# Selected Papers from the 2006 Annual Research Forum of the Linguistic Society of Hong Kong

## 2006年香港語言學學會學術年會論文集

LINGUISTIC SOCIETY OF HONG KONG

Selected Papers

*from the*

2006 Annual Research Forum

*of the*

*Linguistic Society of Hong Kong*

2006 年香港語言學學會學術年會論文集

*edited by*

Colleen Wong

吳杏蓮

Spring 2008

published by *the Linguistic Society of Hong Kong*

香港語言學學會出版

# Contents

# Preface

December of 2006 saw the twentieth anniversary of the Linguistic Society of Hong Kong (LSHK). And we know as well that the Annual Research Forum (ARF) is a central event of the Society. But no one there in the ARF-2006 could tell us how many years had the ARF been on ... only that it is not as old as the Society. The executive committee had decided to publish selected papers based on ARF presentations. So ARF-2006 marks the beginning of LSHK's publication of its own selected papers from the research forum.

Out of over thirty presentations in ARF-2006, about half were written-up afterwards and sent to us for consideration of publication. We then invited colleagues who are experts in the linguistic field to review the papers. The review process was done anonymously.

You will see that the twelve papers in this collection cover various areas in linguistics. We arranged the papers according to the categories in the ARF-2006 programme. So we start with a broad appreciation of linguistics in the daily life, and present to you David Li's comparative study on Chinese-Chinese bilingual interaction, Ge Lingling's survey on jokes, Wong Tak-Sum's study on spelling Cantonese and Zhang Xiaoheng's on Chinese orthography, and Nakayama/Yoshimura's reflection on language immersion experiences.

We then move to Chinese syntax and have four papers offered to you. They are Fan Xiaoqian's headless relative clauses, Cauvis Fung's topic and focus in Cantonese, Sio/Tang's Cantonese indexical element *gam,* and Sun Jing's structure in the context of grammaticalisation.

Three papers related to the acquisition of semantics complete this collection nicely. They are Wong/Man's on aspect, Fan Li's on child first language acquisition, John Wakefield's on second language acquisition.

So you see that both the new and the familiar LSHK people are in touch and in circulation. As editor of this first Selected Papers from the ARF I was touched by how seriously the contributors, fresh graduates or experienced researchers alike, dealt with the reviewers' comments. Above all our heartfelt gratitude goes to the reviewers, for giving LSHK anonymously this precious service amid the busyness of life. We hope they will find the amendments made to their satisfaction. Please enjoy the papers in the New Year, everyone!

*Colleen Wong*
March 2008

# Why '*hi*', '*bye*' and '*sorry*' are Preferred in Chinese-Chinese Bilingual Interactions in Hong Kong and Taiwan: A Comparative Study[1]

David C.S. LI

City University of Hong Kong

## 1. Introduction

This paper reports a number of findings in a large-scale project with the primary objective of comparing motivations of code-mixing (or intra-sentential code-switching) in two Chinese-dominant societies: Hong Kong and Taiwan. The project is a replication of Li and Tse's (2002) experimental study involving 12 English majors in Hong Kong (see http://personal.cityu.edu.hk/~endavidl/index3.htm for details). Participants' self-report data indicate a general preference of *hi, bye* and *sorry* for expressing three speech acts, respectively: (a) greetings, (b) saying good-bye, and (c) apologizing. The main goal of this paper is to report why *hi, bye* and *sorry* are preferred to their putative translation equivalents or indigenized renditions in the bilingual speech of educated Chinese speakers in Hong Kong and Taiwan.

Before reporting on details of the findings, a brief clarification of the status of these expressions as instances of code-mixing or (lexical) borrowing is in order. As is well-known, such a theoretical distinction is crucial for students of contact linguistics. While there is no doubt that the three expressions under scrutiny originated from English, they have been indigenized to a large extent in the Cantonese vernacular of Hongkongers as a result of language contact between speakers of English and Cantonese in this former British colony for decades (Chan and Kwok 1982). With regard to any of these three expressions, in general the closer the speaker's pronunciation approximates phonological norms of some NS-based variety of English (e.g. British English, American English), the stronger the interlocutor's perception of the speaker code-mixing. In contrast, if the actual pronunciation comes close to phonological norms in Cantonese (viz.: $haai^{55}$; $baai^{55}$ / $baai^{55}$ $baai^{33}$; $so^{55}$ $li^{21}$), then it is more appropriately analyzed as borrowing. The theoretical import of such a distinction is clear, but it may not be so clear-cut in informal bilingual interactions between Cantonese-English bilinguals due to considerable inter- as well as intra-speaker variation in pronunciation from one context to another. Among the more salient factors behind linguistic variation are (a) the speaker's level of English proficiency; (b) the speaker's perception of the linguistic repertoire of the interlocutor(s); and (c) the perceived degree of formality of the language situation.

Linguistic variation is not restricted to speech. As is well-known, the expressions *hi*, *bye* and *sorry* occasionally occur in informal sections of Hong Kong Chinese newspapers and magazines such as columns, comic strips and adverts, where language use tends to follow norms of informal speech, which helps explain why script mixing is so common. Both *hi* and *bye* have evolved Chinese written forms (嗨, *haai⁵⁵* / *hai*; 拜 *baai⁵⁵* / *bai*; 拜拜 *baai⁵⁵ baai³³*), which are often used in free variation with their respective English counterparts, sometimes within the same text by the same writer. *Sorry*, on the other hand, tends to appear in its original English spelling, probably because no Chinese morpho-syllables are considered to be suitable orthographic renditions of $so^{55}$ and $li^{21}$. In terms of the theoretical distinction between code-mixing and borrowing, therefore, the former come closer to being candidates of borrowing than the latter.

## 2. Methodologies, participants and procedures

A total of 108 student participants (65 in Taiwan and 43 in Hong Kong) were asked to speak only the dominant community language for one day (i.e. Mandarin and Cantonese, respectively), and record their personal experiences in a diary in a language of their choice (Chinese or English). Participants were students at three universities: National Donghwa University and National Chengchi University in Taiwan, and City University of Hong Kong. All except one Aboriginal student are ethnic Chinese. Their major disciplines and numbers are presented in Table 1. Participants were self-selected by responding to an advert on campus. Each participant was required to attend a briefing, undergo the one-day experiment, write a diary, and attend a focus group interview. All participants were given a modest remuneration for their participation.

| University and date of the experiment | Student participants' major discipline | No. of students |
|---|---|---|
| National Donghwa University, Hualien, Taiwan (TDU) (December 7, 2003) | Chinese majors (Ch) | 9 |
| | English majors (En) | 8 |
| | Science / Technology / Engineering majors (Sc) | 8 |
| | Business / Economics / Marketing majors (Bu) | 8 |
| National Chengchi University, Taipei, Taiwan (TCU) (December 13, 2003) | Chinese majors (Ch) | 8 |
| | English majors (En) | 8 |
| | Psychology majors (Ps) | 8 |
| | Business / Economics / Marketing majors (Bu) | 8 |
| City University of Hong Kong, Hong Kong (HCU) (January 6, 2004) | Chinese majors (Ch) | 8 |
| | English majors (En) | 9 |
| | Science / Technology / Engineering majors (Sc) | 9 |
| | Business / Economics / Marketing majors (Bu) | 7 |
| | Psychology majors (Ps) | 10 |
| | Total | 108 |

**Table 1**: Participants and their major disciplines at each of the three universities

At the briefing at each of the three research sites, the rationale of the study was explained in the local dominant community language, clear instructions were given, followed by a short try-out and trouble-shooting. At each university the one-day experiment took place during a school holiday so as to minimize disruption to student participants' school work. They were instructed to send their reflective diaries to the researchers electronically the next day or the day after. The diaries were then analyzed inductively and collated thematically for each group. The result is a list of quotations structured under different sub-themes serving as stimulus material for a focus group discussion, which took place within 72 hours after the day of the experiment. The focus groups, each lasting for about two hours, were audio- and video-recorded with the participants' consent, and subsequently transcribed for data analysis. The data reported below are extracted from both reflective diaries and focus groups.

## 3. Results

### 3.1 *Hi* for greeting

One of the problems encountered by both Hong Kong and Taiwanese participants was the choice of greeting expression. The following are greeting expressions found

in our data:

| Cantonese: | *haai*[55] (嗨), *nei*[23] *hou*[35] (你好), *zou*[35] *san*[21] (早晨), *wai*[33] (喂), non-verbal such as hand gesture |
| Mandarin: | *nǐ hǎo* (你好), *nǐ zǎo* (你早), *zǎo* (早), *zǎo ān* (早安), *wǔ ān* (午安), *wǎn ān* (晚安) |
| Southern Min: | *za a* (functionally equivalent to Mandarin zǎo) |
| English: | *hi, hello, hey man, good morning* |

By far the most frequently mentioned greeting expressions are (Mandarin) *nǐ hǎo*, (Cantonese) *nei*[23] *hou*[35], and (English) *hi*. Most participants in Taiwan and Hong Kong reported being inconvenienced by the artificial 'one day with only Mandarin/Cantonese' rule of speaking. As shown in both the diary and focus group data, 'hi' (also commonly written in Chinese as '嗨', especially in Taiwan) was preferred to corresponding Chinese greeting expressions if the context is informal, not only in their interactions with close friends and peers, but also new acquaintances of the same age group. The corresponding expressions in Mandarin (*nǐ hǎo*) or Cantonese (*nei*[23] *hou*[35]) are generally perceived as being overly formal and signaling a marked social distance. For example:[2]

| (1) | 在學校看到同學，我直覺是打招呼說了一聲 'hi'。我突然想到今天只能說國語，不過還是脫口而出，現在我們打招呼常說 'hello' 或是 'hi'，幾乎不會說 '早安'、'午安'、'晚安'，或是 '你好'，所以光是打招呼，我就會不自覺說了 'hi'。 (TDU Sc M8) | When [I] saw my classmates, I instinctively greeted [them] with '*hi*'. Then I suddenly recalled that I was supposed to speak only Mandarin today, but still [that English word *hi*] slipped out of [my] mouth unconsciously. Nowadays we often greet each other with '*hello*' or '*hi*'; rarely [would we] say '*zǎo'ān*', '*wǔ'ān*', '*wǎn'ān*', or '*nǐ hǎo*'; that is why on greeting alone, I unconsciously uttered '*hi*'. |
| (2) | 連打招呼常發生的陷阱：Hi，我都改成你好，只是覺得很奇怪，好像我跟那些同學都不熟的樣子或是我是一個外國人剛剛學中文的感覺。(TCU Ch F6) | Even the trap in greeting: *hi*, I had it converted to *nǐ hǎo*, but it gave me a very strange feeling: it was as if those classmates and I were no longer so close, or I sounded like a foreigner learning beginner-level Chinese. |

(3) When I want to say 'hi' to my friend as greeting, I have control myself and translate it into '你好'. My friend used the mystery eyes to stare at me and ask me what happened to me today. We seldom use Chinese terms such as '你好' or '再見' but saying 'hi' and 'bye' instead. We have used to speak English for greeting. (HCU Bu M7)

One Hong Kong participant, F7 from the science group, attributes that perception to a common association between writing *nei²³ hou³⁵* in primary school compositions:

| | |
|---|---|
| (4) '你好' 似乎只會在小學作文時代才會用得著，說了出來好像太過文雅，不夠親切。 (HCU Sc F7) | '*Nei²³ hou³⁵*' sounds like what primary school pupils would write in compositions; it seems far too literary, not intimate enough. |

Similar perceptions are widely shared by participants in various focus groups, as shown in the following excerpts. Thus F1 in (5) makes the point that when we meet a friend of similar age, the natural greeting expression would be *hi*, rather than *nǐ hǎo*:[2]

| | | | | |
|---|---|---|---|---|
| (5) | 240 | F1 | 比較難爲情的話… 我覺得就有些打招呼呀 / 問候語用英文表達是比較自然這樣 // 像你跟人家..人家迎面走來 / 你跟人家 hi / 你當然不會說..[…] 你好 // (TCU En) | Somewhat embarrassing [cases]… I feel some greeting / greeting expressions in English sound more natural // like [if] you cross someone.. someone coming your way / you [would] say *hi* to that person / you certainly won't say […] *nǐ hǎo* // |

Another Taiwanese participant F6 reports her roommate's strange reaction when she responded to her greeting *hi* with *nǐ hǎo*, which gave her roommate a feeling of meeting a new acquaintance, as if they had just shaken hands (6):

| | | | | |
|---|---|---|---|---|
| (6) | 52 | F6 | 像我去對面寢室進去的時候 / 她就很奇怪 / 那個同學就很習慣說 / hi / 然後我就..我就說你好 […] 她就用很奇怪的眼神看我 / 好像我兩人剛見面 / 然後剛握完手說你好 / 是啊 / 這種感覺很生硬 / 其實很生硬 // (TCU Ch) | Like when I entered the bedroom / she felt very strange / that classmate said *hi* [to me] as usual / then I responded *nǐ hǎo* […] then she stared at me with startling eyes / as if the two of us had just met / and then just finished shaking hands / like that / a kind of stiff / very stiff [relationship] // |

*Hi* is similarly preferred by Hong Kong participants because $nei^{23}\ hou^{35}$ sounds far too formal to them, as shown in M3's remark in (7):

| | | | | |
|---|---|---|---|---|
| (7) | 52 | M3 | 誒可能我平時生活習慣都可能會夾雜啲英文 / 咁尤其是話講一啲誒詞語譬如誒好似話…打招呼 / 通常都會話叫'hi'呀 / 或者'good morning'咁樣 / 咁但係又好少無端端啲誒就算係話同一個中國人講 / 咁我都會 say hi 囉 / 啲唔會話突然間話'你好'咁樣 / 啲會覺得感覺好怪囉呢個就 // | Eh maybe it is a habit of mine to code-mix some English / so for some expressions eh such as… greeting / usually [I] would say *hi* / or *good morning* like that / but rarely for no reason eh even when talking to a fellow Chinese person / I would *say hi* you know / eh [I] won't out of the blue say $nei^{23}\ hou^{35}$ / [I] would find that very strange indeed // |
| | 53 | DL | 咁嗰日有無講'你好'呀? | So did you say $nei^{23}\ hou^{35}$ on that day? |
| | 54 | M3 | 無乜 / ..啲我會覺得好怪 / 啲因為我平平時遇朋友嘅時候都係'hi' / 會感覺自然啲/ 因為大家朋友之間都係用用'hi'呀或者 // (HCU Bu) | Not really / ..eh I would find it very odd / because usually when meeting friends [I] usually say *hi* / [that] makes me feel more natural / because between friends we all say *hi* maybe // |

