大學). (eds.) 1995. *Hanyu Fangyan Cihui* 漢語方言詞匯 [*The Vocabulary of Chinese Dialects*] (2nd edition). Beijing: Language and Culture Press.

Peking University (北京大學). (eds.) 2003. *Hanyu Fangyin Zihui* 漢語方音字匯 [*The Phonology of Chinese Dialects*] (2nd edition). Beijing: Language and Culture Press.

Sagart, Laurent. 1988. On Gan-Hakka. *Tsing Hua Journal of Chinese Studies*, New Series 18.1:141-160.

Sun, Yizhi (孫宜志). 2007. *Jiangxi Gan Fangyan Yuyin Yanjiu* 江西贛方言語音研究 [*Studies in the Phonology of Gan Dialects in Jiangxi Province*]. Beijing: Language and Culture Press.

Ting, Pang-Hsin (丁邦新). 1975. *Wei Jin Yinyun Yanjiu* 魏晉音韻研究 [*Chinese Phonology of the Wei-Chin Period: Reconstruction of the Finals as Reflected in Poetry*]. Taipei: Institute of History and Philology, Academia Sinica.

Ting, Pang-Hsin (丁邦新). 1995. Chongjian Hanyu zhonggu yinxi de yixie xiangfa 重建漢語中古音系的一些想法 [Thoughts on reconstructing middle Chinese phonology]. *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1995.6:414-419.

Tung, T'ung-ho (董同龢). 1959. Sige Min fangyan 四個閩方言 [Four Southern Min dialects]. *Bulletin of the Institute of History and Philology Academia Sinica*

30:729-1042.

Wang, Ping (汪平). 1996. *Suzhou Fangyan Yuyin Yanjiu* 蘇州方言語音研究 [*Studies in Suzhou Phonology*]. Wuhan: Huazhong University of Science & Technology Press.

Wu, Rui-wen (吳瑞文). 2004. Tan tan youbie yu xiandai fangyan 覃談有別與現代方言 [‘Tan tan youbie’ and modern dialect]. *Shengyun Luncong* 聲韻論叢 [*Chinese Phonology*] 13:147-186. Taipei: Student Book.

Xiang, Mengbing (項夢冰). 1997. *Liancheng Kejiahua Yufa Yanjiu* 連城客家話語法研究 [*Grammatical Study of Liancheng Hakka*]. Beijing: Language and Culture Press.

Xie, Liuwen (謝留文). 1998. *Yudu Fangyan Cidian* 于都方言詞典 [*A Dictionary of the Yudu Dialect*]. Nanjing: Jiangsu Education Press.

Xie, Liuwen (謝留文). 2003. *Kejia Fangyan Yuyin Yanjiu* 客家方言語音研究 [*Studies in Hakka Phonology*]. Beijing: China Social Sciences Press.

Xiong, Zhenghui (熊正輝). 1994. *Nanchang Fangyan Cidian* 南昌方言詞典 [*A Dictionary of the Nanchang Dialect*]. Nanjing: Jiangsu Education Press.

Xu, Baohua (許寶華), and Rujie You (游汝杰). 1988. Fangzhi suo jian Shanghai fangyan chutan 方志所見上海方言初探 [Preliminary studies in the Shanghai dialect as recorded in local gazetters]. *Wuyu Luncong* 吳語論叢 [*Essays on Wu Dialects*], 184-192. Shanghai: Shanghai Education Press.

Xu, Yue (徐越). 2007. *Zhebei Hangjiahu Fangyan Yuyin Yanjiu* 浙北杭嘉湖方言語音研究 [*Studies in the Phonology of Hangjiahu Dialect*]. Beijing: China Social Sciences Press.

Yan, Sen (顏森). 1986. Jiangxi fangyan de fenqu (gao) 江西方言的分區（稿）[The grouping of Jiangxi dialects]. *Fangyan* 方言 [*Dialect*] 1986.1:19-38.

Yan, Sen (顏森). 1995. *Lichuan Fangyan Cidian* 黎川方言詞典 [*A Dictionary of Lichuan Dialect*]. Nanjing: Jiangsu Education Press.

Ye, Xiangling (葉祥苓). 1988. *Suzhou Fangyanzhi* 蘇州方言志 [*A Survey of Suzhou Dialect*]. Nanjing: Jiangsu Education Press.

You, Rujie (游汝杰). 1995. Wuyu li de rencheng daici 吳語裡的人稱代詞 [Personal pronouns in Wu dialect]. *Wuyu yu Minyu de Bijiao Yanjiu* 吳語與閩語的比較研究 [*Comparative Studies among Wu and Min Dialects*], 32-49. Shanghai: Shanghai

Education Press.

Zhan, Bohui (詹伯慧), and Yat-shing Cheung (張日昇). 1988. *Zhujiang Sanjiaozhou Fangyan Cihui Duizhao* 珠江三角洲方言詞匯對照 [*Comparison of the Lexicon in the Dialects of the Pearl River Delta*]. Guangzhou: Guangdong People's Publishing House.

Zhan, Bohui (詹伯慧), and Yat-shing Cheung (張日昇). 1994. *Yuebei Shi Xianshi Yue Fangyan Diaocha Baogao* 粵北十縣市粵方言調查報告 [*A Survey of Yue Dialects in North Guangdong*]. Guangzhou: Ji'nan University Press.

Tonglu Gazetteer Editorial Committee, and Beijing Normal University. 1992. *Tonglu Fangyanzhi* 桐廬方言志 [*A Survey of Tonglu Dialect*]. Beijing: Language and Culture Press.

Zhou, Zumo (周祖謨). 1966. *Wen Xue Ji* 問學集. Beijing: Zhonghua Book Company.

Zhou, Zumo (周祖謨). 1988. *Zhou Zumo Yuyan Wenshi Lunji* 周祖謨語言文史論集 [*Selected Essays of Zhou Zumo*]. Hangzhou: Zhejiang Ancient Books Publishing House.

The 魚/虞 Distinction and Modern Chinese Dialects

Tsu-Lin Mei

Cornell University

In Mei (1994, 2000), I proposed a hypothesis to the effect that modern Chinese dialects can be classified into two groups, those with a Jiangdong (江東) stratum and those without. Mandarin, Hakka and Yue are descendents of the Hebei dialect of Nan-Bei Chao, and in these dialects, the rhyme categories 魚 and 虞 have merged. Wu, Gan and Min each preserved a Jiangdong stratum, and in this stratum, the rhyme categories 魚 and 虞 are kept distinct. The purpose of this paper is to substantiate this hypothesis.

The second section presents data from two Gan dialects and four Wu dialects and shows that Gan and Wu dialects share three distinctive features: 1) The rhyme categories 魚 and 虞 are kept distinct. 2) The second person pronoun is etymologically 汝 'you'. 3) 儂 'person, human being' occurs as a suffix in personal pronouns in Wu, Gan (and Min) dialects, e.g. 我儂, 汝濃, 渠儂.

The third section shows, in two steps, that the rhyme categories 魚 and 虞 have merged in Hakka. First, in Hakka dialects in Guangdong, 魚 and 虞 are no longer distinct. Second, in Hakka dialects in southern Jiangxi (贛南) such as 于都 and 寧都, there is a stratum in which 魚 and 虞 are kept distinct, but this stratum was borrowed from the Gan dialects in southern Jiangxi.

The fourth section shows that there are two 3rd person pronouns 渠 in Chinese dialects, one with unrounded vowel, Middle Chinese *gie2*, the other with rounded vowel, MC *giu2*. This section also tries to explain how 渠 'he, she, it' Old Chinese *gjag can split into MC *gie2* and MC *giu2*.

The fifth section introduces a 6 x 6 matrix which shows how the major Chinese dialects were formed.

Keywords: the 魚/虞 distinction, chronological strata, comparative method, Chinese historical dialectology

Human Language Resources and Linguistic Typology
Papers from the Fourth International Conference on Sinology
Taipei: Academia Sinica, 2013

Human Language Resources:
Their Role in Research, Development and Application

Dafydd Gibbon
Universität Bielefeld and COCOSDA

A key issue in the language sciences and technologies is the provision of an adequate clear and computer accessible empirical foundation for research, description and development, with suitable high quality corpora, corpus analyses, lexica, grammars and software tools. Documentary Linguistics has emerged as a new branch of applied linguistics in the past 15 years, concerned with language documentation and maintenance, especially for endangered languages. In the Human Language Technologies, 'Human Language Resources' for text, speech and gesture are the basis for applications such as internet information dissemination and retrieval, machine translation, automatic speech recognition and synthesis.

The present contribution surveys the motivations for developing Human Language Resources gathered on a 'fair play' basis, argues for the need for data and corpus analyses based on a comprehensive semiotic architecture, and finally provides a linguistic perspective on practical ways in which to use computational methods to use Human Language Resources, ranging from syllable grammars and the analysis of lexical tone to lexicography, the documentation of legacy text manuscripts and to the computational modelling of tone sandhi.

Keywords: documentary linguistics, human language technology, corpus linguistics, lexicography, phonetics

1. The empirical base of language sciences and technologies

1.1 Human Language Resources and Language Documentation

Our word processors, internet software and mobile phones would not work without *Human Language Resources* (*HLR*). The entire World-Wide Web is built on natural text and markup text resources, with other media embedded into the text. And our linguistic models and theories, lexicons and grammars would not be possible without one form or another of *Language Documentation* (*LD*). These two concepts, *HLR* and *LD*, are technological and linguistic sides of the same coin: they express programmes for collating authentic text and speech data, lexicons and grammars, and the methods, formats and tools which are needed for collecting and processing the text and speech data. With this background in mind, the present study falls into two sections: first, a discussion of speech and text resources in the context of scientific and technological development; second, a series of case studies based on projects which the author has conducted.

All the disciplines which deal with language — whether literary or linguistic, whether Sinology or English Studies, whether machine translation or internet search — are to greater or lesser extents dependent on *HLR* and *LD*. In this overview I will concentrate more on the technical side, *HLR*, since *LD* is increasingly adopting the technical methodologies of *HLR*, but I will often combine references to the two as *HLR*/*LD*. This contribution to the *Sinology Congress 2012* is not conceived primarily as a classic academic study of a particular empirical or formal problem, but as a strategic overview and as an information source for the methodologies required in modern empirically founded research in the language sciences and technologies.

Language and speech data are, fortunately, ubiquitous: they are found in libraries, bookshops, CD stores, lexicons, grammars, word processors, web sites,

tweets, blogs, novels, dramas and poems, and of course by direct interviewing of language users. This wide range of data sources is required on the one hand for corpus-based theoretical and descriptive linguistics and for the speech and text technologies such as information retrieval, speech interfaces for computers, and machine translation. But on the other hand these data sources are required for many other sciences. Indeed the 'close reading' study of literary texts is actually far more corpus-based than most varieties of linguistics. Theology and jurisprudence, like many other disciplines, are also dependent on large collections of texts for their work, and the media industry is dependent on language and speech technologies for accessing large collections of speech and text, as well as video media. The need for extensive speech and text data from sources such as those listed above originated in two fields. The first source, concerned with *LD*, is *Documentary Linguistics*, which deals with the documentation and description of the languages of the world, particularly endangered languages, for scientific reasons as well as for the documentation of language and culture heritage. This avenue of text and speech data collation and processing is the *LD* direction.

The second source, *HLR*, emerged mainly from the *Human Language Technologies*, i.e. speech and text technology, which require large quantities of data for statistical analysis in order to create working models for creating text products such as large dictionaries, and speech products such as dictation and reading software. Gradually theoretical and descriptive directions in linguistics are making use of high quality *HLR*/*LD* information from these two sources in order to underpin the traditional linguistic methodology of deduction of hypotheses and testing with intuitively produced or observed opportunistic data.

In a nutshell: modern methodologies in the language sciences, both in general and theoretical linguistics, as well as the speech and text technologies, require extensive data in order to induce and validate their working models and

explanatory theories. The requirement for extensive data — 'big data' — marks a shift from traditional intuition based and deductive methodologies in linguistics to quantitative methods and inductive 'close-reading' corpus-based methodologies.

The field of *HLR*/*LD* has emerged as a general methodological field. Therefore discussion need not be restricted to specific languages or language families, whether Sino-Tibetan languages, Turkic languages, Indo-European languages, Niger-Congo languages or any others. On the contrary: scholars working in these different language domains can profit from and share each other's results. I will accordingly discuss specific *HLR*/*LD* examples from different sources and with different aims, in which I have been involved:

Language structure: the visualization of the complexity of syllable structure constraints in Mandarin.

Speech: the phonetic modelling of spoken language for technological purposes, with reference to the Tibeto-Burman language *Kuki-Thadou* (ISO 639-3 *tcz*), in cooperation with a linguist who is a native speaker.

Text: the documentation of the heritage of a short-lived variety of written language, the Nigerian 'spirit language' isolate *Medefaidrin* (no ISO 639-3 classification) in cooperation with a local linguist and the local community.

Lexicography: the provision of *CESAF* tools for automatizing the acquisition and processing of lexicon data from language corpora, here the Eastern Turkic language *Uyghur* (ISO 639-3 *uig*), for use in human language technologies.

After this I will outline specific technical aspects of creating resources for selected language and speech technologies, lexicon creation and speech synthesis and conclude with brief indications of institutions involved in text and speech resource creation. Finally I offer a commented list of books for further reading, with recommendations for internet search.

1.2 The inductive turn: the need for data

The concept — and the problem — of *Natural Resources* has become one of the permanent issues in politics, media, business and science. The concept of *Human Resources* is one which is at the centre of any work programme, whether a national economy, a business, or a scientific project. The concept of *Financial Resources* is one which concerns most people most of the time. But *Human Language Resources* and *Language Documentation* represent concepts which are perhaps unfamiliar.

The term *Human Language Resources* originates in the more technical areas of speech technology, text technology and computational corpus linguistics, the *Human Language Technologies*, while *Language Documentation* originates in the more traditional area of field linguistics and the linguistic description of the languages of the world. As already noted, the domain of these terms covers the data and methods which we need in order to do our job as empirical language scientists and technologists, whether Sinologists, Anglicists, literary scholars, linguists or engineers in the speech and language technologies.

The thesis underlying these twin methodological disciplines is that the need for *HLR*/*LD* is universal in respect to the empirical description and modelling of any language in any of the language disciplines. A definition of the domain of these two terms will be a useful reference point. I propose the following definition as being very close to consensual in the field:

> Both *Human Language Resources* (*HLR*) and *Language Documentation* (*LD*) provide the foundation for empirical and theoretical language studies and technological applications, while differing in focussing on applications versus descriptions, and consist of

(a) *corpora* of text and speech data,
(b) *tools* for efficiently processing, storing and communicating these corpora,
(c) *information* such as transcriptions, glosses, lexicons and grammars derived from the corpora by means of tools.

The *HLR*/*LD* domains have arisen because of increasing awareness not only that there is a need for more inductive approaches to language description and modelling, but also that text and speech, and the principles underlying them, are too varied and too complex for the individual human mind to capture all their subtleties of form in the form of deductive rules. This 'inductive turn' in the study of speech and text assigns generalizations such as 'principles', 'parameters', 'rules' a secondary status: they are output of a wide range of inductive methodologies whose input is extensive, authentic, observed data and whose methodology is to some extent rule-based but largely quantitative and statistical.

It has turned out, for example, that automatic translation based on statistical evaluation of very large quantities of parallel or comparable corpora and of existing human translations provides a more effective empirical basis than collections of translation rules, however detailed. Naturally, machine translation goes wrong, as anyone can test using the translation machines on the internet, because adequate contexts are missing, but then human translation also goes wrong when an adequate context is missing. Anyone can translate simple texts and spoken utterances, but not everyone can translate technical texts or literature, first, because we lack the contexts for wide ranges of language activities, and second, because the world is categorized differently in the vocabulary and the grammar of different languages.

Another area where extensive *HLR*/*LD* are required is automatic information retrieval and language understanding, closely related to machine translation: the more data for training the system, the better the result. I do not know if anyone has quantified the amount of data to which a child is exposed in the first seven years of life, by which time his command of the language has more or less stabilized — it may well be of the same order of magnitude as that which would be required for the development and statistical training of modern speech and text technology systems.

In speech technology it is impossible to capture all the subtle variations in individual utterances by means of rules, and so for automatic speech recognition, as used in dictation software and in menu control, the systems have to be trained with statistical methods, using very large amounts of data. The same applies to speech synthesis, which is used in many contexts from public announcements to satellite navigation software in vehicles and to reading software for people with sight impairments.

As already indicated in the introduction, *HLR*/*LD* apply to any discipline in which language studies are foremost, such as Sinology, English Studies, Comparative Literature Studies, Oral Literature Studies, Cultural Studies, Linguistics, Phonetics, Speech Technology. Indeed, as we see in this Congress, there is increasing cooperation between disciplines such as these in the area of *HLR*/*LD*, as well as increasing cooperation between scholars working on entirely different languages, with the common goal of providing high quality empirical foundations for their work.

1.3 Deployment zones for *HLR*/*LD*

We find *HLR*/*LD* information in many places, though they may not be

traditionally referred to in such terms. Libraries, for example, embody *HLR*/*LD*: they contain corpora of data in the form of texts and recordings; they provide generic conceptual tools such as cataloguing systems for describing and for searching these data; they provide specific 'metadata' information in the form of catalogue entries. The Internet has emerged as the largest and most easily accessible source of *HLR*/*LD* which has ever been created, though a wild one, for which methods of taming are constantly being developed.

From the point of view of the literary scholar, perhaps also of the philosopher, perhaps it sounds a little odd to characterize books as 'data'. Colleagues in Departments of Literature all over the world may shudder at the 'naive' notion that the works of fiction, of poetry, of drama which they study with hermeneutic methods are a corpus of 'data' or 'resources' — these terms could call up controversial notions of positivism in the philosophy of science, or perhaps more straightforward associations with applied sciences such as engineering and economics. Nevertheless, without works of literature in their concrete manifestations, and without collections of these works in libraries, the study of literature would be poor indeed. So I will risk the scorn of some literary colleagues by using a rather general concept of *HLR*/*LD* repository. For colleagues who work in rhetoric, or stylistics, or in empirical literary studies, and who study the forms of literary language, this usage will seem uncontroversial. And I will risk the scorn of some linguistic colleagues by insisting on the prime role of *HLR*/*LD* in the language sciences, but will attempt to appease them by conceding that our minds are potential Human Language Resources — but not *HLR*/*LD* of a type which is sharable independently of the individual researcher unless recorded as a corpus. The tendency of post-structuralist generative linguistics from the 1960s to the 1990s to assign corpus data a secondary role and to concentrate on modelling intuited data is even more extremely hermeneutic than the interpretative

methodologies of literary studies and conversation analysis: the data are not 'given' — the original Latin meaning of the term — independently of the observer who discovers them, but constructed intuitively. In varieties of linguistics which study corpora of speech and text, whether qualitatively or quantitatively, the need for *HLR*/*LD* will seem uncontroversial.

The first disciplines concerned with *HLR*/*LD*, though not as massively as in modern times, were *translation* — whether on the Rosetta Stone, or on paper and parchment — and lexicography. From the 19th to the 21st centuries, lexicography in Western Europe gradually turned from being primarily the art of the lexicographer into a modern digital technology based on corpora of many millions of words of text and transcriptions of spoken language. Translation proceeded at the same time in a similar direction with the aid of databases as translation memories and machine learning techniques.

The key to this development lies in the increasing importance of extensive *sharable data* for empirical language studies, the need for which has been generally recognized since the 18th century in connection with the development of empiricism and detailed studies of the natural world (though sporadically recognized in many cultures long before that).

In the day-to-day use of *HLR*/*LD* there are also practical considerations, as in the following criterial properties of *HLR*/*LD* which are constantly under discussion in the field:

1. **Independent**. The data are ontologically independent of the observer by virtue of being recorded on storage media and identically accessible to many observers. This is not true of the intuited opportunistic data which are often used in linguistics, typically: "Can you say...?".

2. **Reusable**. The data are usable many times, perhaps by many analysts, and not just once by an individual scholar.
3. **Sustainable**. The data are not *ad hoc* individual observations but have an extensive lifetime. This is an obvious condition for endangered languages — if sustainable forms of data are not available, then if the language dies, the research on it dies. For this purpose, the notion of *sustainable repository* is crucial. In general, the structure of funded projects does not provide for sustainable repositories. These are a supraregional, national or international responsibility.
4. **Searchable**. This condition ensures that the data which are relevant for a given research purpose can be found. The condition also implies that the data are provided with adequate descriptive 'metadata', i.e. catalogue data.
5. **Interoperable**. The resources must be usable in a variety of environments, for example on different computer systems. What is true of cars must be true of *HLR*/*LD*.
6. **Standardized**. The role of a standard format, category system or tool, is to facilitate interoperability and exchange of data. This applies whether the standard is proprietary (such as the *de facto* norms of WAV files in the audio domain, or PDF and DOC files in the text domain) or official (such as HTML norms). Sometimes a crippling effect of over-standardization on innovation and creativity is feared, but this not generally justified: standards are, in the medium and long terms, mutable.
7. **Trustworthy**. The *validity* of *HLR*/*LD* must be checkable, the method of its creation must be *available* and *replicable*, the *HLR*/*LD* must be *accessible*, and storage of *HLR*/*LD* must be *reliable*.

All of these criteria are, from the scientific point of view, language and culture independent, and also discipline independent in regard to the scientific and technological language disciplines. The criteria also apply to data from other kinds of communication such as conversational gestures, and the sign languages used by the speech and hearing impaired.

In the following sections I discuss scenarios for the creation of *Human Language Resources*, taking a broad perspective in view of the immense number of varieties of human languages and then the case studies mentioned previously.

2. Scenarios for human language resource creation

2.1 The architecture of language

In a sense, *HLR*/*LD* are primarily concerned with the *forms* of texts and speech rather than their *functions*, i.e. their *meaning* and their *purpose* — no functions without forms. But of course interest goes beyond the forms of texts and speech.

Figure 1 shows a model of the minimal requirement for a unit of documentation in a human language resource: the language sign, a representation of observable forms of text or speech, a representation of the structure of each form (whether simple or complex, paradigmatic or syntagmatic), and a representation of its meaning.

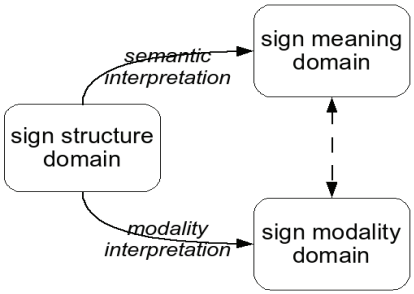


Figure 1: The units of language documentation: signs

The triadic structure is more adequate than the common dyadic structure of *significans* and *significatum* in many areas of linguistics, or in distinctions between the *logical structure* and the *rendering structure* of representations on the World Wide Web. The core of the sign, often interpreted as a mental representation, is abstract: the simple or complex *structure* of the sign. The structure of the sign has two interpretations in terms of the real world within which the sign user is located. The first interpretation is the *modality* interpretation, either acoustic (phonetic interpretation) or visual (as in gesture or in writing). The second interpretation is the *semantic* interpretation in terms of abstract concepts or objects in the environment.

But the triadic sign representation is insufficient in two ways. First, the triadic sign representation is too general. More detail, as shown in Figure 2, the *Rank Interpretation Model* (*RIM*) of language architecture, is needed in order to gain an understanding of the breadth and depth of HLR/LD.

Second, the triadic sign representation is incomplete: there is no representation of the situational semantic or pragmatic context. The situational semantics and pragmatics of these forms — meanings in context and the background knowledge and strategies of their users — can only be documented indirectly

and less explicitly than the textual and spoken forms of language. The situational semantic and pragmatic functionality of text and speech in use can be partially documented by use of additional textual characterization and by video recordings, however, and this practice is increasingly in use, with a growing interest in systematic transcription and annotation practices for communicative movement and gesture as well as speech and text.

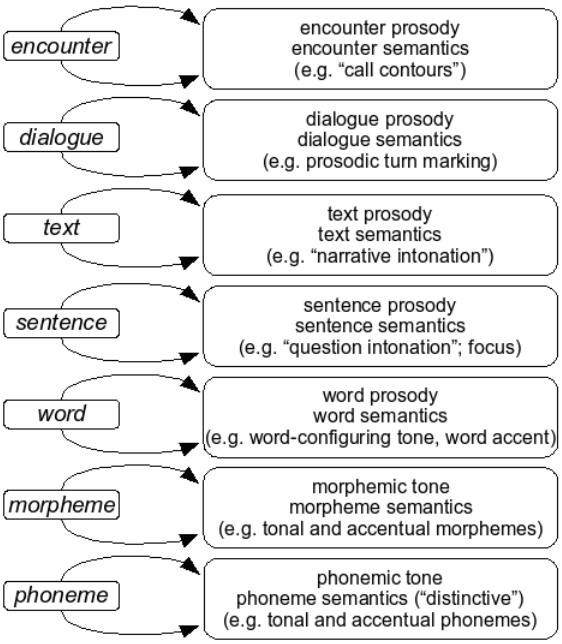


Figure 2: The Rank Interpretation Model (RIM) of the architecture of language

Signs are highly complex, though at each level there is a need for modality and semantic interpretations. Figure 2 is related to and differs from conventional models of language in many ways. Essentially, it combines the notion of the *rank* of language sign units of different sizes, a characteristic of functionalist

views of language, with a notion of *semantic interpretation* taken originally from formal logic and *phonetic* interpretation from linguistics. At each rank, from the phoneme through morpheme, word and sentence to the dialogue contribution and the encounter (and many sublevels between these), the units thus receive a *semantic interpretation* and a *modality interpretation*. The usual modality interpretation in linguistics is *phonetic interpretation*, but there is increasing interest in the documentation of writing in different scripts (including handwriting) conversational gesture and also of sign languages, which are comparable in complexity to spoken languages, into language documentation and human language resources. This extension of the Rank Interpretation Model is shown in Figure 3.

The purpose of proposing this integrated *Multimodal Rank Interpretation Model* (*MRIM*) here is to underline the fact that *HLR*/*LD* are not only about the properties of phonemes, words and sentences, but about the roles of all these items in a structured rank hierarchy, about the modality interpretation of these items in terms of speech, gesture and writing at different rank levels, and about the semantic interpretation of these signs at different rank levels. The inter-modality correspondence tendencies of intonation, accent and rhythm with punctuation, font choice and layout at the supra-word level are obvious. Perhaps less obvious is the inter-modality correspondence of prosodies such as accents and intonations with the accent-like and intonation-like rising and falling and rhythmical movements of hands, head, facial features. The correspondence of iconically demonstrating or deictic pointing gestures with words is more obvious.

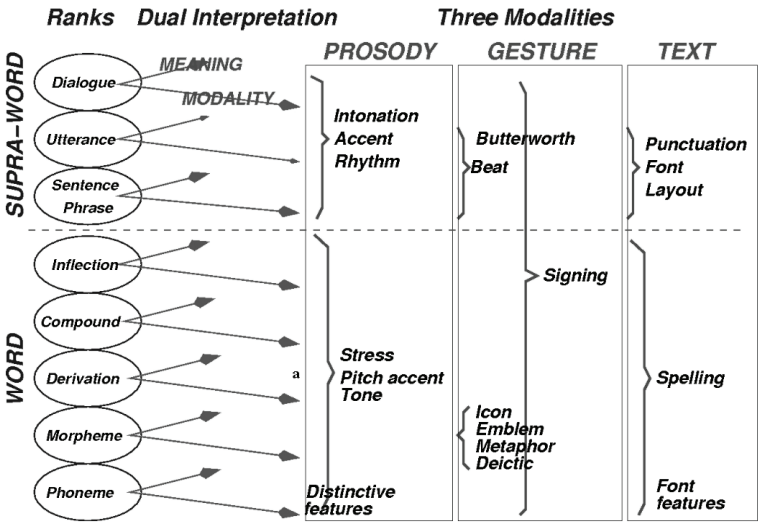


Figure 3: Multimodal Rank Interpretation Model (MRIM)

2.2 Varieties of text and speech

We have seen rapid developments in the *HLR*/*LD* area under the influence of modern communication technologies: in all languages, when SMS text messages are used, the limitations — but also the potential — of the technology enforce changes in the use of language. The development of writing technologies 3000 years ago enforced not only changes in the use of language but changes in language itself. There are many scenarios for the use of language in text and speech, and for many reasons we want to document and harness these resources, for heritage preservation, for the advancement of education in the languages of the world, and for deployment with new technologies. It is this practical need for actual usage which prompts the use of ‘text’ and ‘speech’, rather than ‘language’.

The word ‘language’ itself has a wide range of meanings, across the enormous number of languages, dialects, speech varieties, written registers with

which we communicate naturally as humans, and the specializations of written registers of language which we use in mathematics, logic, the formal sciences and technology, for description, programming computers, for describing knitting patterns and chess games, for transcribing the pronunciation of languages, and for technical communication. The word 'language' refers, also, to the gesture sign languages of the visually and vocally impaired. The word 'language' is also used for the communication systems of birds, apes, whales and other animals. And the word 'language' is used metaphorically for many other things in nature which are said to 'speak' to us.

In these broad senses, there are uncountably many languages and varieties of languages. But for *HLR* documentation — as for any other activity — we only have limited means, which compete with the means required by other activities. Resource scarcity means rationing, selection, the setting of priorities. There is a 'magic number' which is often cited for the number of human languages which have developed until the present day: the *Ethnologue*[1] catalogue of the world's languages lists 6909 languages and regional variants; this catalogue underlies the international language name codes standardized in ISO 639-3, and is based on the activities of multitudes of linguists around the world. Aleksandr Kibrik once characterized language from this point of view as *the spoken communication of a village*. One might even say, from the point of view of migrating populations, *the spoken communication used in a family*.

The number 6909 is a simplification, in relation to the varieties outlined in the present discussion, and it is already an enormous reduction in the number of language varieties in the broad sense of the term, but still a rather large, though not astronomical number.

[1] URLs: <www.ethnologue.com>, <www.ethnologue.org>.

So we need additional, more pragmatic criteria for selecting languages for creating human language resources. For *HLR*, in a large technology company the answer to selection lies in the standard languages of the world and other languages spoken by large populations, which promise a market for technology. But for *LD* in the humanities, the criterion is different: the priority lies with the languages which are on the verge of disappearing, the 'endangered languages'. Endangerment-based criteria for selecting languages for resource-creation are used by funding agencies; the main criterion (apart from dwindling community size) is whether an affluent linguist or group of linguists is interested in these languages.

The reasons for these pragmatic criteria are obvious: there are economic priorities — we have limited means. For this reason, a group of linguists in West Africa and I have formulated more detailed criteria for both *HLR* and *LD*, and formulated this as the *CESAF* list of *HLR*/*LD* criterial properties:

- **Comprehensive**: all levels of linguistic analysis, relative to the task.
- **Efficient**: maximal automatization of data acquisition and processing.
- **State of the art**: data structures, algorithms; web applications, mobile intranet servers.
- **Affordable**: browser, any desktop/laptop/tablet/cellphone.
- **Fair**: open source payback to data donors.

In the following discussion I will deal with both the heritage type of motivation and the technological type of motivation for selecting languages for resource creation, and provide examples, concentrating on the following areas: the cycle of language birth-development-endangerment-death.

2.3 The cycle of language birth, development, endangerment, death

Behind the many language varieties which have already been outlined is an ever-changing cycle of language change: language *birth* (e.g. pidgins, creoles, artificial languages), language *change* (and ultimately *birth*) by medium and long term social and internal processes, language *endangerment* and language *death*, and sometimes also language *re-birth* as in the case of *Ivrit* (Modern Hebrew) in Israel, or, to a marginal extent, *Kernewek* (Cornish) in the United Kingdom.

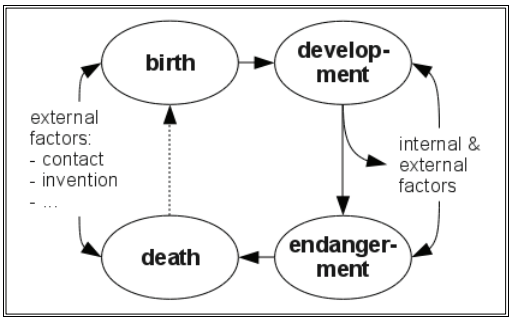


Figure 4: The cycle of language birth – development – endangerment – death

To clarify the background I propose a model of cycle of language *birth – development – endangerment – death*, shown in Figure 4, and discuss a number of details: the first decision in creating human language resources is to decide for which phase in this cycle *HLR*/*LD* are to be provided — a technology company would concentrate on different phases than a cultural history project.

Language birth itself has many facets: the slow evolution of human languages from unknown sources around 200,000 years ago with the evolution of *homo sapiens sapiens*, the splitting of languages into new languages both within Africa and after the out-of-Africa migration 70,000 years ago, and again

after the last Ice Age 10,000 years ago and the development of agriculture, metal processing and large urban civilizations. Later facets of language birth and development include the mutation of colonial languages into regional languages, pidgins and creoles, from the first empires in Europe and Asia 5000 years ago to several Indo-European languages and to Chinese in the modern age. And 'natural' languages may be invented: *Esperanto* is the most famous example, based on the structural typology and vocabulary of European languages, and later in this paper a different kind of 'invented language' will be described. Sometimes older phases of languages which have developed into new languages are retained, mainly in written form, in formal and ritualized contexts such as religious language: Classical Chinese, Sanskrit, Greek, Latin, Old Church Slavonic, Classical Arabic, with legacy documents with special requirements for manuscript documentation.

Language development is part of a many-faceted process of language change. On the one hand, it involves the standardization of languages for sociolinguistic, cultural and political reasons, in which factors which influence the stability of societies are subject to the control of academies, education and media. On the other hand, it involves the de-standardization of languages which were once standards, as in the cases of Classical Chinese, Sanskrit, Greek, Latin, Classical Arabic, which have already been mentioned, but also in many other cases in which the language in its standard form has disappeared, such as the East Germanic Language Gothic, or the languages of South and Central American civilizations.

A different development of this kind is where dialects become standards and then give way to other dialects: this has the case with the dialects of Britain since the Middle Ages, where the creation of a standard was pushed by Norman French colonists, Norman French itself being a variety of mediaeval French which had been influenced by Scandinavian colonists. German is another case in

point: the dominance of South-West Germany in the Middle Ages gave way over the centuries to the dominance of East German varieties under the influence of Luther's Bible translation, and later still to the dominance of North German dialects under the influence of Prussia. We can see the grain of truth behind Max Weinreich's joking definition of a standard language: *A language is a dialect with an army and a navy*. Evidently, this is the kind of language variety preferred on political and macro-economic grounds in societies with a strong need for a unified identity.

The point is that languages are not defined by mutual comprehensibility alone, but by political identity, and both of these can change rapidly. The very different dialects of Austria, Switzerland and Germany are referred to for historical political reasons as 'German'. In contrast, the languages Dutch and Flemish, on the other hand, are linguistically very similar, both to each other and, less so, to German, but have different political histories and are treated as different languages. Both are very similar to the neighbouring German dialects of the Rhineland, which in turn are very different from many other German dialects. Evidently the motivation for defining a language is political as well as linguistic.

In principle, though with major historical differences, these considerations also apply to the development of Chinese through the centuries and under different political entities in various parts of Eastern Asia and in the Chinese diaspora.

Language endangerment is a function of language development: the languages of some communities thrive as the communities themselves thrive and influence others, while the languages of other communities decline as the communities decline and are influenced by others. The immediate source of language endangerment is when languages are no longer transmitted by parents

to their children. But this complex process has reasons. The reasons may be migration and diaspora, due to climate change such as is currently going on in many parts of the world. Or, in the case of a language like Yiddish, endangerment and at least partial extinction may be due to persecution and destruction of the community. Or the reasons may be due to political repression, of which there are many examples in all continents. The most common reason, indirectly linked to these other reasons, is economic survival and progress. My first contact with this was in a taxi in Dublin in the 1980s: I asked the taxi driver if he spoke Irish. He replied "Of course!". I then asked if his children spoke Irish. His reply: "Of course not! I forbid them to speak it. They'll never get rich if they speak Irish!" And in fact my own history — an English-speaking descendant of Welsh and English families — is coloured by some of these reasons.

Language death is the consequence of extreme language endangerment, via a moribund stage in which very few speakers of the language remain, and is easily defined: there are no more users of the language, and potential insights into human cognition and culture which the language might have provided have disappeared forever.

2.4 Preliminary conclusion: the variety of language resources

The preceding discussion has ranged widely over the different varieties of text and speech for which *HLR*/*LD* are collected. First, the complexity of language signs themselves was outlined with reference to the *Multimodal Rank Interpretation Model* of the architecture of communication in text, speech and gesture. Then varieties of text and speech were sketched, and finally the cycle of language birth, development, endangerment and death was outlined. Each of these criteria enters into decisions about what kinds of *HLR*/*LD* to create.

3. Case studies

The following discussions of aspects of HLR/LD creation are based on projects and bilateral communications with scholars of different linguistic interests, but with the goal of developing high quality empirical data with appropriate methodologies for use in various linguistic and technological applications. Resources are being collected on a grand scale for many languages, for technological applications in large communities to language documentation of endangered languages. But resources are not just data: they also include enhancements of date with appropriate tools. The following case studies illustrate selected resource enhancement projects for written and spoken language.