Later on in the same focus group [HCU Bu], when the topic of language choice for greeting was brought up again, M4 echoes the same impression of 'strangeness', while M7 finds $nei^{23}\ hou^{35}$ much less intimate, as if it was a customer service staff addressing a client:

| (8) | 511 | DL | ..感覺如何呢? '你好' ? | ..how did you feel about [saying] *nei²³ hou³⁵*? |
|---|---|---|---|---|
| | 512 | M4 | 怪囉 / 即平時講開 'hi' // | kind of strange / for usually [I] say *hi* // |
| | 513 | M7 | 我覺得你好似係嗰啲客戶服務先用㗎咋 / 即打嗰啲 / '你好' / '郭先生' / 呢度係..咁 / 即佢反而你唔係咁親切先會用'你好' // | Saying *nei²³ hou³⁵* makes me feel like a customer service staff / like calling / *nei²³ hou³⁵* / Mr. Kwok / here is.. [like that] / it is used when the relationship is not so intimate // |
| | 514 | M2 | 如果同你識嗰啲人 / 可能如果夾硬講咪'早晨呀'咁囉 / 可以..未必一定用'你好'呢個字嘅 // | With people you know / maybe to say [something I] would say *zou³⁵ san²¹* ['good morning'] like that / one need not use '*nei²³ hou³⁵*' // |
| | 515 | M7 | 'hi'我覺得佢表示到好親切嘅感覺 // | [With] *hi* I feel it conveys a strong sense of intimacy // |

Likewise, in another focus group in Hong Kong, F4 reports that rather than saying *nei²³ hou³⁵* which sounds too formal to her, he used non-verbal communication (e.g. hand-waving, head-nodding) as a functional substitute for saying *hi* on the day of the experiment:

| (9) | 475 | DL | 係喎..如果你..譬如果日唔講得'hi'啦 / 而[...]無理由就咁掂行掂過咁呀 / 你會..你用咩講法呢咁? | By the way.. if you.. suppose you could not say *hi* on that day / but [...] [you] can't possibly just walk by / what would you say then? |
|---|---|---|---|---|
| | 476 | F4 | 揮揮手咁樣 [ F8: 係呀係呀] // | [I] waved my hand like this [F8: yes yes ] |
| | 477 | DL | 哦:::即係所謂 non non verbal / | Oh::: using so-called non-verbal / |
| | 478 | F4 | 即係用個頭咁樣咁樣 / [HAHAHA] / 即係諤一諤咁樣 // (HCU Ch) | That is, using head movement like this / [HAHAHA] / that is, nodding like this // |

There is thus clear evidence that educated Chinese-English bilinguals in Hong Kong and Taiwan generally prefer using *hi* as the greeting expression in social interactions with friends and peers. Another interesting observation is that with few exceptions, all participants perceived *hi* as an English rather than an indigenized Chinese expression (viz.: 嗨), suggesting that the theoretical distinction between

code-mixing and borrowing, which matters so much to students of contact linguistics, has little or no linguistic reality in the mental lexicons of Chinese-English bilinguals.

## 3.2 English for saying good-bye

A similar problem was encountered by our participants with regard to the speech act of saying good-bye. The relevant expressions mentioned by participants are as follows:

| | |
|---|---|
| Cantonese: | $zoi^{35} gin^{21}$ (再見) |
| Mandarin: | $zàijiàn$ (再見), $bai$ (拜), $wŏ xiān zŏu le$ (我先走了) |
| Southern Min: | choice of expression depends on context and speaker's role |
| English: | $bye, bye bye$ |

As in the choice of greeting expression among friends and peers who are bilingual in Chinese and English, there seems a problem of stylistic incongruence between the functional, but not quite semantic, equivalents. While *bye-bye* (variant: *bye*) and 再見 ($zàijiàn, zoi^{33} gin^{33}$) are widely perceived as translation equivalents, the former sounds informal and easy-going, the latter formal and distant. This was probably the reason why one Taiwanese participant found *bye bye* so irresistible on the day of the experiment:

| | | | | |
|---|---|---|---|---|
| (10) | 327 | F5 | 我是 bye bye 完全避不掉 / 我從早上接電話到晚上 / 我都一直避不開這個字 / 對呀 / 就算結束之前 / 再接一通 / 然後還是..一直還會很直接的說 bye bye // (TDU Bu F5) | For me, *bye bye* was unavoidable / receiving calls from early morning to evening / I could not avoid using this word / right / even right before the end [of the experiment] / [I] received another call / still.. I kept saying *bye bye* // |

Stylistic incongruence also explains why, when some participants used *zàijiàn* with peers, they were taken aback, as if it signaled a deterioration of their relationship:

| (11) | 沒想到國語中的打招呼用語在日常生活中已漸漸被英文取代。還有類似的如道別說再見的時候，我們現在通常會用"bye-bye"來取代中文的"再見"。老一輩的人或許還是常用再見，但今天我刻意使用這個來道別時，對方還有些錯愕，可能感覺說起來沒有"bye-bye"那麼的平易近人,那麼的熟悉吧！（TCU En F1） | [I] didn't realize that Mandarin greeting expressions in everyday life have gradually been replaced by English [counterparts]. Similar examples such as saying good bye, we'd usually say '*bye-bye*' instead of Mandarin *zàijiàn*. Elderly people may still say *zàijiàn*, but today when I said *zàijiàn* to bid farewell, the addressee was somewhat taken aback, maybe it's because [*zàijiàn*] does not sound as intimate and familiar compared with *bye-bye*! |
| --- | --- | --- |

There is general consensus among Taiwanese participants that *zàijiàn* is commonly used with teachers, partly because one needs to be polite when talking to them:

| (12) | 563 | F6 | 你會講老師再見 / 然後小孩 bye 呀 // (TDU Ch) | You would say *lǎoshī zàijiàn* ('good-bye teacher') / and then with children [you would say] *bye* // |
| --- | --- | --- | --- | --- |

But age and/or the level of acquaintance may have a bearing on the choice between *zàijiàn* and *bye bye*. Thus F6 in (13) indicates that to younger teachers who are of a similar age to them, she would probably say *bye*, while M1 in (14) said he might use this expression with teachers he knows well:

| (13) | 679 | F6 | 我對那個..跟我年紀差有點多 / 譬如說已經五..是五六十歲的老師的話 / 我就會跟他說 '那我先走了' / 那或 '再見' / 會比用比較尊敬的語氣來講 / 可是如果是遇到剛從像像我的國中老我的國中導師是 / 他我們低年級的時候他就剛剛剛從師範大學畢業 / 所以他是比較跟我們年紀相近的 / 那所以我在跟他講的時候 / 我會說 / '那我先走了 *bye*' 這樣子 [ . . . ] // | To [teachers who are] much older than me / say teachers [who are] already 50 or 60 years old / I would say to them *wǒ xiān zǒu le* ['I leave now'] / or '*zàijiàn*' / [I] would use a polite tone / but if I come across a teacher [who] like my teacher of Chinese in upper secondary school / he just graduated from Normal University when he taught our junior class / his age was close to ours / so when talking to him / I would say *wǒ xiān zǒu le, bye* like that [...] // (TCU Ch) |
| --- | --- | --- | --- | --- |

| (14) | 343 | M1 | 其實..我個人的感覺 / 因爲我那天有去找系主任 / 那..個人感覺是 / 當你叫不熟的老師 / 你就可能會 '老師再見' / 那如果比較熟的話 / 就可能用 '老師 bye bye' / 就直接就帶過了 / 那其實我覺得 / 那對我來說 / 可能是熟悉度的關係 // (TDU Bu) | Actually.. my own feeling / because that day I went to see the Department Head / my personal feeling was / when addressing teachers you are not familiar with / you would say *lǎoshī zàijiàn* / but to teachers I am more acquainted with / [I] would probably say *bye bye* / directly / in fact to me I think / it is a matter of the level of acquaintance // |

Two Taiwanese participants, M8 (15) and F5 in (16), independently pointed out that they would say *zàijiàn* when getting angry with their interlocutor:

| (15) | 1035 | M8 | 我覺得'再見'是那種..比如說你很生氣 / 然後..'再見' // (TCU Ps) | I feel that *zàijiàn* is kind of.. for instance when you get angry / then.. *zàijiàn* // |

| (16) | 770 | DL | 有有沒有一些場合 / 你會跟對話人講'再見'的？ 能想得到嗎？ | Are there any situations / you would say *zàijiàn* to [your] interlocutor? Can [you] think of any? |
| | 771 | F5 | 我生氣的時候 / 就講'再見'.. // | When I get angry / then [I] would say *zàijiàn* // |
| | 772 | DL | [HAHAHA] / 生氣的時候 / 生氣的時候會講 / 有沒有想過是爲什麼？ | [HAHAHA] / when [you] get angry / [you] would say it when angry / any idea why? |
| | 773 | F5 | 比較..義正嚴詞吧 // (TCU Bu) | More.. serious and stern [with the interlocutor] |

Finally, one Taiwanese participant, M7 in (17), suggests that *zàijiàn* is typically used for bidding farewell, for instance seeing somebody off at the airport:

| (17) | 352 | DL | 甚麼情況之下你們會說'再見'呢? [...] 剛才提到..那個老師不熟的話 / 其他人呢? | Under what circumstances would you say *zàijiàn*? [...] Just now [someone] said [if] not acquainted with the teacher / [What do] others [think]? |
| | 353 | M7 | 生離死別吧 // [...] 就是在機場呀 / 或者是 / 要離開很久很久很久 // (TDU Bu) | [Situations involving] life and death maybe // [...] like at the airport / or [people] leaving somewhere for a long long time // |

A similar observation was made by Hong Kong participants in general: *bye bye* is preferred in informal social interactions, for they found *zàijiàn* much too formal, as F9 in (18) puts it:

| (18) | 474 | F9 | *bye bye* 都幾多講 // 但係講 '再見' 呢 / 就誂喺正式嘅場合啦 // (HCU En) | [I prefer to] say *bye bye* // as for *zoi³³ gin³³*, it is usually said in formal situations // |

As in the case of greeting expressions, the participants in Hong Kong and Taiwan all regarded *bye* and *bye bye* as English rather than Chinese expressions, despite the fact that, much like 沙律 (*saa⁵⁵ leot³⁵*, 'salad') and 沙拉 (*sā lā*, 'salad'), 沙發 (*sā fa*, 'sofa') and 梳化 (*so⁵⁵ faa³⁵*, 'sofa'), both *bye* and *bye bye* have indigenized and evolved written forms in Chinese (i.e. Cantonese: 拜/拜拜, *baai⁵⁵ / baai⁵⁵ baai³³*; Mandarin: 拜, *bai*). This again suggests that Chinese participants are unconcerned about the theoretical distinction between code-mixing and borrowing, which is so important for students of contact linguistics.

## 3.3 English for making apologies

Unlike greeting and saying good-bye, there is less linguistic variation in the range of expressions (particles) mentioned by our participants for the speech act of apologizing:

| Cantonese: | *deoi³³ m²¹ zyu²²* (對不起), *m²¹ hou³⁵ ji³³ si³³* (唔好意思) |
| Mandarin: | *duìbùqǐ* (對唔住), *bùhǎo yìsi* (不好意思), *bàoqiàn* (抱歉) |
| Southern Min: | not found in data |
| English: | *sorry, excuse me* |

Our data suggest that participants in Hong Kong and Taiwan alike are aware that in their respective dominant community languages, Mandarin and Cantonese, there exists a functional contrast between two Chinese expressions used for apologizing: *duìbùqǐ* (對不起) / *deoi$^{33}$ m$^{21}$ zyu$^{22}$* (對唔住) and *bùhǎo yìsi* (不好意思) / *m$^{21}$ hou$^{35}$ ji$^{33}$ si$^{33}$* (唔好意思). The former is used for offences which are perceived as more serious; the latter for offences perceived as minor or less serious. Both are generally regarded as functional equivalents of *sorry* in English. Interestingly, in social interactions between Chinese-English bilinguals in Hong Kong and Taiwan, *sorry* seems to be preferred to *bùhǎo yìsi* (不好意思) / *m$^{21}$ hou$^{35}$ ji$^{33}$ si$^{33}$* (唔好意思). Consequently, the indigenous contrast seems to have gradually given way to a bilingual contrast, such that *duìbùqǐ* (對不起) and *deoi$^{33}$ m$^{21}$ zyu$^{22}$* (對唔住) are reserved for expressing apologies for serious offences, whereas for less serious offences, the speech act of apology is gradually taken over by 'sorry'. For example:

| (19) | 日常見到朋友都會打聲招呼 'hello', 不小心在街上碰撞到一個陌生人, 也會從心很自然地說聲 'sorry', …有時說英文是比較好和自然的。(HCU Sc F7) | Usually [when I] see [my] friends [I] will say 'hello'; [when I] bump into a stranger by accident, instinctively and naturally I would say 'sorry', …sometimes speaking in English is better and sounds more natural. |
|---|---|---|

| (20) | 487 | M4 | 就撞到人就 'sorry' / [...] // | [If I] bump into somebody, [I'd say] *sorry* [...] // |
|---|---|---|---|---|
| | 488 | DL | 爲其麽會那麽普遍呢? | Why is [*sorry*] so common? |
| | 489 | M3 | 那就要看程度 / 像我的話 / 誒 'sorry' 這樣 'sorry 抱歉'這樣子 / 還是 '不好意思' / 可是隨着那個就是程度越來越嚴重 / 會開始慢慢偏向 '對不起' / [...] 應該是覺得心理上這個覺得很愧疚才會這樣 // (TDU Sc) | That depends on degree [of seriousness] / like me / eh *sorry* like this or *sorry bàoqiàn* like that / or maybe *bùhǎo yìsi* / but as the degree of seriousness gets more and more marked / [I] would gradually shift to *duìbùqǐ* / [...] that applies when one is psychologically guilty // |

| (21) | 225 | DL | 唔好意思 / 諗同'對唔住'有乜分別呢? | $m^{21} hou^{35} ji^{33} si^{33}$ / and $deoi^{33} m^{21} zyu^{22}$ what is their difference? |
|---|---|---|---|---|
| | 226 | F1 | 可能冇個程度上嘅唔同 // | Probably a difference in terms of degree // |
| | 227 | M2 | 好似錯啲咁 / '對唔住' // [...] '唔好意思'好似無咁嚴重咁 / '對唔住'好似唔知做錯啲咩 // | Seems more serious / $deoi^{33} m^{21} zyu^{22}$ // [...] $m^{21} hou^{35} ji^{33} si^{33}$ does not sound so serious / $deoi^{33} m^{21} zyu^{22}$ seems to imply a [serious] blunder // |
| | 228 | F8 | 係囉 / 大錯咁 // | Yes / a big blunder / |
| | 229 | M4 | 喺問路你唔會話 '對唔住' / 咁請問你' / '呀唔好意思呀想問吓你諗點樣點樣去'咁樣../ 你唔會話 '對唔住'.. // (HCU En) | Like asking for directions you wouldn't say $deoi^{33} m^{21} zyu^{22}$ / so [you would say] $cing^{35} man^{22} nei^{23}$ ['please may I ask you'] / $m^{21} hou^{35} ji^{33} si^{33}$ how to go.. / you won't say $deoi^{33} m^{21} zyu^{22}$ // |

In one instructive example, a Hong Kong participant F7 was late for an appointment with a classmate (also a participant in the 'one-day' experiment: F5). She honored the Cantonese-only rule of speaking by apologizing in Cantonese (i.e. $deoi^{33} m^{21} zyu^{22}$, 對唔住); and she addressed F5 by her Chinese given-name instead of her adopted English first name ('Karen') – something that F7 would never do normally. It turned out that F7's apology in Chinese was perceived by F5 as being very sincere, for it reflected a deeper sense of regret than if she had apologized using *sorry* (22):

| (22) | 約了朋友今天早上十一點在銅鑼灣地鐵站等，但我要遲到了。在巴士行駛途中，我致電給我的朋友。我已想好避開用英文，把平日叫她'Karen'改爲叫她中文名的末二字'君怡'。電話一接通，我即"君怡，對唔住呀！"…若是平日的話，我必是: 'sorry 呀 Karen!'的說。如此一個中文版的道歉，使她覺得我的歉意很深，比原有的誠意更有誠意。(HCU Ch F7) | [I] had an appointment with a friend this morning at MTR Causeway Bay, but I was going to be late. On the bus, I called up my friend. I already thought about avoiding using English, by calling her 'Kwan Yee' [Chinese given name] instead of *Karen* [adopted English name]. When she answered the phone, I said 'Kwan Yee, $deoi^{33} m^{21} zyu^{22} aa^{33}$!'… Normally, I would say '*sorry aa^{33} Karen*'. Such a Chinese version of apology made her feel that my regret was deep, much deeper and more sincere than the original [code-mixed] apology. |
|---|---|

The choice between *duìbùqǐ* (對不起) and *deoi³³ m²¹ zyu²²* (對唔住) on one hand, and *sorry* on the other, is not always clear. Thus in those cases where the perceived degree of seriousness is ambivalent, the speaker may invoke both (saying, e.g. 'sorry *deoi³³ m²¹ zyu²²*'), as in (23) where the Hong Kong participant M3 makes the point how he would apologize to someone on whose toes he stepped on unintentionally (compare '*sorry bàoqiàn*' said by M3 in 20):

| (23) | 244 | M3 | 似係誽受 excuse me 同埋 sorry 嘅影響囉 / 譬如你誽你問你誽阻人哋一分鐘問人時間 / 你通英文講話'excuse me / 我可唔可以阻你一分鐘呀'咁樣囉 // 譬如你踩親人會講 'sorry 對唔住' 咁樣囉 / 我覺得係呢個係英文潛移默化左喎 // (HCU En) | It seems to be the influence of 'excuse me' and 'sorry' / e.g. [if] you ask someone and keep that person for a minute / you can say in English excuse me / may I keep you for a minute like that // if you step on someone [you] would say '*sorry deoi³³ m²¹ zyu²²*' like that / I feel this is influence from English subconsciously // |

## 4. Discussion and conclusion

As is different from other approaches to the research question 'Why do bilinguals code-switch?', which tend to rely on the researcher's interpretation of a given body of code-switching data, the approach in this study allows the researcher to tap into code-switchers' awareness of the reasons why they would want to code-switch in context-specific interactions with others. Since Blom and Gumperz's (1972) classic study of code-switching in Hemnesberget, Norway, the practice of asking speakers to provide information on their language use patterns has been judged to be unreliable. Li and Tse's (2002) experimental study, however, demonstrates that code-switchers' own accounts are highly informative regarding specific reasons why they wish to switch to another language. The key lies in methodology: the breaching technique 'revelation through disruption' in ethnomethodology (Garfinkel 1967) turned out to be very helpful in raising participants' consciousness of language choice decisions as they go about their day-to-day verbal interactions with others. This experimental design, replicated on a larger scale in Hong Kong and Taiwan, successfully elicited useful data regarding Chinese-English bilinguals' context-specific concerns for speaker meanings – meanings which are otherwise inaccessible to the researcher. It is therefore a central claim of this study that informants' views toward specific reasons

why they would want to code-switch are researchable, and that their reflective feedback and spontaneous comments constitute useful data for triangulation purposes – provided their awareness has been raised beforehand through some research design characterized by consciousness-raising.

The data presented in this study form a small subset of a much larger data set. They focus on participants' general preference of English expressions (*hi*, *bye*, and *sorry*) for the speech acts of greeting, saying good-bye and apologizing respectively, as well as their subjective reasons behind those preferences. The analysis shows that their views are not isolated, but widely shared by fellow Chinese-English bilinguals in other groups. Of concern to us here is the reason why *hi*, *bye* and *sorry* are preferred by a majority of educated bilinguals in Hong Kong and Taiwan, especially when interacting with friends and peers. As shown in the above analysis, *hi* and *bye* are largely preferred for their informality. In situations judged to be casual and where the marking of politeness is a low-level concern, both expressions are generally perceived as more appropriate than their Chinese functional equivalents 你好 (*nǐ hǎo, nei$^{23}$ hou$^{35}$*) and 再見 (*zài jiàn, zoi$^{33}$ gin$^{33}$*), respectively, for the latter are seen as far too formal and, in the eyes of some participants, carry other unwanted associations. Thus many participants pointed out that 你好 gives them the impression of new acquaintances meeting for the first time, and for this reason, highly marked in terms of degree of politeness and formality (see 1-9). As for the preference of *bye*, some participants indicated that to them 再見 is more appropriately used for bidding farewell, as what people would say to each other on the assumption that both sides won't be meeting again in the near future (17). To a few other participants, 再見 uttered with the right prosodic features suggesting indignation would be the appropriate expression if they lost temper and would like to indicate the additional metaphorical meaning: 'let us part ways for good' (15-16). There is thus some evidence showing that *hi* and *bye* are preferred to their respective Chinese functional equivalents due to the bilingual speaker's wish to use an informal exponent for the speech acts of greeting and saying good-bye – meanings which cannot be obtained in the Chinese exponents without leading to a problem of stylistic incongruence. In short, for the speech acts of greeting and saying good-bye, the English words *hi* and *bye* filled a functional gap in Cantonese and Mandarin (and possibly other Chinese varieties as well) in regard to the stylistic opposition between 'formal' and informal'.

There is thus strong evidence suggesting that sustained contact with English has brought about a shift in the linguistic repertoire of educated Chinese bilinguals in Hong Kong and Taiwan with regard to the linguistic exponents of the speech acts of

greeting and saying good-bye. Prior to contact with English, only one general and normative expression unmarked for the degree of formality was used (i.e. 你好, *nǐ hǎo, nei²³ hou³⁵*; 再見, *zài jiàn, zoi³³ gin³³*). As a result of sustained contact with English, a bipolar, differentiated system consisting of two exponents has evolved: one exponent (from English) is marked for 'low degree of formality' (i.e. *hi* and *bye*), while the Chinese exponents (i.e. 你好 and 再見) are marked for 'high degree of formality' and are reserved for formal contexts. It is conceivable that the locus of this systemic shift was first invoked in bilingual interactions among educated Chinese in the form of intra-sentential code-switching or code-mixing, before being spread far and wide in the local communities, probably through individual bilinguals' social networks and subsequently the pervasive influence of popular culture and mass media.

While the spread of *hi* and *bye* in Chinese communities was probably driven by a collective need to fill functional gaps in the local community language, the spread of and community-wide preference for *sorry* seems to be different: as shown in many participants' remarks above, there already existed a bipolar contrast between two indigenous expressions, one marking 'high degree of offence' (i.e. *duìbùqǐ*, 對不起; *deoi³³ m²¹ zyu²²*, 對唔住) and the other 'low degree of offence' (i.e. *bùhǎo yìsi*, 不好意思; *m²¹ hou³⁵ ji³³ si³³*, 唔好意思). The fact that many participants were aware of this contrast suggests that the bipolar system continues to be viable in the local communities. What seems to have happened is that the linguistic exponent marking 'low degree of offence' is increasingly supplanted by a more informal-sounding *sorry*. One consequence is that over time, in the repertoire of educated Chinese-English bilinguals, the linguistic exponents for the speech act of apologizing have shifted from an indigenous bipolar system to a three-way contrast roughly as follows:

| The speech act of apologizing | Perceived degree of offence | | |
|---|---|---|---|
| | **'high'** | **'fuzzy': 'not high'** | **'low'** |
| Exponents in Mandarin (Taiwan): | duìbùqǐ (對不起) | sorry | bùhǎo yìsi (不好意思) |
| Exponents in Cantonese (HK): | deoi³³ m²¹ zyu²² (對唔住) | sorry [variant: 'sor'] | m²¹ hou³⁵ ji³³ si³³ (唔好意思) |

In sum, the dominant local language of Chinese-English bilinguals in both Hong Kong and Taiwan shows considerable influence from English, as indicated in their preference of English expressions for realizing the speech acts greeting, saying good-bye, and apologizing. There are two possible reasons for this: (a) extended

exposure to English in all forms of audio-visual entertainment from USA and UK, especially songs, TV programs, films, electronic games and above all, the Internet; and (b) the bilinguals' wish to fill a functional – more specifically stylistic – gap in the expression of the three speech acts in question. Thus the popular use of *hi* and *bye* (variant: *bye bye*) constitutes clear evidence of such a gap in the stock of Chinese exponents for marking the first two speech acts as stylistically informal, while the ubiquity of *sorry* in the bilingual speech of educated Chinese may be accounted for by the inherently subjective and fuzzy nature of offences with regard to the functional contrast between 'serious' (i.e. 'high degree of offence') and 'minor' (i.e. 'low degree of offence'), hence the need for a fuzzy, third linguistic exponent ('not so high degree of offence').

While *hi* and *bye* were treated in the above analysis as English elements, it should be noted that their status as English or Chinese words (i.e. borrowings in the latter case) received relatively little attention by the participants in this study. A few Hong Kong participants did point out that it was unclear to them whether *bye bye* was an English or Chinese word, on the grounds that they had already been rendered into written Chinese (i.e. 拜拜; compare 嗨 'hi' in written Chinese):

| (24) | 462 | F5 | ..其實唔知 bye bye 究竟當唔當中文喎? 因為..中文可以寫個拜拜 // | ..Actually it is not clear whether *bye bye* is Chinese? Because.. [in] Chinese [one] can write 拜拜 // |
|------|-----|----|------|------|
|      | 463 | F9 | 所以已經係屬於一種新嘅詞彙.. / 係啦 / 係已經屬於新詞彙 // (HCU EN) | So it is already a type of new vocabulary.. / yes / already a new [Cantonese] vocabulary [item] // |

| (25) | 345 | F4 | 係呀..跟住我都喺度諗..諗..其實 'bye-bye'..諗..因為中文都有 [F7:都有得寫] '拜拜'兩個字咁.. // | Yeah.. then I thought.. eh.. actually *bye bye*.. eh.. because [in] Chinese has [F7: can also be written] 拜拜 the two characters.. // |
|------|-----|----|------|------|
|      | 346 | F3 | 咁所以我就..係囉我都諗 / 'bye bye' 呢個係中文定係英文呢 // (HCU Ch) | So I.. yeah I also think / *bye bye* [I'm not sure] whether it is Chinese or English // |

Prior to the one-day experiment, no attempt had been made to clarify the fuzzy boundary between code-mixing (i.e. when the English element is pronounced like an English word, e.g. *printer*) and lexical borrowing (i.e. when the element of English origin is pronounced after phonological norms of the local community language, e.g.

*pin$^{55}$ taa$^{35}$*, 'printer'). What seems certain is that regardless of how *hi* and *bye* were pronounced, they were perceived by the majority of participants as English elements. That is, while segments of bilingual speech involving *hi* and *bye* (to a lesser extent, *sorry*) may be analyzed as either code-mixing or borrowing depending on their degree of approximation to phonological norms of the embedded language (EL, here English) or matrix language (ML, here Mandarin or Cantonese), they are generally regarded by bilingual Chinese participants as English expressions. This is the main justification why the examples of bilingual speech concerning the use of *hi*, *bye* and *sorry* are analyzed as instances of code-mixing (or intra-sentential code-switching) rather than borrowing. I believe this point has important implications for further research in the theoretical distinction between code-mixing and borrowing in contact linguistics.

Finally, the findings in this study suggest that in language-contact situations, linguistic exponents of the embedded language have the potential to enrich a functional domain in the matrix language through the transfer of stylistically marked EL linguistic exponents into the ML. Depending on one's ideological inclination, the English expressions *hi*, *bye* and *sorry* may be seen as a welcome trend or an unwanted intrusion 'contaminating' the Chinese language. One thing is certain, however: such contact-induced changes are extremely resilient and will remain an integral part of the linguistic repertoire of Chinese communities in Hong Kong and Taiwan for a long time to come.

## Notes

[2] For readers who are not literate in modern written Chinese and/or written Cantonese, diary data written in Chinese are rendered in English by the author (DL). All focus groups took place in Mandarin or Cantonese; all excerpts cited from focus groups are similarly accompanied by an idiomatic English translation (also by DL).

## References

Blom, Jan-Petter and John J. Gumperz. (1972). Social meaning in linguistic structure: Code-switching in Norway. In John J. Gumperz and Dell Hymes (eds.) *Directions in sociolinguistics. The ethnography of communication*. New York: Holt, Rinehart and Winston, Inc. 407-434.

Chan, Mimi and Helen Kwok. (1982). *A study of lexical borrowing from English in Hong Kong Chinese*. Hong Kong: The University of Hong Kong.

Garfinkel, Harold. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Li, David C.S. and Elly C.Y. Tse. (2002). One day in the life of a 'purist'. *International Journal of Bilingualism 6*: 147-202.