3.1 Speech resource visualization: Mandarin syllable complexity

Large quantities of text and speech resources are available for Mandarin Chinese, and their deployment in many areas of applied linguistics and speech technology is extensive. In particular, syllables in Mandarin and their properties such as tone are well researched and the subject of continuing research. I propose a technique here for adding to existing knowledge, using a visualization strategy which is related to the 'big data' visualization and sonification techniques which are being developed currently in order to gain insights into the structure of data prior to further analysis.

The structure of Mandarin syllables is often portrayed in the form of a 'pinyin table', as in Table 1 (He 2004). The Table visualizes a binary syntagmatic relation between the initials (onset consonants) and finals (vowels and vowel-nasal sequences, i.e. rhymes). The relation licences existing lexical syllables associated with characters. The table contains 399 characters; other tables with slightly different numbers of syllables may be found, e.g. Aidaoguangci (a nickname)

2010. The 437 gaps in the table have different explanations: partly systematically excluded due to Mandarin phonological structure; partly ‘accidental gaps’ which are possible but not used in formal contexts (though some may be found in informal spoken language); partly complementary distribution of allophones which have independent representations.

Table 1: Table of Mandarin lexical syllable cooccurrences (adapted from He 2004) representing a binary structural relation between initials and finals

	b	**p**	**m**	**f**	**d**	**t**	**n**	**l**	**g**	**k**	**h**	**j**	**q**	**x**	**zh**	**ch**	**sh**	**r**	**z**	**c**	**s**	
-i															+	+	+	+	+	+	+	
a	+	+	+	+	+	+	+	+	+	+	+				+	+	+		+	+	+	+
o	+	+	+	+																		+
e			+		+	+	+	+	+	+	+				+	+	+	+	+	+	+	+
ê																						+
er																						+
ai	+	+	+		+	+	+	+	+	+	+				+	+	+		+	+	+	+
ei	+	+	+	+	+		+	+	+		+						+		+			
ao	+	+	+		+	+	+	+	+	+	+				+	+	+	+	+	+	+	+
ou		+	+	+	+	+	+	+	+	+	+				+	+	+	+	+	+	+	+
an	+	+	+	+	+	+	+	+	+	+	+				+	+	+	+	+	+	+	+
en	+	+	+	+			+		+	+	+				+	+	+	+	+	+	+	+
ang	+	+	+	+	+	+	+	+	+	+	+				+	+	+	+	+	+	+	+
eng	+	+	+	+	+	+	+	+	+	+	+				+	+	+	+	+	+	+	
ong					+	+	+	+	+	+	+				+	+		+	+	+	+	
i	+	+	+		+	+	+	+				+	+	+								+
ia								+				+	+	+								+
ie	+	+	+		+		+	+				+	+	+								+
iao	+	+	+		+	+	+	+				+	+	+								+
iou					+		+	+				+	+	+								+
ian	+	+	+		+	+	+	+				+	+	+								+

	b	p	m	f	d	t	n	l	g	k	h	j	q	x	zh	ch	sh	r	z	c	s	
in	+	+	+				+	+				+	+	+								+
iang							+	+				+	+	+								+
ing	+	+	+		+	+	+	+				+	+	+								+
iong												+	+	+								+
u	+	+	+	+	+	+	+	+	+	+	+				+	+	+	+	+	+	+	+
ua									+	+	+				+		+					+
uo					+	+	+	+	+	+	+				+	+	+	+	+	+	+	+
uai									+	+	+				+	+	+					+
uei					+	+			+	+	+				+	+	+	+	+	+	+	+
uan					+	+	+	+	+	+	+				+	+	+	+	+	+	+	+
uen					+	+		+	+	+	+				+	+	+	+	+	+	+	+
uang									+	+	+				+	+	+					+
ueng																						+
ü							+	+				+	+	+								+
üe							+	+				+	+	+								+
üan												+	+	+								+
ün												+	+	+								+

Examination of the syllable finals shows that they are syntagmatically complex, so their internal cooccurrence constraints are not expressed in the table. It is important to know these constraints for a number of reasons, which will be detailed below. There are a number of analytic steps which can be taken in order to bring out the constraint patterning:

1. Introduction of a glottal stop consonant as initial for syllables which the Table notes as being vowel-initial. The motivation for this is that vowel and vowel-nasal syllables start with a glottal stop when spoken in isolation.

2. Treatment of the initial high vowels of diphthongs and triphthongs as independent segments, specifically as approximants or glides [j], [w], [ɥ].
3. Treatment of nasals [n], [ŋ] as independent segments.

In principle, final high vowels in diphthongs and triphthongs could also be treated in this way, but this step is not taken here. Syllable contractions are also not taken into account here. The steps discussed here are not new in themselves, but the resource enhancement treatment proposed here appears to be novel in the context of Chinese syllable modelling.

The problem which arises is how to represent the non-binary relation which a more detailed analysis of segment sequencing constraints in syllables requires. It is possible with embedded tables, or with trees for individual items, but a more informative and expressive technique is to use directed graphs, which express more complex syntagmatic relations in an intuitively insightful way.

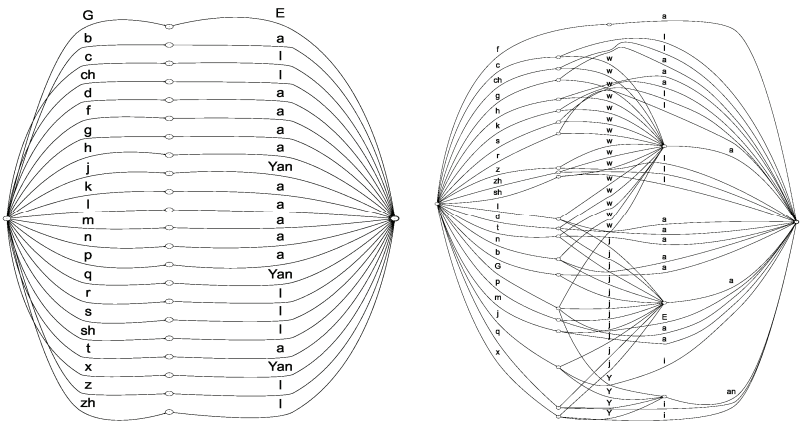


Figure 5: Visualizations of two models of Mandarin syllable structure, using directed graphs. Left: model of binary relation as expressed in the ‘pinyin table’. Right: model with glides [j], [w], [ɥ] as independent segments.

In Figure 5 automatically generated visualizations of two models of Mandarin syllables are shown, using left-right directed graphs. In each graph, there are differing numbers of parallel edges between pairs of nodes. These are pruned, in order to make the structure clearer, with only one edge representing the set of parallel edges in each case; the important issue here is the structure, not the specific symbol inventaries on the edges. The names of the nodes are not significant. The graphs are deliberately rendered in such a way as to emphasize structure at the expense of detail; the underlying structures are highly detailed, however.

The graphs were constructed from the table in a four-stage process: first, the nodes and edges were created, based on Table 1, and, second, extended with a node insertion algorithm on the basis of the additional segmentations listed above (implemented in *Python*). Third, the graph was visualized using the *Graphviz* graph construction and optimization package. Fourth, a graph traversal algorithm (also implemented in *Python*) was used to generate the entire syllable set for testing purposes by starting at the leftmost node, traversing the edges to the next node on the right, picking up the edge label symbol on the way, and concatenating these symbols until the final node on the right is reached.

On the left in Figure 5 is a straightforward model of the binary relation expressed in the 'pinyin table', with no further analysis of syllable finals; the only nodes intervening between the initial and the terminal node are those which link the initial and final edges, the long vertical series of nodes in the centre. The model on the left captures exactly the syllables defined in the Table. The model on the right in Figure 5 expresses the constraints for approximants as separate segments. The series of three nodes for the approximants can be seen between the initial-final linking nodes and the terminal node. The explicit handling of approximants leads to an enriched understanding of internal syllable constraints.

The extended network with the interpolated nodes and edges for approximants overgeneralizes a little, i.e. produces syllables which are not in the basic inventory of 399 syllables defined by the Table.

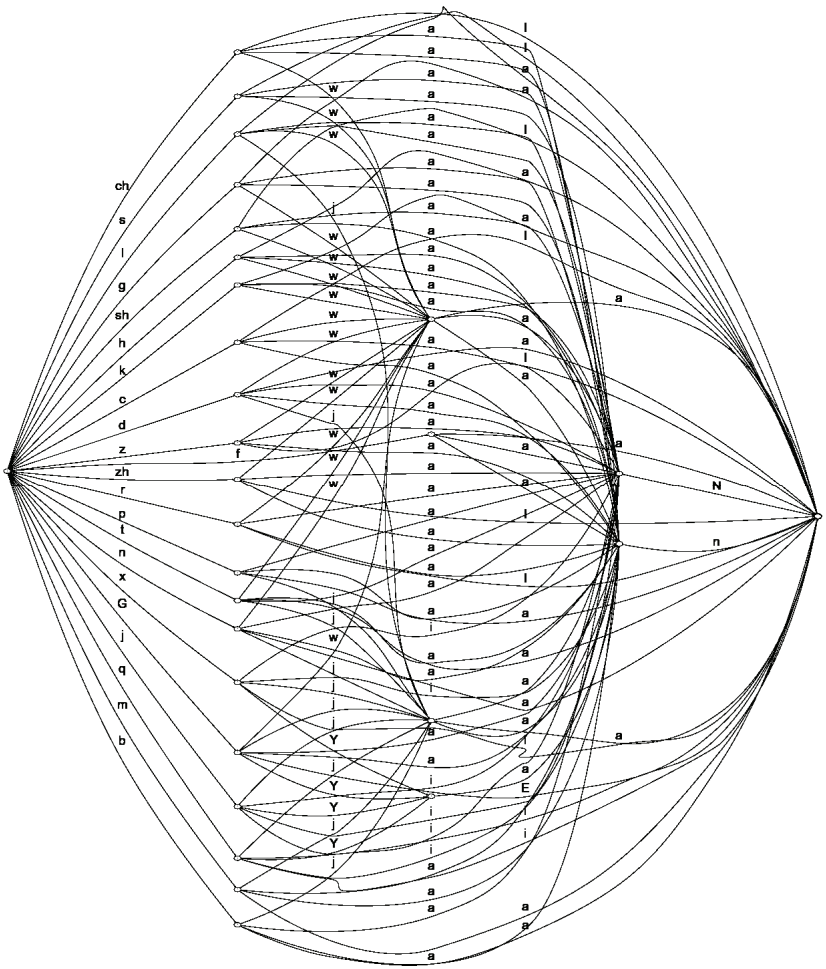


Figure 6: Visualization of Mandarin syllables with explicit incorporation of constraints on approximants and nasals as independent segments

When the final nasals [n], [ŋ] are treated as separate segments, a more complex constraint network shown in Figure 6 is induced from the basic resource data in the Table. The nodes linking the vowels to the nasals appear as a vertical series of two, immediately to the left of the terminal node. The graph also overgeneralizes slightly, taking over the same overgeneralizations introduced by the treatment of approximants. Adding the nasals did not increase the number of overgeneralizations.

It is legitimate to ask what are the actual advantages of such visualizations. These advantages are both theoretical and practical. On the theoretical side, they permit a more sophisticated appreciation of the typological properties of the language, supporting comparison of dialects and the study of language history. On the practical side an enhancement of basic resource data is needed when theories of sound production and perception and of the sequentiality of speech acquisition (with child and adult learning) and speech loss (due to accident or illness) are examined. Also on the practical side, a detailed understanding of sound patterns is a prerequisite for the design of corpus markup for the speech-corpus-based technologies of speech recognition and speech synthesis. In these contexts, the graphs can be also interpreted as finite automata (state machines) for symbolic or stochastic processing purposes.

3.2 Spoken language resources: developing tone models for Kuki-Thadou

The long-term intention of providing *HLR*/*LD* for the Tibeto-Burman language Kuki-Thadou (Thadou) is to provide an empirical foundation not only for traditional applications in education, but also to examine the problems which need to be handled in providing *HLR*/*LD* for linguistics and speech technology

in a Sino-Tibetan language, here specifically a Tibeto-Burman language on the India-Myanmar border. The particular features of interest are phonemic tone and its phonetic correlates.

The typical *HLR*/*LD* information collection procedure encompasses lexicon compilation, and simultaneous phonemic analysis, supplemented by tonemic analysis. The phonemes of Thadou, as elicited by these methods, are shown in Table 2.

Table 2: Thadou phonology and tonology

left: consonants; top right: vowels; bottom right: tones

	bilabial	**labio-dental**	**alveolar**	**palatal**	**velar**	**laryn-geal**
plosive	p ph b		t th d		k kh g	ʔ
nasal	m		n		ŋ	
fricative		v	s z		[x]	h
apic. affr.			ts			
lat. appr.			l			
lat. fric.			ɬ			
appr.	w			j		

	Front	**Central**	**Back**
close	i		u
mid	e	ə	o
open		a	

	Gibbon	**Hyman**	**Gloss**
sá	H	HL	animal
sǎ	LH	LH (H)	hot
sà	L	L	thick

The starting point for the analysis is the classic method of ‘ear phonetics’. But for speech technology, in particular speech synthesis, a more precise model is required, of the fundamental frequency (F0) of tones in isolation (citation contexts) and of tones in sequence, and this is supplemented and supported by instrumental measurements and perceptual tests. For application purposes, whether in language teaching or in speech technology, it is important to document each phoneme in context.

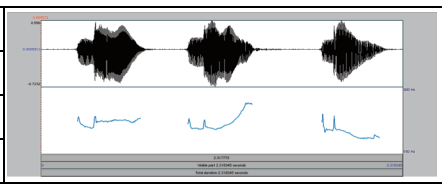
Among the consonants, unlike the different two-way distinctions between plosive types in Chinese or English, Thadou has a three-way distinction between unvoiced, aspirated and voiced plosives, presumably an areal phenomenon related to the neighbouring Indo-Aryan languages which have such distinctions.

The property of phonemic tone is shared by Thadou with other Sino-Tibetan languages. As with segmental phonemes, standard vocabulary and minimal contrast methods are used to determine the set of tones. As with segmental phonemes, varying analytic methodologies and varying theoretical assumptions can sometimes lead to slightly different inventories or characterizations of the tones, as in the difference between the Hyman and Gibbon tone inventories in Table 2.

The next step, at least for the purposes of *HLR*/*LD* in speech technology, is to determine the physical correlates of the tones. In medical applications, physiological correlates are investigated, and in the more common speech technology applications the acoustic correlates are examined. The usual procedure is to use software for examining the waveform and pitch traces of utterances, to annotate the occurrences of tones in the utterance, and to extract the acoustic measurements of frequency and duration associated with these tones (cf. Table 3).

Table 3: Thadou tone model with basic descriptive statistics

Tone	N	min	max	mean	sd	offset	slope
H	18	200	244	221	0.29	221	-0.03
LH	17	215	237	220	7.07	209	1.3
L	18	192	213	203	6.3	215	-1.31



The three tones of Thadou, spoken in isolation, show clear phonetic properties. The H (high) and L (low) tones are clearly different, with barely

overlapping ranges and clearly different means of 221 Hz and 203 Hz respectively; the small, non-overlapping standard deviations show, in relation to these means, that the difference in frequency is significant. The LH (low-high) contour tone is not so different from the H tone, but the shape is very different, as shown by the considerable positive slope of 1.3, in relation to the marginally negative slope of -0.03 shown by the H tone and the clearly negative slope of -1.31 shown by the L tone.

These are measurements on citation forms in isolation. Tones in context may behave rather differently. Figure 7 shows an example of tone spreading in an authentic utterance, and in a model of this tone spreading in a speech synthesis example of this utterance.

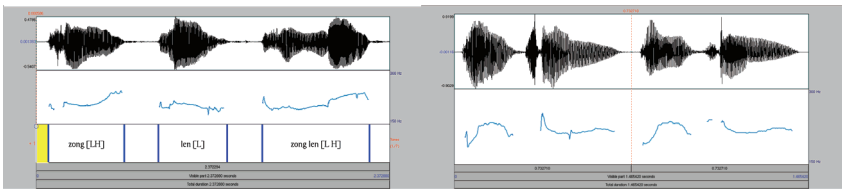


Figure 1: Thadou tone spreading (left). Tone original and synthesis (right).

The application of *HLR*/*LD* procedures such as these provides a documentary foundation for further linguistic description and technological application.

3.3 Written language resources: documenting Medefaidrin texts

Customarily one thinks of written language as books, newspapers, the internet. But in *HLR*, for many purposes handwriting and manuscript documents are equally important. Examples in technology are scanning with optical character recognition, the recognition of handwriting gestures on touchscreens, and in

forensics and historical philology writer identification. In heritage documentation, oftentimes available data are in manuscript form.

In the case of manuscripts, which are particularly important in the study of local and minority languages of many kinds, especially endangered languages, there are many more stages of documentation on the physical level, comparable in many ways with speech, than with electronically available written text. Manuscripts may be 'noisy', i.e. faded, smudged, stained; they may be subject to 'fast writing' phenomena of assimilation and reduction, like 'fast speech' phenomena; the media may be subject to physical degradation such as discolouring and decomposition through dryness or damp. The first stages of documentation are, therefore, as with speech, segmentation, transcription and annotation, together with vocabulary collation and analysis and contrastive study of the characters, punctuation marks and layout conventions of the text, and the grammatical conventions, which may differ from the conventions of spoken language.

Figure 8 shows documentary fragments pertaining to the case of the Nigerian 'spirit language' Medefaidrin.

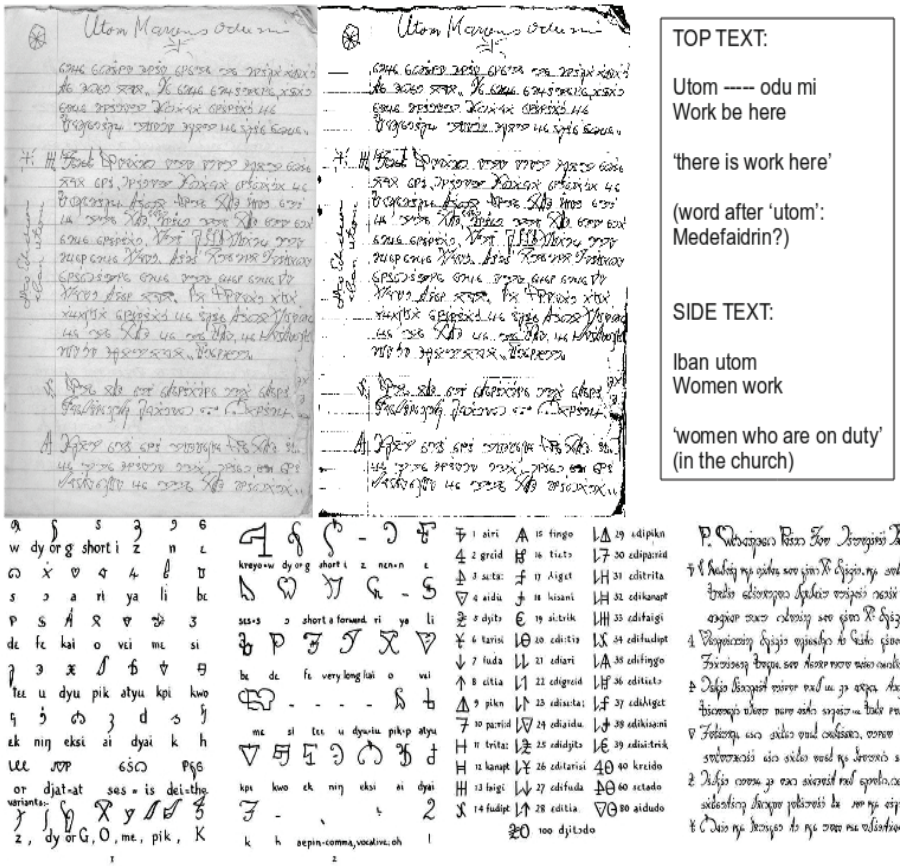


Figure 8: Samples of phases in Medefaidrin 'spirit language' manuscript analysis

Medefaidrin is spoken and written by the Oberi Okaime religous community near the Calabar River in South Eastern Nigeria. The language is said to be a 'spirit language' which was revealed to the originator of the language by divine inspiration (Gibbon et al. 2010).

A number of phases of Medefaidrin documentation at different levels of analysis are shown (top left clockwise to bottom left): the original scanned pages of a notebook; the same page, optically enhanced for readability; metadata about

the marginalia of the page; number system inventory; alphabet inventory. The phases are somewhat analogous to the phonetic analysis of speech signals: noisy signals are filtered to extract the essential information, then the signals are tokenized and inventarized. The scanning and filtering destroys some important background information, however, which is only of indirect interest to the linguist, but may be of immense interest to the historian: the state of preservation of the document, and its age — only the original document can provide direct evidence for this information, though circumstantial evidence and historical reconstruction may contribute.

The divine inspiration of Medefaidrin is not entirely immune to influence from the contemporary prevailing colonial language of missionaries in the area, English, as individual characters in the script, the number system and text layout show. An interesting feature of number formation is, however, the vigesimal (base 20) system, as inspection of the number system in Figure 8 shows, of which only vestiges in English (‘score’, 20; biblical ‘three score’, 60) and French (‘quatre-vingt’, 80).

3.4 Lexicography: prelexical resources for Uyghur

One central area of *Human Language Resources* — perhaps the central area — is the lexicon or dictionary, of which there are many forms. The lexicographic data acquisition and product creation process has often been seen as an art, but there is a logic to the process, and tools are employed at each stage. The stages of lexicon acquisition, each stage involving abstraction and generalization from corpus data (perhaps with lateral input from other dictionaries) is illustrated in Figure 9.

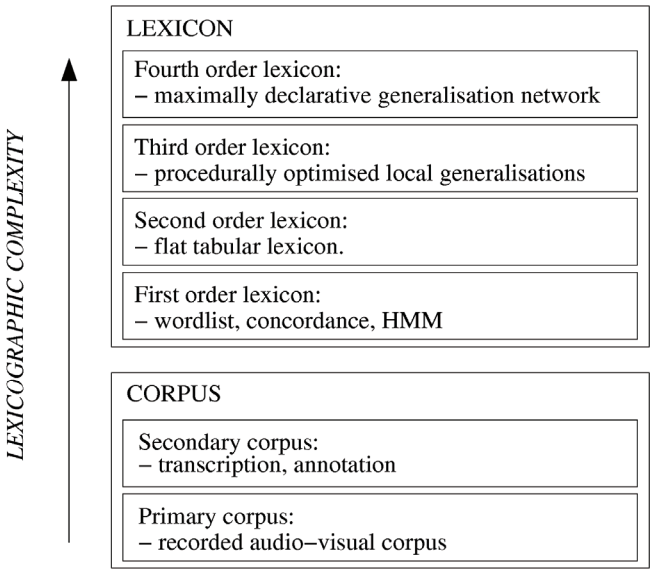


Figure 9: Hierarchical Lexicon Generalization Model

The core resources are at the corpus level: both the primary corpus level of recorded written documents and speech recordings and the next level of abstraction, transcriptions and annotations — such as part-of-speech (POS) tagging. The next level involves a jump from a corpus of *tokens* to the first order lexicon, consisting of tables and lists of *types*: word frequency lists, pronunciation lexicons, concordances, interlinear glossing structures. The second order lexicon requires the integration of all this information into a highly redundant flat tabular lexicon in which every entry contains a full specification of lexical information — which is practically never fully constructed, because of the extreme redundancy. The third order lexicon has the familiar shapes of semasiological (e.g. alphabetically organized) and onomasiological (e.g. synonym based) dictionaries. The fourth order lexicon is mainly of theoretical linguistic and lexicological interest: it involves a maximum of generalization over as many entries in the lexicon as

possible, and thereby a minimum of idiosyncratic information in each lexical entry. The third and fourth order lexicons have different structural levels: the *megastructure* (metadata about the creation process and the product), the *macrostructure* (overall organization, for example as semasiological or onomasiological dictionary), *microstructure* (which organizes the information for each lexical entry in terms of data categories) and the *mesostructure* (links to the generalizations about grammar, pronunciation etc., and to word fields in the dictionary as well as to corpus examples).

At the first order lexicon level, lists of 'prelexical' units of many different kinds are compiled. For linguists, the most useful kinds of lists are word lists, in particular word frequency lists, with a type/token frequency ratio to indicate to what extent a corpus is saturated (i.e. where decreasing numbers of new units occur). In lexicon creation, information about contextual word properties is required, and a popular tool for aiding this procedure is the concordance, i.e. a list of words accompanied by the contexts in which they occur, usually either lines of text, or sentences. A concordance is already a prototype dictionary, though missing other kinds of information such as pronunciation, part of speech, definition. For speech technology, other kinds of unit such ad diphones and digraphs are needed (for decoding speech) as well as words and their contexts (for language models).

Examples of prelexical data at the first order lexicon lexicon level for the Eastern Turkic language Uyghur are shown in Figure 10 (cf. Gibbon 2012).

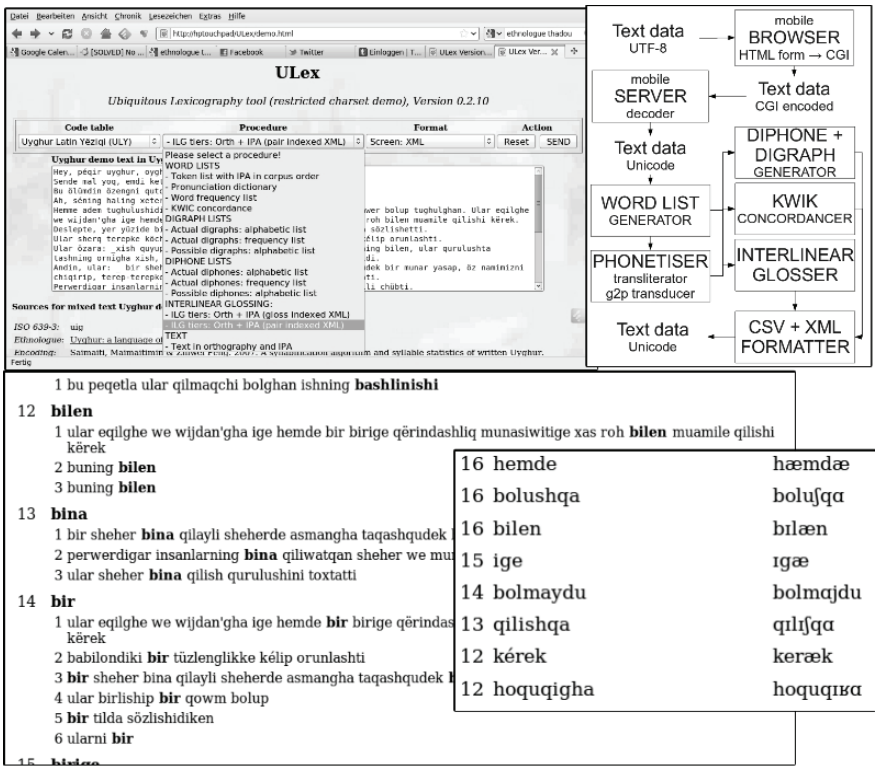


Figure 10: Collage of pre-lexical corpus input interface (top left), data flow (top right) and automatically created outputs (KWIC concordance, bottom left; frequency list with IPA transcription, bottom right).

The first order lexicon extracts show a user interface for corpus input, the procedure for processing this input, and two examples of the output obtained: a KeyWord In Context (KWIC) concordance to aid lexicological and grammatical analysis, and a frequency list with IPA transliteration to aid the development of pronunciation lexicons and speech technology systems. All of these prelexical data structures are generated automatically from the corpus.

3.5 Formal issues: Finite Machines as resources for tone modelling and prediction

The modelling of the phonetic properties of tonal resources is only part of the story. *Human Language Resources* are not only data and not only lexicons, but also grammars and prosodic models. In order to be able to use these grammars and prosodic models in a technological context, for example in reading software for the blind and in speech synthesis software for the vocally impaired, precise models are required.

Fortunately, we know by now that the mathematics underlying these models is relatively simple: in the Chomsky Hierarchy of Formal Grammars, from the most complex (Type 0) transformational operations through context-sensitive (Type 1) grammars for cross-serial dependencies, and context-free (Type 2) grammars for classic tree constructions, the simplest type (Type 3), the regular or linear grammars, are adequate for phonological, prosodic, and most morphological patternings. These grammars can be operationally modelled by Finite Machines, i.e. Finite State Automata and Finite State Transducers. Without going into detail, these are models of finite devices, i.e. devices which only need a finite memory for their operation. They are exceptionally easy to compute; nevertheless, by iteration (repetition) they can still describe infinite numbers of patterns.

For Niger-Congo languages I discovered in the mid-1980s that the phonetic interpretation of tone sequences of terraced tone languages is straightforwardly modelled by means of Finite State Transducers (Figure 11, left).

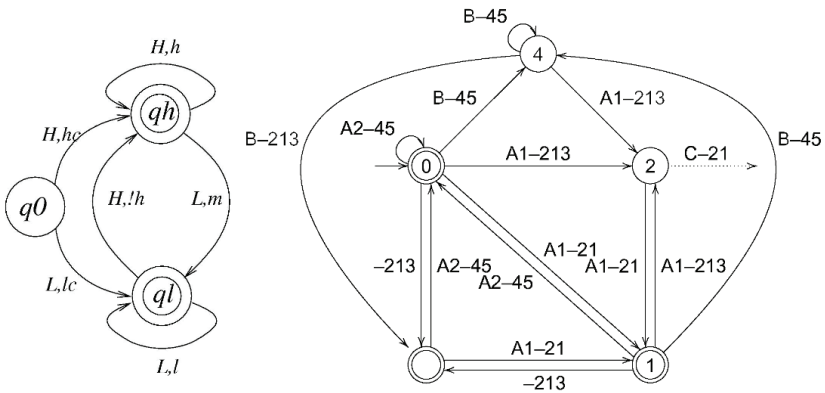


Figure 11: Left: general Finite Machine for phonetic tone realisation in two-tone Niger-Congo languages (Gibbon 2007). Right: Jansche's Finite Machine for lexical tone sandhi in Tianjin Mandarin (1998).

Starting a sequence at 'q0', a high tone moves to the high state, 'qh', where further high tones circle around until a low tone appears, when the low tone moves to the low state, 'ql', and so on. Similarly, if the sequence starts with a low tone it moves to the low state 'ql', if more low tones follow they cycle around until a high tone appears, and moves to 'qh'. On each type of transition a different kind of phonetic operation takes place: from 'ql' to 'qh' downstepping (assimilation of a high tone to a preceding low tone) occurs; from 'qh' to 'ql', low tone assimilation to the preceding high tone occurs. Languages differ in the details, but the general model for 2-tone languages remains the same.

Starting from this idea, Jansche developed a Finite Machine for describing the tone sandhi of the Tianjin variety of Mandarin (Figure 11, right). Without going into the details, it is obvious the typology of the two types of tone system is very different. The key to understanding the typological difference lies in noticing that the output of the Niger-Congo system is phonetically defined, while in the Mandarin system it is phonologically defined.

Nevertheless, the same formal system of Finite State Transducers is adequate for both systems: for describing both similarities and differences, and permits the creation of *HLR*/*LD* for computational purposes in speech technology, in particular in building prosodic models for speech synthesizers. Indeed, Finite Machine models are ubiquitous in *HLR*/*LD* for speech technology. The *Hidden Markov Models* (*HMMs*) in speech technology are statistically enhanced Finite Machines.

4. Technical resource creation procedures

4.1 A general resource for text and speech: the lexicon

An example of a specific lexicon from a linguistic point of view was given as a case study in the previous section. Lexicon resource processing consists of the following main steps, which require appropriate software tools (note that the procedures listed after tokenization are not necessarily conducted in the order given):

1. **Tokenization**. Individual word tokens are identified, including abbreviations, numbers, prices, dates, punctuation, identification of complex layout objects such as tables.
2. **POS (part of speech) tagging**. Each token is provided with a label (or set of labels) constituting a hypothesis about its part of speech; the European EAGLES (Expert Advisory Groups for Language Engineering Systems) developed a standard POS tagset for European languages, which has been extended and applied to other languages (these sets are in flux; consult the internet for up-to-date details).
3. **Word token and word type list creation**. A list of (possibly inflected) word types is extracted from the set of tokens, often also in conjunction

with the word token frequencies.

4. **Lemmatization**. A list of lemmas is created from the list of word types, involving stemming in the simplest case, and morphological analysis in the general case.
5. **Concordancing**. A context dictionary consisting of a list of items (types, lemmas, tags, etc.) and the contexts in which they occur in the texts. The best known kind of concordance is the KWIC (KeyWord In Context), a simple list of words and their left and right context strings.
6. **Word sketching**. Extraction of a maximum of (grammatical and other kinds of similarity) information about lemmas based on their distribution in the texts.
7. **Dictionary database compilation**. Semi-automatic (moderated) entry of information into the lexical database.
8. **Manual editing** of lexicon articles (definitions, etc.).
9. **Production**. Selection, organization and formatting of lexical information for the intended dictionary megastructure.

These procedures apply, with suitable modifications, to the compilation of other types of dictionary, including dictionaries for use in multilingual, speech-based and multimodal communication systems. The steps 1-7 above pertain to three of the stages shown in Figure 9: secondary corpus processing, and first and second order lexicons. Steps 8 and 9 pertain to third order lexicons. The fourth order lexicon is a product of research and not discussed further here.

4.2 Resource creation procedures for text-to-speech synthesis

A *Text-To-Speech* (*TTS*) synthesis system requires resources for developing the following subcomponents which resemble the *HLR*/*LD* components outlined

in the case studies in §3, with additional smaller components and more detailed steps:

1. **Text parser**. The text parser is a special case of the kind of parser which is used in text processing in general, enhanced with phonetization and prosodic modelling information, and will not be discussed further here. The text is pre-processed in order to extract implicit information:
    1. The spelling and ultimately the pronunciation of special text components such as abbreviations and numbers must be extracted.
    2. A pronunciation lexicon, usually with additional pronunciation rules, is required.
    3. A parser is needed for disambiguating the structure by picking the correct word readings from the lexicon and delimiting the phrasing of sentences.
    4. A grapheme-to-phoneme (phonetization) component is used to derive a transcription of the speech sounds for input to the speech processing component.
    5. A prosody module is needed for deriving intonation, accentuation and timing patterns for input to the speech processing component.
2. **Signal processing component**: conversion from an interface with parsed and phonetized text with added prosodic information into a synthetic speech signal. For the signal processing component there are several different speech synthesis paradigms, including the following main types, for which paradigm specific resources are required:
    1. **Pre-recorded 'canned' speech**. Canned speech is typically used in straightforward information service environments such as satellite navigation systems for vehicles, and for railway station announce-

ments. Systems such as these use a restricted set of utterance templates which permit substitution of station names and times, but also permit a combinatorially large set of new utterances to be synthesized. Canned speech is in principle very comprehensible and very natural, provided that the template units are carefully designed and produced, with close attention paid to the correct prosody (intonation and accentuation), and to appropriate transitions between canned speech units.

2. **Concatenative speech synthesis**. Small units, such as phonemes, diphones, demi-syllables and sometimes larger units, are concatenated to form words and sentences. There are three main approaches, each of which requires different kinds of resource:
    1. **Diphone synthesis** is one of the first kinds of concatenative speech synthesis, and is still used. In diphone synthesis, pre-recorded speech samples containing all the diphones in the sound system of the language are used, which are concatenated in order to reproduce the patterns of the input syllable and word sequences. A diphone is essentially a pair of phonemes (speech sounds; see below).
    2. **Unit selection synthesis**, a popular variety of speech synthesis, and in general more natural than diphone synthesis, is based on selecting continuous units from a large recorded corpus. The corpus is designed to contain all the phonemes, generally all the diphones, and perhaps all the triphones (sequences of three phonemes). Units are concatenated after calculating the best possible fit (cost, weight).

3. **Hidden Markov Model (HMM) synthesis**, a recent development based on stochastic modelling of unit sequences, trained on a suitable corpus.

3. **Formant speech synthesis**. Formant synthesis is one of the earliest kinds of speech synthesis, and is based on the spectral structure of speech sounds. An acoustic signal is reconstructed from empirical information about vowels, consonants, and the pitch, intensity and duration patterning of the intended synthetic speech signal. In principle, this approach is the most flexible and parametrizable in terms of linguistic and phonetic properties, but is more difficult to use in practical systems than concatenative techniques.

Each of these components and procedures requires different *HLR*/*LD* steps. The text analysis component shares many features with lexicography, and the initial stages of the signal processing component shares many features, such as transcription and annotation, with linguistic speech documentation.

5. Conclusion: the institutional perspective

It should have become very clear that *HLR*/*LD* creation can only be a transdisciplinary, collaborative, international effort. The traditional division of linguistic disciplines according to languages and language families is of course well justified, in permitting intensive study of history and typology of these various language domains. But in the field of *HLR*/*LD* creation procedures, all languages are equal and the same fundamental formal, computational methods apply to them all. The task is huge, and empirically justified progress can only come from cooperation. However, in the prioritization of languages for *HLR*/*LD* creation all languages are not equal: for practical reasons, the human language

technologies set different priorities from linguistics, even though the problems and methods are very closely related.

The need for transdisciplinary and transnational cooperation becomes very clear if the criteria listed earlier are considered. High quality resources and documentation must be *independent*, *reusable*, *sustainable*, *searchable*, *inter-operable*, *standardized*, and *trustworthy*. There are many institutions world-wide concerned with compliance with these quality standards in production of language resources and documentation of the kinds discussed in the present contribution. It is only possible to provide a few relevant categories of such institutions here, and to mention a few initiatives:

1. **Language and speech resource agencies**: in many regions, the two most well known being perhaps the *Linguistic Data Consortium*, in Philadelphia, and the *European Language Resources Association* (*ELRA*) with its operational wing, the *Evaluations and Language Resources Distribution Agency* (*ELDA*), in Paris.
2. **Academic institutions**:
    1. Research institutions such as *Academia Sinica*, *CASS* Beijing, the *School of African and Oriental Studies*, London, the *Institute for Language and Internet Technology* at Eastern Michigan University, and Institutes of Technology in many countries.
    2. Empirical linguistic and human language technology teaching institutions world-wide, too numerous to mention.
3. **Academic networks**:
    1. Conferences, in particular the *Language Resources and Evaluation Conference* (*LREC*), which takes place every two years.

2. Coordination initiatives in the speech technology field, in particular the *International Coordinating Committee for Speech Databases and Assessment* (*COCOSDA*), of which I am currently Convenor, and, very prominently *Oriental COCOSDA*, which originated as the Asian wing of *COCOSDA*, but is now an autonomous and very successful networking organization with its own conference series.

4. **Funding agencies**: from government agencies in many countries through larger and smaller non-for-profit organizations to the World Bank — again too many to mention here.

It is a fitting conclusion for me, in my role as Convenor of *COCOSDA*, to the present contribution to recall that the achievements in the area of Asian *Human Language Resources* were internationally recognized by the award of the prestigious *Antonio Zampolli Prize* to *Oriental COCOSDA* at the recent *LREC 2012* meeting in Istanbul, 23-25 May 2012. We owe them our congratulations. The integrative international activities of *Oriental COCOSDA* have a vitality and productivity which will produce highly significant results in theory and in application in the great variety of languages of Asia, and will continue feeding into empirical work in speech and text sciences and technologies throughout the world.

Selected further reference

The following list is highly selective, and partly oriented towards linguistics, partly towards technology, with a slight slant towards the languages of Eastern Asia.

With a rapidly developing field like this there is currently no one publication which can cover all aspects of *Human Language Resources* and *Language Documentation*, but there are many online tutorials, guidelines and international standards. Consequently there is no substitute for Internet search, particularly for institutions such as the *International Standards Organization* (*ISO*), the *World-Wide Web Consortium* (*W3C*), the *Text Encoding Initiative*.

Gibbon et al. (1997) deal with basic techniques for creating resources for speech technology, including evaluation methods; Gibbon et al. (2000) extend the scope to multimodal systems. The SAMPA alphabet, which is frequently used in speech transcription in speech technology, is detailed in these publications. Itahashi et al. (2010) and Tseng (2009) pay particular attention to the requirements of Eastern Asian languages and to spontaneous speech, respectively. Koehn (2010) on statistical machine translation and McEnery & Hardie (2012) discuss specific aspects of corpus processing for linguistics and technology, and Atkins et al. (2008) and Van Eynde et al. (2000) provide background to lexical resources. Witt et al. (2009) describe the formal document description languages required for archiving and dissemination of *HLR*/*LD*, and Mehler et al. (2012) provide an overview of all aspects of speech and text technologies, including development, resources and evaluation.

Aidaoguangci (nickname). 2010. Putonghua shengyun peihe biao [A chart for the combination of Mandarin initials and finals]. Forum *Iask*. http://ishare.iask.sina.com.cn/ (Last consulted 2012-09-16.)

Atkins, Beryl T. Sue, and Michael Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford & New York: Oxford University Press.

Gibbon, Dafydd. 2007. Formal is natural: toward an ecological phonology. *Proceedings of ICPhS XVI*, 83-88. Saarbrücken: Universität Bielefeld.

Gibbon, Dafydd. 2012. ULex: new data models and a mobile environment for corpus enrichment. *Proceedings of the Language Resources and Evaluation Conference 2012*, 3392-3398. Paris: ELRA/ELDA.

Gibbon, Dafydd, Roger Moore, and Richard Winski. (eds.) 1997. *Handbook of Standards and Resources for Spoken Language Systems*. Berlin & New York: Mouton de Gruyter.

Gibbon, Dafydd, Inge Mertins, and Roger Moore. (eds.) 2000. *Handbook of Multimodal and Spoken Dialogue Systems Resources*, *Terminology and Product Evaluation*. Dordrecht & Boston: Kluwer.

Gibbon, Dafydd, Pramod Pandey, D. Mary Kim Haokip, and Jolanta Bachan. 2009. Prosodic issues in synthesising Thadou, a Tibeto-burman tone language. *Proceedings of InterSpeech 2009*, 500-503.

Gibbon, Dafydd, Moses Ekpenyong, and Eno-Abasi Urua. 2010. Medefaidrin: resources documenting the birth and death language life-cycle. *Proceedings of the Language Resources and Evaluation Conference 2010*, 2702-2708. Paris: ELRA/ELDA.

He, Hu. 2004. *Xinbian Putonghua Xunlian yu Ceshi Jiaocheng* [*Textbook for Mandarin Training and Testing*]. Lanzhou: Lanzhou University Press.

Itahashi, Shuichi, and Chiu-yu Tseng. (eds.) 2010. *Computer Processing of Asian Spoken Languages*. Los Angeles: Consideration Books.

Jansche, Martin. 1998. A two-level take on Tianjin tone. *Proceedings of the 10th European Summer School in Logic*, *Language and Information*, *Student Session*, 162-174. Saarbrücken: Universität des Saarlandes.

Koehn, Philipp. 2010. *Statistical Machine Translation*. Cambridge & New York: Cambridge University Press.

McEnery, Tony, and Andrew Hardie. 2012. *Corpus Linguistics: Method*, *Theory and Practice*. Cambridge & New York: Cambridge University Press.

Mehler, Alexander, Laurent Romary, and Dafydd Gibbon. (eds.) 2012. *Handbook of Technical Communication*. Berlin: De Gruyter Mouton.

Tseng, Shu-Chuan. (ed.) 2009. *Linguistic Patterns in Spontaneous Speech*. Taipei: Institute of Linguistics, Academia Sinica.

Van Eynde, Frank, and Dafydd Gibbon. 2000. *Lexicon Development for Speech and Language Processing*. Dordrecht & Boston: Kluwer.

Witt, Andreas, and Dieter Metzing. 2009. *Linguistic Modeling of Information and Markup Languages: Contributions to Language Technology*. Dordrecht & London: Springer.

自然語言資源之於學術研究與技術研發應用的重要性

Dafydd Gibbon

畢勒費爾德大學／口語語音資料庫協調暨標準化國際委員會

語言科學與科技研究很重要的關鍵課題是如何爲學術研究與發展提供適當而且可機讀的研究基礎，包含高品質的語料庫資源、語料分析的方法、詞彙庫與語法的建立、與軟體工具的開發。典藏與記錄語言學在過去十五年來已經發展成爲應用語言學門的新次領域，特別是有關瀕危語言的語言記錄與其維護。對於自然語言科技來說，自然語言資源有文字、語音、與手勢等型態。爲網路訊息散播與擷取、機器翻譯、與自動語音辨識與合成等相關領域的應用提供必要的基礎。本文詳述自然語言資源發展的緣起與動機。並且強調以可理解的語言符號架構爲基本設計的資料與語料庫分析必須以計算模型呈現才能提供實證的、具體的語言學研究面向。本文實際以音節語法、聲調分析、與詞彙學等分析方式討論文獻手稿的記錄典藏方法與連讀變調計算模型的研究型態。

關鍵詞：典藏記錄語言學，自然語言科技，語料庫語言學，詞彙學，語音學

《語言資訊和語言類型》
第四屆國際漢學會議論文集
臺北：中央研究院，2013 年

基於大規模中文樹庫的漢語句法知識獲取研究*

詹衛東
北京大學

基於已經標注好的百萬字級漢語樹庫，可以抽取不同層面上的句法知識，為中文信息處理、漢語句法結構的深入研究、以及對外漢語語法教學提供參考信息。對此，本文主要分為三個層次從點到面地展開探討。一是考察具體的一類句法結構的內部結構特徵和分布特徵；二是對具有共性的一大類結構（以非中心擴展結構和非同類成分並列結構為例）的考察；三是對全部句法規則集中右部同型左部根節點不同的潛在歧義結構的考察。

關鍵詞：中文樹庫，句法結構，歧義，非中心擴展結構，非同類並列結構，分布

1. 引言：樹庫加工與利用概述

自 1990 年代以來，樹庫加工及應用在語料庫語言學和自然語言處理領域一直是受到相當重視的研究方向（Marcus et al. 1993，Abeillé 2003，Xue & Xia 2000，周強 2004，Huang et al. 2000）。除用於信息處理技術外，從樹庫中獲取句法知識，也可以爲句法本體研究以及語言教學提供參

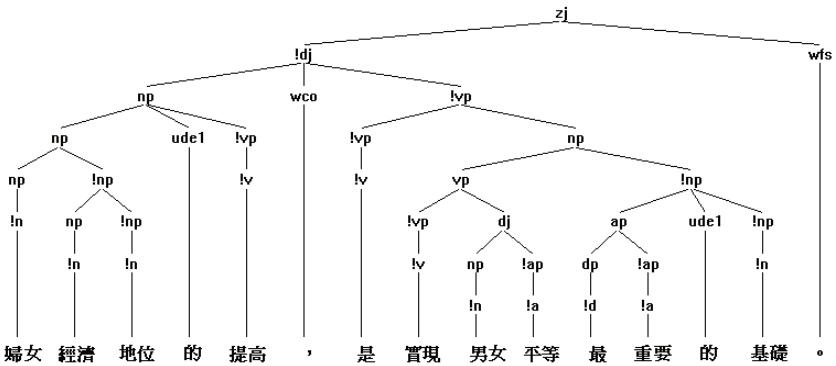
* 本文的研究工作得到霍英東基金項目「大規模中文樹庫構建及其在對外漢語教學中的應用」（項目號：111098）和國家社科基金項目「語言知識資源的可視化技術研究」（項目號：12BYY061）資助。匿名審稿人對本文初稿提出了中肯的修改意見。在此一併致謝。

考。本文介紹近年來北京大學中文系樹庫研究小組在這方面所做的一些工作。北京大學現代漢語樹庫加工採用人機結合的方式：先由程序對原始語料進行斷句、分詞、詞性標注、句法結構分析等處理（這些環節都會碰到不少如何界定及規範的問題，限於篇幅，本文對此不展開討論），然後由人在樹圖編輯軟件環境中逐句進行檢查，修改程序自動分析的錯誤，得到最終的標注了正確語法信息的樹庫。流程可簡要示意如下：

流程	語料示例
(1)原始語料	三、在經濟領域中的平等權利與重要作用 婦女經濟地位的提高，是實現男女平等最重要的基礎。中國政府爲改善和提高婦女的經濟地位作出了卓有成效的努力。……
(2)斷句處理	1. 三、在經濟領域中的平等權利與重要作用 2. 婦女經濟地位的提高，是實現男女平等最重要的基礎。 3. 中國政府爲改善和提高婦女的經濟地位作出了卓有成效的努力。……
(3)分詞和詞性標注	1. …… 2. 婦女/n 經濟/n 地位/n 的/ude1 提高/v ，/wco 是/v 實現/v 男女/n 平等/a 最/d 重要/a 的/ude1 基礎/n 。/wfs 3. ……
(4)句法結構標注	1. …… 2. (zj (!dj (np (np (np (!n (婦女)) !np (np (!n (經濟)) !np (!n (地位)))) ude1 (的) !vp (!v (提高))) wco (，) !vp (!vp (!v (是)) np (vp (!vp (!v (實現)) dj (np (!n (男女)) !ap (!a (平等)))) !np (ap (dp (!d (最)) !ap (!a (重要))) ude1 (的) !np (!n (基礎)))))) wfs (。))) 3. ……

句法結構標注的形式是在計算機中以括號方式標記[1] 在原始句子字符串上進行存儲的。上面流程中加工完成的語料示例第 2 句對應的直觀的樹結構圖如下：

[1] 具體的短語類標記和詞類標記可訪問 http://ccl.pku.edu.cn/doubtfire/Projects/Treebank_Tags.pdf 查詢。有關樹庫部分查詢功能的示例可訪問 http://ccl.pku.edu.cn:8080/WebTreebank/。



〈圖 1〉句法結構樹示例

（語料來源：中國政府白皮書 1994《中國婦女的狀況》）

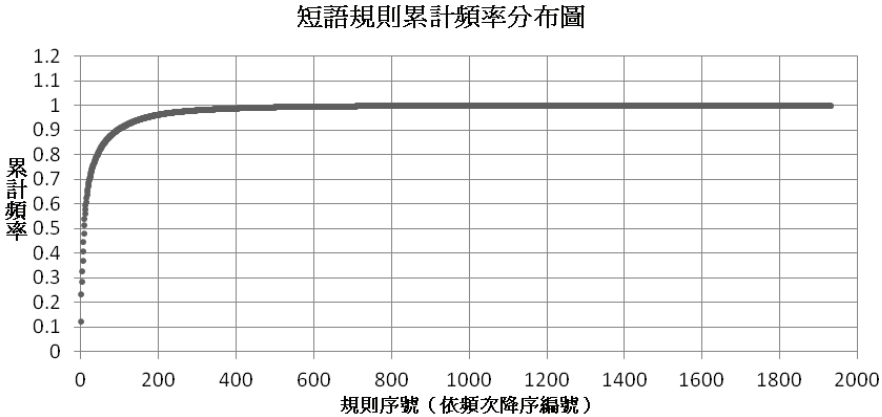
目前北大中文樹庫已經標注的語料規模爲 55,742 句，1,309,719 字，899,365 詞。語料類型包括語文課本 (56.85%)、句型例句 (13.29%)、新聞語料 (13.00%)、科技語料 (9.43%)、政府白皮書 (7.43%)。

從樹庫中可以很容易抽取短語規則。比如〈圖 1〉中「最重要的基礎」對應的規則是 np → ap ude1 !np。其中「最重要」的對應規則是 ap → dp !ap。規則中的「！」標記了其後的成分是該短語結構規則的中心成分。把樹庫中像這樣的規則全部抽取出來，按頻次降序排列，就可以得到如下表所示的現代漢語短語結構規則集：

〈表 1〉現代漢語短語結構規則及其頻次示例

編號	結構規則	結構類型	頻次
1	np -> !n	名詞結構	159778
2	vp -> !v	動詞結構	146250
3	vp -> !vp np	述賓式動詞結構	65751
4	dj -> np !vp	主謂式小句	58724
……	……	……	……
1929	zj -> !dp wfs	獨詞型整句	1
1930	zj -> dj !fj	複句型整句	1

按規則在樹庫中的頻率降序排列，計算規則的累計頻率，可以瞭解短語規則對語料的覆蓋率情況，下圖是短語規則累計頻率分布圖，按累計頻率降序排列後，前 95 條規則（佔 4.9%）覆蓋了樹庫中 90% 的語料，前 446 條規則（佔 23.1%）覆蓋了樹庫中 99% 的語料。剩下的 1,484 條規則（佔 76.9%）覆蓋剩下的 1% 的語料。[2]



〈圖 2〉短語規則累計頻率分布圖

上面這樣的統計數據，可以爲語言教學提供宏觀層面的參考。比如對句法結構可以像對詞的常用性分級一樣，區分不同的層級，安排更爲合理的教學順序。結構規則中近 80% 的低頻規則，有不少涉及到省略、轉指等相對複雜的語言現象，[3] 在語法教學的項目安排中可能就需要考慮更有針

[2] 我們也統計了美國賓州中文樹庫 5.0 版語料 (Xue & Xia 2000, Xue et al. 2005) 的情況。樹庫的規模是 43,385 個詞型 (type)，508,385 個詞例 (token)。規則種數 (type) 為 5,220 條，例數 (token) 為 537,504 條。其中前 116 條規則（佔 2.24%）覆蓋 90% 的語料，前 1,379 條規則（佔 26.42%）覆蓋 99% 的語料。剩下的 3,841 條規則（佔 73.58%）覆蓋剩下 1% 的語料。賓州樹庫的句法結構標注中很多是多分支樹結構，因此規則數目比較多。不過，儘管規則體系差別較大，但按規則百分比來看規則對語料的覆蓋率，兩個樹庫語料統計反映的情況是大致相當的。

[3] 比如「看你把閨女嚇得那個樣子」，「那個樣子」是 np，但出現在補語位置上，違反了一般的句法結構規則要求。再如「他在他父親的公司裡一直呆到他父親去世」。「他父親去

對性的策略。

以上扼要介紹了構建樹庫的流程，以及從樹庫中獲取短語結構規則知識。從樹庫中可以獲取的句法結構知識是多維立體的，除短語結構規則及其頻次外，還包括抽取帶詞性和頻度信息的詞表，兼類詞的分布統計，短語分布環境及其頻次，歧義短語結構及其頻次等等。在這些數據基礎上，還可以就研究者感興趣的問題，做專項信息提取和歸納。限於篇幅，本文分三個層次按照從點到面的順序介紹我們從樹庫中獲取不同類型的句法結構知識的情況。下文第 2 節是考察某一類短語在不同句法位置的差異（以名詞性短語 np 爲例）以及特定句法結構的內部構造特點和外部環境特點（以「把」字結構爲例）；第 3 節是考察漢語中違反中心擴展規約和並列條件的短語結構的情況；第 4 節是定量考察短語結構歧義的情況。

2. 基於樹庫的漢語短語結構個案考察

2.1 np 在主、賓、定、中等句法位置的差異考察

名詞性短語 np 在不同句法結構位置對應的結構規則如下（xp 代表任意短語）：

主語位置	dj → **np** !xp（規則中 xp 通常爲 vp，ap，np，dj 等）	
賓語位置	vp → !vp **np**（動賓）	pp → !p **np**（介賓）
定語位置	np → **np** !xp（定語 1）	np → **np** ude1 xp（定語 2）
中心語位置	np → xp !**np**（中心語 1）	np → xp ude1 !**np**（中心語 2）

世」是一個小句 (dj)，這裡做「呆到」的賓語，實際上是轉指一個時間，起到了相當於 tp（時間詞性短語）的作用。下文還有一些出現頻次為 1 的短語結構例子。這些例子相對於高頻的結構規則來說，某種程度上可以看做是「特例」。儘管母語者對這些組合例子一般來說是感到「習以為常」的，但從系統的角度來講，它們屬於應該特殊對待的研究和教學對象。

上面規則中加粗標記了 np 所在的相應的句法位置。樹庫中 np 在「主、賓、定、中」句法位置出現的總頻次爲 325,747，佔全部 np (392,925) 的 82.9%。np 在這些不同句法位置的寬度（按詞數計寬度）統計結果見〈表 2〉。

〈表 2〉np 在不同句法位置的寬度統計

句法位置	主語	介賓	動賓	定語 1	定語 2	中心語 1	中心語 2
實例數	68279	15247	65751	42291	15372	84491	34316
寬度種數	37	40	60	46	29	36	26
最小寬度	1	1	1	1	1	1	1
最大寬度	54	60	113	112	54	63	53
均值	1.96	2.85	3.23	1.52	1.83	1.69	1.64
方差	3.93	9.65	12.18	4.22	3.74	2.97	2.17

說明：我們把定中結構分爲兩種情況，甲：「定–中」；乙：「定–的–中」。定語 1 和中心語 1 爲甲類中的「定」和「中」位置。定語 2 和中心語 2 爲乙類中的「定」和「中」位置。

如果把結構內包含詞數作爲評價結構複雜性的一個指標的話，np 在定語和中心語位置上的複雜性差異不是很大。np 充當甲類定中結構的定語和中心語的頻次顯著多於乙類定中結構的定語和中心語。原因是乙類定中結構的定語和中心語都可以由非 np 類短語（如 vp，ap）充當，而甲類定中結構的定語和中心語則主要由 np 充當。

np 在主語位置和賓語位置的複雜度差異較明顯，主語位置上的 np 平均詞長不超過 2 個詞，顯著低於介賓和動賓位置上的 np。此外，主語位置上 np 寬度的方差值也顯著低於介賓和動賓位置上的 np，說明主語位置上 np 比賓語位置上 np 的寬度相對更集中。

下面進一步考察 np 在主、賓位置上的內部結構是否存在顯著差異。

〈表 3〉主、賓語位置上的 np 的內部結構及其頻次

語結構規則	主語位置			介賓位置			動後賓語位置		
	序號	頻次	頻率	序號	頻次	頻率	序號	頻次	頻率
np -> !rn	**1**	**23906**	**35.01%**	**2**	**2436**	**15.98%**	**5**	**4788**	**7.28%**
np -> !n	2	15561	22.79%	1	4218	27.66%	1	18852	28.67%
np -> np !np	3	7638	11.19%	3	2227	14.61%	3	7188	10.93%
np -> np ude1 !np	4	3449	5.05%	4	1180	7.7%	4	5313	8.08%
np -> qp !np	**5**	**2630**	**3.85%**	**5**	**775**	**5.08%**	**2**	**8313**	**12.64%**
…	…	…	…	…	…	…	…	…	…
總計	type 數：172 token 數：68279			type 數：128 token 數：15247			type 數：224 token 數：65751		

按頻次降序對不同的結構排序後得到上表的統計結果，從高頻結構的情況可以看到，主語位置的 np 跟介賓位置的 np 性質更爲接近，動後賓語位置上的 np 跟二者相差較大。這在一定程度上印證了以往漢語研究中所觀察到的現象，即舊信息傾向居動詞前位置，新信息傾向居動詞後位置。〈表 3〉中「np → qp !np」規則對應的主要是漢語中的一般的「數＋量＋名」結構，這種結構的 np 在主語和介賓位置的出現頻率都排在第五位，而在動詞後賓語位置則居第二位。這一點，通過〈表 4〉統計的數據可以更清楚地看出。〈表 4〉是詞長爲 3 的 np 在主語、動賓、介賓位置上的內部組成情況。根據〈表 2〉的數據，主語、動賓、介賓三個位置的 np 詞長均值比較接近 3，因此我們重點統計了詞長爲 3 的 np 在這三個位置的內部構成情況。

〈表 4〉詞長為 3 的 np 在主語、動賓、介賓位置上的內部組成情況及其頻次

詞結構規則	主語位置			介賓位置			動後賓語位置		
	序號	頻次	頻率	序號	頻次	頻率	序號	頻次	頻率
np -> m q n	**3**	**760**	**10.05%**	**4**	**180**	**8.26%**	**1**	**2771**	**23.83%**
np -> rn ude1 n	1	1091	14.43%	1	260	11.93%	2	1169	10.05%
np -> n ude1 n	4	730	9.66%	3	217	9.96%	3	853	7.33%
np -> rb q n	**2**	**977**	**12.92%**	**2**	**221**	**10.14%**	**4**	**689**	**5.92%**
np -> a ude1 n	7	161	2.13%	5	117	5.37%	5	666	5.73%

結構規則「np → m q n」是「數＋量＋名」組合，「np → rb q n」是「指示詞＋量＋名」組合。前者一般對應語義上的不定指成分，後者則對應定指成分。在動後賓語位置上，不定指性 np 遠多於定指性 np。而在動前的主語位置和介賓位置，情況則顛倒過來，不過，在動前位置，兩類 np 的數量差異沒有在動後賓語位置相差得那麼大，這主要有兩方面的原因，一是漢語允許「無定 np 主語」（魏紅、儲澤祥 2007），二是形式上的無定 np，在語義上也可以表達定指義或者類指義，如「一個人毀壞了別人的東西，應不應該賠償？」中的「一個人」是無定形式的 np，用於主語位置，語義上並不是表達非定指，而是表達類指。總的來說，從樹庫中獲得的句法數據跟以往漢語研究中從語用角度所觀察到的現象：漢語中舊信息傾向居動詞前位置（主語位置 np 和介賓位置 np 都在謂語動詞前），新信息傾向居動詞後位置 (LaPolla 1995) 是非常吻合的。

2.2 vp 在「把」字結構中的內部構造以及「把」字結構整體分布環境考察

漢語語法學中傳統上關於「把」字結構，即「把 xp vp」中的 vp 的認識主要是它由複雜動詞詞組充任，不能僅僅是動詞的簡單形式。這樣才能滿足整個結構表達「處置」或「致使」語法意義的需要（比如北京大學中

文系現代漢語教研室編寫的《現代漢語》教材「虛詞」章在介紹「把」字結構時就是這樣說明的）。下面我們通過從樹庫中抽取「把」後 vp 實例以及 vp 內部結構規則的方式，進一步來考察這個結構中的 vp 具有哪些結構上的具體特點。

〈表 5〉「把＋xp＋vp」中 vp 的寬度考察

寬度	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**
頻次	864	551	463	235	155	85	58	28	17	13
寬度	**11**	**1**	**13**	**14**	**15**	**17**	**16**	**19**	**22**	**42**
頻次	12	12	8	4	3	2	1	1	1	1

〈表 5〉對「把」字結構中 vp 的寬度進行統計的結果顯示，「把」後 vp 主要由複雜動詞詞組構成，平均寬度（詞長）爲 3.71 個詞，多於全部介詞性短語 (pp) 後的 vp 的平均詞長（3.50 個詞）。這個統計結果佐證了以往人們對「把」後 vp 結構要求具有一定複雜性的語感描寫。同時，〈表 5〉的統計結果也顯示，「把」後 vp 也有單個動詞（詞長爲 1）的情況，北大樹庫中有 12 個這樣的例子，具體的動詞是「公開、分解、抽象化、形式化、平分、消滅、發揚光大、神化、相加、除外、置之度外、還原」。

在考察了「把」後 vp 的寬度之後，下面〈表 6〉給出了「把」後 vp 的具體規則分類情況。

〈表 6〉「把＋xp＋vp」結構中 vp 的構造類型及示例

構造類型	數量	結構規則	示例
述賓式 vp	999 (39.74%)	vp → !vp np vp → !vp sp vp → !vp qp vp → !vp mp …	(把 x) 交給 新幹部 放在 桌子上 放在 第一位 砍去 一半 …

述補式 vp	895 (35.60%)	vp → !v v vp → !v a vp → !v ude3 ap ….	(把 x) 扔 掉 清理 乾淨 布置 得 非常漂亮 …
狀中式 vp	354 (14.08%)	vp → dp !vp vp → pp !vp vp → ap !vp …	(把 x) 也 拋出來 在幾個工作人員中 分配一下 直接 倒到喉嚨裡去 …
附加式 vp	187 (7.44%)	vp → !vp ule vp → !v uzhe …	(把 x) 摔壞 了 珍藏 著 …
連謂式 vp	58 (2.31%)	vp → !vp vp vp → !vp wco vp …	(把 x) 帶回家 放好 變成電信號，再加以放大 …
其他	21 (0.84%)	vp → !v vp → c !vp …	(把 x) 公開 一 剝 …
合計	2514 (100%)	48 種	

從〈表 6〉可以看出，「把」後的 vp 以述賓式構造類型爲最多，這個特點以往語法學中討論「把」字句時注意得不夠。在討論漢語「把」字句的文獻中，有不少是以「把」後的所謂「保留賓語」爲題展開研究的，即從某種程度上來說認爲「把」後 vp 再帶賓語是一種特殊現象。儘管「保留賓語」確實有其自身的特點，但語料調查的結果也顯示，「把」後 vp 主要的結構類型就是述賓結構（包括帶準賓語的情況）。「把」字結構後出現賓語是該結構的用法特點之一。

對一個結構，除考察其內部構成外，還可以看它所處的上下文環境的特點。〈表 7〉列出了「把＋xp＋vp」結構的分布環境的類型。本文關於「分布環境」的定義是：一個結構體 (S) 的分布環境是一個三元組。設樹

(T) 的根節點是 S 的父節點，則 S 的分布環境由 S 的父節點、S 的左鄰節點、右鄰節點三個項目構成。其中左鄰節點和右鄰節點都可以爲空。

〈表 7〉「把＋xp＋vp」的分布環境的類型統計
（按照父節點的類型不同分組）

父節點	數量	左鄰節點	右鄰節點	示例
vp	1504 (59.86%)	dp vp vp wco - …	- - - vp …	連忙 **把它拾起來** 走過去 **把口琴還給錫海** 爬上樹去，**把小鳥放回窩裡** **把門打開** 放狗出去 …
dj	676 (26.89%)	np np wco …	- - …	你 **把它吃了** 古代的埃及人和中國人，**把它用做藥物** …
fj	237 (9.43%)	dj wco …	- …	他一隻手抓住繩子，**把另一隻手伸給水中的孩子**。 …
zj	36 (1.43%)	- …	wfs …	**把瓶子放在桌上**。 …
np	29 (1.15%)	- - …	ude1 !np ude1 …	**把人生溶進偉大事業** 的 人 **把咖啡喝光** 的 …
# [4]	20 (0.80%)	-	-	**把桌子拿出去**
tp	7 (0.28%)	- …	f …	**把羊肉和羊骨粉碎** 後 …
pp	4 (0.16%)	P …	- …	從 **把水放在爐上** 到水開 …
合計	2514 (100%)	82 種		

[4] # 表示父節點為空，這裡意味著「把」字結構獨立成句，佔據一行，且末尾沒有標點。

〈表 7〉反映了「把」字結構的主要用法中直接做謂語是排在第二位的。排在第一位的是「把」字結構跟其他成分組合成更大的 vp，佔到近 60%，頻率是第二位的兩倍多。也就是說，現實中的「把」字結構，其前後往往會有其他的謂詞性成分共現。這個特點，在有關「把」字句的對外漢語教學中應引起注意。當以「把」字句爲視點去看「把」字結構時，往往容易把「把」字結構 vp 直接放在謂語位置上，同時把整個「把」字句跟被字句、主動賓句式放在一個層次上關聯起來，但如果以短語結構的視點去看「把」字結構，會更加全面地看到該結構所在的不同句法位置以及頻率上的差異。

3. 現代漢語非中心擴展結構與非同類並列結構考察

通常情況下，一個短語結構的功能類跟其中心成分的功能類是相同的，這樣的短語規則是所謂的符合「（中心）擴展條件」的規則（記作 HE 規則）。兩個成分構成並列結構，則兩個成分應屬同類短語，這樣形成的並列結構是所謂符合「並列條件」的規則（記作 CC 規則）。當代形式語法理論一般也都強調短語結構規則從形式上應該符合 HE 規則和 CC 規則的要求。沈家煊 (2007) 引用 Lyons (1968:331) 的論述：「N 和 NP 之間，V 和 VP 之間都存在一種必不可少的 (essential) 的聯繫，對哪種語言都一樣。……NP 和 VP 不僅僅是幫助記憶的符號，而是分別表示句法成分 NP 必定是名詞性的，VP 必定是動詞性的，因爲兩者分別以 N 和 V 作爲其必需的主要成分。」他接著說，如果有哪位語言學家提出諸如「NP→V＋VP，NP→V，VP→T（冠詞）＋N」的規則，「那不僅是有悖常情的，在理論上也是站不住的。」這些話是就「擴展條件」而言的，但是也適用於「並列條件」，提出有「NP 和 VP」這樣的並列結構也是有悖常情的，理論上站不住的。

但從樹庫標注的情況來看，我們認爲，實際語料中也有少量的短語結

構，其功能類跟中心成分的功能類是不同的，同時也有少量的並列結構，並列的兩項屬於不同功能類的短語，至少在表層結構形式上是如此。這樣的結構規則我們分別稱爲非中心擴展規則（記作 NHE 規則）和非同類並列規則（記作 NCC 規則）。下面是樹庫中抽取的 HE 規則和 NHE 規則、CC 規則和 NCC 規則各自所佔的比例情況。

〈表 8〉樹庫中 HE 規則、NHE 規則、CC 規則、NCC 規則的數量統計

規則類別	規則種數 (type)	規則例數 (token)	結構示例	示例
全部規則	1,930	1,318,488	dj → np !ap	人 多
HE 規則	1672 (86.63%)	1,048,669 (97.20%)	ap → dp !ap	最 冷
NHE 規則	258 (13.37%)	**30,252 (2.80%)**	np → sp !vp	體內 分布
CC 規則	61 (52.59%)	26,220 (96.37%)	ap → !ap c ap	光榮 而 艱巨
NCC 規則	55 (47.41%)	**987 (3.63%)**	ap → !ap c vp	無知 與 疏忽

說明：上面表中 NHE 及 NCC 規則的統計數據是程序根據規則形式自動判別的，數據會有一定誤差。不過，我們的目的並不是統計出精確的數據做量化分析，而是通過這種方式從實際語料中發現 NHE 規則和 NCC 規則的類型和實例。很顯然，從實例頻次 (token) 的對比來說，NHE 規則和 NCC 規則相對於 HE 規則和 CC 規則（常規情況）來說，都是絕對少數。換言之，真實語料中的大部分短語組合都是符合「中心擴展條件」和「並列結構條件」的，但是我們想強調的是，也確實存在不符合的實例，儘管比例不高，但違反中心擴展條件和並列結構條件的實例也並非特例。下文即通過具體實例的展示和分析來說明語言使用中存在這樣的組合是合理的。

〈表 9〉NHE 規則的內部成分、中心成分考察

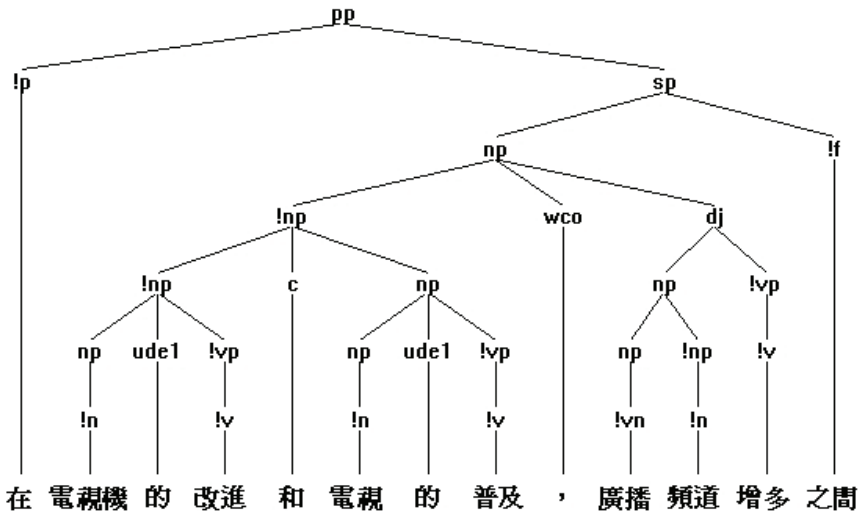
序號	內部構成/中心成分	NHE 規則	示例
1.	跟「的」相關的 NHE	np → np 的 !vp np → pp 的 !vp np → sp 的 !qp	時間 的 推移 在電子產品 可靠性方面 的 應用 他們中 的 三個
2.	跟其他助詞（似的、地）相關的 NHE	ap → !np 似的 ap → !dj 似的 dp → !qp 地 dp → !vp 地	雪片 似的 他是這個地方的主人 似的 一寸一寸 地 有秩序 地
3.	ap 擴展爲 np	np → qp !ap	一點 清涼
4.	dj 擴展爲 np	np → qp !dj	這種 再狹窄發生率降低
5.	qp 擴展爲 np	np → np !qp	這 三 本

上表的例子中絕大多數情況都是通過結構助詞，系統地改變結構的性質，比如「的」「地」等結構助詞，可以系統地使得結構整體的功能不同於其中中心成分的功能。此外，漢語中也存在結構功能不需要標記成分的幫助，直接發生功能轉換的情況，比如例 3、4、5 都是這類情況。陳述性成分、修飾性成分都直接轉爲指稱性成分（關於指稱、陳述概念，參見朱德熙 1982）。

〈表 10〉NCC 規則的內部成分考察

序號	並列項	NCC 規則	示例
1.	ap – vp	ap → !ap c vp	對朋友誠實 和 幫助老人
2.	vp – dj	vp → !vp c dj	地震 與 火山噴發
3.	np – dj	np → !np wco dj	電視機的改進和電視的普及，廣播頻道增多
4.	np – vp	np → !np wco vp	一間紅瓦灰牆的小屋，一排白漆的大柵欄，或許還有三五個人影（，眨眼就消失了。）
5.	dj – tp	dj → !dj c tp	我應該今天開始 還是 明天
6.	ap – dp	ap → !ap c dp	（失戀以後，會是）頹廢 或 奮力