# Presupposition of Jokes in Translation

## GE Lingling

Translation Department
The Chinese University of Hong Kong

## Abstract

Jokes, being rooted in a specific cultural and linguistic context, can pose significant problems to translation. This is one of the factors that explain why jokes often get lost in translation. This paper will discuss data collected from *Wei Cheng* (Fortress Besieged), a novel by Ch'ien Chung-shu (1947/1991), and its English translation by Jeanne Kelly and Nathan K. Mao (1979). We wish to explore how the target English reader might appreciate Chinese jokes from the translated data by examining how successfully the presuppositions in the Chinese source jokes are transferred through the English translation. In theory, we assume that failure to convey the original presuppositions is one of the reasons bringing about failure to understand jokes in cross-cultural communication. The actual tentative findings are: a) some presuppositions in the original Chinese text are sustained in the English translation, b) some are lost, and c) some are partially sustained and partially lost in terms of rendering. Presuppositions perceived in the data are of two types: a) the presupposition is given in the original text and b) the presupposition is presumed in the reader's knowledge.

## 1. Introduction

Joking is part of humor that is regarded as a universal human phenomenon, and funny situations, funny stories, even funny thoughts occur every day to virtually everybody (Raskin 1985:1). Joking, be it written or spoken, acts as a form of entertainment as well as a medium for communication. From the linguistic perspective of humor research, two models are largely accepted: the first model (the script-based semantic theory of humor; SSTH) addresses what makes a text humorous and a cognitive account mobilizing the notion of scripts and script opposition; the second model (the general theory of verbal humor; GTVH) deals with the issue of what makes a text humorous in a comprehensive way. Jokes, as the main study of these two models, involve the following parameters, or knowledge resources: language, situation, narrative strategy, target, logical mechanism and script opposition (Attardo and Raskin 1991:297 and Attardo 2002:176). Attardo further stressed that the production of a joke can be triggered by any knowledge resource, with the rest of them being filled in (Attardo and Raskin 1991:327). Presupposition, a complex phenomenon in language, embodies not only "background assumptions" (Levison 1983:180) but also "common ground" and "mutual knowledge" (Levison 1983:205). It includes linguistic presupposition and non-linguistic ones, both of which are made

every time we speak or write; every text is affected by them and contains linguistic and cultural/contextual triggers (Fawcett 1997:124-125). In translation, the translator's awareness of presupposition in the source text affects his/her translating strategies and translation itself. Therefore, Translators need to know not just what presuppositional information may be lacking in the target culture, but what presupposition exists in that culture which may proactively influence the translation (Fawcett 2001:122). Jokes are translatable, but "they do not always have the same impact" (Newmark 1988:107). How are presuppositions in the original jokes transferred through translation? Do presuppositions in Chinese jokes have the same effect in the English translation? It seems that fewer studies have been made in this domain so far.

Based on a sample of jokes in Wei Cheng (Fortress Besieged) by Ch'ien Chung-shu and its translation by Jeanne Kelly and Nathan K. Mao, the paper will discuss presupposition of jokes in translation in three categories that have been found from the data in section 2, and two types of presupposition observed in these samples will be explored in section 3.

## 2. Data and analysis

The novel *Wei Cheng* (Fortress Besieged) by Ch'ien Chung-shu is perceived as humorous, in which texts full of humor make the depiction of the plot very impressive to readers. Jeanne Keller and Nathan K. Mao's translation is the only one that renders the novel completely. A corpus of comparisons between source texts and target texts is utilized to make the analysis. As we have observed so far, there are three categories. We will discuss them in detail as follows.

## 2.1 Presupposed knowledge sustained in the target text.

According to script-based theory of humor developed by Raskin and Attardo (1984/1991), the mechanism of humor production involves conflicting knowledge representations. A 'script' is such an organized chunk of information about something, a cognitive structure internalized by the speaker which provides him/her with information on how the world is organized, including how one acts in it; in the broadest sense it is an object (real or imaginary), an event, an action, a quality etc. (Raskin 1985:199). Jokes are based on script opposition or incongruity. When the joke teller, the ST writer/the translator, he/she deliberately misleads the hearer, the ST reader/the TT reader into believing that that script is central to the processing of the text, only to reveal again deliberately at the end of the text that the script was in fact incompatible with the one introduced by the script-switch trigger (Attardo 2005: 5).

Jokes are elicited by incongruity. Presupposed knowledge "sustained" here means that the original presupposition is conveyed in the target text. Some examples are

taken from the data, after which P. indicates the page either in the original Chinese text or in the English translation.

(1) 便找到一家<u>門面還像樣的西館</u>。誰知道從冷盤到咖啡，沒有一樣東西可口：<u>上來的湯是涼的，冰淇淋倒是熱的；魚像海軍陸戰隊，已登陸了好幾天；肉像潛水艇士兵，會長時期伏在水裏；除醋外，麵包、牛肉、紅酒無一不酸</u>。P. 20

biàn zhǎo dào yī jiā <u>mén miàn hái xiàng yàng de xī guǎn。shuí zhī dào cóng l ěng pán dào kā fēi, méi yǒu yī yàng dōng xī kě kǒu: shàng lái de tāng shì liá ng de, bīng qi lín dǎo shì rè de; yú xiàng hǎi jūn lù zhàn duì, yǐ dēng lù le hǎ o jǐ tiān; ròu xiàng qián shuǐ tǐng shì bīng, huì cháng shí qī fú zài shuǐ l ǐ; chú cù wai, miàn bāo、niú ròu、hóng jiǔ wú yī bù suān</u>。

They then found <u>a Western-type restaurant that looked respectable enough from the outside; but as it turned out, there wasn't a single thing edible from the cold dishes to the coffee. The soup was cold, and the ice cream was warm. The fish was like Marine Corps. It apparently had already been on land for several days; the meat was like submarine sailors, having been submerged in water for a long time. Besides the vinegar, the bread, the butter, and the red wine were all sour.</u> P. 20

This joke is made on the restaurant by the narrator, which is created by the incongruity of the two scripts. One script is the respectable Western-type restaurant. The other is the terrible thing they had at this restaurant. The sentence "the Western-type restaurant looked respectable" presupposes that there was a restaurant with fine quality and good service, which leads the reader into believing that this is the case, but the following sentences convey meanings quite contradictory. They infer that everything in that restaurant was far from being respectable. Such presupposed knowledge is sustained in the target text which seems to achieve the same effect as the source text.

(2) <u>他仿美國人讀音，維妙維肖</u>，也許鼻音學得太過火了，<u>不像美國人，而像傷風塞鼻子的中國人</u>。P. 49

tā <u>fǎng měi guó rén dú yīn, wéi miào wéi xiào</u>, yě xǔ bí yīn xué dé tai guò huǒ le, <u>bù xiàng měi guó rén, ér xiàng shāng fēng sāi bí zǐ de zhōng guó rén</u>。

He <u>imitated the American accent down to the slightest inflection</u>, though maybe the nasal sound was a little overdone, <u>sounding more like a Chinese with a cold and a stuffy nose, rather than an American speaking.</u> P. 43

The narrator was making a joke on Mr. Zhang in the novel, whose speech was affected by sprinkling nasal sounds, due to his single-minded imitation of the American accent. This joke presupposes that Mr. Zhang's imitation of the American accent was ridiculously overdone. Two scripts are incompatible, for one is that his

imitation of the American accent is perfect and the other is that his nasal sound is like a Chinese with a cold and a stuffy nose. Reading the English translation, we perceive the same presuppositions in the target text.

However, the script-based theory of humor (Raskin and Attardo 1984/1991) cannot explain all verbal humor, for there are some jokes which are created without switching scripts; they are created by using semantic, phonetic and pragmatic techniques (Morreall 2004:393). For example:

(3) 中國是世界上最提倡科學的國家，沒有旁的國度肯這樣給科學家大官做的。外國科學<u>進步</u>，中國科學家<u>進爵</u>。P. 226

zhōng guó shì shì jiè shàng zuì tí chàng kē xué de guó jiā, méi yǒu páng de guó dù kěn zhè yàng gěi kē xué jiā dà guān zuò de。wai guó kē xué <u>jìn bù</u>, zhōng guó kē xué jiā <u>jìn jué</u>。

China is the greatest promoter of science of any country in the world; no other government body is so willing to offer high posts to scientists. As Western science <u>moves forward</u>, Chinese scientists <u>move upward</u>. P. 193

This is a joke on the Chinese government when the novel was written which presupposes that the government paid more attention to offering high posts to a few scientists rather than put more effort to promote science itself. The joke with this presupposition is preserved in the English translation.

## 2.2 Presupposed knowledge lost in the target text

Jokes are context-sensitive, so they are bound to be a problematic area of translation. The original author assumes his readers to be familiar with the presupposition. As pointed out by Fawcett (1998:120), it is frequent that the translators do not possess the awareness the author of the original assumed them to have, therefore presupposed knowledge is lost in the target text. Here "lost" refers to a loss of the original presupposition in the target text. For example:

(4) 大家庭裏做媳婦的女人<u>平時吃飯的肚子要小，受氣的肚子要大</u>；一有了胎，肚子真大了，那時<u>吃飯的肚子可以放大，受氣的肚子可以縮小</u>。P. 141

dà jiā tíng lǐ zuò xí fù de nǚ rén <u>píng shí chī fàn de dù zǐ yào xiǎo, shòu qì de dù  zǐ yào dà</u>; yī yǒu le tāi, dù zǐ zhēn dà le, nà shí <u>chī fàn de dù zǐ k ě yǐ fàng dà,  shòu qì  de dù zǐ kě yǐ suō xiǎo</u>。

In a large family daughters-in-law' <u>usually have to have small stomachs for food but big ones for frustration</u>. Once they become pregnant and their stomachs get big, they can <u>enlarge their stomachs for food and reduce their stomachs for frustration</u>. P. 121

Here is a joke given by the narrator, showing the low embarrassing position of daughters-in-law before 1949 when the People's Republic of China was founded. In the original text exist cultural presuppositions that women are inferior to men at that time, as daughters-in-law, they hold the lowest position in a large family and their position will be raised only when they give birth to a child, especially when they have a son. However, the presupposed knowledge is not obtained from the target text, for it was rendered through literal translation, hence the presupposition in the original text is "lost". The following examples belong to the same category.

(5) 丈夫是女人的職業，沒有丈夫就等於失業，所以該牢牢捧住這飯碗。P. 54

zhàng fū shì nǚ rén de zhí yè, méi yǒu zhàng fū jiù děng yú shī yè, suǒ yǐ g
āi láo láo pěng zhù zhè fàn wǎn。

Husbands are women's careers. Not having a husband is like being unemployed, so she has to hold tightly to her "rice bowl". P. 49

This joke is also made by the narrator on women's status before 1949, which presupposes that women were not independent at that time, and their husbands were their careers, even their whole lives. Without husbands, they could not live on their own. However, the English readers do not share the same cultural background knowledge, so this presupposed knowledge in the original Chinese joke is not possessed by target readers.

(6) 母親笑說：“是要出洋的，學得這樣周到，女人用的東西都會買了。” P. 38

mǔ qīn xiào shuō: "shì yào chū yáng de, xué dé zhè yàng zhōu dào, nǚ rén yò
ng de dōng xī dōu huì mǎi le。"

His mother said with a smile, "It takes going abroad to learn such thoughtfulness. He even knows how to buy things for women." P. 34

This is a joke made on the character Fang Hung-chien. It presupposes that men should not know how to buy things for women in Chinese traditional families. His mother made fun on him that he had learned a lot abroad so that he knew how to buy things for women, but the presupposed knowledge is lost for target readers.

## 2.3 Presupposed knowledge partially sustained and partially lost in the target text

What does presupposed knowledge partially sustained and partially lost in the target text refer to? In this context, on the one hand, the translator perceives some presuppositions embedded in jokes and conveys the observed presuppositions to the target reader, on the other hand, some presuppositions, whether perceived or not, are

not transferred through translation. Therefore, presupposed knowledge is partially sustained and partially lost for the target reader.

(7) 鮑小姐打她一下道："你！蘇東坡的妹妹，才女！"－－"蘇小妹"是同船男學生爲蘇小姐起的個號。"東坡"兩個字給鮑小姐南洋口音念得好像法國話裏的"墳墓"（tombeau）。P. 6

bào xiǎo jiě dǎ tā yī xià dào: "nǐ! sū dōng pō de mèi mèi, cái nǚ!"－－"sū xiǎo m èi" shì tóng chuán nán xué shēng wéi sū xiǎo jiě qǐ de gè hào. "dōng pō "liǎng gè zì gěi bào xiǎo jiě nán yáng kǒu yīn niàn dé hǎo xiàng fǎ guó huà lǐ de "fén mù " （tombeau）。

Miss Pao gave her a cuff, saying, "You! <u>Su Tung-p'o's little sister</u>[5], the girl genius!" "Su Hsiao-Mei" (Su's little sister) was the nickname the men students on broad had given Miss Sue. The words, "Tung-p'o" when pronounced by Miss Pao in her South Seas accent sounded like tombeau, the French word for tomb. P. 7

This joke is made in a conversation between Miss Pao and Miss Sue. In the preceding text, Miss Sue laughed at Miss Pao. Then Miss Pao called the nickname of Miss Sue to make a joke in return rather than praise her for her talent. In the Chinese culture, Su Tung-p'o is a talented poet and his little sister, Su Hsiao-Mei, is viewed as talented as him. This nickname was given to Miss Su due to the fact that she shared the same family name with Su Tung-p'o's little sister and she was talented as she had just got her Ph.D. abroad. Part of the presupposed knowledge is given by the translator in a note as below:

*Note 5: Su Tung-p'o (1036-1101) a celebrated Chinese poet, essayist, painter, and calligrapher of the Sung dynasty. His sister, Su Hsiao-mei, probably a legendary figure, is reputed to have been equally talented. P. 364*

Though the note is considered to be so cumbersome in a joke and simply spoils its humorous effect, one part of the presupposition, that Su Hsiao-Mei is viewed as talented as her brother, is preserved in the target text. However, the other part of the presupposition, that Miss Su shares the same family name with Su Tung-p'o's little sister, is lost for the target reader unless he/she shares this presupposed knowledge.