表中第 3 行是 np 跟 dj 構成並列結構，此例的句法結構如下面〈圖 3〉所示。



〈圖 3〉一個 NCC 規則的結構樹圖

上面例子引出的一個核心問題是：漢語中主謂結構是陳述性成分還是指稱性成分？如果是陳述性成分，則主謂結構跟 np 並列的時候，就違反了「並列結構條件」。如果是指稱性成分，則不違反並列結構條件。但是，當它不違反「並列結構條件」的時候，就要進一步追問，主謂結構的中心語又是什麼呢？如果主謂結構的中心語是謂語 vp，那麼，例中作爲指稱性成分的主謂結構，其功能顯然跟 vp（「增多」）的功能是不一致的，前者是起指稱作用，而後者一般應該是起陳述作用。這樣，就又違反了「中心擴展條件」。看起來，在短語結構的組合過程中，如果在一個層次上要遵守「並列結構條件」，就可能在另一個層次上違反「中心擴展條件」，二者並不總是能兼顧的。實際上，語料中可以觀察到的表層語言現象是：漢語中的主謂結構既可以用於陳述表達功能，也可以用於指稱的表達功能，在表層句法結構上就是表現爲主謂結構可以做主、賓語。

如果違反「並列結構條件」和「中心擴展條件」的 NCC 規則和 NHE 規則都是不可避免的，那很自然會想到一個問題，爲什麼在常規的符合中心擴展條件和並列結構條件的組合之外，語言系統會有「非中心擴展」「非同類並列」的結構存在呢？我們的看法是，這是語言使用中的「簡約」要求使然。語言使用中的「簡約」（或者說經濟）要求使得人們在使用句法結構規則時，常常省略其中的成分，即以部分結構成分代行整體之功能，其中最突出的例子，莫過於漢語中的「的」字結構了，例如：「成套的書」這個組合是常規 np，「書」是中心成分，整個結構符合中心擴展條件，但實際使用時，人們完全可以用「成套的」來替代「成套的書」，語料中後者往往是已知的背景信息，因而中心成分很容易省略。省略後得到的「成套的」這個結構，其中心成分是哪一個，如何讓中心成分跟結構整體功能保持一致，就要讓理論語言學者大傷腦筋了。它造成的「麻煩」不僅僅停留在短語結構規則集多出了 NHE 規則，一個連帶的後果是，非常規結構規則融入常規組合規則中，由此形成的漢語句法結構規則，用於計算機的自動句法分析，會造成更多的系統性的潛在結構歧義。通過考察樹庫中句法組合歧義的分布和數量，可以更深入地瞭解這種影響。

4. 現代漢語句法結構歧義情況的定量考察

本節討論如何從樹庫中抽取有歧義的句法結構。句法結構的歧義有不同的類型，歧義程度也有高低之分。這裡關於句法結構歧義情況的統計和考察僅僅是在組合規則層次上看歧義，還沒有深入到具體的語言實例層次。從樹庫中抽取規則時，可以考慮兩個層面的組合規則，一是通常的短語結構規則（如上文〈表 1〉所示），另一種是以詞性標記串來表達組合規則。[5] 如 vp → p n v v n。對這兩種組合規則，都可以統計規則右部同型而左部根節點不同類的情況。如下〈表 11〉所示：

[5] 本文只考慮了 2 到 8 個詞形成短語結構的情況。

〈表 11〉規則右部同型、左部根節點不同形成的歧義組合示例

	規則左部根節點	規則右部組合模式	實例
短語組合規則	np	qp !ap	兩個 不同
	dj	qp !ap	兩個 不同
詞類組合規則	dj	n v n	幹部 領導 群眾
	np	n v n	政府 領導 幹部
	vp	n v n	科學 種植 西瓜

表中短語組合規則對應的實例就短語本身來看是確實有歧義的，不過在具體語境中往往因上下文的影響而只有一種理解。比如：

例 1　新版本跟上一個版本相比有**兩個不同**，……

例 2　論學歷**兩個不同**，論能力兩個完全一樣。

例 1 中的「兩個不同」應解作 np，即「兩個不同之處」的簡省說法。例 2 中的「兩個不同」則應解作 dj，「兩個」跟「不同」之間是主謂結構關係，「兩個」是「兩個人」的簡省說法。

樹庫中更多的歧義組合是像〈表 11〉中「詞類組合規則」對應的實例所反映的情況，即實例本身並沒有明顯的歧義，但在詞類範疇（或短語範疇）層面，則可以有不同的結構解讀方式，這是計算機在分析句子時會碰到的主要歧義問題。比如「幹部領導群眾」和「政府領導幹部」理論上是可以有歧解的，但兩例各自都沒有明顯的歧義，前者的結構是主謂，短語類應歸爲 dj，後者的結構是定中，短語類應歸爲 np。把這兩例放在更大語境中，可更好地體會它們的區別。例如：

例 3　上級領導下級，**幹部領導群眾**。

例 4　我縣近期組織對**政府領導幹部**進行群眾滿意度調查。

這類「歧義」例子不是針對人而言的歧義，但是對計算機處理來說，則是貨真價實的歧義現象。下面我們就來統計樹庫中同型組合形成的根節點不同的規則總體是一個什麼情況。這可以從一個側面反映按照目前的短語結構標注體系進行短語句法結構組合分析時隱含歧義的程度。[6]

〈表 12〉短語組合規則中右部同型、根節點不同的規則統計[7]

	短語組合規則	同型歧義短語組合		百分比
Type 數	1,930	同型短語組合數	296	15.34%
		同型短語組合規則數	670	34.72%
Token 數	1,318,488		477,142	36.19%

〈表 13〉詞類組合規則中右部同型、根節點不同的規則統計

	詞類組合規則	同型歧義的詞類組合		百分比
Type 數	124,611	同型詞類組合數	3,932	3.16%
		同型詞類組合規則數	8,263	6.63%
Token 數	542,153		240,542	44.47%

[6] 這種統計同時也可以作為檢查語料標注一致性的一種手段。限於篇幅，本文對此不展開討論。需要說明的是，下文給出的統計數據中因此也可能存在一定偏差，即有的同型規則有可能是標注錯誤造成的不一致問題，而非真正的同型歧義組合。

[7] 這裡的右部同型規則只計算了兩分支以上的規則，沒有計算單分支規則（形如 np → !n 這樣的規則）。如果把單分支規則算在內，則同型短語組合數為 327（佔比 16.93%），同型短語組合規則數為 741（佔比 38.43%）。以規則實例 (token) 計，共 819,440 個組合涉及同型組合歧義（佔比 62.15%）。我們按照同樣方式統計了賓州大學中文樹庫的同型短語組合歧義情況，賓州樹庫短語規則 (type) 數為 5,220 條。其中兩分支以上的同型短語 316 個（佔比 6.05%），同型短語組合規則數為 724（佔比 13.85%）。以規則實例 (token) 計，共 253,774 個組合涉及同型組合歧義（佔比 47.21%）。從 type 數來看，賓州中文樹庫的同型歧義情況要顯著低於北大中文樹庫。從 token 數來看，同型歧義組合的比例則高於北大中文樹庫。這大體反映了一方面賓州中文樹庫標注具有更好的內部一致性，另一方面賓州樹庫的短語標記的區分度要更大一些（賓州樹庫的短語標記共 25 個，北大中文樹庫是 17 個）。從這樣的對比來看，北大中文樹庫的短語標記體系還有進一步細分的必要，可以通過短語類的細分來降低同型短語組合規則的比例。此外，語料標注的內部一致性也還需要提高。

下面就進一步從不同的角度來看同型歧義組合中歧義程度相對比較高的情形。這裡主要考慮了三個角度，一是看一個同型組合能形成幾種不同的短語類，即統計同型組合構成的不同根節點數量（以下簡稱「根數」）的多少；二是看同型組合的頻次高低；三是把一個同型組合形成不同短語類看做是一個隨機事件，計算這個隨機事件的信息熵值，比較熵值的大小。

(一) 從同型組合形成的根節點個數多少來看歧義程度

〈表 14〉同型短語組合的不同根數頻次分布

根數	頻次	百分比
2	229	77.36%
3	57	19.26%
4	9	3.04%
5	1	0.34%
合計	**296**	100%

〈表 15〉同型詞類組合的不同根數頻次分布

根數	頻次	百分比
2	3558	90.49%
3	350	8.90%
4	24	0.61%
合計	**3932**	100%

vp 跟 ap 短語組合可能形成根節點數最多達到 5 個，具體每種組合的頻次分布如下表所示。

〈表 16〉根數為 5 的短語組合及其頻次分布示例：vp＋ap 組合

規則左部根節點	規則右部組合模式	頻次	示例
dj	vp !ap	462	發展 很快
ap	vp !ap	85	看著 非常舒服
fj	vp !ap	6	不是星期日 還不著急呢
np	vp !ap	1	（聯繫） 教學 實際
sp	vp !ap	1	過橋 不遠
	合計	**555**	

〈表 17〉根數為 4 的短語組合及其頻次分布示例：ap＋vp 組合

規則左部根節點	規則右部組合模式	頻次	示例
vp	ap !vp	3061	認真 學習
dj	ap !vp	117	快樂 在等待我們
np	ap !vp	33	不同 解釋
fj	ap !vp	1	由於恐懼 而逆來順受
	合計	**3212**	

〈表 18〉根數為 4 的詞類組合及其頻次分布示例：a＋m＋q 組合

規則左部根節點	規則右部組合模式	頻次	示例
qp	a m q	181	近 一千億 元
ap	a m q	93	少 三 票
tp	a m q	7	近 幾十 年
dj	a m q	6	寬 九 米
	合計	**287**	

(二) 從同型組合的頻次高低看歧義程度

同型短語組合中有 96 種頻次超過 200。同型詞類組合中有 139 種頻次超過 200。下面分別列出同型短語組合和同型詞類組合中頻次前 5 位的組合，包括它們能構成的根節點數量，具體是哪些短語類，頻次信息以及示例。

〈表 19〉同型短語組合頻次最高的前 5 個組合

短語類組合	根數	根節點	頻次	合計	示例
!vp np	2	vp	65752	65756	有 天大的困難
		dj	4		是我 聽見的
np !vp	3	dj	58724	60842	我們 正在嘗試
		vp	1448		科學 種植西瓜
		np	670		科學 研究

np !np	2	np	41452	41818	中國 國民經濟
		dj	366		總人口 一千萬人
!vp vp	2	vp	25310	25340	打算 研製新產品
		fj	30		沒有革命的理論 就沒有革命的運動
mp !q	2	qp	19584	19613	兩千多 個
		tp	29		二〇〇六 年

〈表 20〉同型詞類組合頻次最高的前 5 個組合

詞類組合	根數	根節點	頻次	合計	示例
m q	2	qp	18487	18516	三 輛
		tp	29		第五 年
v n	2	vp	14005	16941	送 朋友
		np	2936		輔導 教材
v v	2	vp	16507	16538	推 出去
		dj	31		會談 擱淺
n n	2	np	12979	13015	人民 群眾
		dj	36		小名 鐵蛋
n f	2	sp	5059	5133	樹 後
		tp	74		晚飯 後

〈表 16-20〉分別給出了根數最多和頻次最高的同型短語類組合和同型詞類組合及其實例。從這兩個角度評價同型組合的歧義程度高，有一個明顯的問題，就是同型組合形成的不同短語類頻次分布可能並不均勻，比如〈表 16〉的「vp !ap」的各種組合的頻次就相差很大，〈表 19〉的「!vp np」兩種組合的頻次相差更是懸殊。很顯然，這樣的歧義組合，其歧義程度並不算高。爲了描述同型歧義組合形成不同短語類的頻次分布的均勻程度，可以引入信息熵的概念。

(三) 從同型組合的信息熵值大小看歧義程度

如果把像「vp !ap」這樣的組合形成不同的根節點看做一個隨機事件，就可以用隨機變量的信息熵值來度量一種同型組合形成不同根節點的分布均勻程度，熵越大，分布越均勻，相應的，歧義程度也越高。反之，則分布不均勻，歧義程度也就低一些。熵值計算公式爲：

$$H(S) = -\sum p_i \log_2 p_i$$

公式中 p_i 表示隨機變量 S 可能的取值中第 i 個值出現的概率。對於同型歧義組合規則來說，它組成爲不同短語類的概率可以用各個組合的頻次來估計，比如「vp ap」短語組合爲 dj 的概率爲 462/555，即 0.83。依此計算出各組合規則的概率後，帶入上面的公式，就可以求得「vp ap」短語組合的熵值爲：0.7383。類似地，可以計算得到〈表 17〉中的「ap !vp」短語組合的熵值爲：0.3118。

按照這種方式，可以計算全部同型短語組合和同型詞類組合的信息熵值。[8] 考慮到從詞類組合上升到短語組合的過程中，會減少同型區別的數量（〈表 12〉和〈表 13〉中 token 數的對比），下面就以同型詞類組合的熵值爲例來看歧義程度的差異。因爲是以頻率估算概率，這樣就要求頻率足夠大，才能得到相對比較準確的概率值，但限於目前樹庫的規模，大量的組合都是低頻組合。在 3,932 個同型詞類組合中，頻次在 100 以下的佔 94.15%。對這樣的低頻組合來說，算出來的熵值並不可靠。爲兼顧頻次和熵值，我們在計算出全部同型詞類組合的熵值後，取了頻次在 1,000 以

[8] 通過求全部同型歧義組合的平均熵值，可以在一定程度上評價整個樹庫標注體系的不確定性程度。北大樹庫跟賓州樹庫的同型短語歧義組合的平均熵值都約為 0.57。但考慮到北大樹庫抽取的規則總數是 1,930 條，而賓州樹庫規則總數為 5,220 條，據此估計，賓州樹庫標注體系的確定性程度更高，這在一定程度上反映了賓州樹庫標注內部一致性（尤其是低頻規則的內部一致性）可能優於北大樹庫。

上，熵值在 0.5 以上的組合，共得到 6 個這樣的同型詞類組合。下面是這 6 個組合形成的不同短語類、頻次值、熵值及實例。

〈表 21〉同型詞類組合中頻次及熵值均較高的 6 個組合

詞類組合	根數	根節點	頻次	合計	熵值	示例
n v	3	dj	1838	2777	1.26	前人 開路
		vp	514			全線 崩潰
		np	425			燃料 供應
n v n	3	np	685	1176	1.04	電子 發射 裝置
		dj	482			麥子 需要 春雨
		vp	9			重金 獎勵 發明人
v n ude1 n	2	vp	590	1161	1.00	解釋 工廠 的 困難
		np	571			劃分 句型 的 標準
v n n	2	vp	1291	1622	0.74	解決 技術 問題
		np	331			有 問題 農藥
v n	2	vp	14005	16941	0.67	符合 國情
		np	2936			煉鋼 工人
a v	2	vp	1173	1353	0.67	努力 學習
		ap	149			難 住

上面 6 個組合中涉及的詞類正是名、動、形三大類實詞。這說明目前採用的詞類體系對於句法組合的制約能力有限，對計算機來說，可能造成較嚴重的歧義問題，因此，面向計算機句法分析的需要，詞類的劃分還應加細。這 6 個組合中恰恰包含了漢語中比較經典的句法結構歧義組合「v n ude1 n」（實例「咬死獵人的狗」）。其熵值也基本爲 1。這意味著，如果讓計算機來猜測這個組合該分析爲 vp，還是 np，則命中率就如同扔硬幣猜正反面一樣，只有 50%。

需要說明的是，本文採用計算熵值的方式來評估一個同型組合的歧義程序，只是初步的探索，還不夠成熟。一是如上文已經指出的，受語料規模的限制，用低頻現象去估計隨機事件的概率值，是不可靠的。另一方面，對同型組合的分析深度也是影響熵值的重要因素。比如「v a n」這個同型詞類組合，其實例數爲 743，根節點有 vp，np 兩種，由此計算得到的「v a n」的熵值爲 0.15。但如果考慮「v a n」形成的 vp 和 np 各自內部都有不同的結構情況，再來計算熵值，結果就可能顯著提高。下面兩個表對比了不同分析深度下，對同一個同型詞類組合的熵值計算的差異。

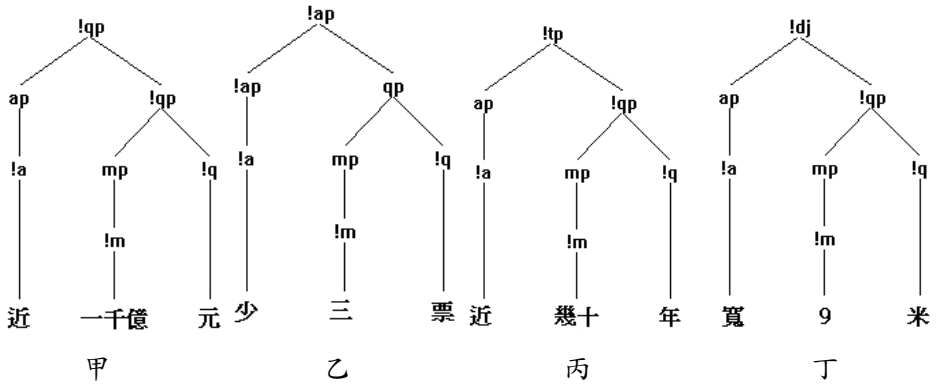
〈表 22〉只考慮根節點差異計算同型詞類組合的熵值

根	詞類組合	頻次	合計	熵
vp	v a n	727	743	0.15
np	v a n	16		

〈表 23〉考慮詞類組合的內部結構差異後計算的熵值

根	詞類組合	內部結構	頻次	合計	熵
vp	v a n	vp (!vp (!v ( )) np (ap (!a ( )) !np (!n ( ))))	608	743	0.80
		vp (!vp (!v ( ) a ( )) np (!n ( )))	119		
np	v a n	np (vp (!v ( )) !np (ap (!a ( )) !np (!n ( ))))	12		
		np (!vp (!v ( ) a ( )) !np (!n ( )))	4		

此外，值得指出的是，按照上述考察方式得出的歧義程度高的組合，也並不一定意味著歧義消解就更困難。以〈表 18〉中介紹的「a m q」的 4 種結構爲例。一方面，這一同型詞類組合的熵值雖然達到 1.25，但該組合形成的 4 類短語有比較清楚的區分條件，因而排除歧義並不困難。下面是「a m q」的 4 種結構。



〈圖 4〉「a m q」的句法結構樹圖

「a m q」的上述 4 種分析方式中，甲和丙可分爲一組來對比，共性爲其中的「a」都由「近」來充當，如果要細分，則甲中的「近」是「接近」，丙中的「近」是「最近」。此外，甲中的量詞 q 爲度量衡單位量詞「元、噸」等，丙中的量詞 q 只能是時間量詞「年、天」等。乙中的 a 只能由「多、少」來充當，丁中的 a 只能由「長、寬、高、重」等少數詞語充當。顯然，「a m q」同型組合中的 a 和 q 都有一定的限制，而且範圍很窄，其結構分析的難度並不大。

5. 結語

本文嘗試基於樹庫語料獲取現代漢語句法結構知識。具體內容分三個層面展開。

第一個層面是對特定的句法結構進行考察。本文選取了兩個對象，一個是 np，考察了 np 在主、賓、定、中等不同句法位置的寬度（詞長）以及內部結構的差異。所得統計結果印證了以往語法研究的定性分析。另一個考察對象是「把」字結構，考察了「把」字結構中 vp 的內部構成情況，以及「把」字結構在句中的分布環境。有兩點發現值得注意：一是

「把」字結構中的 vp 以帶賓語的情況爲最大多數；二是「把」字結構用得最多的並不是直接做謂語，而是跟其他謂詞性成分組合成更複雜的謂詞性結構。

第二個層面是對一類句法結構現象進行考察。本文考察的對象是非中心擴展結構和非同類並列結構。通過分析語料中這兩類結構的實例，我們認爲，這兩類結構是語言中實際存在的，是言語交際中人們出於經濟高效的需要，省略成分或通過借用成分製造出的結構，語法理論設計中關於中心擴展條件和並列結構條件的假設適用於大多數常規情況，但並不能否認實際語言中非常規結構的存在。同時因爲這類結構規則加入到常規結構規則集合中，由此形成的句法結構系統用於計算機分析時，就可能帶來更多的潛在歧義。

第三個層面是對句法結構中的潛在歧義情況做宏觀的定量考察。我們嘗試從同型短語組合和同型詞類組合可能形成的短語類個數、同型組合的頻次、同型組合的熵值等不同角度來衡量一個組合的歧義程度的大小。目前的探討雖對語法研究，特別是計算機自動句法結構分析有一定參考價值，但還是比較初步的，對於句法結構的系統性歧義的考察，還有待在樹庫語料規模擴大，標注信息更爲豐富的基礎上，做出更加可靠的分析。

樹庫資源以往通常是用於數據驅動的計算機自動句法分析 (DOP) 的模型參數訓練，從樹庫中自動抽取可計算的形式化語法模型（如 LFG、HPSG 語法等）。本文則面向語法本體研究，探討從樹庫中獲取不同層次的語法知識。一個樹庫中蘊含的語法知識一部分來自它的標注體系，還有一部分來自基於該標注體系對語言實際材料的標注結果。通過後者獲取的知識，還可以反過來評價前者的設計是否合理，比如如果一個標注體系在標注實際語料後得到的歧義結構的平均熵值較高，就有可能需要回過頭去審視最初的標注體系中各功能標記的設置是否合理。對此，本文還只是做了一些初步的探索，我們希望得到讀者的寶貴意見和建議。

引用文獻

Abeillé, Anne. (ed.) 2003. *Treebanks: Building and Using Parsed Corpora*. Dordrecht & Boston: Kluwer.

Cahill, Aoife, Michael Burke, Martin Forst, Ruth O'donovan, Christian Rohrer, Josef van Genabith, and Andy Way. 2005. Treebank-based acquisition of multilingual unification grammar resources. *Research on Language and Computation* 3.2-3: 247-279.

Huang, Chu-Ren, Feng-Yi Chen, Keh-Jiann Chen, Zhao-ming Gao, and Kuang-Yu Chen. 2000. Sinica Treebank: design criteria, annotation guidelines, and on-line interface. *Proceedings of the Second Chinese Language Processing Workshop*, ed. by Martha Palmer, Mitch Marcus, Aravind Joshi & Fei Xia, 29-37. Hong Kong: Hong Kong University of Science and Technology.

LaPolla, Randy J. 1995. Pragmatic relations and word order in Chinese. *Word Order in Discourse*, ed. by Pamela Downing & Michael Noonan, 299-331. Amsterdam & Philadelphia: John Benjamins.

Lyons, John. 1968. *An Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

Marcus, Mitchell P., Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: the Penn Treebank. *Computational Linguistics* 19.2:313-330.

Shen, Jiaxuan (沈家煊). 2007. Hanyu li de mingci he dongci 漢語裡的名詞和動詞 [Nouns and verbs in Chinese]. *Han-Zangyu Xuebao* 漢藏語學報 [*Journal of Sino-Tibetan Linguistics*] 1:1-16. Beijing: The Commercial Press.

Shen, Jiaxuan (沈家煊). 2009a. Wo kan Hanyu de cilei 我看漢語的詞類 [My view of word classes in Chinese]. *Linguistic Sciences* 8.1:1-12.

Shen, Jiaxuan (沈家煊). 2009b. Wo zhishi jiezhe xiangqian kuale banbu: zai tan Hanyu de mingci he dongci 我只是接著向前跨了半步：再談漢語的名詞和動詞 [Just a small step forward: further remarks on nouns and verbs in Chinese]. *Yuyanxue*

Luncong 語言學論叢 [*Essays on Linguistics*] 40:3-22. Beijing: The Commercial Press.

Wei, Hong (魏紅), and Zexiang Chu (儲澤祥). 2007. 'You ding ju hou' yu xianshixing de wuding NP zhuyuju「有定居後」與現實性的無定 NP 主語句 [On 'postposition of definite subject' and the realistic sentence with indefinite NP as subject]. *Shijie Hanyu Jiaoxue* 世界漢語教學 [*Chinese Teaching in the World*] 2007.3:38-50.

Xue, Nianwen, and Fei Xia. 2000. The Bracketing Guidelines for the Penn Chinese Treebank (3.0). http://www.cis.upenn.edu/~chinese/parseguide.3rd.ch.pdf

Xue, Nianwen, Fei Xia, Fu-Dong Chiou, and Marta Palmer. 2005. The Penn Chinese Treebank: phrase structure annotation of a large corpus. *Natural Language Engineering* 11.2:207-238.

Zhan, Weidong (詹衛東). 2000. Yuyan chengfen de zuhe yu gongneng chuandi 語言成分的組合與功能傳遞 [Combination of linguistic units and inheritance of syntactic features]. *Mianlin Xinshiji Tiaozhan de Xiandai Hanyu Yufa Yanjiu* 面臨新世紀挑戰的現代漢語語法研究 [*Modern Chinese Grammar Studies Meeting Challenges of the New Century*], ed. by Jianming Lu (陸儉明), Yang Shen (沈陽) & Yulin Yuan (袁毓林), 823-833. Ji'nan: Shandong Education Press.

Zhan, Weidong (詹衛東), Baobao Chang (常寶寶), and Shiwen Yu (俞士汶). 1999. Hanyu duanyu jiegou dingjie qiyi leixing fenxi ji fenbu tongji 漢語短語結構定界歧義類型分析及分布統計 [Analysis on types of phrase boundary ambiguity in contemporary Chinese]. *Zhongwen Xinxi Xuebao* 中文信息學報 [*Journal of Chinese Information Processing*] 1999.3:9-17.

Zhou, Qiang (周強). 2004. Hanyu jufa shuku biaozhu tixi 漢語句法樹庫標注體系 [Annotation scheme for Chinese Treebank]. *Zhongwen Xinxi Xuebao* 中文信息學報 [*Journal of Chinese Information Processing*] 2004.4:1-8.

Zhu, Dexi (朱德熙). 1982. *Yufa Jiangyi* 語法講義 [*Lectures on Grammar*]. Beijing: The Commercial Press.

Extracting Syntactic Knowledge from Large-scale Chinese Treebank

Weidong Zhan

Peking University

This paper illustrates three cases of syntactic knowledge extraction on different levels of linguistic analysis from a Chinese treebank which contains more than one million Chinese characters. The first case involves the extraction of the inner structures and the distributions of specified phrases, for instance, noun phrases that used as subject, object and modifier, and verb phrases that used in Chinese *Ba*-construction. The second case concerns the extraction of the non-endocentric constructions and the coordinate constructions whose constituents are asymmetric syntactically. The third case lies in the extraction of ambiguous structures from the treebank and the measure of the degree of ambiguity. In this paper, an ambiguous structure is confined to a construction in which the immediate constituents correspond to two or more reduction rules, which have same components in the right hand side and different categories reduced in the left hand side. The entropy of each ambiguous structure can be calculated by counting its probability. And the entropy value of an ambiguous structure indicates the degree of a structure's ambiguity.

Keywords: Chinese treebank, syntactic annotation, linguistic knowledge extraction

Human Language Resources and Linguistic Typology
Papers from the Fourth International Conference on Sinology
Taipei: Academia Sinica, 2013

Successive Addition Boundary Tone of Chinese Emotional Intonation: Production and Perception*

Aijun Li
Chinese Academy of Social Sciences

'Successive addition boundary tone' (SUABT) proposed by Chao (1933) is employed to convey pragmatic function and found in emotional intonation of Mandarin Chinese. Speakers express 'disgust' or 'anger' through 'falling' SUABT, and 'happiness' or 'surprise' by a kind of 'rising' SUABT. In this case, the boundary tone is composed of two components, i.e. the lexical tone to encode linguistic meaning and the expressive tone to express emotional attitude or pragmatic meaning. In the present study, the intonation patterns and the SUABT analyzed from phonetic and phonological perspective. Then a perceptual experiment was conducted to separate the interaction between the form of Chinese SUABT and its expressive function with the aim to determine whether it can uniquely encode the emotional or pragmatic information. Through the GLM (General Linear Model) analysis on the perceptual results, we proposed that the acoustic features, the length of utterance, and the category of final tone have significant effect on emotional expressiveness. However, the contributions are various in different emotions. The mapping between the

* This research was funded by National 973 project ('Internet-based Language Cognition and Representation', No. 2013CB329301), NSFC Project with No. 60975081, JSPS Ronpaku Program and and CASS innovation project.

form of SUABT and its pragmatic function is a complex many-to-many relation. The results further revealed that SUABT is not the only element speakers employed to express emotions, whereas the other components could be invoked to convey expressive information.

Keywords: Chinese emotional intonation, successive addition boundary tone, production, perception

1. Introduction

The encoding and decoding schemes of emotional speech have not been fully clarified in the interactive communication framework, despite of a substantial body of work that has been conducted in the fields of cognitive psychology, phonetics, speech science and speech technology, etc (e.g. Cowie et al. 2001, Scherer 2003, 2009, Schröder 2004, Tao et al. 2006, Schuller et al. 2009).

The purpose of the present study is to quantitatively clarify the following aspects: (i) the general and fine structure of the acoustic cues of Chinese intonations, especially how tone and intonation are co-encoded in boundary tone to convey expressive information; (ii) the relationship between the acoustic form and function of emotional intonation related with boundary tones.

Within the framework of intonational phonology, intonation is described as a chain of pitch events which are well organized under intonational grammar known as the autosegmental and metrical approach (AM) or the Pierrehumbert model (1980, 1990). The intonation is also considered to be eventually linked with communicative, paralinguistic or pragmatic meaning (Bolinger 1989, Gussenhoven 2002, Ladd 1996, Wu 1995, 2000, Xu et al. 2007, 2011, Hirschberg 2002, Benus et al. 2007, Sagisaka 2012).

Gussenhoven (2002) proposed a relatively more elaborate theory that interprets intonational meaning by a set of “biological codes”: e.g. frequency,

effort, and production.

Liu & Xu (2005) argued that communicative meanings are conveyed through a set of separate functions that are realized by an articulatory system with various biophysical properties. This view of speech melody is summarized into a comprehensive model of tone and intonation, namely, the parallel encoding and target approximation (PENTA) model.

Boundary tone is considered as an important component to convey linguistic and pragmatic information and treated by Pierrehumbert (1980) as a phonological unit for representation of the internal makeup of an intonational phrase. The boundary tone contains two phonological events (refer to Pierrehumbert 1980, Ladd 1996, and Gussenhoven 2004), specifically, H% is employed to indicate a final rise while 'L%' is adopted to indicate the absence of final rise or a default pitch event.

Pierrehumbert (1980) and Xu (2005) proposed that the most obvious F_0 pattern associated with a *yes-no* question is the final rising, which has been attributed to a high boundary tone, i.e. H% in the AM theory. Pierrehumbert & Hirschberg (1990) provided a list of functions of pitch accents and boundary tones with reference to discourse structures. And, Silber-Varod (2011) concentrated on the continuous (C)-boundary inventory in a corpus of spontaneous Israeli Hebrew and investigated the linkage function of the communicative value of the C-boundary tone according to the syntactic relations between the words in the preceding and following positions for each of the C-boundaries.

In regards with Chinese, Chao (1932, 1933, 1968) is one of the pioneers who studied Chinese emotional/expressive intonation. He proposed that the actual melody or pitch movements of a tonal language differ from the mere succession of the language with few fixed tones. It is in fact a resultant of three elements: tone or etymological tone, the neutral intonation, and the expressive

intonation, with the latter two forming the sentential intonation. He emphasized that the expressive intonation depends on the quality of the voice, unusual degrees of stress (or weakness), general pitch of the whole phrase and tempo of speech.

Chao (1933) distinguished at least two types of two addition patterns of tone and intonation: *simultaneous addition* and *successive addition*. The simultaneous addition refers to the tones that are the algebraic sums or the resultants of two factors, specifically, the original lexical tone and the sentence intonation proper. The successive addition refers to the clause that has a rising or falling intonation, which is not added simultaneously to the last syllables but added on successively after the lexical tones. He described the successive addition of the rising ending as (↗) and the falling ending as (↘), specifically they are described in the following formula:

T1: ↗55:= 56:  ↘55:= 551: 

T2: ↗35:= 36:  ↘35:= 351: 

T3: ↗214:= 216:  ↘214:= 2141: 

T4: ↗51:= 513:  ↘51:= 5121: 

Within the above formula, T1~T4 are four lexical tones, 1~5 are tonal values in the tone-letter system (1980), numbers on the left of ':=' are the lexical tone values and the numbers on the right are the addition tone values. 'T1: 55, T2: 35, T3: 214 and T4: 51' are four tones represented in the normal or neutral tonal space. Additionally, Chao adopted '6' to represent extra high pitch in successive addition of rising tone. The most interesting effect is shown with the falling tone (T4) which is resulted in a circumflex tone.

Chao even enumerated forty intonation patterns to demonstrate the forms and the functions of the intonation through grouping them by pitch/duration elements, voice quality and intensity elements (1933, 1968).

Lin (2004, 2011) investigated how ‘simultaneous addition’ is realized in the sentence-final syllables (boundary tone) to convey intonational meanings. And also Jia (2012) proposed a Chinese phonological intonation grammar. Unlike English, the phonological representation of boundary tones ‘H%’ or ‘L%’ indicate the relative ‘H’ or ‘L’ pitch of the initial or final boundary tones, while the boundary syllable will keep its lexical tonal patterns.

Yuan (2006) proposed three mechanisms of question intonation in Mandarin Chinese, i.e. an overall higher phrase curve, higher strengths of sentence final tones, and a tone-dependent mechanism that flattens the falling slope of the final falling tone and steepens the rising slope of the final rising tone. Jiang & Chen (2011) supported the idea that Mandarin interrogative cues distribute over an entire utterance. High pitch at both edges of an intonation phrase should be marked out. Through the employment of a corpus of conversational speech in Mandarin, Callier (2011) investigated the boundary tone of intonational meaning and function in a socio-phonetic way. The verification of the H% or L% boundary tone is approached by clause type, illocutionary force, and speaker gender. Few studies have been conducted for successive addition boundary tone relating to the pragmatic meaning of intonation. Mueller-Liu (2006) listed some successive addition tones in expressive speech to signal the emotion-attitudinal message, and Lu & Lin (2009) also found it in the intonational questions to signal the interrogative mood.

In the present study, we differentiate the expressive speech with emotional speech and attitudinal speech. In previous research, based on Wu Zongji’s intonation theory (1990, 1995, 2000), we found that speakers employ boundary tone H% to express happiness and courtesy in the conversional speech (Li & Wang 2004, Wang et al. 2005), and that sentence stress shifts a lot across the observed emotions (Tao et al. 2006, Li et al. 2008).

In another study (Li et al. 2011), we found that the successive addition boundary tones (SUABT) are employed by speakers to convey expressive information, such as ‘disgust and anger’ which is realized by ‘falling’ successive addition tone, and ‘happiness or surprise’ by a kind of ‘rising’ successive addition tone. As shown in Figure 1, the boundary tones of ‘disgust and anger’ exhibit a falling tail in comparison with that of ‘neutral’ intonation, which keeps the first part as its lexical tone. The falling SUABT in disgust intonation has been acoustically analyzed (Li, Fang, Jia & Dang 2012).

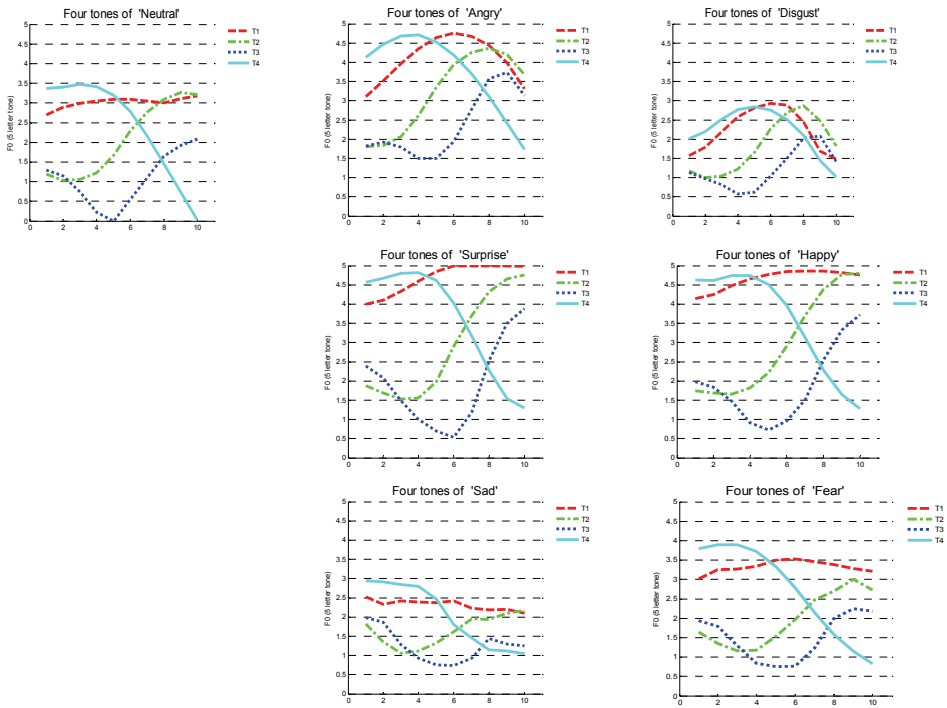


Figure 1: A male speaker’s ‘Neutral’ and six emotional intonations of mono-syllabic utterances with four lexical tones (T1~T4), F_0 is normalized on five tone-letter scale, and tonal duration is normalized into ten points.

In Chao's five-tone letter system, value '6' is adopted to refer to the extremely high tone in SUABT. However, in the present study 5-tone letter scale is applied for all emotional utterances (refer to Figure 1), by comparing the tonal spaces of emotional utterances with that of the neutral utterances. In the present study we still found the same successive addition pattern proposed by Chao (Li et al. 2011).

In the present paper, the overall acoustic patterns of Mandarin emotional intonation are analyzed on more data, specifically, disyllabic utterances and longer utterances. Then, the tone and intonation addition patterns of the boundary tones, especially the successive addition boundary tone (SUABT), are analyzed from both phonetic and phonological perspective. Finally, a perceptual experiment is conducted to separate the interaction between the pragmatic function of the SUABT and its acoustic form, through which to determine whether the SUABT can be independently encoded to express emotion or the mapping between the acoustic form and the expressive function is unique.

2. Materials and recording

The speech data adopted in the present study were selected from the emotional speech corpus Emotion-CASS. A set of one hundred and eleven sentences with various lengths (from 1 to 14 syllables), sentence types (narrative or interrogative) and structures were recorded. The contents of all these sentences were emotionally 'neutral'. The monosyllabic sentences covered the combinations of four lexical tones (T1~T4) and all the vowels. The disyllabic sentences covered sixteen tonal combinations (T1~T4 × T1~T4), among which T3T3 have the same surface pattern as T2T3 due to the phonological tone sandhi. A male and a female professional voice dubbing actors were recruited to produce the

utterances in seven kinds of emotions: Disgust (D), Sad (S), Angry (A), Happy (H), Surprise (SU), Fear (F) and Neutral (N). The sampling rate and resolution were 16 KHz and 16 bits, respectively.

The emotional states of all the utterances were rated in a 5 point scales from (0 to 4) by ten listeners before acoustic analysis (Cao et al. 2012). The confusion matrix is presented in Table 1. The utterances selected in the present study have mean correct recognition scores greater than chance score (4/7≈0.57).

All the utterances were annotated with initial and final boundaries in Praat. The F_0 data of these utterances and segments (syllable, initial and final) were extracted by Praat and they are manually checked to ensure the accuracy of the data. The F_0 values of the tonal bearing part (final) of each syllable were normalized duration by the extraction of 10 points for each syllable. Then, they were transformed into semitone scale with the reference frequency of 75Hz. Finally, all the F_0 values were mapped into a 5-tone letter space (Chao 1980, Wu 2000).

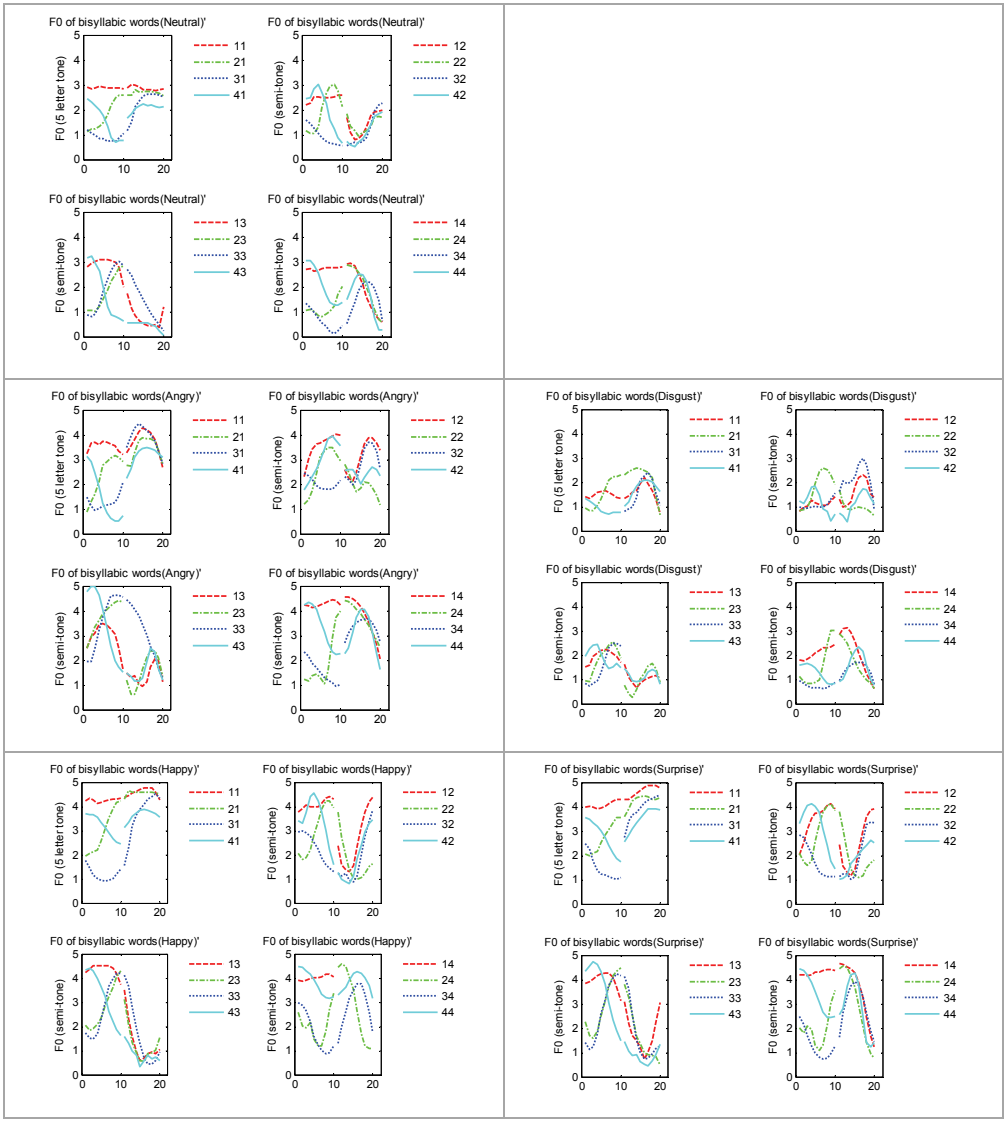
Table 1: the confusion matrix of the intended and perceived seven emotions (*average for 111 utterances in each emotion*).

Intended \ Perceived	Sad	Happy	Fear	Surprise	Angry	Disgust	Neutral
Sad	2.17	0	1.78	0	0	0	0
Happy	0	1.56	0	1.15	0	0.22	0.44
Fear	1.07	0	1.5	0.17	0	0	0.11
Surprise	0.17	0.11	0.22	3.06	0	0.11	0
Angry	0	0	0	0	2.33	1.67	0
Disgust	0.28	0	0.22	0	0.83	1.67	0
Neutral	0.39	0	0	0.17	0.17	0.22	2.06

3. Intonation patterns and SUABT of the emotional utterances

3.1 Disyllabic intonation

In the previous study (Li et al. 2011), we analyzed the F_0 features of these two speaker's emotional intonation for monosyllabic utterances. Compared with the 'Neutral' emotion, the tone pattern varies significantly across seven emotions in tonal range, tonal register and tonal contours. Except for "Angry" emotion, the tonal patterns are similar between the two speakers. 'Fear' has an obvious tremor voice, higher F_0 and narrower pitch range than the 'Neutral' emotion. 'Sad' emotion is characterized by reduced pitch rang and lower F_0. 'Disgust' emotion is similar to 'Sad' emotion, but the boundary tone is a kind of successive addition and performs as falling tone. 'Happy' and 'Surprise' emotions are comparable, with higher pitch range and register, and rising boundary tone. 'Surprise' emotion has higher bottom pitch, but a slightly narrower range in comparison with 'Happy' emotion. It is also expressed a bit faster in speech rate. Both 'Happy' and 'Angry' emotions have higher pitch and faster speech rate from the two speakers. Although these patterns may make them difficult to be recognized, they can be easily distinguished through the boundary tones.


F0 of bisyllabic words(Neutral)'
F0 (5 letter tone)
F0 (semi-tone)
11
21
31
41
12
22
32
42
13
23
33
43
14
24
34
44
F0 of bisyllabic words(Angry)'
F0 of bisyllabic words(Disgust)'
F0 of bisyllabic words(Happy)'
F0 of bisyllabic words(Surprise)'

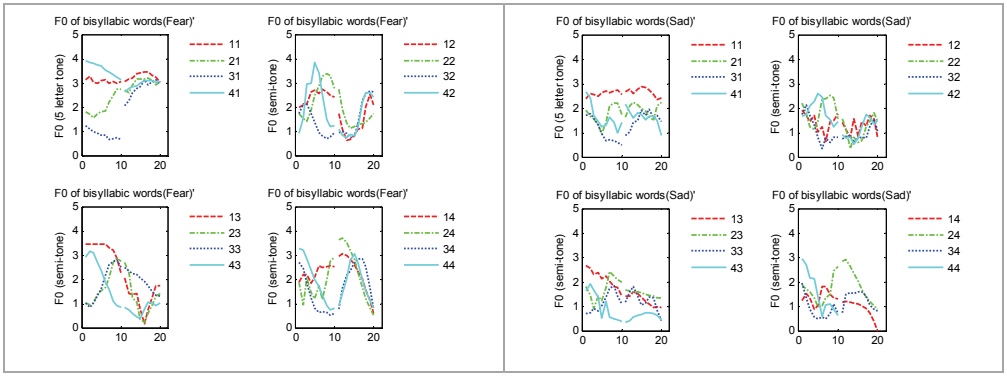


Figure 2: The F_0 Patterns of disyllabic utterances in seven emotions. The disyllabic tones were normalized in twenty points (ten points for each syllable) and the F_0 data were normalized into 5-tone letter scale. The tonal combinations are plotted according to the boundary of tonal category. For example, legend '31' refers to the tonal combination of T3+T1 with the boundary tone as T1. Each emotion has four groups according to the category of boundary tone T1~T4.

Emotional intonations of monosyllabic utterances exhibit that the addition of tone and intonation patterns are different from what we have found in 'Neutral' speech (Wu 1995, 2000, Lin 2004, Lin & Li 2011). The speakers express some emotions like 'Disgust' or 'Anger' by using a kind of falling successive addition tone and 'Happiness' and 'Surprise' by a rising one. The addition part does not belong to the lexical tone of the syllable, which is produced to express the speaker's emotions. The additive tones occupy 1/4~1/2 length of the whole boundary tone without lengthening the duration of the boundary syllable (Li et al. 2011, 2012). The whole boundary tone is longer when it is employed to transmit linguistic information of interrogation compared with statement (Lin 2006). Therefore, the encoding mechanism of the boundary tone may depend on the linguistic and pragmatic function of the intonation.

In the present study, F_0 features of disyllabic utterances were investigated and compared with those of the monosyllabic utterances across seven emotions. The F_0 patterns are grouped according to the tone of the second syllable in each emotion (seven subplots), and were plotted in Figure 2. From the figure, we can obtain that:

(1) The F_0 patterns of 'Neutral' intonation (first row of Figure 2) are identical with the patterns of tonal combination of disyllabic words described by Wu (2000) and the boundary tone patterns the same as described by Lin (2004). In the figure, not all the boundary tones for T3 are presented due to the creaky voices.

(2) Comparing with 'Neutral' intonation, all the boundary tones (T1~T4) of 'Disgust and Anger' (second row in Figure 2), have an additive falling boundary, the typical SUABT with the first part as lexical component and the second part the expressive component to express 'Disgust or Anger', similar as monosyllabic utterance shown in Figure 1.

(3) In comparison with the 'Neutral' intonation, the F_0 patterns of 'Happiness and Surprise' (rows 3 Figure 2) have a rising SUABT with broader F_0 range. The 'rising' feature makes the H tone (T1: HH; T2: LH; T3: LLH) to be raised higher with a steeper slope, while for T4 (HL), the slope become smaller (flatter). In the tonal combinations of 'T2+T2' and 'T2+T3', they have almost overlapped contours caused by the emotional expression.

(4) In regards with 'Neutral' intonation, the F_0 patterns of 'Sad and Fear' intonation (rows 6 and 7 in Figure 2) have some variations on the F_0 range in addition to distinctive quivers observed on F0.

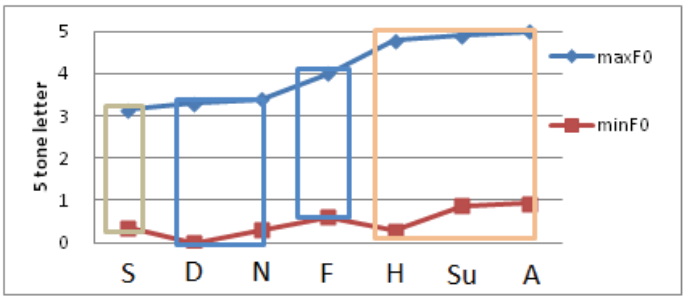


Figure 3: F_0 range variations of disyllabic utterances for seven emotions (male speaker). 'maxF_0' and 'minF_0' are mean maximum and minimum values of F_0 plotted on 5-tone letter scale. "S, D, N, F, H, Su and A" are short for the seven emotions.

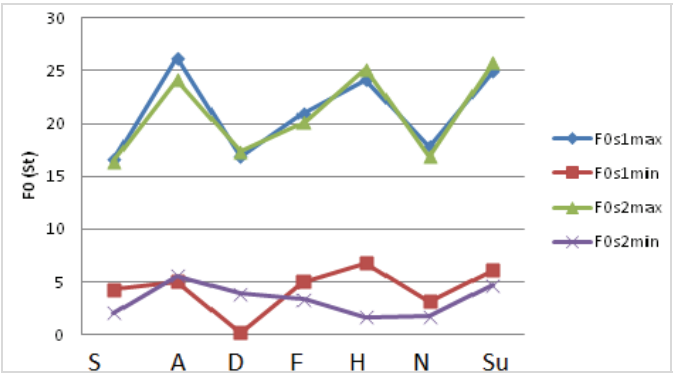


Figure 4: F_0 range variations of the first and the second syllable for seven emotions (male speaker). 'F_0s1max, F_0s2max, F_0s1min and F_0s2min' denote the mean maximum and minimum values of the first and second syllable respectively (on semitone scale). 'S, A, D, F, H, N and Su' are the abbreviation for the seven emotions.

Figure 3 depict the F_0 range variations of the whole disyllabic utterances from the male speaker (the female speakers' data are similar and omitted here).

Figure 3 shows that the F_0 range can be categorized into three levels: Narrow for 'Sad' (range < 3), *Middle* for 'Neutral, Disgust and Fear' (range between 3 and 4) and *Broad* for 'Happy, Surprise and Angry' (range between 4 and 5) intonations. The register of the intonation (maxF_0) can also be categorized into three levels: low register for 'Sad, Disgust' (3~3.5), middle register for 'Fear and Neutral' (3.5~4) and high register for 'Angry, Happy and Surprise' (4~5). Here, the 'Angry' intonation has a broader range, while others keep the same as the monosyllabic intonations (Li et al. 2011).

For the F_0 variations of the two individual syllables (as shown in Figure 4), the maxF_0 of the two syllables are almost the same, however, the 'minF_0' is lower for the second syllable except 'Disgust and Angry' intonation. The 'Happy and Surprise' intonations come with broader F_0 range on the final syllable, but the 'Angry and Disgust' have narrower F_0 range on the final syllable.

3.2 F_0 patterns and SUABT of longer utterances

The study examines the F_0 patterns of longer utterances for seven emotions, as shown in Figure 5 ~ Figure 7. The F_0 ranges and registers keep the same patterns in different contexts as those of the disyllabic utterances. 'Disgust and Sad' intonation has low register, 'Neutral and Fear' have middle register, while 'Happy, Angry and Surprise' have high register. 'Sad' has narrow range, 'Neutral, Disgust and Fear' have middle range, while 'Happy, Angry and Surprise' have broad range.

It seems that the F_0 patterns of SUABT should be depended on the lexical tone of the boundary syllable and the expressive emotion. For boundary syllable whose lexical tone has a 'H' offset feature (T1: HH, T2: LH and T3: LL(H)), referring to the final boundary syllables 'fu1' and 'jing1' (T1:HH) in Figure 6

and Figure 7, a falling tone is added after the H tone to produce a clear turning point in 'Disgust and Angry' intonations; a rising tone is added after the H tone for 'Surprise and Happy' to produce a higher ending. It is described by Chao with tonal value 6 as mentioned in the introduction. In the case of boundary syllable with T4 (HL), where the offset tone is L tone, the boundary tone keeps falling for all emotions but with different registers (refer to Figure 5); and the rising expressive tone sometimes causes a small rising tail as Surprise. It can be seen from 'hui4' and 'sai4' in Figures 5 and 6. However, for longer utterances in Figure 7, T4 keeps falling in seven emotions without showing a circumflex tone suggested by Chao for successive rising addition.

In the figures, 'Fear and Sad' intonation shows a clear tremble F_0 curve. An acoustic analysis on 'Disgust' boundary tones has been conducted for all utterances (Li et al. 2012), the detailed acoustic patterns of SUABT of other emotions will be conducted in the future work.

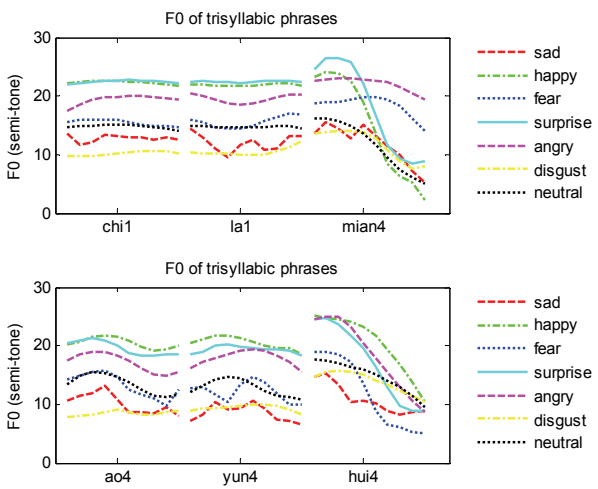


Figure 5: F_0 of tri-syllabic utterances in seven emotions (male speaker)

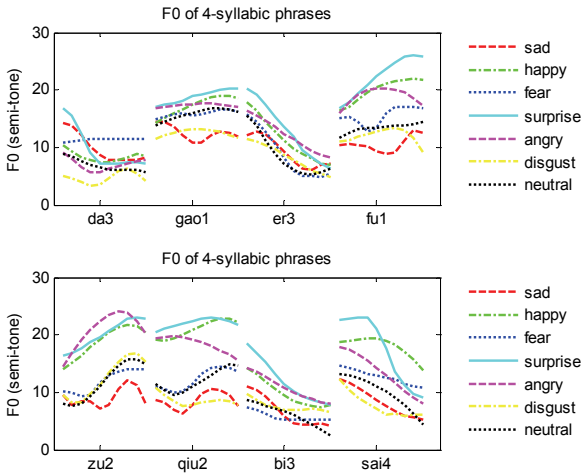


Figure 6: F_0 of quadric-syllabic utterances in seven emotions (male speaker)

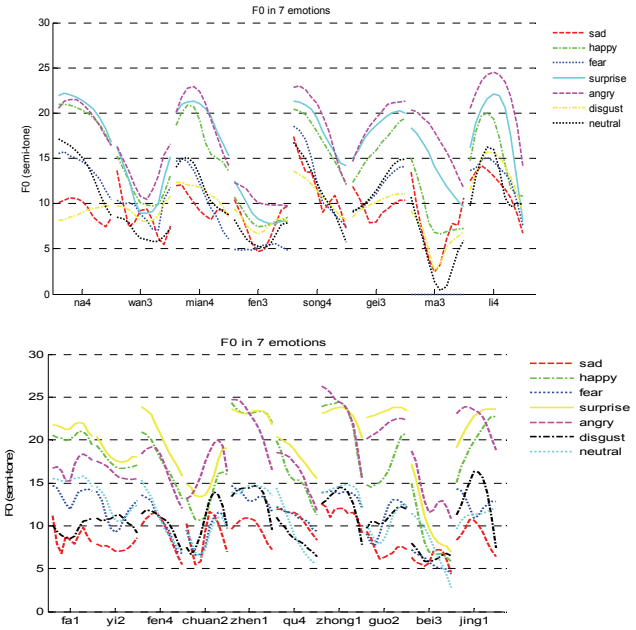


Figure 7: F_0 of utterances with eight syllables (upper panel) and ten syllables (lower panel) in seven emotions (male speaker)

3.3 Phonological representation of overall intonation pattern and the *SUABT*

Table 1 summarizes the overall phonological features of seven emotional intonations and the SUABT. The overall intonation is described by the register and range of F_0 in three levels (H, M and L) based on the previous analysis. F_0 register is determined by the average top value of the F_0, and range feature is determined by the average value of the F_0 range. In five-tone letter space, register features 'H, M and L' denote that the high tone values of F_0 are 5, 4 and 3, respectively. Accordingly, range features, i.e. 'H, M and L' denote the various F_0 v spans which are 5, 4 and 3 respectively. For example, as we mentioned in §3.1, the F_0 of 'Happy' intonation varies between 1 and 5, i.e. the high tone is 5 and the variation of F_0 span is 5, therefore, both the register and the range of 'Happy' intonation are 'H', which indicates that 'Happy' intonation presents a highest register and a widest range.

While the last two columns present features of boundary tones in seven emotional intonations. We propose the adaptation of the combination of the traditional boundary features 'H% or L%', i.e. the relative pitch of boundary tone (Lin 2006, Lin & Li 2011, Jia 2012), and the successive addition tone features 'r' (raising successive addition tone), or 'f' (falling successive addition tone) or 'x' (no successive addition tone) to describe the successive addition boundary tones of Chinese emotional intonations. For example, for monosyllabic and disyllabic intonations as described above, the boundary tones of 'Angry' intonations can be described as 'H-f%', and 'Happy and Surprise' intonations as 'H-r%', respectively. In this regards, we emphasize that the boundary tone features 'H% and L%' are independent pitch events from pitch accents and expressive tone features 'r' or 'l'.

Table 1: F_0 features of emotional intonation and SUABT

Emotions	Range	Register	Additive tone	SUABT
Sad	L	L	x	H%/L-%
Disgust	M	L	f	H-f%/L-f%
Neutral	M	M	x	H-%/L-%
Fear	M	M	x	H-%/L-%
Happy	H	H	r	H-r%L-r%
Surprise	H	H	r	H-r%/L-r%
Angry	H	H	f	H-f%/L-f%

3.4 Schematic representation of SUABT

As shown in Figure 8, a schematic presentation is proposed to describe the SUABT which is composed of two components, e.g. the lexical tone to express lexical meaning, and the expressive tone to express emotional meaning. The expressive tone has two patterns, specifically a rising tone which is realized by the addition of a rising tail after the lexical tone which is employed to express 'Surprise and Happy' emotions in the data; and a falling tone is realized by the addition of a falling tail after the lexical tone with the aim express 'Anger and Disgust' in the data. For falling SUABT: 'T1/T2/T3 + falling expressive tone' presents a clear turning point as shown in Figure 9, and 'T4 + falling expressive tone' keeps the falling pattern; for rising SUABT: 'T1/T2/T3 + rising expressive tone' causes a higher rising of the boundary tone, and 'T4 + rising expressive tone' keeps the falling boundary tone or presents a small rising tail for short utterance.

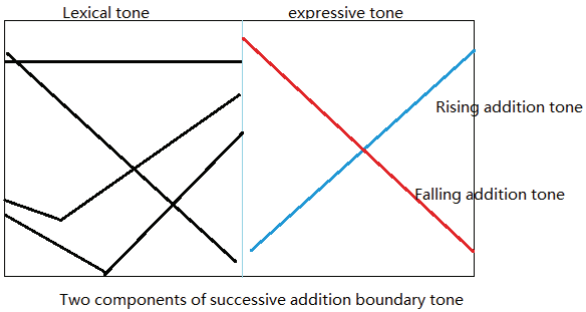


Figure 8: A schematic representation of additive boundary tone with two tonal components: lexical tone and expressive tone

4. Perceptual experiment on the expressiveness of SUABT

The present section explores the 'gradient nature' following Ladd (1996) that the additive boundary tone contributes to the emotional expression. In order to exclude expressive factors other than SUABT, emotionally neutral utterances were adopted as the baseline data and an addition tone was successively added to the final boundary tone of the neutral utterances. The successive falling addition tone was selected in the experiment. There are two reasons for the selection of successive falling addition tone in the experiment. Firstly, it's not difficult to measure the acoustic cues of the lexical and expressive tones in a falling SUABT, because a clear turning point can be observed between the lexical tone and the expressive tone for lexical Tone 1, Tone 2 or Tone 3. Therefore, an acoustic analysis has been carried out for the successive falling addition tone of 'Disgust' intonation (Li et al. 2012). Secondly, as for the exploring of the relation between the form and function of the SUABT, it will be reasonable to control the two components separated by the addition of an expressive tone based on acoustic cues of some emotional boundary.

4.1 Stimuli preparation

Table 2 lists the nine emotionally neutral sentences used in the experiment, which have diverse length and final tone categories. Additive falling tones are simulated based on the variations of the acoustic parameters for 'Disgust' intonation of a male speaker (Li et al. 2012). In order to separate the additive falling boundary tone from the original lexical tone, and to clearly and easily manipulate the lexical and the additive parts, T1~T3 were selected as the boundary lexical tone and T4 was excluded.

Table 2: Nine sentences used in the perceptual experiment

#	Content	Length (sylls)	Final syll.	Final tone
1	一。One	1	yi1	T1
2	姨。Aunt	1	yi2	T2
3	椅。Chair	1	yi3	T3
4	老翁。Old man	2	weng1	T1
5	母羊。Ewe	2	yang2	T2
6	雄蕊。Stamen	2	rui3	T3
7	老周買了五斤海參。 Mr. Zhou bought 2.5kg of sea cucumbers.	8	shen1	T1
8	去年蓋的二層小樓。 Two-story building built last year.	8	lou2	T2
9	長篇小說梅娘曲。long novel Mei Niang Qu	7	qu3	T3

The schematized representation of a falling SUABT is given in Figure 9. The rising part represents the lexical tone while the falling part represents the expressive tone. The acoustic features of the male speaker's 'Disgust' boundary tone were analyzed according to this structure (Li et al. 2012). The features that we are concerned here are the 'duration vs. F_0 slope' of the additive falling tone,

and the duration ratio of the additive falling tone to the preceding lexical tone. The absolute maximum slope value of the additive falling tone is 110st/s (F_0ref = 75Hz); the maximum duration ratio of final falling to the preceding tone is 3.5, while the absolute slope value is less than 65st/s and ratio less than 1.4 within the 95% confidence interval.

When generating the stimuli, the F_0 of the boundary tone (the last syllable) was manipulated by the addition of a falling tail to the neutral boundary tones of the utterances in Table 2. So the new boundary tone are consisted of two parts, i.e. the first part was the original neutral boundary tone, and the second part was the additive falling tone with various lengths and slopes based on the acoustic data measured for the 'Disgust' emotion. The range of additive falling slope *ki* is set from 0 to -80 st/s stepping in -10st/s, $i=1\sim9$. The duration ratio of the final fall to the preceding tone *dj*/*D* is set from 0 to 1.25 stepping in 0.25, $j=1\sim6$.

PSOLA synthesizer was adopted to generate the stimuli by manipulating the boundary tone of the given nine neutral utterances. The lowest F_0 is set to be 75Hz in this process; therefore not all of the '*dj*' and '*ki*' were realized. During manipulation of a falling SUABT, not only the duration but also the shape of the original boundary tone (lexical tone part) is adjusted to fit the falling tone. Finally, we got 425 synthesized stimuli.

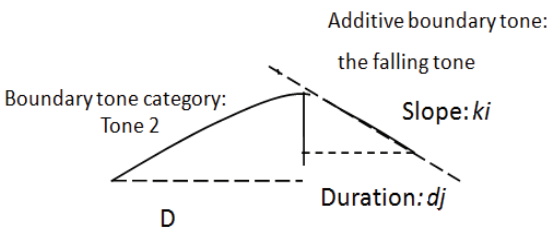


Figure 9: A schematic representation of an additive falling tone with duration 'dj' and slope 'ki'

As shown in Figure 1, Figure 2 and Figure 5 ~ Figure 7, although both 'Disgust and Angry' intonations have falling SUABT, the pitch registers of them are different. The pitch register of 'Disgust' intonation is closer to 'Neutral' than 'Angry' intonation, we hypothesize that stimuli which can generate a greater tendency to be perceived as a 'Disgust' emotion. If the perceived emotions were stronger for 'Disgust' than others, it would support the notion that the SUABT could be encoded alone to express emotional intonations; otherwise the relation between the form and function of SUABT could be 1-to-many or many-to-many.

4.2 Procedure of the perceptual experiment

Twenty university students participated in the perceptual experiment, five male and fifteen female. All of them can speak standard Chinese and have no previous report on hearing problems.

All the stimuli were divided into four groups randomly (107 in 3 groups, 106 in the last group). Before the perceptual experiment, they were trained to use a perceptual program written in Praat script. The subjects were required to judge 20 stimuli chosen from the emotional corpus covering seven emotions, i.e. 'Neutral, Happy, Sad, Angry, Disgust, Surprise, and Fear', and they were asked to evaluate the emotional expression for each stimulus. One or more emotions could be selected from seven emotions. Furthermore, they were allowed to write down other emotional or attitudinal expressions if they perceived them beyond these seven emotions. The perceived score is set to 1 for the selected emotion; otherwise the score is set to 0.

4.3 Analysis on the perceptual experiment results

4.3.1 Perceptual experiment results

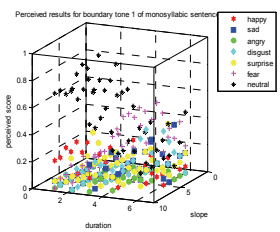
Figure 10 shows the average perceptual results on seven emotions from twenty listeners distributed along the two dimensions of slope and duration ratio which are scaled in corresponding step numbers respectively. The duration ratio steps from 1 to 6 and slope steps are from 1 to 9.

For final boundary T1 in Fig. 10 (1)~(3), the results for mono-syllabic, di-syllabic and long utterances are depicted, respectively. With the increase of the absolute values of duration and slope for the additive falling tone, the perceived scores for ‘Neutral’ emotion are decreased, and they are increased gradually for other emotions. Surprisingly, both the scores for ‘Disgust’ emotion and the scores of ‘Angry and Fear’ emotion perform clear increase.

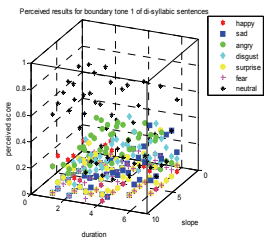
For final boundary T2 (Fig. 10 (4)~(6)), the results for monosyllabic utterances are ‘misunderstood’. It was perceived as ‘Surprise’ emotion rather than Neutral emotion when the absolute values of slope and duration are small. The reason for this is that the sound of ‘yi2’ which is a polyphonic word, and it corresponds to a neutral-tone syllable ‘咦’ when it is read with a rising tone to express ‘Surprise’. However, with the increase of the absolute values of duration and slope of the additive falling tone, ‘Disgust’ emotion is perceived more often than ‘Surprise’ emotion. Meanwhile, the ‘Angry’ emotion can be perceived with rather higher scores.

For final boundary T3 (Fig. 10 (7)~(9)), the perceived results were affected by the slope and duration as well. ‘Disgust, Sad and Angry’ emotions could be perceived with the increase of the absolute values of slope and duration of the falling tone.

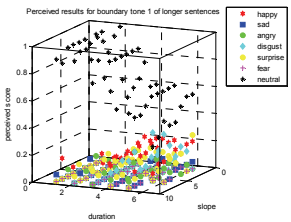
Fig. 10 (10)~(12) depict the results for final tone 1, tone 2 and tone 3 in all utterance conditions respectively. It is clearly shown that with the increase of the step number of duration and slope of the falling tail, the perceived scores decrease for ‘Neutral’ emotion and increase for other emotions, especially for ‘Disgust, Angry, Fear and Sad’ emotions. Fig. 10 (13) shows the overall results.



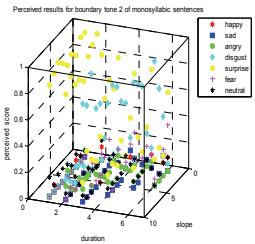
(1)



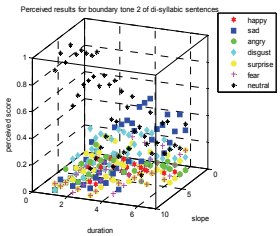
(2)



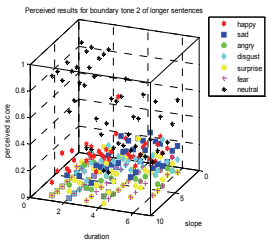
(3)



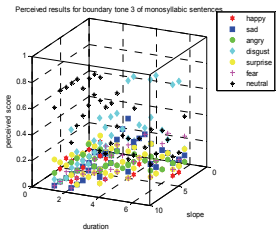
(4)



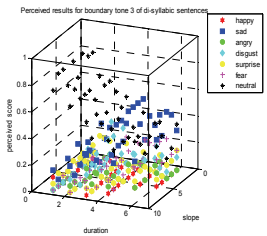
(5)



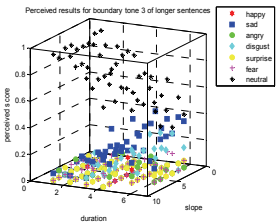
(6)



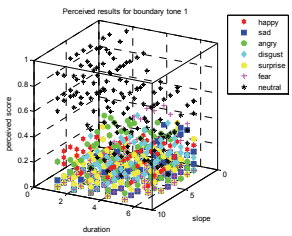
(7)



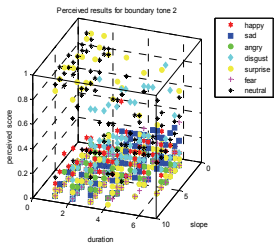
(8)



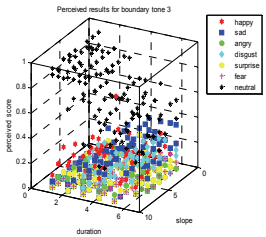
(9)



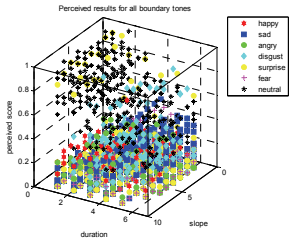
(10)



(11)



(12)



(13)

Figure 10: 3D plots for average perceived scores of seven emotions for boundary tone 1 ~ tone 3 in three different conditions, and general results for boundary tone 1 ~ tone 3, and results for all boundary tones. Slope and duration are scaled in step numbers.

4.3.2 Multifactor GLM analysis

Multifactor GLM analysis was applied to further reveal the factors that they can affect the perception of emotional attitude. The factors include *duration* (1~6) and *slope* (1~9) of the additive falling tone, *boundary tone category* (1~3 for T1~T3), and *utterance type* (1: monosyllabic utterances, 2: disyllabic utterances and 3: longer utterances). Multifactor GLM analysis was applied to further reveal the factors that they can affect the perception of emotional attitude. The factors include duration and slope categories of the additive falling tone, final tone category, and utterance type (length category).

The results are listed in Table 3 which show that: (1) utterance type and the final tone category have significant effect on the emotional perception (*ps=0.0*, *0.0*). The slope and duration of the final falling tone also have significant effect on the perceptual scores except for 'Happy' emotion (*ps=0.0*, *0.0*). (2) Interaction of 'utterance type and final tone category' is significant (all *ps=0.0*), while the interactions of 'utterance type * slope' and 'final tone category * slope' affect the perception of 'Angry, Disgust and Surprise' emotions (all *ps=0.0*), but have no significant effect on the perception of the other four emotions (*ps=0.549*, *0.615*, *0.641*, *0.44*2). 'Utterance type * duration' and 'final tone category * duration' also have significant interaction for the perception of all emotion types (all *ps=0.0*). (3) F-values revealed that: the additive final slope contributes more to the perception of 'Disgust, Angry and Surprise' emotions than other emotions; the additive final duration contributes more to the perception of 'Neutral, Disgust and Sad' emotions than other emotions; the final tone category contributes more to the perception of 'Surprise, Neutral and Disgust and Sad' emotions than other emotions; and lastly, utterance length contributes more for the perception of 'Neutral, Surprise and Disgust' emotions than other emotions.