(8) 鮑小姐睡了一天多才起床，雖跟方鴻漸在一起玩，<u>不像以前那樣脫略形骸</u>，也許因爲不日到香港，先得<u>把身心收拾整潔</u>，作爲見未婚夫的準備。P. 23

bào xiǎo jiě shuì le yī tiān duō cái qǐ chuáng, suī gēn fāng hóng jiàn zài yī qǐ wán, <u>bù xiàng yǐ qián nà yàng tuō lüè xíng hái,</u> yě xǔ yīn wéi bù rì dào xiāng gǎng, xiān dé <u>bǎ shēn xīn shōu shí zhěng jié,</u> zuò wéi jiàn wèi hūn fū de zhǔn bèi 。

Miss Pao stayed in bed for a day or two; then she finally got up. She still toyed with Fang <u>but not freely as before.</u> Perhaps because they would be reaching Hong Kong in

a few days, she had <u>to cleanse her mind and body</u> in preparation for meeting her fiancé. P. 22

Miss Pao dressed herself and behaved too freely till she realized that she would meet his fiancé soon. The joke is made by the narrator on Miss Pao for her changing dressing and behavior, which presupposes that Miss Pao was not wearing or acting in the same way as she did in the past. In the previous statement of the novel, Miss Pao's exposed body was considered to be an insult to the body politic of the Chinese society. The presupposition that she had to cleanse her mind and body is preserved through literal translation, but the presupposition that Miss Pao's exposed body constituted an insult to the body politic of the Chinese society is lost in the semantic translation "but not freely as before".

## 3. Two types of presupposition

Jokes are created not only by switching scripts but also by using semantic and pragmatic techniques. The quoted examples illustrate how presupposition plays a vital role in comprehending jokes to some degree. From the above data analysis, we have observed two types of presupposition. The first type is given by the author in the original text as shown in the examples (1)-(3), where the texts denoting the presuppositions are:

(1)' ......門面還像樣的西館 ......
......mén miàn hái xiàng yàng de xī guǎn......
...a Western-type restaurant that looked respectable enough from the outside...

(2)' ......鼻音學得太過火了......
......bí yīn xué dé tai guò huǒ le......
...the nasal sound was a little overdone...

(3)' ......沒有旁的國度肯這樣給科學家大官做的......
......méi yǒu páng de guó dù kěn zhè yàng gěi kē xué jiā dà guān zuò de......
...no other government body is so willing to offer high posts to scientists...

As we see, the presuppositions are stated clearly in the original text. The reader is assumed to share the author's awareness of what is presupposed in the text. The presuppositions are elicited by the actual language used in the text and they are unlikely to be affected in translation (Fawcett 2001:115/118). In the English translation, we observe that presuppositions in (1)', (2)', (3)' are indeed translated. It is obvious that both source and target texts share the presuppositions in these instances. As a result, we assume that the target reader will be able to understand the jokes as well as the source reader will.

Now consider the second type of presupposition, which is presumed in the reader's knowledge. In other words, the author presumes that the reader has the presupposed knowledge in order to understand the joke. This type is observed in the examples (4), (5), (6) and (7), where the author presumes that the reader of the source text has the knowledge to obtain these cultural presuppositions.

The cultural presupposition is occasionally but not commonly shared in both source texts and target texts, which is focused on hereafter. In translation, the point is that it is unnecessary to supply information for which there is no demand, since translators believe it to be shared, and if it is assumed that the cultural presupposition is not shared which poses acute problems, translators need to take "a delicate balancing act" (Fawcett 2001:121), either leaving the presuppositions in the dark by not supplying what is required in the source text, or finding out "the optimal translation technique, to pass on the information with a minimum of disruption" (Fawcett 2001:121), if translators are able to perceive the presuppositions.

Jokes in the examples (4), (5) and (6) are triggered by the cultural presuppositions. As mentioned above, in old China, women are regarded as inferior to men in that women consider husbands as their whole life, women must do all the housework for husbands and women can only raise their status a bit when they are pregnant or give birth to a son. The author presumes that the reader has such cultural background and it is not doubtful that most Chinese readers can appreciate the joke. While in the English rendering, as we have discussed, presuppositions in (4), (5) and (6) are not sustained in the target text, although translators must, in any event, possess the knowledge which the author presupposes his readers to have, even though a lot of translation is done without such knowledge (Fawcett 2001:121). We are not sure whether the translators had the presupposed knowledge or not, but the real situation is that the translators didn't transfer the intended presupposition to the English culture in the version and this cultural presupposition is less likely for the target reader to share as the source reader. Therefore, we can deduce that the target reader will be unable to understand the jokes because the cultural presupposition is lacking in the target text and translators use omission rather than rewriting or annotating in the translation.

An interesting instance of this type can be seen in the example (7), where part of the presupposition is translated and part of it is not transferred in the version. In (7), most Chinese do not have to be told who Su Tung-p'o is and the author presumes that the source reader has the presupposed knowledge. The presupposition consists of two parts. On the one hand, in the Chinese culture, Su Tung-p'o's little sister is viewed as talented as her brother, on the other hand, Chinese people tend to give a person a nickname based on the same family names. The translators were aware that the first part of the presupposition was lacking in the target culture. It is for that reason that translators gave a note explaining one part of the presupposition, even though there is a considerable loss in understanding the joke because the target reader is frustrated to read an annotation. However, the other part of the presupposition is not rendered. If

the target reader does not share the cultural knowledge that a nickname is given based on the same family name, omission of this part of the presupposition in the translation may affect the target reader's comprehension of the joke.

## 4. Discussion and concluding remarks

Presupposition of jokes in translation affects the target reader's appreciation of humor in cross-cultural communication. The data discussed above provide a perspective from which the relationship between authors, translators and readers can be seen. Authors have their assumptions about the world, the society and culture in which they live and especially about readers of the source text. However, in rendering, translators themselves, though their translation is based on the original text, bring a set of assumptions that may in some cases coincide with those of the authors while in others there will be great differences. This interplay of coincidences and differences may to some extent explain why some presuppositions are sustained, some are lost and some are partially sustained and partially lost.

In the present study, the samples are collected randomly and the analyses are preliminary as well. Besides the three categories mentioned, is it possible that the presupposed knowledge in the source text is shifted, half shifted and half sustained, or half shifted and half lost in the target text? In our data, we found only two types of presupposition so far. In translation the presupposition the author gives in the original text is either conveyed or lost in the target text, and the presupposition that the author presumes that the reader should obtain is mostly lost. Is there any other circumstance in rendering we have not investigated yet? In addition to the two types, are there any other types of presupposition existing? If there are, what happens in the version? Are the presuppositions totally sustained, lost or shifted? This study has simply focused on presuppositions in a sample of jokes collected from one version that is the only one that has been translated from Chinese to English up to the present. Some variables may affect the findings of the research. Different versions by different translators may bring about considerable different interpretations of presuppositions in jokes. Moreover, data analysis mentioned above is more subjective rather than objective, for the comparison is only based on my own viewpoint without certain theoretical framework so far. Enlarging the scope of samples and analyzing data based on a more acceptable criterion will lead to a better understanding of this issue. This will be the subject of a subsequent research project.

# References

Attardo, S. and Raskin, V. (1991). Script theory revisited: Joke similarity and joke representation model. *Humor: International Journal of Humor Research* 4(3-4), 293-347.

Attardo, S. (1994). *Linguistic Theories of Humour.* Berlin: Mouton de Gruyter.

Attardo, S. (2002). Translation and humour: An approach based on the general theory of verbal humour (GTVH). *The Translator* 8-2, 173-194.

Attardo, S. (2005). Homor. In J. Verchueren, J-O. Östman and J. Blommaert (eds.) *Handbook of Pragmatics.* Amsterdam: John Benjamins.

Ch'ien Chung-shu. (1991). *Wei Cheng.* Chengdu: Sichuan Culture and Art Press.

Fawcett, P. (1997). *Translation and Language: Linguistic Theories Explained.* Manchester: St. Jerome Publishing.

Fawcett, P. (2001). Presupposition and translation. In Hickey, L. (ed.) *The Pragmatics of Translation. Shanghai:* Shanghai Foreign Language Education Press, 114-123.

Kelly, J. & Mao, N.K. (trans.) (1979). *Fortress Besieged.* Bloomington and London: Indiana University Press.

Levison, S. C. (1983). *Pragmatics.* Cambridge: Cambridge University Press.

Morreall, J. (2004). Verbal humor without switching script and without non-bona fide communication. *Humor* 17-4 393-400.

Newmark, P. (1988). *A Textbook of Translation.* New York: Prentice Hall.

Raskin, V. (1985). *Semantic Mechanisms of Humor.* Dordrecht-Boston-Lancaster: D. Reidel.

# A Look at the Phonological System of Cantonese in the Mid-19<sup>th</sup> Century, with a Focus on the Sibilants: A Quantitative Analysis of the Cantonese Romanisation Scheme Used by the British

WONG Tak-Sum

Division of Humanities

The Hong Kong University of Science and Technology

## Abstract

The Cantonese romanisation scheme used under the 145-year of British rule in Hong Kong shows most of the tokens in the romanisation scheme pronounced in Received Pronunciation match with the pronunciations of present-day Cantonese. However, in some cases there are deviations. This phenomenon essentially reflects the historical pronunciations of Cantonese in the mid-19<sup>th</sup> century.

In this research, the names of the streets with their transliterations were considered. The unsuitable ones were first systematically filtered out and the rest are used in the database. Results show that almost for each of the onsets and rhymes in modern standard Hong Kong Cantonese (HKC), at least one but at most six transliterations are used. By using a quantitative approach with logical inference and comparison with other romanisation schemes, the phonetic values represented by transliterations, including those no longer exist in contemporary HKC, were identified without relying much on previous studies in Cantonese historical phonology. This shows that the phonological system of HKC has been undergoing simplification. The onset sets [tʃ]-[tʃʰ]-[ʃ] and [ts]-[tsʰ]-[s] contrasted in the mid-19<sup>th</sup> century Cantonese have merged to form the only set [ts]-[tsʰ]-[s].

## 1. Introduction

Before the return of Hong Kong to China in 1997, Hong Kong had been under the British colonial rule for around 150 years. Up until 1974, the only official language used by the Hong Kong government was English. Therefore, a Cantonese romanisation scheme was required to be used for the transliteration of the Chinese names for people, streets, roads, etc. in official documents.

Due to the requirement of the transliteration of names in official documents, the British government adopted a romanisation scheme at the very beginning of their governance. It can be seen that most of the tokens in the romanisation scheme pronounced in Received Pronunciation (RP, *henceforth*) match with the pronunciations of present-day Cantonese. However, in some cases there are deviations. This phenomenon essentially reflects the historical pronunciations of Cantonese in the

mid-19th century.

Due to the nature of the Chinese script and the tradition of despising dialectal literature, there is always a shortage of materials for studying historical phonology of Chinese, not to mention the various southern dialects. The research in historical phonology of the southern dialects mainly relies on the materials left by missionaries and textbooks for dialectal teaching purposes.

## 1.1 Related work

Previous researches concerning the geographical names in Hong Kong mainly focus on cultural aspects. These include Lun's work (1993), which discusses the use of Chinese characters of street names in Hong Kong and Yiu's (1998), which has explored thoroughly the cultural aspect of geographical names in Hong Kong. Wong (2000) has also researched on the use of Chinese characters of place names in the Yue dialect. Recently, Lee and Kataoka (2005) have reported on the impact of the romanised Cantonese on place and personal names in Hong Kong.

Previous works in Cantonese historical phonology are affluent though not abundant. These include Chen and Newman's (1984), which describes the phonological changes from Middle Chinese to modern Cantonese; and Li (1997), which illustrates the development of vowels in the last few centuries. Some researched on the recent changes in Cantonese over the last hundred years, such as Chen and Mo (1989), Li (1989) and Li (1997). Others investigated the historical phonology of Cantonese in early Qing by making use of the dialectal

rhyming-dictionary *Fenyuncuoyao* 分韻撮要 published in the mid-Qing dynasty as

in Lau and Cheung (2001), Peng (1989) and Peng (1992). Cheung (2003) also made use of the Cantonese textbooks left by missionaries to examine changes in Hong Kong Cantonese (HKC, *henceforth*) from the late 19th century up to the present day. More recently Hashimoto (2005) has tried to reconstruct the phonological system of proto-Yue by comparing pronunciations of various modern Yue dialects. Last but not least, Bauer (1979), Bauer (1983), Bauer (1986) and Zee (1990) research on changes and variations in recent HKC. A microscopic phonological analysis in the romanisation of the geographic names with sibilant-initial in Hong Kong is briefly mentioned in Zee (1990) but a systematic and quantitative investigation is lacking.

## 1.2 A micro-history of the Romanisation Scheme

As mentioned previously, English was the only official language used by the colonial government until very late in their governance; the Cantonese romanisation

schemes commonly used in the late 19<sup>th</sup> century include the ones derived by Morrison, Ernst Johann Eitel, Dyer, Ball, Bernhard F. Meyer and Theodore F. Wemepe; Williams and Bridgman; Chalmers (Lee and Kataoka 2005). Did the colonial government derive a system for itself; or did it adopt one of the schemes mentioned above?

According to Lee and Kataoka (2005), the official gazetteer published by the colonial government in 1960 refers to the scheme as using the "Eitel/Dyer-Ball" system. They finally concluded that the scheme was generally based on the Eitel's system with some revisions adopted from the Ball's system and Meyer and Wemepe's system. The research conducted by Zee (1990) and Lee & Kataoka (2005) point out that the romanisation scheme represents the historical sound system of Cantonese.

### 1.3 This paper

As mentioned in section 0, a systematic and quantitative investigation on the phonology of the romanisation scheme is lacking. Therefore, this research tried to fill in this gap by conducting a comprehensive survey of the romanisation scheme.

In this paper, the Cantonese romanisation scheme used by the government will be constructed by generalising the transliteration of the street names listed in the book *Hong Kong Guide 2005* published by the Hong Kong Government. The generalised scheme is then compared with the pronunciations of contemporary HKC. After that, the differences will be compared with the previous studies of Cantonese historical phonology, and historical pronunciations reflected will be explained by means of Chinese historical phonology.

This paper is organised as follows. In section 0, a research methodology is outlined. In section 3, the results of the statistical analysis and the scheme constructed are given with an account of the analysis. Conclusions and future work are given in section 4.

### 2. Research Methodology – Setting up the Database

In this research, the names of streets with their transliterations in the book *Hong Kong Guide 2005* (pp. 338 – 349) published by the Hong Kong Government are considered. However, not all of them are suitable for the romanisation analysis in this research. In this section, the criteria of filtering out the unsuitable names and the setting up of the database are described.