Table 3: Part of the multifactor GLM results

Source	Dependent Variable	df	F	Sig.	Source	Dependent Variable	df	F	Sig.
utterance type	Happy	2	37.022	.000	utterance type * slope	Happy	16	.917	.549
	Sad	2	62.534	.000		Sad	16	.861	.615
	Angry	2	135.212	.000		Angry	16	2.871	.000
	Disgust	2	267.876	.000		Disgust	16	5.593	.000
	Surprise	2	310.222	.000		Surprise	16	4.695	.000
	Fear	2	135.378	.000		Fear	16	.839	.641
	Neutral	2	413.666	.000		Neutral	16	1.010	.442

Source	Dependent Variable	df	F	Sig.	Source	Dependent Variable	df	F	Sig.
final tone	Happy	2	72.782	.000	boundary tone * slope	Happy	16	1.242	.226
	Sad	2	79.623	.000		Sad	16	1.583	.065
	Angry	2	33.401	.000		Angry	16	1.859	.020
	Disgust	2	102.618	.000		Disgust	16	3.476	.000
	Surprise	2	247.964	.000		Surprise	16	2.202	.004
	Fear	2	29.937	.000		Fear	16	1.316	.176
	Neutral	2	122.520	.000		Neutral	16	1.451	.109
slope	Happy	8	.548	.821	utterance type * duration	Happy	10	8.978	.000
	Sad	8	6.935	.000		Sad	10	5.287	.000
	Angry	8	10.897	.000		Angry	10	6.686	.000
	Disgust	8	50.393	.000		Disgust	10	23.688	.000
	Surprise	8	11.391	.000		Surprise	10	27.102	.000
	Fear	8	3.378	.001		Fear	10	10.862	.000
	Neutral	8	5.620	.000		Neutral	10	6.343	.000
duration	Happy	5	.759	.579	boundary tone * duration	Happy	10	6.973	.000
	Sad	5	49.516	.000		Sad	10	4.544	.000
	Angry	5	20.372	.000		Angry	10	2.411	.007
	Disgust	5	187.108	.000		Disgust	10	11.359	.000
	Surprise	5	22.576	.000		Surprise	10	19.877	.000
	Fear	5	32.240	.000		Fear	10	4.700	.000
	Neutral	5	213.446	.000		Neutral	10	2.208	.015
utterance type * boundary tone	Happy	4	26.968	.000	slope * duration	Happy	39	.744	.878
	Sad	4	27.428	.000		Sad	39	2.377	.000
	Angry	4	70.675	.000		Angry	39	2.291	.000
	Disgust	4	67.570	.000		Disgust	39	6.628	.000
	Surprise	4	255.789	.000		Surprise	39	1.759	.002
	Fear	4	72.631	.000		Fear	39	1.138	.256
	Neutral	4	56.886	.000		Neutral	39	1.135	.260

4.3.3 Logistic regression analysis

In order to model the perceptual results, we conducted a logistic regression analysis. The co-variations include the *duration* and *slope* of the final additive tone, *utterance length*, *and boundary tone category*. Dependent variations are the perceptual scores of seven emotions. Table 4 gives the estimated parameters in logistic regression formulae: (1) for each emotion. Then from these P formulae, the estimation surfaces were plotted in Figure 11 for 'Neutral' emotion with one of the other 6 emotions by settings 'boundary tone' = 1, 2, 3 and 'utterance type' = 1, 2, 3 (mono-syllabic, bi-syllabic and long utterances) respectively.

$$(1) \quad P = \frac{e^{b0+b1X1+b2X2+\ldots+bkXk}}{1+e^{b0+b1X1+b2X2+\ldots+bkXk}}$$

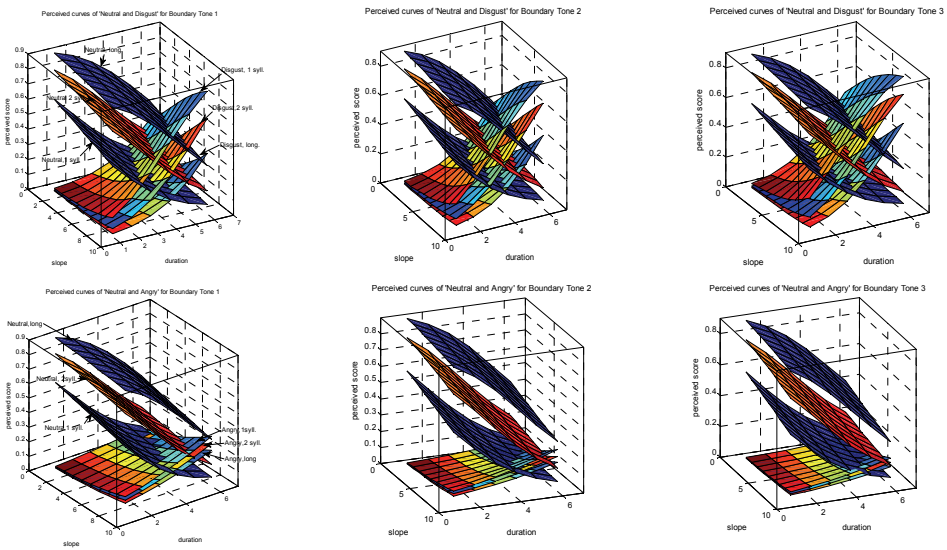
Table 4: Parameters estimated in logistic regression formulae and the correct class rate for the model

Emotions	$b_0+b_1X1+b_2X2+\ldots b_kX4$	Correct Class %
Neutral	0.123+0.975*utterancetype-0.083*boundarytone-0.068*slope-0.515*duration	68.8%
Angry	-3.302-0.232*utterancetype-0.424*boundarytone+0.200*slope+0.252*duration	93.6%
Disgust	-4.610-0.943*utterancetype+0.381*boundarytone+0.286*slope+0.681*duration	86.1%
Sad	-5.334-0.129*utterancetype+0.766*boundarytone-0.066*slope+0.421*duration	89.7%
Fear	-1.537-1.203*utterancetype+0.166*boundarytone-0.054*slope+0.402*duration	93.9%
Happy	-2.386+0.374*utterancetype-0.535*finaltone	92.8%
Surprise	0.913-1.186*utterancetype-0.092*slope-0.174*duration	89.1%

It is shown in Figure 11 that with the increase of duration and slope of the falling tone, the perceived scores of 'Neutral' emotion decrease, while the perceived scores of other emotions increase. The longer the utterance is, the higher the perceived score for 'Neutral' emotion, and the lower the perceived score for

other emotions. We observe similar tendencies for all six other emotions, but the tendency is more apparent for Disgust, Angry, Fear and Sad emotions.

The results further demonstrate that the change of the acoustic form of the additive tone attached after a ‘Neutral’ intonation will cause the pragmatic expression of the intonation. For the falling tail per se, ‘Disgust, Angry, Fear and Sad’ emotions are most likely to be perceived with the increasing slope and duration of the falling tone. In the same situation, the longer utterances are more likely to keep the ‘Neutral’ emotion than the shorter ones. In other words, the boundary tone has less effect on emotional expression for longer utterances than for shorter ones.



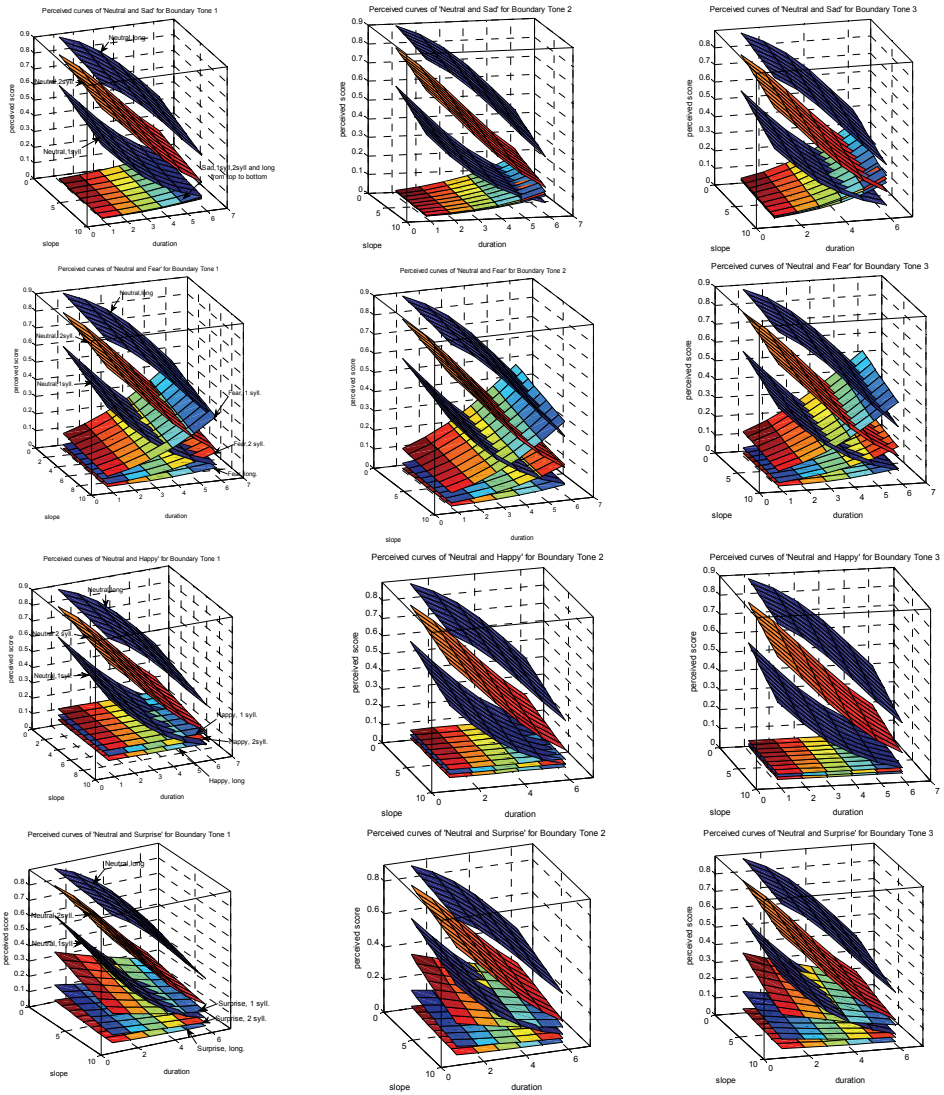


Figure 11: Simulated perception surfaces by the adoption of Logistic Regression formulae. Each row plots the surfs of 'Neutral' vs. one of the emotions for 3 utterance length (type) and 3 tone categories (in column).

5. Conclusion and discussion

The present study analyzed the F_0 patterns phonetically and phonologically for the 'successive addition boundary tone' (SUABT) in both monosyllabic and disyllabic utterances in order to separate the interaction between the form of Chinese SUABT and its expressive function. It further examines whether it can uniquely encode the emotional or pragmatic information. We also examined longer utterances.

We proposed that the SUABT is composed of two components, i.e. the lexical tone to express linguist meaning, and the expressive tone to express pragmatic meaning.

The overall patterns of emotional intonation can be represented phonologically by F_0 register and range in three levels, and the SUABT is described by traditional boundary tone features H% and L% with the addition of the successive tonal features 'r% or f%'. 'Neutral and Fear' intonations have both 'M' range and 'M' register, but 'Fear' intonation always has quiver voice. 'Sad' intonation has both 'L' range and register. 'Happy and Surprise' intonations have both 'H' range and register. 'Disgust' intonation has 'M' range and 'L' register. 'Angry' intonation has 'H' ranges and 'H' register.

New features 'r' (rising) and 'f' (falling) are proposed to describe the falling or rising features of expressive component in SUABT. In this case the boundary tone can be described by H% or L% (relative register of the boundary tone) plus an expressive tone feature 'r' or 'f': H/L-r/f/x%.

Although both rising and following SUABT are observed in the emotional intonations, currently only the falling SUABT in disgust boundary tone were elaborately analyzed (Li et al. 2012), i.e. the duration ratio of the two boundary tone components and the slope of the falling expressive tone are statistically

analyzed. Therefore, the perceptual experiment was only conducted for falling SUABT to clarify the form and function of the falling SUABT. To this end, the stimuli were obtained by elaboration of the boundary tones of nine emotionally neutral utterances into a falling SUABT, where the duration ratio of the expressive tone to the lexical tone and the slope of expressive tone are varied in six and nine steps respectively.

After the conduction of a multi-factor GLM analysis and a logistic regression analysis on the perceptual experiment results from twenty listeners, we found that the final falling additive tone can change the emotional expression of intonation. However, the perceived emotions vary with the change of the slope and duration of the additive final fall. Except for monosyllabic utterances of 'Disgust' emotion, none of the scores can go higher than the scores for 'Neutral' intonation, besides, many stimuli have more than one perceived results (Li et al. 2011). This result suggests that the successive addition boundary tone can be partly encoded to express emotional attitude, but the mapping from the form of the successive addition boundary tone to the pragmatic function is a one-to-many relation. In other words, the successive addition boundary tone cannot be encoded independently to express a specific emotion. On the other hand, from the simulated surface of perceived emotions in Figure 11, we know that one could perceive some common emotions produced with different acoustic features of SUABT, which indicates that the relation between the acoustic form of the SUABT and the pragmatic function of it is 'many to many'.

The factors affecting the perceptual score of seven emotions include the slope and the duration of the final successive fall, the utterance type and the final tone category. However, each factor makes different contributions to different emotions. The final tone factor could extend Yuan's tone-dependent mechanism to pragmatic aspect.

Shorter utterances are affected more obvious by the final successive addition tone than longer utterances. A reasonable interpretation is that longer utterances could employ richer components to express both linguistic and pragmatic formation than shorter ones, especially monosyllabic utterances. The components may include emotional focus, VQ (vice quality), speech rate, the boundary tone category and the intensity, etc. This reveals that both boundary tone and other components will be encoded in parallel to express linguistic and pragmatic information, as suggested by Xu's (2005) PENTA model.

The falling SUABT of Chinese may supply an interpretation why a dynamic lowering tone will be perceived as an anger man (Xu & Chuenwattanapranithi 2007).

However, the present study only investigates the form and function of the falling SUABT, the acoustic cues of raising SUABT should be analyzed in detail and simulated to verify the present results in the future.

The next generation human-machine or human-robot interactive system will be expected to behave as a "real human". Therefore, the future work will also concern with the spontaneous speech full of rich emotional expressions. Another question may be raised: whether SUABT could convey other emotions or attitudes in spontaneous speech. The answer is definitely yes. It can be obtained the intonation contour in Figure 12, clipped from our children language acquisition corpus, a disyllabic utterance, uttered by a 5 year-old girl, where a falling SUABT is employed to convey the imperative attitude.

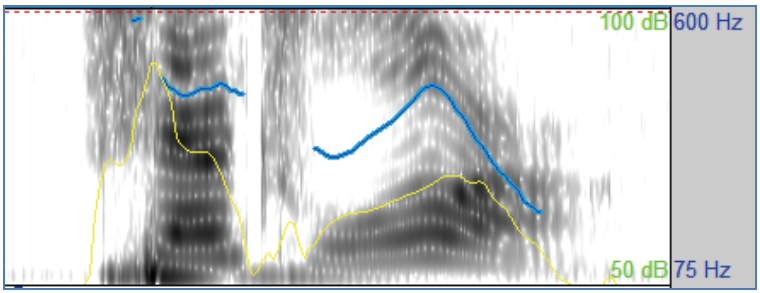


Figure 12: The intonation (blue line) of an imperative utterance 'Delete it!' (删除! shan1 chu2!) uttered by a 5 year-old girl. The boundary tone is T2 with a falling SUABT.

To sum up, the coming problem is to collect a speech corpus with natural and real emotional expressions. Fortunately, some clever methods have been introduced to collect such kinds of expressive databases (Douglas-Cowie et al. 2003, Mori et al. 2011). Based on the corpus, the intonation patterns should be investigated at the discourse level in a framework such as HPG (Hierarchical Prosodic Phrase Grouping) proposed by Tseng (2008): 'all the discourse prosody is higher level specification as well as cross-phrase association in addition to discrete intonation pattern'.

References

Benus, Stefan, Agustín Gravano, and Julia Hirschberg. 2007. Prosody, emotions, and… 'whatever'. *Proceedings of InterSpeech 2007*, 2629-2632.

Bolinger, Dwight. 1989. *Intonation and its Uses: Melody in Grammar and Discourse*. Stanford: Stanford University Press.

Callier, Patrick. 2011. On the edge: the sociophonetics of boundary tones and final lengthening in Mandarin Chinese. *eVox* 5.1:16-36.

Cao, Mengxue, Aijun Li, Qiang Fang, Jianguo Wei, and Chan Song. 2012. Acoustic and articulatory analysis on Japanese vowels in emotional speech. *Proceedings of ISCSLP 2012*, 40-44. Hong Kong: Chinese University of Hong Kong.

Chao, Yuen Ren. 1932. A preliminary study of English intonation (with American variants) and its Chinese equivalents. *Qingzhu Cai Yuanpei Xiansheng Liushiwu Sui Lunwenji* [*Festschrift in Honor of Professor Yuan-Pei Tsai on his 65th Birthday*], 105-156. Beiping: Institute of History and Philology, Academia Sinica.

Chao, Yuen Ren. 1933. Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology Academia Sinica* 4.2:121-134.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chao, Yuen Ren. 1980. A system of "tone-letters". *Fangyan* [*Dialect*] 1980.2:81-82.

Cowie, Roddy et al. 2001. Emotion recognition in human-computer interaction. 2001: *IEEE Signal Process*. Magazine 18.1:32-80.

Douglas-Cowie, Ellen, Nick Campbell, Roddy Cowie, and Peter Roach. 2003. Emotional speech: towards a new generation of databases. *Speech Communication* 40.1-2: 33-60.

Gussenhoven, Carlos. 2002. Intonation and interpretation: phonetics and phonology. *Speech Prosody 2002*, 47-57. Aix-en-Provence: Université de Provence.

Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge & New York: Cambridge University Press.

Hirschberg, Julia. 2002. Communication and prosody: functional aspects of prosody.

Speech Communication 36.1-2:31-43.

Jia, Yuan. 2012. *A Phonetic and Phonological Analysis of Focus in Standard Chinese*. Beijing: China Social Sciences Press.

Jiang, Ping, and Aishu Chen. 2011. Representation of Mandarin intonations: boundary tone revisited. *Proceedings of the 23rd North American Conference on Chinese Linguistics* (*NACCL-23*), ed. by Zhuo Jing-Schmidt, 97-109. Eugene: University of Oregon.

Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge & New York: Cambridge University Press.

Li, Aijun. 2008. Stress pattern of emotional speech. *Chinese Journal of Phonetics* 1:132-140. Beijing: The Commercial Press.

Li, Aijun, and Haibo Wang. 2004. Friendly speech analysis and perception in Standard Chinese. *Proceedings of ICSLP 2004*, *Paper in Spec4301o*. Jeju: The Acoustical Society of Korea.

Li, Aijun, Qiang Fang, and Jianwu Dang. 2011. Emotional intonation in a tone language: experimental evidence from Chinese. *Proceedings of ICPhS XVII*, 1198-1201. Hong Kong: City University of Hong Kong.

Li, Aijun, Qiang Fang, Yuan Jia, and Jianwu Dang. 2012a. More targets? Simulating emotional intonation of Mandarin with PENTA. *Proceedings of ISCSLP 2012*, 217-275. Hong Kong: Chinese University of Hong Kong.

Li, Aijun, Qiang Fang, Yuan Jia, and Jianwu Dang. 2012b. Successive addition boundary tone in Chinese disgust intonation. Paper presented at the 24th North American Conference on Chinese Linguistics (NACCL-24), June 8-10, 2012. San Francisco: University of San Francisco.

Lin, Maocan. 2004. Boundary tone of Chinese intonation and its pitch pattern. *From Traditional Phonetics to Modern Speech Processing*, ed. by Gunnar Fant, Hiroya Fujisaki, Jianfen Cao & Yi Xu, 309-327. Beijing: Foreign Language Teaching and Research Press.

Lin, Maocan. 2006. Interrogative mood and boundary tone in Chinese. *Zhongguo Yuwen* [*Studies of the Chinese Language*] 2006.4:364-379.

Lin, Maocan, and Zhiqiang Li. 2011. Focus and boundary in Chinese intonation. *Proceedings of ICPhS XVII*, 1246-1249. Hong Kong: City University of Hong Kong.

Liu, Fang, and Yi Xu. 2005. Parallel encoding of focus and interrogative meaning in Mandarin Chinese. *Phonetica* 62.2-4:70-87.

Lu, Jianming, and Maocan Lin. 2009. Raising tail of the question vs. the modal particle. *Frontiers in Phonetics and Speech Science*, ed. by Gunnar Fant, Hiroya Fujisaki & Jiaxuan Shen, 49-56. Beijing: The Commercial Press.

Mori, Hiroki, Tomoyuki Satake, Makoto Nakamurab, and Hideki Kasuya. 2011. Constructing a spoken dialogue corpus for studying paralinguistic information in expressive conversation and analyzing its statistical/acoustic characteristics. *Speech Communication* 53.1:36-50.

Mueller-Liu, Patricia. 2006. Signaling affect in Mandarin Chinese: the role of utterance-final non-lexical edge tones. *Speech Prosody 2006*, PS6-3-48. Dresden: Dresden University of Technology.

Pierrehumbert, Janet Breckenridge. 1980. *The Phonology and Phonetics of English Intonation*. Cambridge: MIT dissertation.

Pierrehumbert, Janet, and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. *Intentions in Communication*, ed. by Philip R. Cohen, Jerry L. Morgan & Martha E. Pollack, 271-311. Cambridge: MIT Press.

Sagisaka, Yoshinori. 2012. Modeling prosody variations for communicative speech and the second language towards trans-disciplinary scientific understanding. *Keynote Speech of Speech Prosody 2012*, 187-193. Shanghai: Tongji University.

Scherer, Klaus R. 2003. Vocal communication of emotion: a review of research paradigms. *Speech Communication* 40.1-2:227-256.

Scherer, Klaus R. 2009. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364:3459-3474.

Schröder, Marc. 2004. *Speech and Emotion Research: An Overview of Research Frameworks and a Dimensional Approach to Emotional Speech Synthesis*. Saarbrücken: Institut fur Phonetik, Universität des Saarlandes.

Schuller, Björn, Stefan Steidl, and Anton Batliner. 2009. The INTERSPEECH 2009 emotion challenge. *Proceedings of InterSpeech 2009*, 312-315.

Silber-Varod, Vered. 2011. Dependencies over prosodic boundary tones in spontaneous spoken Hebrew. *Proceedings of Depling 2011* (*International Conference on Dependency Linguistics*), 241-250. Barcelona: Pompeu Fabra University.

Tao, Jianhua, Yongguo Kang, and Aijun Li. 2006. Prosody conversion from neutral speech to emotional speech. *IEEE Transactions on Audio*, *Speech*, *and Language Processing* 14:1145-1154.

Tseng, Chiu-yu. 2006. Higher level organization and discourse prosody. *The Second International Symposium on Tonal Aspects of Languages*, 23-34. La Rochelle, France.

Tseng, Chiu-yu. 2008. Corpus phonetic investigations of discourse prosody and higher level information. *Language and Linguistics* 9.3:659-719.

Wang, Haibo, Aijun Li, and Qiang Fang. 2005. F0 contour of prosodic word in happy speech of Mandarin. *Affective Computing and Intelligent Interaction*, ed. by Jianhua Tao, Tieniu Tan & Rosalind W. Picard, 433-440. Berlin & Heidelberg: Springer.

Wu, Zongji. 1990. Basic contour patterns of intonation of Standard Chinese. *Essays in Honor of Professor Wang Li*, 54-73. Beijing: The Commercial Press.

Wu, Zongji. 1995. Predictability of different attitudinal intonation in Standard Chinese. *Proceedings of the 13th International Congress of Phonetic Sciences*, Vol. 2. Stockholm, Sweden.

Wu, Zongji. 2000. From traditional Chinese phonology to modern speech processing: realization of tone and intonation in Standard Chinese. Keynote speech of *International Conference of Spoken Language Processing 2000*. Beijing: Institute of Automation and Institute of Acoustics, CAS; Beijing Jiaotong University.

Xu, Yi, and Suthathip Chuenwattanapranithi. 2007. Perceiving anger and joy in speech through the size code. *Proceedings of ICPhS XVI*, 2105-2108.

Xu, Yi, Andrew Kelly, and Cameron Smillie. 2011. Emotional expressions as communicative signals. *Prosody and Iconicity*, ed. by Sylvie Hancil & Daniel Hirst, 33-60. Amsterdam & Philadelphia: John Benjamins.

Xu, Yi. 2005. Speech melody as articulatorily implemented communicative functions.

Speech Communication 46.3-4:220-251.

Yuan, Jiahong. 2006. Mechanisms of question intonation in Mandarin. *Chinese Spoken Language Processing*, ed. by Qiang Huo, 19-30. Berlin: Springer.

漢語情感語調的連續疊加邊界調：產出與感知

李愛軍

中國社會科學院

趙元任先生指出漢語的語調實際上是字調和語調本身的“代數和”，並指出存在兩種聲調疊加的模式：同時疊加 (simultaneous addition) 和連續疊加 (successive addition)。連續疊加是指“這種上升和下降不會同時加載最後的音節上，而是在字調讀完以後再接上去的”(1933)。本研究發現漢語普通話發音人使用這種連續疊加邊界調來表達情感語氣，比如在憤怒和厭惡的情感語調末尾出現了下降的連續疊加邊界調 (successive addition boundary tone: SUABT)；在高興和驚訝語調末尾出現了上升的連續疊加邊界調。這個連續疊加的邊界調包括兩個成分：第一個成分是聲調 (lexical tone)，對邊界音節的聲調進行編碼，傳遞語言學信息；第二個成分我們稱之爲情感表達調 (expressive tone)，傳遞情感語氣和態度等語用功能。本文對連續疊加邊界調的聲學表現進行分析，並給出音系的表達。同時，在對厭惡語調的下降後續疊加邊界調的兩個成分進行聲學分析基礎上，設計感知實驗，來考察連續疊加邊界調下降成分的聲學變化與整個語調情感表達的關係，是否連續疊加邊界調能夠對情感語調的表達進行獨立編碼，也即情感語調的形式與語用表達功能之間的關係。實驗刺激是在中性語調的邊界調後增加一個下降成分（改變下降的斜率和下降成分的時長），使邊界調在原調（除了去聲外的三個調）基礎上增加一個下降尾巴。通過對感知結果的線性回歸分析，我們發現連續疊加成分的聲學特性、邊界調的聲調類型以及句子的長短等因素對情感語氣的感知都有顯著的影響，但是對感知不同的情感影響不同。而且，連續疊加邊界調的聲學形式和語用表達功能之間不是一對一的關係，而是多對多的關係。也說明了，發音人不會僅僅利用連續疊加邊界調這一種手段對情感語氣表達進行編碼，而是同時使用其他手段來傳遞情感語氣。

關鍵詞：漢語語調，情感語調，邊界調，連續疊加邊界調

Human Language Resources and Linguistic Typology
Papers from the Fourth International Conference on Sinology
Taipei: Academia Sinica, 2013

A Conceptual Model of Chinese Illocution, Emotion and Prosody

Yueguo Gu
Chinese Academy of Social Sciences

This paper presents a conceptual model for empirically studying the tripartite interaction between illocution, emotion and prosody in Chinese. The illocutionary act, taken as given, is defined as a property exemplified by the speaker in producing an illocutionary act-token in a given social situation. Technically the act-type property is formally constructed and represented as an octet. Occurrent emotions are anatomized in three tiers, background, primary and social. Prosody is construed as the psychological attributes of sound such as pitch, loudness and duration. The tripartite interactions are empirically studied in performance units of illocutionary acts taken from the Spoken Chinese Corpus of Situated Discourse (SCCSD). Also discussed are two assumptions about child discourse: (1) that there exists a perfect match between what is said, what is thought of, and what is felt (the STF-Match Assumption), and (2) that if there are prosodic cues produced, they are genuine, that is, the cues truly reflect what is said, what is thought of, and what is felt (the PG Assumption).

Keywords: Chinese illocutionary acts, Chinese emotions, Chinese prosody, the STF-Match Assumption

1. Preliminary remarks

This paper examines the tripartite interaction between Chinese illocution, emotion and prosody. As our pragmatic intuition tells us, when the speaker explicitly apologizes to the hearer by saying *wo xiang ni daoqian* (我向你道歉 'I to you apologize'), if it is sincerely performed, there is an emotional state accompanying the act, that is, the speaker feels remorseful. Furthermore, the utterance is articulated with the intonation that is appropriate for the act. Our everyday experience also indicates to us, that not all apologies are sincere, because the speakers are heard as being *you kou wu xin* (有口無心 'have mouth no heart'). Surely speakers themselves know which cases of performance are sincere, and which are not. Hearers, though less certain than speakers, can still detect the differences through *cha yan guan se* (察言觀色 'examine words, look at facial colour') and *ting hua ting yin* (聽話聽音 'listen to talk listen to sounds').

In a sense this paper presents our attempt to study systematically and empirically our everyday pragmatic competence of *cha yan guan se* and *ting hua ting yin*. *Yan* (言 'words') is examined in terms of the illocutionary act/force; *se* (色 'facial colour') is looked at in terms of emotion revealed in facial expression; and *yin* (音 'sounds') here is construed as prosody. There is a pragmatic correlation between the illocutionary act/force (*yan*), emotion/facial expression (*se*), and prosody (*yin*). Emotional states accompanying illocutionary acts are bound to be shown, manipulatively or non-manipulatively alike, in facial expressions and prosodies. That the study is systematic and empirical is in the sense that it is based on the Spoken Chinese Corpus of Situated Discourse (SCCSD).[1]

[1] The SCCSD is a collection of spontaneous activities in real-life social situations audio- and video-taped. It is labeled as a multimodal corpus, for it consists of three types of texts,

In what follows, we first, as background preparation, spell out the working definition of illocution (see §2). Illocution in this paper is taken as given. The research focus is placed on (1) the interaction between illocution and emotion (see §3), and (2) the interaction between illocution-emotion and prosody (§4). The research methodology and procedure are discussed in §5. The literature review will be carried out on the run, but very sparingly for lack of space. The paper concludes with discussions of some remaining theoretical issues.

Before we go to details, some explanation about "a conceptual model" is in order here. The notion is borrowed from software engineering, in which conceptual modeling, very important methodology now in the field (see e.g. Borgida et al. 2009, Parsons et al. 2010), is the first thing application developers do before they write programme codes. For them, conceptual modeling is content analysis proper. The programme codes constitute a way to implement the content analysis in order to materialize an application. In this paper, our conceptual model is analogous to the product of conceptual modeling. That is, the analytic framework to be presented in this paper remains conceptual, and it will not spell out the details about how the framework is to be implemented at a large scale in future corpus development (see also §8 below).

2. Illocution: some background preparation

2.1 Conceptual analysis

The illocutionary act is formulated on the basis of the insight drawn by Austin (1980[1962]) that, given the right circumstances, in saying U (=utterance), the speaker is also performing an Act, called an illocutionary act. Since its

orthographic, audio, and video, which are synchronized currently in 5-minute chunks. For details see Gu (顧曰國) 2002; also www.multimodalgu.com.

formation, the illocutionary act has remained the hard core of pragmatics and an enormous amount of literature has been accumulated. We may view the illocutionary act as a three-dimensional phenomenon: the Act-dimension, the U-dimension, and the H-dimension. Thus viewed, the illocutionary act is a subset of all the acts performed by humans, admitting the fact that there are many human acts that are performed without involving any speech at all, and furthermore, given an illocutionary act, e.g. greeting, it can be performed by nonverbal means. The U-dimension is a subset of all the utterances that are associated with the performance of the illocutionary act. Given an illocutionary act, it can be performed explicitly using performative utterances with the syntactic formula: *wo* + (*qita chengfen*) + *zhishi dongci* + (*qita chengfen*) (我+(其他成分)+直施動詞+(其他成分)). For example,

(1) the act of *xuanshi* (宣誓 'swear')

<u>我</u>代表全體裁判員和工作人員<u>宣誓</u>：恪守體育職業道德，

wo	daibiao	quanti	caipanyuan	he	gongzuo	renyuan
(I	on behalf	all	referee	and	working	staff

xuanshi:	keshou	tiyu	zhiye	daode
(swear:	uphold	sports	profession	ethics)

(2) the act of *gantan* (感歎 'feel')

可是<u>我</u> <u>感歎</u>我可大不如你們……

keshi	wo	gantan	wo	ke	da	buru	nimen
(but	I	feel	I	indeed	very	no match	you)

Or, it can be performed implicitly, e.g. (3) and (4):

(3) the act of *zhaoren* (找人 'looking for somebody by way of asking')
王……同學來了嗎
Wang tongxue lai le ma
(Wang classmate come ASP Particle)

(4) the act of *yaoqiu* (要求 'requesting')
往中間坐一坐，謝謝呀
wang zhongjian zuo yi zuo, xiexie ya
(to middle sit one sit thanks Particle)

Finally the H-dimension is no less essential in that the illocutionary act is social in nature, and that it involves the hearer's participation and partnership in providing "uptake" (Austin's term). This is particularly acute in such Chinese illocutionary acts as *cai quan xingling* (猜拳行令 'number-guessing game'). Minimally it takes two persons to do this. S and H, two drinkers, each have to do two things simultaneously: (a) the verbal utterance, e.g. *yi dingzhong* (一定中 'one hit right'), *liang xianghao* (兩相好 'two making a good match'), *san liushun* (三六順 'three going well with six'), *siji facai* (四季發財 'four seasons all making fortunes'), *wuzi dengke* (五子登科 'five sons all becoming officials'), etc.; (b) the nonverbal gesture showing a figure, e.g. using one finger to indicate Figure 1, two fingers for Figure 2, etc.

The three dimensions have been elucidated in various ways. Regarding the Act-dimension, Austin's analysis of felicities and infelicities, and Searle's successful conditions analysis (1969, 1979) have laid a foundation for the subsequent studies in the last half century. The U-dimension has attracted a great deal of attention from linguists (e.g. Leech 1983, Levinson, 2000) as well as from philosophers (e.g. Grice 1989, Recanati 2004), to name but only few. As we know, the utterance has a sort of intrinsic meaning of its own (e.g. a sentence-type

meaning). When it is used to perform an illocutionary act, it seems to have acquired an extra meaning, viz. illocutionary force. In case of performative utterances, what is literally uttered makes the force transparent, whereas in non-performative utterances, the force has to be worked out through inferences (e.g. via conversational implicature). The study of illocutionary force indicating devices (IFID) thus has become an important research topic (e.g. Zaefferer 1998, Wilson & Sperber 1998). The H-dimension, generally touched upon marginally by the forerunners, is taken up e.g. by Clark & Carlson (1982).

Now let us continue using the example of *wo daibiao quanti caipanyuan he gongzuo renyuan xuanshi*. In Austin's view, it is a "total speech act", out of which an illocutionary act, i.e. *xuanshi* (swearing, or pledging) is extracted through abstraction. While subscribing to the multiplier thesis about human action in general (see Davis 1984 for further discussion), we apply Goldman's distinction between act-types and act-tokens (1970:10) to the illocutionary act, thus giving rise to the distinction between the illocutionary act-type, and the illocutionary act-token respectively. The illocutionary act-type is simply an act-property, a property such as swearing an oath, greeting somebody, promising, etc. To describe a particular speaker as performing an illocutionary act of swearing, is equivalent to saying that the speaker has exemplified a swearing act-property at a given social situation. *Wo daibiao quanti caipanyuan he gongzuo renyuan xuanshi* is a swearing act-token by which the speaker has exemplified the swearing act-property. When a sportsman also swears on behalf of all the players, he produces another swearing act-token, e.g. *wo daibiao quanti yundongyuan xuanshi* (我代表全體運動員宣誓 'I swear on behalf of all sportsmen'). The two act-tokens are not identical because they do not "involve the same agent, the same property, and the same time" (Goldman's words, ibid). When we say that the referee and the sportsman both have performed the

illocutionary act of the same type, namely swearing, what we are actually doing is idealizing and approximating by way of ignoring the differences the speakers exemplified in the act-tokens.

In a nutshell, an illocutionary act, in this paper, is defined as a property exemplified by the speaker in producing an illocutionary act-token in a given social situation. We define this property more formally adopting the formula from set theory.

> The Illocutionary act/force {S { }, H {}, performativity {}, the essential content {}, the intentional states {}, the emotional states {}, the occasion {}, the interdependency {}}[2]

Note that the order of the elements is irrelevant; {} signals that each element itself can have subsets of its own. Here are the further elaborations.

S { } —> {individual, collective}, meaning that for every illocutionary act there is an agency substantiated in the speaker role that can be individual or collective;

H {} —> {addressee, audience, addressee + audience}, meaning that for every illocutionary act there is a co-agency substantiated in the hearer role that can be addressee, or audience, or both;

Performativity {} —> {performative, non-performative}, meaning that for every illocutionary act there is a performance unit that is performed performatively or non-performatively;

[2] Goldman uses three determinants, agent, property, and time. In footnote 13, he mentions a fourth determinant, the *way* (1970:11). In our case, even four determinants are not adequate. We propose eight.

The essential content {} —> {explicit, implicit}, meaning that for every illocutionary act there is an essential content which can be made explicit or remain implicit;

The intentional states {} —> {attitude, belief, attitude + belief}, meaning for every illocutionary act there is some attitude, or belief or attitude and belief associated with it; attitude defined as "a relatively enduring organization of beliefs around an object or situation predisposing one to respond in some preferential manner" (Rokeach (1970[1968]:112)

The emotional states {} —> {background emotion, primary emotion, social emotion, background + primary + social emotions}, meaning that for every illocutionary act there is some background emotion, or primary emotion, or social emotion, or all of them; see §3 below for detailed discussion.

The occasion {} —> {situated space, situated time, situated activity}, meaning that for every illocutionary act there are spatial-temporal, and situational properties that are bound to and inherited from the particular, specific occasion of occurrence;

The interdependency {} —> {situated interdependency, non-situated interdependency, situated + non-situated interdependency}, meaning that for every illocutionary act there is an interdependency relation with what is happening at the here-and-now behavior setting, and with what happens beyond the here-and-now, or with both.

The octet serves as a logical basis for our project of Chinese speech act ontology using OWL (the project still in progress). Technically an illocutionary

act is a specific configuration of specific values[3] being assigned to those members by the speaker/actor while performing the act. Logically, any performance by any speaker/actor gives rise to a specific configuration, thus resulting in an illocutionary act being performed. In China, there are now over 1.3 billion speakers/actors, who, while speaking, will assign a set of 1.3 billion specific values to those members, thus generating 1.3 billion instances of illocutionary acts.

The octet — The act/force {S { }, H {}, performativity {}, the essential content {}, the intentional states {}, the emotional states {}, the occasion {}, the interdependency {}} — is a highly abstract pattern formulated on the bases of studies of a limited number of instances. It is open to revision of course, but is sufficiently adequate for the present purpose.

2.2 Performances of illocutionary acts

As discussed above, an illocutionary act is defined as a property exemplified by the speaker in producing an illocutionary act-token in a given social situation. Accordingly, when the speaker says something to the addressee and/or to an audience, if her or his saying-act-token exemplifies the properties as shown in the octet, s/he performs an illocutionary act of a certain type. Now let us demonstrate how this exemplification theory of illocution and the octet are applied to real data. Here is the performance of an illocutionary act of *gandao* (感到 'feel') taken from the Opening Ceremony of Beiwai Online Education Institute. The speaker, a University president, adopts a performative utterance *wo…gandao…* (我…感到… 'I…feel…'). The actual performance unit reads like this (note that the orthographic representation is used here for easy reading):

[3] The term *value* is a technical term borrowed from software engineering. Simplistically speaking, a value here means a token instance functioning as a member of the set in question.

(5) The performance of *gandao* (感到)

爲此我爲你們成爲我們學校第一屆網絡學院//網絡教育學院的學生和老師//感到高興和自豪.

weici wo wei nimen chengwei women xuexiao
(thanks to I for you become our university
diyijie wangluo xueyuan // wangluo jiaoyu xueyuan
first internet college internet education college
de xuesheng he laoshi // gandao gaoxing he zihao
of student and teacher feel happy and proud)

The conceptual analysis of *gandao* using the octet is shown as Figure 1.

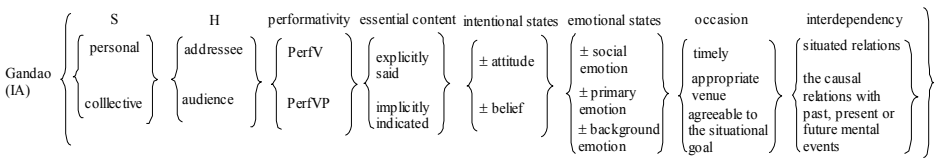


Figure 1: Conceptual analysis of *gandao*

Note: ± means positive or negative, e.g. positive or negative social emotions, etc.

The elements shown in Figure 1, in theory, all can be verbalized by the speaker. In practice, only some are verbalized, and the remaining is taken for granted. In case of *gandao*, the element that must be verbalized is the ***intentional states*** {+/−attitude, +/− belief}. This is the essential part of the act *gandao*. If the speaker indeed only verbalizes this element, s/he has adopted a non-performative and implicit strategy. If the speaker, on the other hand, chooses to verbalize S (i.e. *wo* 我 in Chinese), V (i.e. *gandao*), etc., as the real-life speaker above does, s/he adopts a performative and explicit strategy.

In (1) above, *weici* verbalizes the interdependency relation with the previous discourse; *wei nimen chengwei women xuexiao diyijie wangluo xueyuan wangluo jiaoyu xueyuan de xuesheng he laoshi* makes the interdependency relation explicit, i.e. the cause for feeling happy and proud; *gaoxing he zihao* (i.e. emotional states, see further discussion below), is the essential content of the act.

Situationally speaking, knowledge about *weici*, *wo*, *wei nimen chengwei women xuexiao diyijie wangluo xueyuan wangluo jiaoyu xueyuan de xuesheng he laoshi*, can be constructed by the hearer and taken for granted. The fact that they are verbalized and made seemingly redundant shows that human interaction goes far beyond sheer information exchange. Emotional satisfaction, for instance, needs both explicit verbalization and so-called "redundant" information.

So much for the background preparation. Now we turn to the major concern of the paper.

3. Emotion

3.1 Early treatment in Confucianism

As we know, *qinggan* (情感 'emotion') is one of the pivotal concepts in Chinese culture. Meng Peiyuan, a contemporary philosopher, argues that "情感是全部儒學理論的基本構成部分，甚至是儒學理論的出發點." (Emotion constitutes a fundamental part of all Confucianism. It can even be said that it originates from it. 蒙培元 2002:2; see also 蒙培元 1998). In 禮記, one of the earliest works of Confucianism, emotion is explicitly stated as follows: "何謂人情? 喜怒哀懼愛惡欲七者，弗學而能." (What is emotion? There are seven types: Happiness, anger, sorrow, fear, love, hate, desire, which are possessed not through learning.) These seven emotions, assumed to be in-born, need to be cultivated via the leadership of the sage king: "故聖人所以治人七情，修十義,

講信修睦，尚辭讓，去爭奪，舍禮何以治之?” (Therefore, the sage king cultivates people’s seven emotions, ameliorates ten principles, educate people about trust and harmony, promote modesty and politeness, and dissolve conflicts. Without resorting to rites, how can all these be achieved? see 禮記禮運.) This shows that the cultivation of seven emotions is part of the *li*-programme, i.e. the execution of sovereign rule through rites and rituals. Later in 禮記樂記, emotion is further dealt with in connection with music.

> 樂者，音之所由生也；其本在人心之感於物也。是故其哀心感者，其聲噍以殺。其樂心感者，其聲嘽以緩。其喜心感者，其聲發以散。其怒心感者，其聲粗以厲。其敬心感者，其聲直以廉。其愛心感者，其聲和以柔。六者，非性也，感於物而後動。是故先王慎所以感之者。故禮以道其志，樂以和其聲，政以一其行，刑以防其奸。
>
> (With regard to music, it is composed of sounds. It originates from the human heart-mind being touched by the outside world. Hence, the heart-mind feeling sorrowful gives birth to anxious and tumbling sounds; the heart-mind feeling pleased gives birth to broad and relaxed sounds; the heart-mind feeling happy gives birth to high-pitched and jubilant sounds; the heart-mind feeling angry gives birth to rude and sharp sounds; the heart-mind feeling respectful gives birth to pious and articulated sounds; the heart-mind feeling affectionate gives birth to thoughtful and gentle sounds. All the six cases are not inborn, but responses triggered by the interactions with the outside world. That is why the sage king is very prudent about emotions. Hence he makes rites to foster social consciousness, music to harmonize sounds, rules to regulate behavior, and laws to prevent evils.)

Note that *yue* 樂 in the ancient times includes (1) oral texts to be sung (詩經 being a collection of such texts); (2) music played by instruments; and (3) dance. So *yin* "音" in the extract includes oral texts, or music, or both. One of the distinctive features of oral texts for singing is their prosody. The singing texts, according to the extract, are originated in the heart-mind[4] that interacts with all sorts of things. As a result, the heart-mind generates different emotions. The heart-mind, charged with varied emotions, will produce corresponding sounds (i.e. prosodies in the terminology of this paper):

哀心 (sorrowful heart-mind)	聲噍以殺 (sounds anxious and heavily low)
樂心 (pleased heart-mind)	聲嘽以緩 (sounds broad and relaxed)
喜心 (happy heart-mind)	聲發以散 (sounds high-pitched and jubilant)
怒心 (angry heart-mind)	聲粗以厲 (sounds rude and sharp)
敬心 (respectful heart-mind)	聲直以廉 (sounds pious and articulated)
愛心 (loving heart-mind)	聲和以柔 (sounds thoughtful and gentle)

There exists a long tradition of philosophical treatment of emotions by various schools of thoughts. It is impossible for this paper to review them here. Apart from Meng's works cited above, readers may very well consult Santangelo (2003), Yao & Tu (2011).

3.2 Emotion in western literature

In the Western literature, there is a general consensus about the distinction between studying emotion as disposition (i.e. long-term emotional properties as

[4] The term heart-mind is deliberately adopted here in order to show its difference from the term brain-mind. As indicated in the text, the Chinese philosophy emphasizes *qing yu zhi de jiaorong* (情與智的交融 'emotion and intellect integration'), hence the heart-mind.

part of personality), and as occurrence (i.e. currently emotional responses to internal and external events). This distinction is upheld in this paper. The Chinese philosophical treatment of emotions can be said to be largely focused on dispositional emotion. The emotional disposition of course affects emotional occurrence, but the investigation of this interaction requires a longitudinal study, and it has to be left unexplored herein. The emotions presently dealt with will be occurrent emotions, or using our preferred term, "situated emotions".

As well recognized in the research community, emotion is multifaceted. First, it is evolutionary, that is, animals also have emotions (see e.g. Darwin 1899). Second, it is physiological, that is, emotions involve physiological and neurochemical reactions; Pananicolaou (1989:xi) succinctly states: "Emotions are caused by stimuli, by the brain, by the limbic brain, by the right hemisphere of the brain. They are due to environmental influences, to neurotransmitters to hormones or due to 'apperceptive masses'." Third, the majority of emotions are cognition-mediated (see Dalgleish & Power 1999, Maiese 2011 among others). Fourth and finally, the majority of emotions are socially acquired (see Elster 1999 for instance).

Following Damasio (1996, 1999) and Damasio & Meyer (2009), we treat emotions in three tiers: Tier 1, background emotions; Tier 2, primary emotions; and Tier 3, social emotions.[5] Background emotions refer to emotional states such as anxiety (焦慮), fatigue (疲憊), discouragement (沒神), malaise (病態), relaxation (放鬆), well-being (康健), enthusiasm (興致), and energy (精神). As pointed out by Damasio & Meyer (2009:5), background emotions continuously underscore all human conscious and purposeful actions.

[5] The notion of tier is used here to indicate that the three types of emotion can occur simultaneously. It also implies the degree of conscious control. The social emotions are more likely to be consciously controlled.

> Background emotions are composite expressions of those regulatory actions [of human body] as they unfold and intersect moment by moment in our lives. I imagine background emotions as the largely unpredictable result of several concurrent regulatory processes engaged within the vast playground that our organisms resemble. These include metabolic adjustments associated with whatever internal need is arising or has just been satisfied; and with whatever external situation is now being appraised and handled by other emotions, appetites, or intellectual calculation. The ever-changing result of this cauldron of interactions is our "state of being," good, bad, or somewhere in-between. When asked "how we feel," we consult this "state of being" and answer accordingly. (Damasio 2003:44)

Background emotions, among other things, affect the prosodic key, i.e. the pitch range of the speaker (see §6.3 below for further discussion).

Primary emotions refer to emotions such as disgust, fear, happiness, sadness, surprise, and anger. These are found to be universal by some researchers. The seven emotions as defined in 禮記 (see above) can be regarded as primary emotions, though it is an open question whether they are in-born, or developed postnatally.[6] Social emotions, on the other hand, as the name indicates, are emotions learned socially.[7] Classification of social emotions varies as much as cultures diversify. This is quite understandable, for different cultures evaluate social emotions differently. Against the Chinese cultural background, we have

---

[6] Chinese medicine also recognizes seven emotions, but slightly differently from 禮記, which are *xi* (喜 'happy'), *nu* (怒 'angry, furious'), *you* (憂 'worry'), *si* (思 'miss'), *bei* (悲 'sad'), *kong* (恐 'shocked'), *jing* (驚 'frightened').

[7] Some researchers differentiate social emotions from other emotions in terms of private vs. non-private dimensions. We do not subject to this position.

mapped out social emotions in three sub-categories: (1) positive social emotions, (2) negative social emotions, and (3) neutral social emotions, as shown in Figure 2.

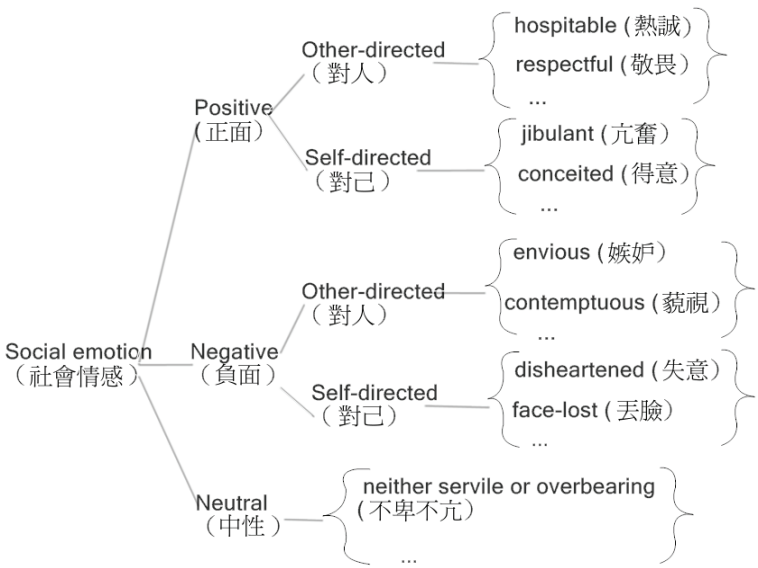


Figure 2: Social emotion: a classification

The treatment of emotions in terms of three tiers is meant to indicate that an experiencer may experience the three types of emotion simultaneously. Since we adopt Damasio's position, the background emotion is always present in situated discourse, while primary and social emotions are events-related. It is not uncommon that the emotions experienced in real-life situations are mixed ones rather than simple and pure ones. As Budd (1985:6) observes, the fact that someone experiences a certain thought with pleasure can distress him and, conversely, the fact that he experiences a certain thought with pain can please him. Furthermore, someone can be displeased that he experiences pride or he can derive pleasure from being sad (see §6 for further discussion).

Ekman (1994:56) maintains that "emotions can be very brief, typically lasting a matter of seconds or at most minutes. When we speak of an emotion lasting for hours, we probably are summating the recurrent emotion episodes within that time period." The emotions Ekman has in mind are primary and social situated emotions in our classification, and his observation does not apply to the background emotions. Ekman's time dimension (a matter of seconds or at most minutes) creates a verifiable hypothesis in our analysis (see further discussion §7 below).

4. Illocution and emotion

4.1 Dual, complementary perspectives of investigation

The complex interactions between illocution and emotion can be dealt with in two complementary perspectives. The first is from the illocution perspective or the hearer perspective: How does illocution function as antecedents to emotional experiences? In everyday parlance, there is a saying *yijuhua neng ba ren shuo xiaoqilai*, *ye neng ba ren shuo bengqilai* (一句話能把人說笑起來, 也能把人說蹦起來). That is, using the terminology of this paper, an illocutionary act performed with a certain force can make people happy with laughter or jump with outrage.

The second is from the emotion perspective or the speaker perspective. In this connection there are three cases: (1) emotion underscoring illocution; (2) emotion enabling or derailing illocution, and (3) emotion merging with illocution. The first is concerned with the way the performance and perception of the illocutionary act/force is affected by background emotions. For example, when greeting (打招呼) the hearer, if the speaker is in a state of discouragement (沒神), or anxiety (焦慮), this will make the act of greeting infelicitous (meaning not

perfect), and sends a signal to the hearer that something has gone amiss. Similarly, when receiving the speaker's greeting, if the hearer is found in the same type of emotional states, this will just as well affect the perception of the act/force.

The second is derived from Searle's speech act theory (1969:60). Searle's sincerity condition (or rule, 1969:63) states that in performing some illocutionary acts (e.g. promise, order, assert, etc.), there must be some appropriate intentional states accompanying the performance. With promise, for example, the sincerity condition is that the speaker intends to do the act promised. Without this intention, the promise is insincere. In our analysis, this intentional state is associated with the social emotion of *zhencheng* (真誠, a subclass of 熱誠). Searle confines himself to the intentional states, which, in turn, are co-determined with the propositional content. Because of this, he writes: "Greetings are a much simpler kind of speech act. … In the utterance of 'Hello' there is no propositional content and no sincerity condition." (1969:64). In our analysis, we are no longer confined to intentional states only. We believe that emotional states are no less important than intentional states in affecting the felicity of the illocutionary act performance. Thus, greeting may have no sincerity condition in Searle's sense, but it has emotional states with it in our analysis. The default emotion accompanying the act of greeting is the primary emotion of happiness. The absence of this emotion will render greeting infelicitous.[8]

The third case refers to the fact that there are illocutionary acts of emotions themselves, e.g. *yanwu* (厭惡 'disgust'), *jingya* (驚訝 'surprise'), *ai* (愛 'love'), *diuren* (丟人 'lose face'), *hen* (恨 'hate'), to name but only few. With many people, these emotions, if genuine, are difficult to hide. They merge

[8] Note that the background emotion underlying greeting, e.g. anxiety accompanying the speaker's greeting, is not conceptually associated with greeting. The primary emotion, e.g. "being happy or pleased" is conceptually considered as part of the act of greeting.

seamlessly with the performance of the illocutionary acts of emotions in question. The illocutionary act of *gandao* discussed above arguably merges with the mixed emotions of being happy (i.e. primary emotion) and being proud (i.e. other-directed social emotion). The mixed emotions were actually verbalized by the speaker (i.e. *gandao gaoxing he zihao*) in (5) above.

4.2 Occurrent emotions, emotion episodes and activity types

Emotions associated with performances of specific illocutionary acts are occurrent emotions. We hypothesize that occurrent emotions (mainly social emotions) are co-extensive with illocutionary performances. This co-extension hypothesis is verifiable through empirical investigation of actual performance data. Ekman's view that emotions last in seconds or at most minutes (see quote above) lends support to this hypothesis.

As Frijda and others have pointed out, some occurrent emotions may last much longer than Ekman shows.

> When subjects were asked to report a recent emotional incident, 50 percent of these incidents was described as having lasted for longer than one hour, and 22 percent for longer than 24 hours…It was clear from the descriptions of these incidents that the affective states focused on the emotional object or event during those entire intervals, either continuously or intermittently. If intermittently, the successive phases were felt to belong together and not represent a succession of individual emotions. (Frijda 1994a:61)

Frijda introduces the notion of "emotion episode" to talk about his finding. This paper cannot provide any fresh evidence to verify or falsify Frijda's finding.

But we feel it is very plausible to associate Frijda's emotion episode with the activity type. For instance, attending a wedding ceremony may give the attendee an emotional state of happiness that may indeed last quite long, in spite of the fact that the attendee may be momentarily upset, e.g. by being forced to drink too much alcohol. Instead of overstretching Frijda's notion of emotion episode, we shall use *atmospheric emotion* (氣氛情感) to describe the general emotional pattern that is dominant in a given social situation/activity. Hatfield, Cacioppo & Rapson (1994) show us that emotions are contagious. It is perhaps particularly true of atmospheric emotion, which is even normative, as it would be totally out of place or face-losing if the wedding ceremony attendee publicly exhibits the emotional state of sadness instead of happiness.[9] A more telling example is found lately in Xi'an, China. Mr. Yang, a provincial government official in charge of public safety, was found beaming at a horrific scene of car crashes resulting in 36 deaths (see Figure 3).

Figure 3: A beaming face at a tragic scene

[9] Note that some mothers burst into sobbing, seeing their daughters being carried away by bridegrooms. This is the case of mixed emotions of happiness and sadness. In folk terminology, it is *xi zhong bei* (喜中悲 'sorrow in happiness').

The atmospheric emotion is extremely sad. The official in charge shows a beaming face, which reveals that he is quite happy. As pointed out above, the atmospheric emotion can be normative, as it is the case with the car crash scene. The incompatibility of the beaming face with the atmospheric emotion has ignited Web-based outrage all over China.

4.3 The illocutionary act, personal interest, psychological states and mediated emotion

There exists a distance in terms of personal interest between the speaker/hearer and the illocutionary act to be performed. The personal interest distance can be described in a scale: from the most beneficial to the neutral and to the most harmful. For instance, what the speaker talks about has nothing whatever to do with him/herself, nor with the hearer, e.g. about a scientific fact, a mathematic formula, etc. In this case, the speaker can be absolutely neutral, that is, if s/he chooses to do so. On the other hand, what is talked about may mean life or death to the speaker/hearer, as found in some court trial cases. In this case, the speaker's or hearer's personal interest is put at stake. In performing interest-neutral illocutionary acts, the speaker/actor and H can remain emotionally neutral from the acts. In performing interest-biased illocutionary acts, on the other hand, emotional neutrality is difficult to maintain.

Although they do not use the term "personal interest distance" associated with illocutionary acts, ancient Chinese philosophers notice the phenomenon when they emphasize *zhi yan* in order to *zhi ren* ("知言" 以 "知人" i.e. to know the person by listening to what he says). Confucius observes: *bu zhi yan*, *wu yi zhi ren* (不知言, 無以知人, unable to know the words, no way to know the

person 論語堯曰).[10] Zhou Yi (周易) analyzes the interconnections between speakers' psychological states, and what and how they say things.

> 將叛者其辭慚，中心疑者其辭枝，吉人之辭寡，躁人之辭多，誣善之人其辭遊，失其守者其辭屈。《周易繫辭下》
> (Those who are meditating a revolt speak in a shameful tone; those who are doubtful about the course diverge in topics; those who are kind-hearted talk little; those who are coarse talk a lot. Those who slander what is good talk emptily and unsubstantially; those who are losing what they ought to uphold speak in a crooked manner.)

This "reading people by their talks" (i.e. 知人) can be translated into our octet scheme as follows:

Performance	S		U	Psychological state	
can 慚	panzhe	叛者	qi ci 其辭	jiangpan	將叛
zhi 枝	zhongxin yizhe	中心疑者	--	xinyi	心疑
gua 寡	jiren	吉人	--	xianshan	賢善
duo 多	zaoren	躁人	--	xinzao	心躁
you 游	wushan zhi ren	誣善之人	--	wuxian	誣陷
qu 屈	shi qi shouzhe	失其守者	--	shishou	失守

The personal interest distance lies in between S and U above. The varied personal interest distances generate the corresponding psychological states. The psychological states, if hidden at the heart-mind, will remain undetected. Since the speakers have to utter words (i.e. *ci* 辭), what they say will betray their

[10] Mencius has this to say about 知言: "詖辭知其所蔽, 淫辭知其所陷, 邪辭知其所離, 遁辭知其所窮." (孟子公孫丑上) (When words are one-sided, I know how the mind of the speaker is clouded over. When words are extravagant, I know how the mind is fallen and sunk. When words are all-depraved, I know how the mind has departed from principles. When words are evasive, I know how the mind is at its wit's end. James Legge's translation)

psychological states. It is important to note here that words are public means of communication, and the psychological states betrayed by them are thus exposed to public for evaluation. The performance properties (e.g. *can* 慚，*zhi* 枝 etc.) are the results of the internal interaction between the psychological states and the potential social evaluations of them. Take the first case (*jian pan zhe* 將叛者) for example. The psychological state of intending to betray others is to the best personal interest of the speaker. But this personal interest cannot stand the public inspection, for the personal interest conflicts public interest. Hence what the speaker says and the manner s/he says it reveal the property of *can* 慚, i.e. shame and anxiety.

The shame or anxiety or shameful anxiety (i.e. mixed emotions) is in our analysis emotional phenomenon. It is cognition-mediated: The speaker (i.e. would-be traitor) believes that it is immoral to betray friends. Without this belief as mediation, the speaker/traitor would not experience shame or anxiety.

Here in order is a note about the intentional state and the emotional state in our octet. In Searle's analytic scheme, the intentional state does not differentiate beliefs that mediate emotions, because emotion is ignored in his conceptualization. In our octet scheme, the intentional state is intended to cover emotion-mediating beliefs, besides those beliefs that are conceptually necessary for the illocutionary act to be an act of certain type.

To further demonstrate the analysis let us look at a real-life example of *zhuhe* (祝賀 'congratulate'). It is taken from the Beijing Sports Games for the Retired.

(6) *Zhuhe*

我代表大會組委會//向獲得前八名的隊//獲得優勝的隊//向體育道德風尚獎的隊//表示//熱烈的祝賀//<聽眾鼓掌>

wo daibiao dahui zuweihui // xiang huode qianbaming
(I on behalf of conference committee to obtain first eight
de dui // huode yousheng de dui // xiang tiyu daode
(of delegation obtain best win of delegation to sports ethics
fengshang jiang de dui // biaoshi // relie de
fashion award of delegation express very warm of
zhuhe //
congratulation < ting zhong gu zhang > (audience applause)

In the octet framework, the speaker (S) performing *zhuhe* may do so as an individual person or as a persona representing a collectivity. The hearer must include the recipient or the persona of the recipient. *Zhuhe* can be performed using the performative or non-performative utterances. The essential content of *zhuhe* is information about why *zhuhe* is being made. The intentional and the emotional states of *zhuhe*, if felicitous, are all required to be positive (see further discussion below). The occasion for *zhuhe* is that it should be timely, at an appropriate venue, and agreeable to the social goal of the activity. The inter-dependency relation presupposes that the recipient has done something that is considered glorious, face-enhancing, reputable, etc. The conceptual analysis of *zhuhe* is graphically represented in Figure 4.

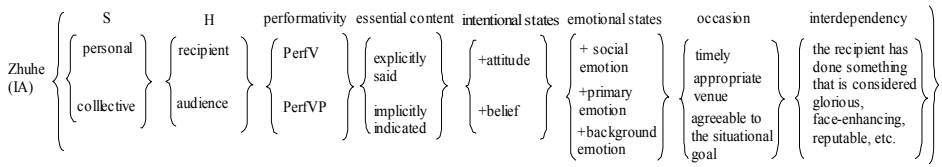


Figure 4: Conceptual analysis of *zhuhe*

Now let us zoom in to examine at length the intentional and emotional states. The intentional state includes belief and attitude.[11] The belief conceptually necessary for *zhuhe* to be such at all is the speaker's belief that the congratulatee has achieved something positive. The attitude, as relatively enduring belief (see §2.1 above), is that the speaker believes that it is socially desirable to say something appreciative about the congratulatee's achievement (i.e. *huode qianbaming*, *huode yousheng*). It is this attitude that mediates the speaker's emotions which include feeling happy (primary emotion), and feeling proud (social emotion). The emotions are signaled through the choice of words *relie* (熱烈), *zhuhe* (祝賀) and prosodies (see also §5 below).

Note also that in (6), the speaker makes his persona role explicit (*wo daibiao dahui zuweihui*). The hearer (H) is explicitly divided into two classes: the recipient i.e. those teams that are honored, and the audience. The act is performed using an explicit performative formula. The essential content is provided by the phrases *huode qianbaming*, *huode yousheng*, and *tiyu daode fengshang jiang* respectively. The speaker's act of *zhuhe* is performed as part of the activities of the social situation. Its satisfactions with the occasion and interdependency conditions are made transparent and taken for granted by the participants.

4.4 Lexical and syntactic representations of emotions in reported illocutionary acts

The emotions accompanying illocutionary acts may be made explicit in a lexical and syntactical way, as shown in (7).

[11] Recall that in the octet the order of presentation is irrelevant. "Attitude and belief" is the same as "belief and attitude".

(7) 我們<得恭恭敬敬地>請人家。

women	<dei	gonggong	jingjing di >	qing	renjia
(we	must	very	respectfully	invite	person

我們領導求賢<若渴>。

women	lingdao	qiu	xian	<ruo	ke >
(we	leader	beg	the virtuous	like	thirst)

Dei gonggongjingjing di and *ruoke* are emotional state indicators that accompany the felicitous performance of the illocutionary acts of *qing renjia* and *qiu xian* respectively.

There are of course many other ways of indicating emotions. Compare: *ya huan kai men* (丫鬟開門, the maid opened the door) with *ya huan <zhengzhe> kai men* (丫鬟<爭著>開門, the maids <all wanted to> open the door); *qishen jielai* (起身接來, stood up to receive) with *<mang> qishen <ying> shanglai* (<忙>起身<迎>上來 'hurriedly stood up to welcome'). The use of address terms is commonly used to indicate emotions: *baobao* (寶寶 'darling darling') , *guaiguai* (乖乖 'little dear') are regularly found in mothers addressing their babies.

5. Prosody, intonation-group and performance unit

5.1 Prosody in Austin's original conceptualization

As we know, the native environment of speech acts is live speech, which is materialized in articulated sound streams. When Austin conceptualizes his speech act theory, it is the speech sounds not orthographic texts that he has in mind. We may recall that Austin abstracts a series of acts from a total speech act (see Austin 1980[1962]:92-100), which are rearranged as follows:

The total speech act
the perlocutionary act
the illocutionary act
the locutionary act
the rhetic act (rheme)
the phatic act (pheme)
the phonetic act (phone)

Reading from bottom upwards, the phonetic act is an act of "uttering certain noises". But this is not enough for saying something in the full sense of "say". The speaker has to do simultaneously a phatic act, that is, the noises uttered must belong to and as belong to a certain vocabulary, … "with a certain intonation" (Austin 1980[1962]:92), etc. The illocutionary act, the major concern of this paper, if performed successfully, presupposes the successful performance of all the acts under it.

Linguistically, speech sounds are studied in separate branches, e.g. articulatory phonetics, perceptual phonetics, acoustic phonetics, segmental phonology, intonation or prosody, etc. These branches roughly cover Austin's phonetic and phatic acts. It is interesting to note that Austin does not associate intonation with the illocutionary act, but with the phatic act. This does not mean that intonation is regarded as being irrelevant to the illocutionary act. On the contrary, Austin discusses "spoken language" features such as "tone of voice", "cadence", "emphasis" thus (see Austin 1980[1962]:74):

It's going to charge ! (a warning);
It's going to charge ? (a question);
It's going to charge !? (a protest);

One and the same “It’s going to charge” can be uttered to perform three different illocutionary acts, thanks to the differences in tone of voice, cadence, and emphasis. He points out that the marks in “written language” such as !, ?, !? cannot adequately capture the spoken language features.

Even with a mono-syllable performance, there is an intonation involved in the production, and the intonation may play the decisive role in indicating what illocutionary force is intended by the speaker. Now compare the following:

(8) *zou* (走‘walk’)

Zou (走, mid key, fall-rise, normal tempo): has the force of *zhishi* (指使, instruct)

Zou (走, high key, fall-rise, fast tempo): has the force of *cuicu* (催促, hurry up)

Zou (走, mid key, breathy, lengthening, fall-rise): has the force of *yaoqing* (邀請‘invite’)

Note that in view of the illocutionary act performance, the mono-syllable *zou* is complete and fully functional.

Works on intonation and prosody (e.g. Crystal 1969, Bolinger 1986, Cruttenden 2002, Ladd 2008) deal with intonational or prosodic issues of general nature. These works discuss in passing intonations associated with questions, directives, commands, etc., which are arguably classes of illocutionary acts. Focused studies of illocution and prosody are rather rare. As far as I know, Wennerstrom (2001) is the only author who dedicates a chapter to it.

5.2 Prosody: working definition

In the literature, the terms intonation and prosody are given different

interpretations. In this paper we adopt Crystal's notion of "prosodic systems" (1969:5-6), including

> "the psychological attributes of sound … as pitch, loudness and duration, … 'Intonation', … is viewed as the product of a conflation of different prosodic systems of pitch contrasts; 'stress' is referable to variations in the loudness parameter. Other prosodic systems comprise independently varying vocal effects based on combinations of these three parameters in specific ways (for example, thythmicality), or on contrasts in silence (the system of silent pause)."[12]

Cook (2002:101) adopts a similar view about prosody: "Prosody itself is normally divided into three distinct parts: (i) the rhythm or tempo of speech, (ii) changes in the intensity or loudness of the voice, and (iii) changes in pitch (i.e. intonation)."

It is important to bear in mind that prosodic systems are language-specific, and the functions the systems play are also language-specific. It remains an open question whether there exist some prosodic features that are universal across all languages. For example, rising pitch at the end of an utterance seems to be generally interpreted as "forward-looking". When angry, for another example, the key tends to be higher than normal. So this paper makes a default distinction between *prosodic systems as potential*, and *prosodic systems as choices of behavior*. In the former case, we can build up a pool of potential prosodic resources gathering information from any natural language. In the latter case, we focus on choices of prosodic systems that are found in real use (i.e. in our case,

[12] The notion of prosody adopted here is definitely different from the Chinese concept of *yinyun xue* (音韻學). It is close to *jielü xue* (節律學) as found in 吳潔敏, 朱宏達 (2001) 漢語節律學. 北京: 語文出版社.

found in the SCCSD).

5.3 Performance unit and intonation-group

As we know, an illocutionary act can be performed by uttering a mono-syllable such as *zou* (走), *hao* (好, good), *chi* (吃, eat), or as long as multiple utterances in a turn or a sequence involving turn-taking exchanges. For ease of reference, we propose to use "performance unit" as an abstract umbrella term to cover all the cases discussed above. In other words, for a speaker to perform any illocutionary act, it takes ***just one performance unit*** to implement the act.[13] The length of the performance unit varies from one act to another, and from one instance to another.

Since prosody is the major concern of this paper, we shall use "intonation-group" as the basic descriptive unit to analyze the performance unit. So *zou*'s in (8) above are all mono-syllabic intonation-groups. This represents the minimal size of a performance unit or an intonation-group. However, the performance unit of *jieshao* (介紹), in contrast, can be quite long, such as (9) taken from the Opening Ceremony of Sports Games for the Retired:

(9) *jieshao*

呃現在呀//呃向他們<口誤: 你們>介紹一下子//今天出席我們大會的領導同志//啊//今天有些領導同志啊//都有活動//啊//沒有來得了//啊//現在今天到會的呀//有我們大會組委會副主任//啊//東城區副區長//朱 x x 同志// ……

e	xianzai	ya //	e	xiang	tamen	<kouwu:	nimen >
(eh	now	Particle	eh	to	them	<error	you>

[13] A performance unit is an act-token.

jieshao yixiazi // jintian chuxi women dahui
introduce once today attend our conference
de lingdao tongzhi // a // jintian youxie lingdao
of leaders comrades Particle today some leaders
tongzhi a // dou you huodong // a //
comrades Particle all have activity Particle
meiyou lai deliao // a // xianzai jintian
didn't come ASP Particle now today
daohui de ya // you women dahui zuweihui
have come ASP Particle have our conference committee
fuzhuren // a // dongchengqu fuquzhang // Zhu xx tongzhi //
key members Particle east district deputy leader Zhu xx comrade

Only the introduction of the first VIP was included above. The act of *jieshao* does not finish until the last VIP has been introduced. Theoretically, the linear structure of the performance unit of *jieshao* flows in the way as shown in Figure 5.

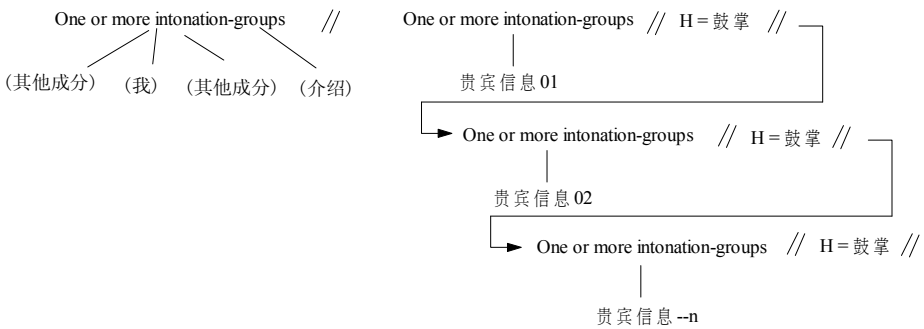


Figure 5: Theoretical structure of the performance unit of *jieshao*

The boundary of the intonation-group is crystal clear when the performance of the illocutionary act is performed at a single breath within a single turn of speaking. When the performance of an illocutionary act takes multiple utterances involving multiple turn-takings, as mentioned above, the boundaries between multiple intonation-groups become problematic.

> Judgement that an intonation-group boundary is present would in an ideal situation be based on 'external criteria', i.e. on phonetic cues present at the actual boundary. But in practice such phonetic cues (e.g. pause) may be either ambiguous or not present at all. Therefore 'internal criteria', must also play a part, …The assignment of intonation-group boundaries is therefore something of a circular business; we establish some intonation-groups in cases where all the external criteria conspire to make the assignment of a boundary relatively certain; we note the sorts of internal intonational structure occurring in such cases and this enables us to make decisions in those cases where the external criteria are less unambiguous. And in some difficult cases, we take grammatical or semantic criteria into account, i.e. when regular correspondences between intonation and grammar/semantics have been established in cases where boundary assignment is clear, we may lean heavily on such correspondences when assigning boundaries in the difficult cases. (Cruttenden 2002:29-30)

The present study does indeed "lean heavily" on pragmatic correspondences when assigning boundaries between intonation-groups. As suggested above, in this study, the illocutionary act/force is taken for granted. That is, the investigation starts with the assumption that there is no problem with identifying what the

illocutionary act/force is, although in actuality it is not that rosy at all. In other words, the performance of the illocutionary act/force sets up the most outer boundary (i.e. equal to the boundary of the performance unit of the act) for the intonation-group(s).