Generally speaking, most of the English names for the streets in Hong Kong are transliterations of Chinese character-by-character, that is, syllable-by-syllable. However, in some cases, especially those established at the beginning of the colonial

governance, the Chinese names are translation of the meanings of the English names. For instance, *Queen's Road*, which is one of the main roads on Hong Kong Island, has the Chinese name "皇后大道" Huanghou Da Dao. The English name is obviously not a transliteration of the Chinese one. Such road names are not suitable for analysis and have been filtered out. The easiest way to identify this kind of names is by the number of words and characters in the English and Chinese names respectively. When the number of characters of a street in Chinese is found to be not equivalent to the number of words in the English name, it is most likely that the English name is not a transliteration version of the Chinese one. For instance, with the name *Queen's Road*, the number of words (two words) is not equivalent to the number of characters in its Chinese name "皇后大道" (four characters). Thus most of the semantically translated names are filtered out.

However, even though the origin of the English name of a street is from Chinese, it is always the case that the proper name is transliterated by pronunciation while the general name such as *Street* (街) and *Road* (道) is translated by meaning. In this way, if it is found that the words in the English name of a street is a word in the lexicon of English and its meaning matches the corresponding morpheme(s) of the Chinese name, that word and the corresponding Chinese morpheme(s) are discarded; and only the remaining parts are used. For example, the Chinese name of *Tat Chee Avenue* is "達之路". The word *Avenue* is an English word and is matched with the morpheme "路" in the Chinese name. Therefore, only the part *Tat Chee* / "達之" is analysed and *Avenue* / "路" is not analysed.

After collecting all the suitable tokens in the book *Hong Kong Guide 2005*, variations for the transliterations of each of the initials and rhymes in contemporary HKC are listed. The occurrences for each of the variations are counted for analysis and variations are to be accounted for by means of Chinese historical phonology and sociolinguistics.

## 3. Results and Discussion

### 3.1 Results of statistical analyses

After the filter process, 6153 tokens were collected, as shown in Table 1:

## Table 1: Onsets of the Transliterations

| Standard Cantonese Onsets | Young Cantonese Onsets | Romanisation Onsets | Count | Percentage |
|---|---|---|---|---|
| [ø] | ø ~ [ŋ] | (null) | 195 | 100.00% |
| [f] | [f] | f | 364 | 99.45% |
| | | w | 2 | 0.55% |
| [h] | [h] | h | 543 | 100.00% |
| [j] | [j] | (null) | 12 | 2.32% |
| | | y | 506 | 97.68% |
| [k] | [k] | ch | 1 | 0.29% |
| | | k | 343 | 99.71% |
| [kʷ] | [kʷ] | kw | 101 | 100.00% |
| [kʰ] | [kʰ] | k | 72 | 100.00% |
| [kʷʰ] | [kʷʰ] | kw | 57 | 100.00% |
| [l] | [l] | l | 486 | 100.00% |
| [m] | [m] | m | 292 | 100.00% |
| [n] | [l] | l | 2 | 2.67% |
| | | n | 73 | 97.33% |
| [ŋ] | [ŋ] ~ ø | ng | 47 | 100.00% |
| [p] | [p] | b | 4 | 2.06% |
| | | p | 190 | 97.94% |
| [pʰ] | [pʰ] | p | 147 | 100.00% |
| [s] | [s] | ch | 2 | 0.30% |
| | | hs | 1 | 0.15% |
| | | s | 243 | 36.11% |
| | | sh | 425 | 63.15% |
| | | t | 1 | 0.15% |
| | | ts | 1 | 0.15% |
| [t] | [t] | d | 14 | 4.03% |
| | | t | 333 | 95.97% |
| [tʰ] | [tʰ] | t | 332 | 100.00% |
| [ts] | [ts] | ch | 223 | 78.25% |
| | | ts | 62 | 21.75% |
| [tsʰ] | [tsʰ] | ch | 299 | 62.29% |
| | | ts | 181 | 37.71% |
| [w] | [w] | (null) | 1 | 0.17% |
| | | h | 1 | 0.17% |
| | | m | 2 | 0.33% |
| | | w | 595 | 99.33% |
| **Total** | | | 6153 | |

## Table 2: Rhymes of the Transliterations

| Standard Cantonese Rhymes | Some Young Rhymes | Romanisation Rhymes | Count | Percentage |
|---|---|---|---|---|
| [a] | [a] | a | 207 | 78.41% |
| | | ah | 55 | 20.83% |
| | | ak | 1 | 0.38% |
| | | ar | 1 | 0.38% |
| [ai] | [ai] | ai | 194 | 100.00% |
| [akˀ] | [atˀ] | ak | 45 | 100.00% |
| [am] | [am] | am | 43 | 100.00% |
| [an] | [an] | an | 190 | 99.48% |
| | | en | 1 | 0.52% |
| [aŋ] | [an] | ang | 44 | 100.00% |
| [apˀ] | [apˀ] | ap | 13 | 100.00% |
| [atˀ] | [atˀ] | at | 66 | 100.00% |
| [au] | [au] | au | 23 | 92.00% |
| | | ow | 2 | 8.00% |
| [ɐi] | [ɐi] | ai | 315 | 98.75% |
| | | ei | 3 | 0.94% |
| | | i | 1 | 0.31% |
| [ɐkˀ] | [ɐtˀ] | ak | 86 | 97.73% |
| | | ei | 1 | 1.14% |
| | | uck | 1 | 1.14% |
| [ɐm] | [ɐm] | am | 131 | 82.91% |
| | | an | 2 | 1.27% |
| | | en | 3 | 1.90% |
| | | om | 10 | 6.33% |
| | | um | 12 | 7.59% |
| [ɐn] | [ɐn] | an | 298 | 94.30% |
| | | en | 1 | 0.32% |
| | | in | 1 | 0.32% |
| | | un | 16 | 5.06% |
| [ɐŋ] | [ɐn] | ang | 74 | 97.37% |
| | | eng | 1 | 1.32% |
| | | ung | 1 | 1.32% |
| [ɐpˀ] | [ɐpˀ] | ap | 3 | 18.75% |
| | | op | 13 | 81.25% |
| [ɐtˀ] | [ɐtˀ] | at | 33 | 91.67% |
| | | ut | 3 | 8.33% |
| [au] | [au] | au | 177 | 95.16% |
| | | iu | 1 | 0.54% |
| | | ou | 2 | 1.08% |
| | | ow | 6 | 3.23% |
| [ɛ] | [ɛ] | e | 15 | 100.00% |

| | | | | |
|---|---|---|---|---|
| [ei] | [ei] | ee | 35 | 20.59% |
| | | ei | 104 | 61.18% |
| | | i | 27 | 15.88% |
| | | ze | 4 | 2.35% |
| [ek˺] | [et˺] | ek | 57 | 95.00% |
| | | ik | 3 | 5.00% |
| [eŋ] | [en] | eng | 9 | 100.00% |
| [ey] | [ey] | u | 1 | 1.05% |
| | | ui | 94 | 98.95% |
| [en] | [en] | un | 109 | 100.00% |
| [et˺] | [et˺] | / | 0 | / |
| [i] | [i] | e | 3 | 2.14% |
| | | ee | 32 | 22.86% |
| | | i | 85 | 60.71% |
| | | z | 15 | 10.71% |
| | | ze | 5 | 3.57% |
| [ɪk˺] | [ɪt˺] | ek | 6 | 13.95% |
| | | ick | 9 | 20.93% |
| | | ik | 27 | 62.79% |
| | | itt | 1 | 2.33% |
| [im] | [im] | im | 11 | 100.00% |
| [in] | [in] | en | 1 | 0.41% |
| | | ien | 1 | 0.41% |
| | | in | 241 | 99.18% |
| [ɪŋ] | [m] | eng | 17 | 2.54% |
| | | ing | 652 | 97.46% |
| [ip˺] | [ip˺] | ip | 70 | 100.00% |
| [it˺] | [it˺] | i | 1 | 12.50% |
| | | it | 7 | 87.50% |
| [iu] | [iu] | iu | 89 | 100.00% |
| [m̩] | [m̩] | / | 0 | / |
| [ŋ̩] | [m̩] | ng | 7 | 100.00% |
| [ɔ] | [ɔ] | o | 133 | 95.68% |
| | | oh | 2 | 1.44% |
| | | oo | 4 | 2.88% |
| [œ] | [œ] | / | 0 | / |
| [œk˺] | [œt˺] | euk | 4 | 80.00% |
| | | ork | 1 | 20.00% |
| [œŋ] | [œn] | ang | 1 | 0.43% |

| | | | | |
|---|---|---|---|---|
| | | eong | 38 | 16.17% |
| | | eung | 195 | 82.98% |
| | | ung | 1 | 0.43% |
| [ɔi] | [ɔi] | ai | 1 | 0.88% |
| | | oi | 112 | 99.12% |
| [ɔk˺] | [ɔt˺] | ok | 144 | 100.00% |
| [ɔn] | [ɔn] | on | 141 | 100.00% |
| [ɔŋ] | [ɔn] | eong | 1 | 0.36% |
| | | iang | 4 | 1.43% |
| | | ong | 275 | 98.21% |
| [ɔt˺] | [ɔt˺] | / | 0 | / |
| [ou] | [ou] | ao | 1 | 0.54% |
| | | o | 181 | 97.84% |
| | | oa | 1 | 0.54% |
| | | oo | 1 | 0.54% |
| | | u | 1 | 0.54% |
| [u] | [u] | oo | 9 | 8.91% |
| | | u | 92 | 91.09% |
| [ui] | [ui] | ay | 1 | 1.82% |
| | | ui | 54 | 98.18% |
| [uk˺] | [ut˺] | ok | 1 | 0.65% |
| | | ook | 7 | 4.58% |
| | | uk | 145 | 94.77% |
| [un] | [un] | an | 1 | 1.79% |
| | | oon | 6 | 10.71% |
| | | un | 49 | 87.50% |
| [uŋ] | [uŋ] | ong | 1 | 0.23% |
| | | oong | 1 | 0.23% |
| | | ung | 438 | 99.55% |
| [ut˺] | [ut˺] | ood | 1 | 100.00% |
| [y] | [y] | u | 34 | 50.75% |
| | | ue | 33 | 49.25% |
| [yn] | [yn] | uen | 196 | 93.33% |
| | | un | 12 | 5.71% |
| | | ung | 2 | 0.95% |
| [yt˺] | [yt˺] | uet | 10 | 100.00% |
| Total | | | 6153 | |

**Table 1** shows the onsets of the transliterations of the street names while Table 2 lists the rhymes. In each of the tables, column one lists the onsets/rhymes in modern standard HKC. Column three shows the possible symbols used for the transliteration of the Chinese characters of the onsets/rhymes in column one while column four shows the frequency count. Column five shows the percentage of frequency of that transliteration used for the romanisation among the same onset/rhyme in

contemporary HKC. For instance for the onset [p], four tokens are transliterated as *b*, which constitutes 2.06% while 190 tokens are transliterated as *p*, which constitutes 97.94% of the total count; while for the onset [k$^w$], the only way of transliteration is *kw*, which constitutes 100%.

In addition, in **Table 1**, column two shows the corresponding onsets used by the young and middle-aged speakers while in Table 2, column two shows the corresponding rhymes used by some young speakers with an exception of the rhyme [m̩], which is also adopted by most middle-age speakers to replace [ŋ̍]. These pronunciations are suggested in Cheung (2003). It should be noted that in the second row of **Table 1**, the symbol "(∅)" means zero-onset while (null) means no symbols are used for the transliteration of this onset.

It can be seen that among the onsets and rhymes in modern standard HKC, some are represented by only one transliteration such as the zero-onset and the voiceless unaspirated labio-velar stop onset [k$^w$] but most of them are represented by at least two transliterations, for example, the voiceless unaspirated bilabial stop onset [p] and the rhyme [a].

## 3.2 Some observations

Table 1 and Table 2 show that the number of transliterations representing one onset/rhyme is quite different. For some onsets/rhymes such as the zero-initial and the rhyme [ai], only one transliteration is used to represent each of them. For some rhymes such as [œ] and [ɐt̚], no transliterations are found. The maximum number of transliterations representing one onset/rhyme (the onset [s]) is six, including *ch*, *hs*, *s*, *sh*, *t*, and *ts*.

It is worth noting that for some of the transliterations representing one contemporary HKC onset or rhyme, the frequency count and percentage is quite low, *e.g.* only 2.07% (4 counts) of the onset [p] is represented by the transliteration *b* and only 0.38% (1 count) of the rhyme [a] is represented by the transliteration *ak*. On the other hand, the frequency count and percentage for some transliterations is quite high, e.g. the transliteration *ch* constitutes 62.29% (299 counts) while the transliteration *ts* constitutes 37.31% (181 counts) among all the transliterations representing the onset [ts$^h$]. It is likely that the transliterations in the list of low frequency of usage are exceptions, perhaps due to mistakes or some unspecified reasons. The transliterations with high frequency of usage are systematic and consistent transliterations in denoting the onset or rhyme system.

More interestingly, we find that some cases have more than one transliteration

for the representation of the same onset and rhyme, e.g. 36.11% of the onset [s] are represented by the transliteration *s* while 63.15% are represented by the transliteration *sh*. In these cases, both transliterations are consistently used in high percentages to represent the same onset or rhyme.

These tokens with a high percentage of usage will be discussed next.

## 3.3 Rhymes not transliterated

The rhymes in modern standard Cantonese with no transliterations represented are: [m̩], [ɵt˙], [œ], and [ɔt˙]. These rhymes probably existed in HKC at the time when the romanisation scheme was designed since they are also found in early-*Qing* Cantonese reconstructed by Lau and Cheung (2001). However, as a matter of fact, these rhymes are not used frequently in contemporary Cantonese and they are represented by a few characters in common use. The following lists the characters used commonly with these rhymes found in *A Chinese Talking Syllabary of the Cantonese Dialect: an Electronic Repository*. It can be seen that the number of commonly used characters for each rhyme is at most five, so it is not surprising that no street names are rhymed using these characters.

[m̩]: 唔
[œ]: 朵, 唾, 鋸, 靴
[ɵt˙]: 訥, 率, 律, 栗
[ɔt˙]: 割, 葛, 喝, 渴, 褐

On the other hand, all onsets are represented by at least one transliteration. This is probably due to the small number of onsets in modern standard Cantonese, *i.e.* twenty.