5.4 Prosody: research methodology

As well documented in the literature, as far as research methodology of prosody is concerned, there exists

> a duality of approaches to the study of prosodic features, in particular to the study of pitch patterns. The most common labels attached to the two approaches are the AUDITORY and the INSTRUMENTAL. (Cruttenden 2002:5)

As Cruttenden points out, the proponents of the two approaches criticize each other for failures, or incompetence, etc. Readers interested in knowing more about it are cordially referred to Cruttenden (2002:5-6), also to Ladd, Scherer & Silverman (1986), and Ladd (2008:18-31). What is to be said here is that we follow Cruttenden by trying "to use the best of both approaches" (Cruttenden 2002:6). In the present case, we are clearly on the auditory side, and the acoustic as well as video image analyses using Praat and Elan are moves towards the instrumental side.

A word about the Praat-based acoustic analysis is in order here. Our data, i.e. sound streams, are very "dirty" in comparison with the sound streams recorded in the sound-proof lab. Our sound streams are recorded in the naturally occurring real-life social situations. The information extracted using Praat is "dirty" too. As a measure to amend the situation, cleaner sound streams are also

used, including the following: (a) audio books of story-telling recorded in the sound-proof studios (i.e. *pingshu*, 評書); (b) TV series of *pinzui Zhang Da Ming de xingfu shenghuo* (貧嘴張大明的幸福生活 'The Talkative Zhang Da Ming's Happy Life', a 20-set series) and *bianjibu de gushi* (編輯部的故事 'Stories from An Editorial Office', a 25-set series). The cleaner data are used to construct idealized exemplars, which serve as contrastive models against which cases found in the dirty data are examined. Such differentiated treatment of the data is obviously justifiable as a research strategy.

Prosodic transcription is notoriously difficult and often varied from one researcher to another. This paper will rely heavily on the graphics representations drawn by Praat.

6. Emotion and prosody in illocution

6.1 Prosody conveys emotion

Emotional behavior is represented in many different ways, two of which this paper is concerned with are prosody and facial expression. In this section we discuss the first, postponing the second to the next section.

As mentioned in §3 above, the correlation of the emotional states of the heart-mind with certain prosodic properties was described as early as 禮記. In modern literature, the correlation is noted by researchers of varied academic backgrounds. Bolinger, a linguist, for example, observes (1986:viii) that intonation responds to "feeling more directly than does any other well-systematized part of spoken language." Damasio, a neuroscientist, writes (1999:92):

> Words and sentences, from the simple "Yes," "No," and "Hello" to "Good Morning" or "Good-bye," are usually uttered with a background

> emotional inflection. The inflection is an instance of prosody, the musical, tonal accompaniment to the speech sounds that constitute the words. Prosody can express not just background emotions, but specific emotions as well.

Linguists as well as laymen will readily endorse this neuroscientist's observation. The crux of the matter is not about the status of correlation, but about in what ways the two are correlated, i.e. the correspondences of specific emotions with particular prosodic patterns.

6.2 Two fundamental assumptions

The STF-Match Assumption

The assumption goes like this: In child discourse there exists a perfect *match* between what is ***s***aid, what is ***t***hought of, and what is ***f***elt (the STF-Match Assumption for short). For example, when a child utters: *wo yao laolao* (我要姥姥, i.e. I want Granny) the child truly wants the presence of *laolao*, truly thinks of *laolao*, and truly feels the miss of *laolao* (see further analysis below).

The Prosodic Genuineness Assumption

Closely associated with the STF-Match Assumption is the prosodic genuineness assumption (the PGA for short) about child discourse: If there are prosodic cues produced, the prosodic cues are genuine, that is, the cues truly reflect what is said, what is thought of, and what is felt.

Let us demonstrate how the two assumptions operate in real-life discourse. The instance is taken from the sub-corpus of child discourse in the SCCSD. The child was about 4 years. Upon being told that Granny, who had looked after him

for 3 years, had gone to the United States, he was seized with a fit of crying. While crying, he repeatedly uttered: *wo yao laolao*.

Wo yao … is one of the illocutionary acts children acquire at their early years of life, and even before the appearance of audible speech, they can perform it nonverbally just as effectively. There are three features that need to be highlighted here. (1) The speaker/actor believes (i.e. the intentional state) that what is wanted is available, and cherishes a positive attitude to the object desired. (2) The primary and social emotions, if involved, are positive, not negative.[14] The primary emotion associated with *wo yao* is positive, such as feeling happy, pleased, and so on. (3) The interdependency is that what is wanted used to be desirable and available. Figure 6 shows the conceptual analysis of *yao* in the octet format.

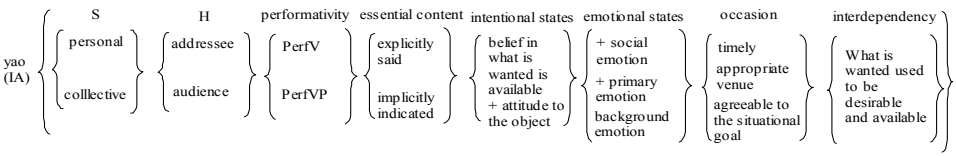


Figure 6: Conceptual analysis of *yao*

The current case is an analysis of a frustrated request. The child speaker/actor is aware that the request cannot be met, but persists in making it. *Wo yao laolao*, a performative utterance, verbalizes *S*, *performativity*, and the *essential content* respectively. What is interesting in this case is that the speaker's intentional states, i.e. his belief, and attitudinal attachment, his primary emotion, and the interdependency have turned out to go amiss, and blocked by the changed circumstances. Should such things happen to an adult, s/he would give up the

[14] Note that in children, social emotions (e.g. feel shameful, arrogant, contemptuous, etc.) are acquired gradually.

request, and opt for alternatives.[15] Since the speaker/actor is a child, he pursues the act all the more vehemently in spite of the fact that he was aware that his want was unlikely to be satisfied. He felt frustration, and was seized with crying.

Prosodically *wo yao laolao* is a four-syllable intonation-group. What is somewhat unique is that it sits on bursts of crying. Anyone listening to the original sound stream will be convinced immediately that the child speaker not only genuinely, but extremely strongly *yao laolao*. As pointed out in the conceptual analysis above, the primary emotion associated with *yao* is the positive *xi*. Thanks to the frustration, the positive primary emotion is transformed into a negative primary emotion, 悲.[16] The background emotion underlying the utterance is *diqi zu* (底氣足, i.e. deep-seated air full). The child speaker/actor has exercised all his energy, and sweated a lot. Figure 7 shows the ELan-Praat-based analysis.

[15] In the adult case, because the adult is more re-adjustable, the original intentional, and emotional states are subsequently abandoned.

[16] *Xi* (喜) and *bei* (悲) are opposite classes of emotion. The desired *xi*, if denied, is transformed into the opposite *bei*.

Figure 7: Prosodic analysis of *yao* by a child speaker

The STF-Match Assumption is part of the felicity conditions of an illocutionary act performance. It corresponds to 3 elements in our octet analytic framework of illocution, viz., *the essential content*, *the intentional states* and *the emotional states* respectively.

Experiences show that the STF-Match Assumption breaks down quite often in adult discourse (most noticeably in politicians' discourse!), resulting in infelicitous performances of illocutionary acts. Experiences also show that the PGA also breaks down quite often in adult discourse, resulting in ultra-default interpretations. All the break-downs of the STF-Match Assumption, and the PGA in adult discourse are generally motivated, except for those found in pathological discourse.

6.3 background emotions and personal keys

Now as mentioned above, we shall use "keys" as index to background emotions. In everyday parlance, when the speaker is in good health, s/he is described as *diqi zu* or *shengyin hongliang* (聲音洪亮, i.e. voice high-pitched). The notion of key, borrowed from Brazil (1997), is used to capture this everyday notion of *diqi zu* or *shengyin hongliang*. Doctors practicing Chinese medicine adopt this prosodic feature as their diagnostic aids.[17] Poor health or emotional states affect the speaker's key. The speaker's keys can be arranged in a scale: normal key (associated with the speaker's states of being healthy and calm); high key (e.g. being agitated, enthusiastic, energetic); low key (e.g. poor health, discouraged).

6.4 Primary/social emotions and prosody

Primary and social emotions will be examined, as mentioned above, in their associations with the performances of illocutionary acts. As pointed out above, in this paper, illocution is taken for granted. Our research is concerned with two complementary questions: (1) from the speaker perspective, given an illocutionary act, what is the function emotion plays in its performance? and (2) from the hearer perspective, given an illocutionary act performed, what is the emotion of which the act functions as an antecedent? The issue of prosody arises in connection with both questions, since prosody is somehow intrinsically associated with emotional representations in real-life discourse.

[17] Chinese doctors practicing traditional Chinese medicine in their initial phase of diagnosis use four diagnostic methods, *wang wen wen qie* (望聞問切 'look smell ask touch'). *Wen* (問) includes smelling and listening to the patient.

The two questions both are taken as empirical ones. That is, to investigate the tripartite interaction between illocution, emotion and prosody, all we assume is that the interaction is real. What we do not know is the details, which can be discovered through systematic investigations. The discovery procedure is shown in Figure 8.

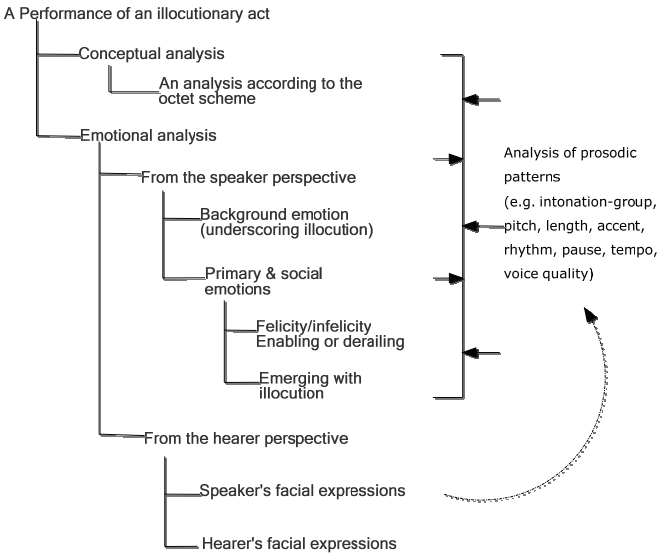


Figure 8: A discovery procedure for the tripartite interaction

Given a performance of an illocutionary act, first it is put to a conceptual analysis, that is, an analysis according to the octet scheme. Then it is an emotional analysis from the speaker and hearer perspectives. The emotional analysis from the speaker perspective, as demonstrated above, includes the analysis of background emotions that underscore human conscious action in general, and the analysis of primary and social emotions. This latter analysis involves a decision about whether emotion merges with the illocutionary act or affects the felicity

conditions of the act. Prosodic analysis is to look for prosodic patterns that correlate with the illocutionary act/force and emotions.

The three analyses, viz., conceptual, emotional and prosodic, can be carried out sequentially or in parallel. In actual practice, there are many back and forth mixed steps.

The emotional analysis from the hearer perspective will only be touched upon very briefly in this paper. The data will be confined to facial expression only. As indicated in Figure 8, facial expression analysis of emotions feeds well into prosodic analysis.

It is important to bear in mind that, for illocutionary acts being looked at as single instances, the analyses are descriptive in nature. The prosodic patterns we look for only emerge when we compare and contrast instances. For this purpose, we are compiling a prosodic corpus of illocutionary force, which consists of prosodic instances organized in terms of illocutionary act categories.

The discovery procedure can be effectively facilitated by using Praat and Elan. Praat is specially tailored for sound-based analysis. The performance instance (i.e. a sound stream as a performance unit) is anatomized into a stack of abstract tiers, as shown in Figure 9.

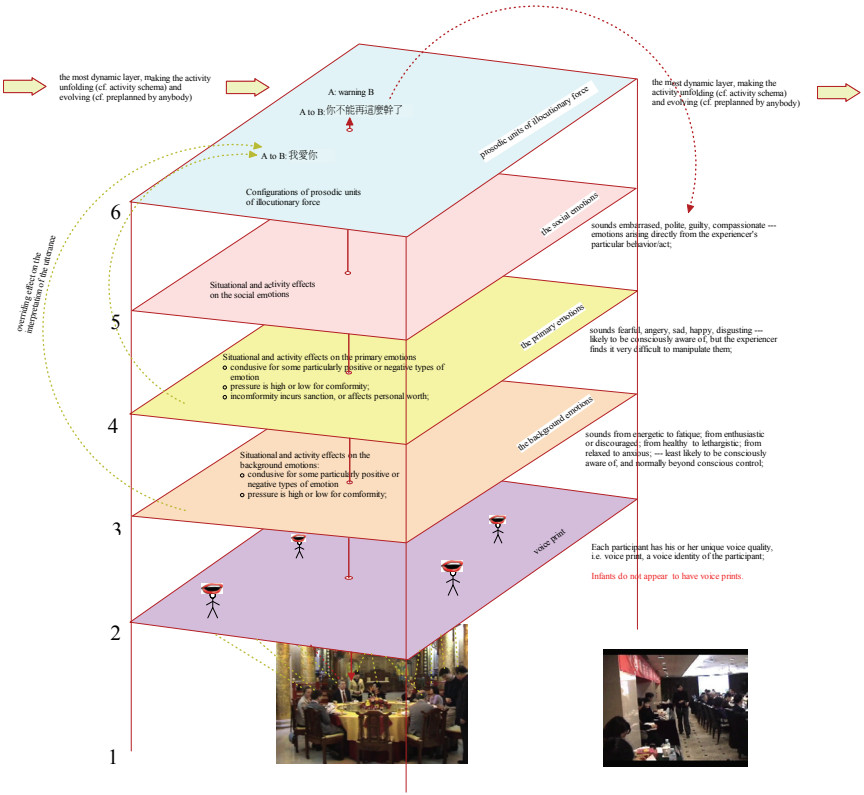


Figure 9: A stack of abstract analytic tiers

Reading from bottom upwards, the first is the "social situation" tier, which is treated as a background tier.[18]

The second tier is about voiceprint. The tier name "voiceprint" is a bit tongue in cheek. It is actually used here as another way of indicating turn-taking. It is preferred here simply to remind us that from the prosodic point of view, speakers' voices can be uniquely identified by the human ear except for infants whose voiceprint is not distinctively mature yet.

[18] It is a "background" only in the light of the present study. Elsewhere (see Gu 2006, 2009, 2012) it is the primary object of investigation itself. The data about this tier is drawn from the video image streams.

The third to fifth tiers correspond to three tiers of emotions talked about in §3 above. The sixth and top tier is about the illocutionary act/force. The central line with duel arrowheads cutting across all the tiers is used to remind us of the fact that the tiers are all inter-dependent and interactive in real life situations.

The data obtained through Praat-enabled analysis then is imported to Elan for video-based analysis (readers interested in video-based analysis are advised to Gu 2006, 2009). As pointed out above, video-based analysis in this paper is only concerned with the correlation between emotion and facial expression, to which we turn.

7. Facial expressions and the triadic interactions of illocution, emotion and prosody

"The study of facial expression has been central to the field of emotion since its inception, and is the continued focus of theoretical controversy and ongoing research" (Keltner & Ekman 2000:236). Research on the recognition of facial expressions of emotion has been long in sociopsychology and personality psychology. Ekman & Rosenberg (2005) assemble an anthology of some best works concerning facial expression measurements. In this paper, facial expressions of emotion are treated as supplementary information in our examination of the role emotion plays in the performance of illocutionary acts, and emotional effects on prosody.

7.1 Illocutionary acts/forces as antecedents to emotions exhibited in facial expressions

As discussed in §3 above, occurrent emotions are event-related. To hearers, the speaker's performance of the illocutionary act is an external event, of which

the hearer's perception and interpretation generates an internal cognitive event that may trigger emotional reactions of the hearer. This chain of interactions is demonstrated in our analysis of emotional responses found in face-to-face classroom discourse. Figure 10 is a screen shot of Elan-based analysis of students' emotional responses to the teacher's illocutionary acts. It will take too much space to go to any details here. Suffice it to give a summative account about our initial investigation. One and the same illocutionary act performed by the teacher is correlated with a range of facial expressions exhibited by students. Figure 11 summarizes such correlation in terms of 1-minute interval, and 7-minute interval respectively. (Note that only four out of 120 students are shown here in order to save space).

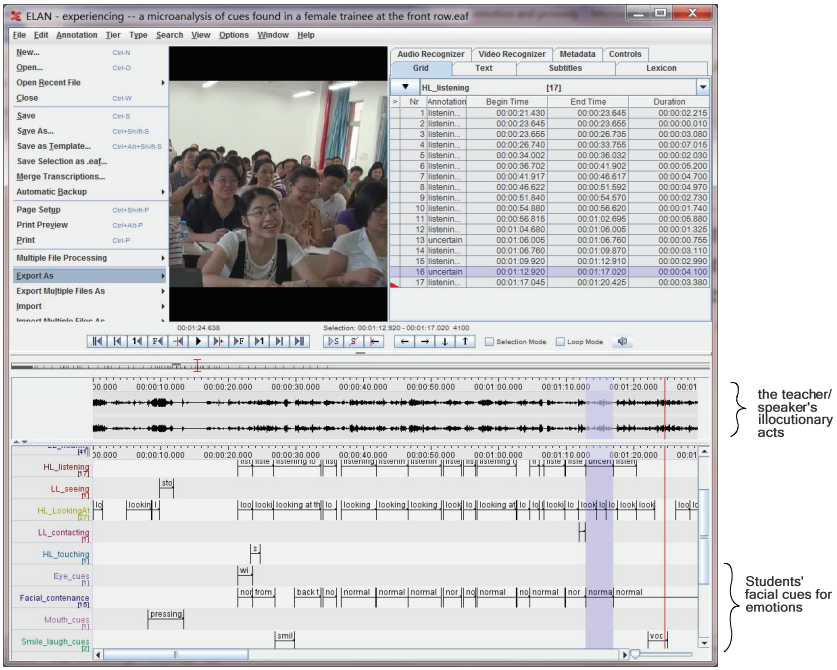


Figure 10: Elan-based analysis of the hearer's emotional responses

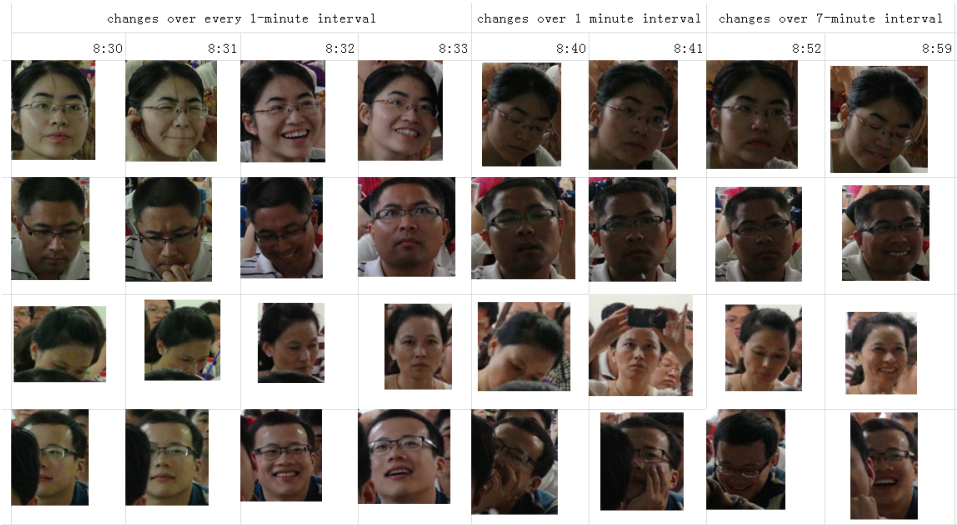


Figure 11: Students' facial expressions in response to the teacher's illocutionary acts

(Note: At 8:31, the loud speaker let off a piercing noise, which triggered the response of painful facial expressions by some students, e.g. the second facial expression at the first row.)

In our another study of facial expressions of emotion in classroom, there is evidence that indicates that some students consciously control their facial expressions associated with emotions. As we know, normally, emotions that are likely to be controlled consciously by adults tend to be those primary and social emotions that are negative. However, there are students who also consciously control the exhibition of their positive emotions. We interpret such control as evidence of the interaction between the personal pattern of dispositional emotion (associated with personality) and the exhibition of occurrent emotions.

7.2 Facial expression as illocutionary force indicating device

Facial expression as illocutionary force indicating device (IFID) is understood in this paper as an indirect IFID. That is, facial expression is correlated with emotion, and it is emotion that affects the perception and interpretation of what the illocutionary act/force a given performance unit conveys. One of the advantages of facial expression as IFID, though indirect in nature, is that it is visual, and difficult to manipulate. It is extremely helpful in case of implicit performance. In performative utterances, it may override what is explicitly declared to be. Take *mama shengqi le* (媽媽生氣了 'mummy is angry') for example. Toddlers do not take this performative utterance seriously until they check it out against Mummy's facial expression, apart from the prosodic cues such as loudness, pitch level and tempo.

8. Some theoretical issues and future research

The paper's vowed objective is to present a conceptual model for investigating, systematically and empirically, our everyday pragmatic practice of *cha yan guan se* and *ting hua ting yin* in terms of the tripartite interaction between illocution, emotion and prosody. An illocutionary act is defined as a performance token-act-unit the speaker performs in exemplifying the property of an illocutionary act-type, which is considered to be social in nature. Based on this definition, the analysis of the property exemplified in the performance token-act-unit constitutes the essence of the whole investigation. The fact that it counts as the whole investigation is due to the fact that emotion, as argued above, is an intrinsic element of illocution. We hold a very strong position that there is no performance of illocutionary acts without involving emotions. Prosody comes in also as an intrinsic element of performance, simply because live speech is the

native natural environment of illocutionary acts. Illocutionary acts found in literature, newspapers, and so on are reported illocutionary acts. They are not "real" to us.

The act-type property exemplified in performance is conceptually analyzed in terms of an octet scheme. The octet scheme in turn is semi-formalized by adopting the formula from set theory. This will facilitate the construction of illocutionary act ontology coded in the OWL.

The paper suffers from two drawbacks. One is its failure to map out some general prosodic patterns of illocutionary classes. The paper initially made some attempts at it, but gave up. The major obstacle to this objective lies in the poor quality of the sound streams. As pointed out in §5 above, the data we deal with are very "dirty", and given that there are some general prosodic patterns emerging from our analysis, the validity of the patterns will be doubtful. In order to overcome this drawback, as mentioned above, a project is already in progress, namely to apply the analytic framework of this paper to much cleaner data so that a contrastive study can be made between the dirty and the clean. Prosodic patterns of illocutionary force based on the two sets of results, hopefully, will be reliable.

The second drawback is the failure to spell out the details of segmentation and annotation in the Praat-Elan-enabled environment. The omission is mainly due to two considerations. First, literature on Praat/Elan-based analysis of various linguistic phenomena is already available. Second, the author of this paper has also published two papers (already cited above) concerning the segmentation and annotation of video data streams. So there is not much to be gained when literature is available for consultation elsewhere.

Finally, the conceptual modeling, as adopted in this paper, is not constrained, as found in software engineering, by the worries about whether the conceptual modeling can be implemented or not. In corpus linguistics, and human-machine

interaction (HMI) research, there has been increasing interest in annotating spoken data in terms of speech acts (often called dialogue acts), as well as prosodic annotation. The spoken data fall into two general categories: (1) spontaneous spoken discourse, and (2) spoken discourse recorded in a laboratory. The first itself includes two sub-categories, (a) generic, i.e. spoken corpus compiled for the generic purpose of understanding natural language in general, (b) restricted, i.e. spoken discourse samples taken from restricted domains, e.g. telephone inquiries, travel agency, taxi service, and so forth. The lab-produced spoken discourse also belongs to the restricted sub-category.

It is important to bear in mind the differences in research objectives in annotating spoken discourse, spontaneous or restricted alike. The generic purpose of understanding how natural language works is human intelligence-oriented, whereas the purpose of developing a restricted domain is often aimed at machine intelligence. The two are complementary in perspective, but quite different in conceptualization and workflow. The present study is in the first group. When we look at the data, we are not worried about whether it is technologically feasible, given our research goals. We are interested in what the state-of-the-art technology can be mobilized to provide assistance. In other words, we are not constrained by technology, but ready to put up challenges for technology to grow. This is the motivation of focusing a conceptual model. Its validity is measured in terms of whether it truly reflects what happens in the real world rather than whether it can be implemented by machines. This also explains why the literature on computing dialogue acts, emotions and prosodies found in AI, and HMC, is ignored in this paper.

References

Austin, John Langshaw. 1980[1962]. *How to do Things with Words*. Oxford: Oxford University Press.

Bolinger, Dwight. 1986. *Intonation and its Parts: Melody in Spoken English*. Stanford: Stanford University Press.

Borgida, Alexander T., Vinay K. Chaudhri, Paolo Giorgini, and Eric S. Yu. (eds.) 2009. *Conceptual Modeling: Foundations and Applications*. Berlin: Springer.

Brazil, David. 1997. *The Communicative Value of Intonation in English*. Cambridge & New York: Cambridge University Press.

Budd, Malcolm. 1985. *Music and the Emotions*. London: Routledge & Kegan Paul.

Clark, Herbert H., and Thomas B. Carlson. 1982. Hearers and speech acts. *Language* 58.2:332-373.

Cook, Norman D. 2002. *Tone of Voice and Mind: The Connections between Intonation, Emotion, Cognition and Consciousness*. Amsterdam & Philadelphia: John Benjamins.

Cruttenden, Alan. 2002. *Intonation* (2nd edition). Beijing: Peking University Press; New York: Cambridge University Press.

Crystal, David. 1969. *Prosodic Systems and Intonation in English*. New York: Cambridge University Press.

Dalgleish, Tim, and Mick J. Power. (eds.) 1999. *Handbook of Cognition and Emotion*. Chichester: John Wiley & Sons.

Damasio, Antonio. 1996. *Descartes' Error: Emotion, Reason and the Human Brain*. London: Papermac.

Damasio, Antonio. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. London: William Heinemann.

Damasio, Antonio. 2003. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. London: William Heinemann.

Damasio, Antonio, and Kaspar Meyer. 2009. Consciousness: an overview of the phenomenon and of its possible neural basis. *The Neurology of Consciousness:*

Cognitive Neuroscience and Neuropathology, ed. by Steven Laureys & Giulio Tononi, 3-14. Amsterdam: Academic Press Elsevier.

Darwin, Charles. 1899. *The Expression of Emotion in Man and Animals*. New York: D. Appleton and Company.

Davis, Steven. 1984. Speech acts and action theory. *Journal of Pragmatics* 8.4:469-487

Ekman, Paul. 1994. Moods, emotions, and traits. *The Nature of Emotion: Fundamental Questions*, ed. by Paul Ekman & Richard J. Davidson, 56-58. New York: Oxford University Press.

Ekman, Paul, and Richard J. Davidson. (eds.) 1994. *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press.

Ekman, Paul, and Erika L. Rosenberg. (eds.) 2005. *What the Face Reveals* (2nd edition). Oxford: Oxford University Press.

Elster, Jon. 1999. *Alchemies of the Mind: Rationality and the Emotions*. Cambridge: Cambridge University Press.

Frijda, Nico H. 1994a. Varieties of affect: emotions and episodes, moods, and sentiments. *The Nature of Emotion: Fundamental Questions*, ed. by Paul Ekman & Richard J. Davidson, 59-67. New York: Oxford University Press.

Frijda, Nico H. 1994b. Emotions are functional, most of the time. *The Nature of Emotion: Fundamental Questions*, ed. by Paul Ekman & Richard J. Davidson, 112-122. New York: Oxford University Press.

Gaines, Philip. 2001. Intonation and speech act theory. *The Music of Everyday Speech: Prosody and Discourse Analysis*, ed. by Ann K. Wennerstrom, 131-165. Oxford & New York: Oxford University Press.

Goldman, Alvin I. 1970. *A Theory of Human Action*. Englewood Cliffs: Prentice-Hall.

Grice, Herbert Paul. 1989. *Studies in the Way of Words*. Cambridge: Harvard University Press.

Gu, Yueguo. 2002a. Towards an understanding of workplace discourse. *Research and Practice in Professional Discourse*, ed. by Christopher N. Candlin, 137-185. Hong Kong: City University of Hong Kong Press.

Gu, Yueguo. 2002b. Beijing diqu xianchang jixi huayu yuliaoku de quyang yu

daibiaoxing wenti [Issues concerning sampling and representativeness in compiling the situated discourse in Beijing area]. *Quanqiuhua yu Ershiyi Shiji* [*Globalization and the 21st Century*], 484-500. Beijing: China Social Sciences Press.

Gu, Yueguo. 2006. Multimodal text analysis: a corpus linguistic approach to situated discourse. *Text and Talk* 26.2:127-167.

Gu, Yueguo. 2009. From the real-life situation to video stream data-mining. *International Journal of Corpus Linguistics* 14.4:433-466.

Gussenhoven, Carlos, and Tomas Riad. (eds.) 2007. *Tones and Tunes*, Vol. 2: *Experimental Studies in Word and Sentence Prosody*. Berlin & New York: Mouton de Gruyter.

Hatfield, Elaine, John T. Cacioppo, and Richard L. Rapson. 1994. *Emotional Contagion*. Cambridge & New York: Cambridge University Press.

Johns-Lewis, Catherine. (ed.) 1986. *Intonation in Discourse*. London: Croom Helm.

Johnstone, Tom, and Klaus R. Scherer. 2000. Vocal communication of emotion. *Handbook of Emotions* (2nd edition), ed. by Michael Lewis & Jeannette M. Haviland-Jones, 220-235. New York: The Guilford Press.

Kasher, Asa. (ed.) 1998. *Pragmatics: Critical Concepts*, Vol. 2: *Speech Act Theory and Particular Speech Acts*. London: Routledge.

Keeley, Maureen P. 2005. The nonverbal perception scale. *The Sourcebook of Nonverbal Measures*, ed. by Valerie Manusov, 93-104. Mahwah: Lawrence Erlbaum Associates.

Keltner, Dacher, and Paul Ekman. 2000. Facial expression of emotion. *Handbook of Emotions* (2nd edition), ed. by Michael Lewis & Jeannette M. Haviland-Jones, 236-251. New York: The Guilford Press.

Ladd, D. Robert. 2008. *Intonational Phonology* (2nd edition). Cambridge & New York: Cambridge University Press.

Ladd, D. Robert, K. R. Scherer, and Kim Silverman. 1986. An integrated approach to studying intonation and attitude. *Intonation in Discourse*, ed. by Catherine Johns-Lewis, 125-138. London: Croom Helm.

Leech, Geoffrey N. 1983. *The Principles of Pragmatics*. London: Longman.

Levinson, Stephen C. 2000. *Presumptive Meaning*. Cambridge: MIT Press.

Lewis, Michael, and Jeannette M. Haviland-Jones. (eds.) 2000. *Handbook of Emotions*

(2nd edition). New York: The Guilford Press.

Li Ji Jin Zhu Jin Yi [*Today's Translation and Annotations of the Book of Rites*] (Part 1 & Part 2), by Wang Mengou. Taipei: The Commercial Press, 1978.

Maiese, Michelle. 2011. *Embodiment*, *Emotion*, *and Cognition*. London: Palgrave.

Mencius. Beijing: Zhonghua Book Company.

Meng, Peiyuan. 1998. *Xinling Chaoyue yu Jingjie* [*Spiritual Transcendence and Cultivation*]. Beijing: People's Publishing House.

Meng, Peiyuan. 2002. *Qinggan yu Lixing* [*Emotion and Rationality*]. Beijing: China Social Sciences Press.

Pananicolaou, Andrew C. 1989. *Emotion: A Reconsideration of the Somatic Theory*. New York: Gordon and Breach Science Publishers.

Parsons, Jeffrey, Motoshi Saeki, Peretz Shoval, Carson Woo, and Yair Wand. (eds.) 2010. *Conceptual Modeling: ER 2010*. Berlin: Springer.

Recanati, Francois. 2004. *Literal Meaning*. Cambridge: Cambridge University Press.

Rokeach, Milton. 1970[1968]. *Beliefs*, *Attitudes and Values*. New York: Jossey-Bass.

Santangelo, Paolo. 2003. *Sentimental Education in Chinese History: An Interdisciplinary Textual Research on Ming and Qing Sources*. Leiden: Koninklijke Brill NV.

Searle, John R. 1969. *Speech Acts*. Cambridge: Cambridge University Press.

Searle, John R. 1979. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.

The Book of Changes. Beijing: Zhonghua Book Company.

Wennerstrom, Ann K. 2001. *The Music of Everyday Speech: Prosody and Discourse Analysis*. Oxford: Oxford University Press.

Wilson, Deirdre, and Dan Sperber. 1998. Mood and the analysis of non-declarative sentences. *Pragmatics: Critical Concepts*, Vol. 2: *Speech Act Theory and Particular Speech Acts*, ed. by Asa Kasher, 268-289. London: Routledge.

Wu, Jiemin, and Hongda Zhu. 2001. *Hanyu Jielüxue* [*A Study of Chinese Prosody*]. Beijing: Language and Culture Press.

Yao, Xinzhong, and Wei-ming Tu. (eds.) 2011. *Confucian Studies*, 4 Vols. London: Routledge.

Zaefferer, Dietmar. 1998. On a formal treatment of illocutionary force indicators. *Pragmatics: Critical Concepts*, Vol. 2: *Speech Act Theory and Particular Speech Acts*, ed. by Asa Kasher, 250-265. London: Routledge.

論漢語施事行爲、情感與節律研究的概念模型

顧曰國

中國社會科學院

本文旨在研究漢語中施事行爲、情感與節律三者之間的互動關係。施事行爲是本文默認的出發點，定義爲一種人類行爲特質，說話人在社會場景裡產出的具體行爲是這種行爲的實例。施事行爲特質視爲由 8 本基本成員構成的集合來作形式化的描述。情感指現場即時情感，分爲背景情感、基本情感和社會情感三個層次。節律指由音高、音強、重音、節奏等心理特徵組成的系統。三者之間的互動關係通過來自“漢語現場即席話語語料庫”中提取的施事行爲實例進行調查分析。本文還討論了幼兒話語裡“言、思、情整一”和“節律顯真情”兩個真前提。

關鍵詞：施事行為，情感，節律，言、思、情整一，節律顯真情

《第四屆國際漢學會議論文集》分冊書名表

中央研究院歷史語言研究所策畫

李宗焜主編《出土材料與新視野》
邢義田、劉增貴主編《古代庶民社會》
柳立言主編《近世中國之變與不變》
祝平一主編《衛生與醫療》
康豹（Paul R. Katz）、劉淑芬主編《信仰、實踐與文化調適》
陳光祖、臧振華主編《東亞考古的新發現》
陳熙遠主編《覆案的歷史：檔案考掘與清史研究》

中央研究院民族學研究所策畫

余安邦主編《身體、主體性與文化療癒：跨域的搓揉與交纏》
張　珣主編《漢人民眾宗教研究：田野與理論的結合》
葉光輝主編《華人的心理與行爲：全球化脈絡下的研究反思》

中央研究院近代史研究所策畫

呂妙芬主編《近世中國的儒學與書籍：家庭、宗教、物質的網絡》
沙培德（Peter Zarrow）、張哲嘉主編《近代中國新知識的建構》
潘光哲主編《近代中國的政治與外交》

中央研究院中國文哲研究所策畫

林玫儀主編《文學經典的傳播與詮釋》
彭小妍主編《跨文化實踐：現代華文文學文化》
楊貞德主編《視域交會中的儒學：近代的發展》
鍾彩鈞主編《東亞視域中的儒學：傳統的詮釋》

中央研究院臺灣史研究所策畫

謝國興主編《邊區歷史與主體性形塑》

中央研究院語言學研究所策畫

鄭秋豫主編《語言資訊和語言類型》

中央研究院人文社會科學研究中心策畫

朱德蘭主編《跨越海洋的交換》

國家圖書館出版品預行編目（CIP）資料

語言資訊和語言類型 / 鄭秋豫主編.-- [臺北市]：
中研院，2013.11
面； 公分. --（中央研究院國際漢學會議論文集. 第四屆）
部分內容為英文
ISBN 978-986-03-8697-4（精裝）
1. 漢語 2. 語言學 3. 文集
030.8 102022276

中央研究院第四屆國際漢學會議論文集

語言資訊和語言類型

有著作權，侵害必究

發 行 人： 翁啓惠
總 編 輯： 黃進興

主 編： 鄭秋豫
出版發行： 中央研究院
印 刷 者： 文匯印刷資訊處理有限公司
總 經 銷： 聯經出版事業股份有限公司
經 銷 商： 台灣語言學學會
四分溪書坊
樂學書局有限公司
萬卷樓圖書股份有限公司
定 價： 單冊 新台幣 500 元
出版日期： 2013 年 11 月

GPN 1010202510 ISBN 978-986-03-8697-4 (精)

本刊不負責來稿內容物之版權問題，由作者自負文責。