## 3.4 Onsets/Rhymes represented by only one transliteration

The onsets and rhymes with only one transliteration are as follows:

Onsets: ∅ *(null)*, [kʷ] *kw*, [h] *h*, [kʰ] *k*, [kʷʰ] *kw*, [l] *l*, [m] *m*, [ŋ] *ng*, [pʰ] *p*, [tʰ] *t*
Rhymes: [ai] *ai*, [ak˙] *ak*, [am] *am*, [aŋ] *ang*, [ap˙] *ap*, [at˙] *at*, [ɛ] *e*, [eŋ] *eng*, [ɵn] *un*,
    [im] *im*, [ip˙] *ip*, [iu] *iu*, [ŋ̩] *ng*, [ɔk˙] *ok*, [ɔn] *on*, [ut˙] *ood*, [yt˙] *uet*

The one-to-one mapping shows that one English spelling pattern is used to represent the above onsets/rhymes. Counter examples include the rhyme [i], which can be spelt as both *ee* and *i*; and also [u], which can be spelt as both *oo* and *u* in contemporary English.

The common transliteration *kw* between [kʷ] and [kʷʰ] reflects that aspiration is not a distinctive feature in English.

## 3.5 Frequency counts of transliteration

Below are the transliterations with frequency counts in percentage. Those below 10% are excluded (see section 3.2 above).

Table 3: Transliterations of Onsets
in High Frequency of Usage

| Modern Cantonese Onsets | Young Cantonese Onsets | Romanisation Onsets | Count | Percentage |
|---|---|---|---|---|
| [p] | [p] | p | 190 | 97.94% |
| [tʃʰ] | [tʃʰ] | ch | 299 | 62.29% |
| [tʃʰ] | [tʃʰ] | ts | 181 | 37.71% |
| [t] | [t] | t | 332 | 95.68% |
| [f] | [f] | f | 364 | 99.45% |
| [k] | [k] | k | 343 | 99.71% |
| [j] | [j] | y | 506 | 97.68% |
| [n] | [l] | n | 73 | 97.33% |
| [ʃ] | [ʃ] | s | 243 | 36.11% |
| [ʃ] | [ʃ] | sh | 425 | 63.15% |
| [w] | [w] | w | 595 | 99.33% |
| [tʃ] | [tʃ] | ch | 223 | 78.25% |

| [ɐt̚] | [ɐt̚] | at | 33 | 91.67% |
|---|---|---|---|---|
| [ɐu] | [ɐu] | au | 177 | 95.16% |
| [ei] | [ei] | ee | 35 | 20.59% |
| [ei] | [ei] | ei | 104 | 61.18% |
| [ei] | [ei] | i | 27 | 15.88% |
| [ek̚] | [et̚] | ek | 57 | 95.00% |
| [ey] | [ey] | ui | 94 | 98.95% |
| [i] | [i] | ee | 32 | 22.86% |
| [i] | [i] | i | 85 | 60.71% |
| [i] | [i] | z | 15 | 10.71% |
| [ɪk̚] | [ɪt̚] | ik | 27 | 64.29% |
| [in] | [in] | in | 241 | 99.18% |
| [ɪŋ] | [ɪn] | ing | 652 | 97.46% |
| [ɔ] | [ɔ] | o | 133 | 95.68% |
| [œŋ] | [œn] | eong | 38 | 16.17% |

| [tʃ] | [tʃ] | ts | 62 | 21.75% |
|---|---|---|---|---|

Table 4: Transliterations of Rhymes
in High Frequency of Usage

| Modern Cantonese Rhymes | Some Young Rhymes | Romanisation Rhymes | Count | Percentage |
|---|---|---|---|---|
| [a] | [a] | a | 207 | 78.41% |
| [a] | [a] | ah | 55 | 20.83% |
| [an] | [an] | an | 190 | 99.48% |
| [au] | [au] | au | 23 | 92.00% |
| [ɐi] | [ɐi] | ai | 315 | 98.75% |
| [ɐk̚] | [ɐt̚] | ak | 86 | 97.73% |
| [ɐm] | [ɐm] | am | 131 | 82.91% |
| [ɐn] | [ɐn] | an | 298 | 94.30% |
| [ɐŋ] | [ɐn] | ang | 74 | 97.37% |
| [ɐp̚] | [ɐp̚] | op | 13 | 81.25% |
| [œŋ] | [œn] | eung | 195 | 82.98% |
| [ɔi] | [ɔi] | oi | 112 | 99.12% |
| [ɔŋ] | [ɔn] | ong | 275 | 98.21% |
| [ou] | [ou] | o | 181 | 97.84% |
| [u] | [u] | u | 92 | 91.09% |
| [ui] | [ui] | ui | 54 | 98.18% |
| [uk̚] | [ut̚] | uk | 145 | 94.77% |
| [un] | [un] | un | 49 | 87.50% |
| [uŋ] | [un] | ung | 438 | 99.55% |
| [y] | [y] | u | 34 | 50.75% |
| [y] | [y] | ue | 33 | 49.25% |
| [yn] | [yn] | uen | 196 | 93.33% |

### 3.5.1 Onsets/rhymes with mono-transliteration of high frequency of usage

The onsets/rhymes with only one transliteration in **Table 3** and Table 4 can be grouped into three:

A. [p] *p*, [t] *t*, [f] *f*, [k] *k*, [j] *y*, [n] *n*, [w] *w*, [an] *an*, [akˀ] *ak*, [am] *am*, [ekˀ] *ek*, [ıkˀ] *ik*, [in] *in*, [ıŋ] *ing*, [ɔ] *o*, [ɔi] *oi*, [ɔŋ] *ong*, [ou] *o*, [u] *u*, [ʊkˀ] *uk*, [un] *un*, [ʊŋ] *ung*, [au] *au*

B. [ɐn] *an*, [ɐŋ] *ang*, [ɐpˀ] *op*, [ɐtˀ] *at*, [ɐu] *au*,

C. [ɐi] *ai*, [ɵy] *ui*, [ui] *ui*, [yn] *uen*.

Group A shows that the onsets/rhymes existing in English of which the transliterations in RP (*or* one of the RP-pronunciations for multi-pronounced alphabets, *henceforth*) are very similar to the corresponding phonetic value in Cantonese. Group B shows that the rhymes exist in English but the transliterations in RP are not similar to the phonetic value. Group C shows the rhymes that do not exist in English.

In the view of the high frequency of usage in the romanisation scheme, the above three groups of transliterations for the onsets/rhymes should be considered as "standard" that are used in the scheme. It is also worth noting that most of the Group A transliterations when pronounced in RP are very similar to the Cantonese-pronunciation. In Group B, the phonetic value [ɐ] in English, which is always referred to the phoneme /ʌ/, is usually written as *u* in the English orthography. However, probably due to the fact that *u* in English can also be pronounced as [uː] and [juː]; *a*, which is also close to [ɐ] phonetically, is used instead. The Cantonese rhymes in Group C do not exist in English and so no comparison can be made.

### 3.5.2   Onsets with multi-transliterations: [tsʰ], [s] and [ts]

Due to the limitation in space, only three of the onsets/rhymes in **Table 3** and **Table 4** with multi-transliterations will be discussed in this paper: [tsʰ], [s] and [ts].

The frequency counts for these onsets listed in **Table 3** are copied here:

Table 3: Partial Frequency Counts of the Transliterations of [tsʰ], [s] and [ts]

| Modern Cantonese Onsets | Young Cantonese Onsets | Roman-isation Onsets | Count | Percentage |
|---|---|---|---|---|
| [tsʰ] | [tsʰ] | ch | 299 | 62.29% |
| [tsʰ] | [tsʰ] | ts | 181 | 37.71% |
| [ts] | [ts] | ch | 223 | 78.25% |
| [ts] | [ts] | ts | 62 | 21.75% |
| [s] | [s] | sh | 425 | 63.15% |
| [s] | [s] | s | 243 | 36.11% |

It can be clearly seen in **Table**   that the only set of (shared) transliterations for [ts] and [tsʰ] consists of *ch* and *ts*.   The sum of the percentages is 100%. For [s], *sh* and *s* are used, and the sum of percentages almost reaches 100%.   This means that there are hardly any exceptions.

One may think that different transliterations for the same onset may be used in different historical periods. However, it is found that both transliterations of each of the onsets are used for the street names throughout history until now. Some may think that alternative forms of the transliterations are random results. However, the transliteration set *ch-ch-sh* is parallel to the set *ts-ts-s* in terms of percentage and in terms of orthography. In terms of percentage, the distribution is similar in that those of the first set are around 60% to 80% while those of the second set are around 20% to 40%. In terms of orthography, the spellings are also similar in that the *ch-ch-sh* set always contains the letter *h* while the *ts-ts-s* set always the letter *s*. In English the pronunciations of *ch*, *sh* and *s* are different.

The parallel distribution of the frequency counts and the consistent variation of spelling shows that the choice of transliterations is not random.   There is a systematic representation in that the sets *ch-ch-sh* and *ts-ts-s* each represents a different set of phonemes. However, the problem to find out the phonetic value of each transliteration remains.

In RP orthography, *ch* is pronounced with the alveolar-palatal affricate /tʃ/ while *sh* is pronounced with the alveolar-palatal fricative /ʃ/; *s* is pronounced with the alveolar fricative /s/ while the spelling *ts* does not exist in English as an onset. Hence, each of the transliterations should represent a phone similar to that pronounced in English where *ch* should also represent the aspirated counterpart..

Although the spelling *ts* does not exist as onset in English, we can get some hints from the romanisation scheme of other languages containing *ts*. In Hepburn's romanisation system for Japanese, which is also based on English phonology, the *kana* つ[tsɯ̹] is transliterated as *tsu*, also with the onset *ts*. *ts* in this romanisation system represents the alveolar affricate [ts] in contemporary Japanese. This can be used as a reasonable prediction of the corresponding onset in Cantonese since the transliteration *s* in the same set also represents the homogeneous /s/.

In modern standard HKC, there is only one set of sibilants: [ts], [tsʰ] and [s]. Using the above deduction, one could say that the transliteration sets *ch-ch-sh* and *ts-ts-s* probably represent the two phonemic sets [tʃ]-[tʃʰ]-[ʃ] and [ts]-[tsʰ]-[s], in which the contrast has been neutralised later to form the only set [ts]-[tsʰ]-[s] (described in Zee (1999) as [tʂ-, tʂʰ-, ʂ-] in narrow transcription) in modern standard

HKC. As the aspirations of the affricates are distinctive in modern standard HKC, as well as for other obstruents, it was probably also distinct in the past although this was not indicated in the romanisation scheme.

In conclusion, the choice of the alternative forms was not based on random selection, but with the intention to represent the different historic phonemes at that time when the romanisation scheme was created.

In fact, Zee (1990) has investigated the historical sibilants in HKC and presented a microscopic analysis of the sibilants in the romanisation of the geographical names in Hong Kong. However, no quantitative analysis was conducted so it is not known how general the situation was. With the quantitative analysis presented in this paper, it is clear that the different transliteration sets are systematic representations of two different sets of phonemes in historical Cantonese.

After deducing the phonetic values of the two sets of sibilants, one needs to identify the historical source. According to Cheung (2003), two sets of sibilants existed in early 20th century. Generally, one set was developed from the ancient alveolar sibilants (chitouyin 齒頭音, *or* jing-series 精組); while another set (set *B*) was developed from the ancient alveolo-palatal dentals (sheshangyin 舌上音, *or* zhi-series 知組) and ancient palatal sibilants (zhengchiyin 正齒音, *or* zhao-series 照組, including divisions 2 and 3). According to Chen and Mo (1989), in early Cantonese, the characters in zhi rhyme group 止攝 of the zhao-series division 2 (照二) also share the same onset with those in jing-series (set *A*).

Assuming the number of characters among different onset groups is comparable, the size of set *A* should be smaller than set *B*. According to Table 3, the number of tokens in the set *ch-ch-sh* is double that in the set *ts-ts-s*. As the naming of streets is governed by semantics, which should be independent of phonology, we may consider that the tokens are selected randomly from each set. Therefore the set *ch-ch-sh* should correspond to set *B* (larger in size, generally zhi- and zhao-series) while the set *ts-ts-s* should correspond to set *A* (smaller in size, generally jing-series). The historical sources of the two sets of transliterations are clear up to this point.

It is worth noting that although two transliterations are found in each of the onsets [ts] and [ts$^h$], six are found for the onset [s]. On the other hand, it can be seen in **Table 1** that other than the transliterations *s* and *sh*, the rest only constitute a tiny amount. Although these can be considered as random error[1] statistically, explanations are worth seeking from sociolinguistics. These exceptions are list below:

---

[1] In social science applications, it is common to consider event occurrence higher then five percent as significant.

Table 4: Exceptions in the Romanisation of the onset [s]

| Character | Romanisation | Pronunciation in Modern Cantonese | Full name of the Street |
|---|---|---|---|
| 誠 | Cheng | [sɪŋ] | 誠明徑 |
| 崇 | Chung | [suŋ] | 崇基路 |
| 西 | Hsi | [sɐi] | 江西街 |
| 藪 | Tau | [sɐu] | 鶴藪道 |
| 崇 | Tsung | [suŋ] | 崇文街 |

In the first two cases, the paths are located in the Chinese University of Hong Kong. The romanisation follows the buildings nearby: "誠明館" *Cheng Ming Building* and "崇基書院" *Chung Chi College,* which are actually Mandarin readings in the Wade-Giles system.

In the third case, some streets in that location are romanised with the Chinese Postal Map System such as "安徽街" *Anhui Street* and "浙江街" *Chi Kiang Street.* Therefore, this can be regarded as a regional feature of romanisation, probably because the early residents in this district were Mandarin-spoken.

In the fourth case, it can be seen that the romanisation used for the names of the places near this road is the same such as "鶴藪圍" *Hok Tau Wai.* According to Lau (1995), the place "鶴藪" *Hesou* is in the Hakka dialect so the onset is different.

In the last case, the romanisation of the character "崇" is the same as that of the buildings on that street such as "崇文閣" *Tsung Man Court.* It is possible that this is a mistake made by the officials when naming this road. However, dialectal factor may be involved since this district, namely Aberdeen, was not well developed at the beginning of the British governance.

Therefore, it can be seen that it is possible for the romanisation to follow dialect pronunciations other than Cantonese. It is also possible that mistakes are involved while naming. However, these exceptions are rare and they do not contradict the fact that the system generally follows Cantonese pronunciations.

## 4. Conclusions and Future Work

In this research, the Cantonese romanisation scheme used by the British government was reconstructed by generalising the transliterations of the street names listed in the book *Hong Kong Guide 2005* published by the Hong Kong Government. By using a quantitative approach with logical inference and comparison with other

romanisation schemes, the phonetic values represented by the transliterations, including those no longer existing in contemporary HKC are found out independently, without relying much on previous studies in Cantonese historical phonology.

The generalised scheme has then been compared with previous studies in Cantonese historical phonology and the historical sound changes are explained with reference to Chinese historical phonology. The results show that the phonological system of HKC has been undergoing simplification. The onset sets $[t\int]$-$[t\int^h]$-$[\int]$ and $[ts]$-$[ts^h]$-$[s]$ contrasted in the mid-19th century Cantonese are merged to form the one set $[ts]$-$[ts^h]$-$[s]$.

As seen in Table 2, for some of the rhymes, the frequency of the transliteration is relatively small while some do not appear in the street names. Therefore, as future work, the place names in Hong Kong will also be inserted to the database so as to increase the accuracy of the statistics. In addition, other transliterations of high frequency count not appearing in this paper, along with the sound changes will also be explained. Last but not least, the romanisation scheme will also be compared with other romanisation schemes such as Eitel's system, Ball's system and Meyer & Wemepe's system.

## Notes

I hereby wish to thank my mother for assisting in typing parts of the data used in this research. In addition, valuable comments from Dr. Kwok Bit-Chee, Dr. Sun Jingtao and Prof. Samuel Cheung Hung-nin on this paper are appreciated.

## References

Bauer, Robert S. (1979). Alveolarisation in Cantonese: a Case of Lexical Diffusion. *Journal of Chinese Linguistics* 7: 132 – 141.

----- (1983). Cantonese Sound Change across subgroups of the Hong Kong Speech Community. *Journal of Chinese Linguistics* 11(2): 301 – 354.

----- (1986). The Micro-history of a Sound Change in Progress in Hong Kong Cantonese. *Journal of Chinese Linguistics* 14(1): 1 – 42.

Chen, Matthew Y. 陳淵泉 and Newman John. (1984). From Middle Chinese to Modern Cantonese (Parts 1 – 3). *Journal of Chinese Linguistics* 12(1): 148 – 198, 12(2): 334 – 388 and 13(1):122 – 170.

Chen, Wancheng 陳萬成 and Mo, Huixian 莫慧嫻. (1989). 私 si 師 shi 詩 shi in Late Cantonese Speech. In Zhan, Bohui 詹伯慧 (ed.). pp. 72 – 78.

Cheung, Hung-nin Samuel 張洪年. (2003). 21<sup>st</sup> Century Hong Kong Cantonese: the Formation of a New Phonological System. In *Di Ba Jie Guoji Yue Fangyan Yantao Hui Lunwenji* 第八屆國際粵方言研討會論文集 (The Proceedings of the 8<sup>th</sup> International Conference in the Dialect Yu7), ed. Zhan Bohui 詹伯慧, pp. 129 – 152. Peking: China Social Science Press.

Eiltel, Ernest John. (2001). *A Chinese-English Dictionary in the Cantonese Dialect.* London: Ganesha. First revised and enlarged by Immanuel Gottlieb Genähr. Hong Kong : Kelly & Walsh, 1910 – 1911.

Hashimoto, Oi-Kan Yue 余藹芹. (2005). Yue Yu Zuyu de Gouni 粵語祖語的構擬 (The Re-construction of the Proto-Language of the Yu7 Dialect). Paper Presented at *the Tenth International Conference on the Yu7 Dialect*, The Chinese University of Hong Kong, Hong Kong, 12<sup>th</sup> – 14<sup>th</sup> December, 2005.

Hong Kong Government. (1960). *A Gazetteer of Place Names in Hong Kong, Kowloon, and the New Territories.* Hong Kong: Government Printer.

Lau, Chun Fat 劉鎮發. (1995). *Xianggang* de Ke Fangyan Diming 香港的客方言地名 (The Place Names of Hakka Dialect in Hong Kong). *Kejia Yanjiu Jikan* 7 客家研究輯刊.

Lau, Chun Fat 劉鎮發 and Cheung, Kwan Hin 張羣顯. (2001). Qing Chu de Yue Yin Yinxi — 'Fenyuncuoyao' de Shengyun Xitong 清初的粵音音系──《分韻撮要》的聲韻系統 (The Phonological System of Cantonese at Early Qing — the Phonological System of *Fenyuncuoyao*). In *Di Ba Jie Guoji Yue Fangyan Yantao Hui Lunwenji* 第八屆國際粵方言研討會論文集 (The Proceedings of the 8<sup>th</sup> International Conference in the Dialect *Yue*), ed. Zhan Bohiu 詹伯慧, pp. 206 – 223. Peking: China Social Science Press.

Lee, Yin Ping Cream 李燕萍 and Kataoka, Shin 片岡新. (2005). Legacies of Missionary Romanisations: Place and Personal Names in Hong Kong. Paper presented at *the Third International Conference on Missionary Linguistics*, the Hong Kong University of Science and Technology and Instituto Inter-Universitário de Macau, Hong Kong and Macau, 12<sup>th</sup> – 15<sup>th</sup> March, 2005.

Li, Xinkui 李新魁. (1989). Centuries of Development of Cantonese Dialect Vowels. In Zhan, Bohui 詹伯慧 (ed.). pp. 35 – 40.

----- (1997). Yi Bai Nian Qian de Guangzhou Yin 一百年前的廣州音 (The Pronunciation of Cantonese One Hundred Years Ago). In *Li Xinkui Yinyunxue Lunji* 李新魁音韻學論集 (Collection of Discussions of Phonology by Li Xinkui), ed. Li Xinkui 李新魁, pp. 429 – 436. Swatow: Shantou Daxue Chubanshe.

Lun, Suen Caesar 藺蓀. (1993). The Use of Chinese Characters in Hong Kong Street Names. Paper presented at *the Fourth International Conference on Cantonese and Other Yue Dialects*, City Polytechnic of Hong Kong, Hong Kong, 17th – 19th December, 1993.

Meyer, Bernhard F. and Wempe, Theodore F. (1947). *The Students' Cantonese-English Dictionary*. 3rd ed. Maryknoll, New York: Catholic Foreign Mission Society of America.

Morrison, Robert D. D. (1828). Vocabulary of the Canton Dialect (廣東省土話字彙). Macao: East India Company's Press. Reprinted in 2001 by London: Ganesha Publishing Limited.

Peng, Xiaochuan 彭小川. (1989). The System of Consonants in the Cantonese Book 'Outline Rhyming'. In Zhan, Bohui 詹伯慧 (ed.). pp. 72 – 78.

----- (1992). Yueyu Yunshu 'Fenyuncuoyao' Ji Qi Yunmu Xitong 粵語韻書《分韻撮要》及其韻母系統 (The Rhyming Dictionary *Fenyuncuoyao* and its Rhyming System). *Jinan Xuebao* 暨南學報, Vol. 4, pp. 153 – 159.

Survey and Mapping Office, Lands Department. (2005). *Hong Kong Guide 2005 (Photomap Edition)*. Hong Kong: Hong Kong SAR Government.

Williams, Samuel Wells. (2001). *A Syllabic Dictionary of the Chinese Language: Arranged according to the "Wu-fang Yuen Yin", with the Pronunciation as Heard in Peking, Canton, Amoy, and Shanghai*. London: Ganesha. First ed. published in Shanghai: American Presbyterian Mission Press, 1889

----- (2001). *A Tonic Dictionary of the Chinese Language in the Canton Dialect*. London: Ganesha. Originally published in Canton 1856.

Wong, Pui Kwong 王培光. (2000). On Chinese Characters in the Place Names of Hong Kong Yue Dialect. *The Dialect* 2: 186 – 190.

Yiu, Kau Tsoi 饒玖才. (1998). *Xianggang Diming Tansuo* 香港地名探索 (Exploring the Geographic Names in Hong Kong). Hong Kong: Cosmos Books.

Zee, Eric 徐雲揚. (1999). Change and Variation in the Syllable-Initial and Syllable-Final Consonants in Hong Kong Cantonese. *Journal of Chinese Linguistics* 27(1): 120 – 167.

----- (1999). Chinese (Hong Kong Cantonese). In *Handbook of the International Phonetic Association: a Guide to the Use of the International Phonetic Alphabet*, pp. 58 – 60. Cambridge: Cambridge University Press.

Zhan, Bohui 詹伯慧(ed.). (1989). *A Collection of Thesis from the 2nd International Cantonese Dialect Discussion*. Canton: Jinan University Press.

----- (ed.) (2003) *Di Ba Jie Guoji Yue Fangyan Yantao Hui Lunwenji* 第八屆國際粵方言研討會論文集 (The Proceedings of the 8th International Conference in the Dialect Yu7), Peking: China Social Science Press.

## 摘要

　　香港於受英國殖民政府管治初期，唯一之官方語言為英語；因此，當時需要一個廣州話拼音系統以音譯漢語之人名、街名及地名等。由於官方文件之需要，殖民政府早已採用了一個拼音系統。然以英語之拼音法來讀該系統的大部份音節之讀音跟當代香港粵語之發音非常近似，反映其系統性；但有些則存在差異，反映出十九世紀中葉之廣州話之語音面貌。

　　本研究以香港之街道名稱為材料，首先把不適用者有系統地去除，其他的則用於設立資料庫。結果顯示於標準香港粵語，差不多每個聲母及韻母都有代表其之音譯，數目從一到六個不等。此外，亦利用定量方法加上邏輯分析及比較其他拼音系統，找出這些音譯所代表之音值，包括於當代香港粵語已不復存在者，而不大依賴其他之歷史廣州話研究。結果顯示香港粵語之音韻系統已趨簡化，於十九世紀中葉廣州話對立之聲母[tʃ]-[tʃʰ]-[ʃ] 及 [ts]-[tsʰ]-[s]，於當代香港粵語已合併成一套[ts]-[tsʰ]-[s]。

# 两岸汉字笔顺规则问题讨论

张小衡
香港理工大学

## 1．前言

笔顺对于汉字的键盘输入，手写输入，排序检索和语言教学等都具有很重要的作用。应该特别指出的是，当前方兴未艾的对外汉语教学更是需要笔顺知识的支持。

关于汉字笔顺，中国两岸四地现行的官方标准共有两套：一套来自内地，另一套来自台湾，香港和澳门都没有官方标准[1]。内地的标准共两个文献，即《现代汉语通用字笔顺规范》(国家语委 1997) 和《GB13000.1 字符集汉字笔顺规范》(国家语委 1999)；台湾的标准只有一个文献，即《常用国字标准字体笔顺手册》(台湾教育部 1996)。由于这两套标准文献都以列表的方式逐字规定笔顺，不便学习和应用，所以与之对应的笔顺规则应运而生。本文专题讨论这些规则。通过分析和比较，指出存在的问题，并提供一些解决方法。

## 2．两岸笔顺规则比较

内地"笔顺规则"这个概念在台湾的标准中称为"笔顺法则"。本文为了讨论方便，在不需要区分两地的情况时，常会将他们统称为"笔顺规则"或简称为"规则"。

本研究选用《常用国字标准字体笔顺手册》中的《笔顺基本法则说明》(共17条) 作为台湾笔顺规则的代表。因为《笔顺基本法则说明》是官方标准文献的一部分，而且文中还特别说明这套法则是在标准笔顺字表的基础上归纳出来的，最具权威。内地的笔顺规范文献没有提供笔顺规则，我们选用的是《现代汉语模范字典》的《汉字书写笔顺规则表》 (许嘉璐，2000，p713，共 16 条规则)。因为该书的主编是前国家语委主任，还特别注明其中的汉字书写笔顺规则是根据国家语委汉字规范制作的。社会上颇有影响的内地版《语言文字规范使用指南》(李行健，费锦昌，2001，p227)中的《汉字笔顺规则表》和香港版《商务学生字典》(盛九畴，2003，p488)中的《汉字书写笔顺规则表》可能都抄自《现代汉语模范字典》，因为三个笔顺规则表的内容完全相同。

为方便分析讨论，我们将上述的两套笔顺规则合为一个对照表，根据他们之间在内容上的涵盖关系作对齐和归类处理，并加注说明。结果如附表所示。

表中的第一项和第二项内容分别是台湾《常用国字标准字体笔顺手册》的笔顺基本法则和例字的原文，第三和第四项是《现代汉语模范字典》中的笔顺规则和例字原文。最后一列是两地比较与说明，指出同一行上的笔顺规则和法则的基本内容是完全相同，部分相同还是为内地或者台湾所独有。两地的笔顺规则作了对齐处理，使得（基本）内容完全相同或部分相同的规则和法则处于同一行。如果一条规则/法则与多条法则/规则对齐，则只在第一次出现时给出内容，再次出现时只标示规则/法则号码。如果某一规则/法则为内地或台湾一方所独有，则只

---

[1] 香港教育署和香港教育学院曾经为小学发表过一些指引，但不属政府标准，而且内容简单，影响不大。

在该方的项目上填写内容。

例如：第一行上的台湾笔顺规则是法则一"自左至右"，与之对应的是内地笔顺规则 4"从左到右"，两者含义"完全相同"。又如，法则十"点在上或左上的先写，点在下、在内或右上的，则后写。" 与规则 8"点在正上或右上先写点。"的内容部分相同。而法则十一"凡从戈之字，先写横画，最后写点、撇。"则是台湾特有的笔顺。

在内地的 16 条笔顺规则和台湾的 17 条笔顺法则之间，双方含义相同的有 7 对，包括内地规则"先横后竖"、"先撇后捺"、"从上到下"、"从左到右"、"先外后里"、"先外后里再封口"和"缺口朝上的先里后外"。台湾独有的规则 2 条："凡从'戈'之字，先写横画，最后写点、撇。"和"横画在中间而地位突出者，最后写"。其余的都属部分含义相同。

可见，两套规则的内容大体是一致的。内地的七条基本规则(先横后竖，先撇后捺，从上到下，从左到右，先外后里，先外后里再封口，先中间后两边)中，前六条都有内容相同的台湾法则与之对应。而最后一条"先中间后两边"也有三条台湾法则与其部分相同，即法则六"竖画在上或在中而不与其它笔画相交者，先写竖画。"，法则十三"横、竖相交，横画左右相称的结构，通常先写横、竖、再写左右相称之笔画。"和法则十七"凡字的上半或下方，左右夹中，且两旁相称或相同的结构，通常先写中间，再写左右。"。

总的来说，台湾规则比较详细，内地的规则比较简要，两套规则各有千秋。可以说，他们是当前两地最具权威，影响最大的笔顺规则。然而，这两套笔顺规则都存在改善的空间。问题分为两类标注在附表中的相应位置：一条规则的问题写在该规则的内容之下（并加括号）。一条规则的例外字（其标准笔顺与该规则/法则相悖）写在该规则的例字下面（并加括号）。下一节将较深入地分析探讨几个比较重要的问题。

## 3．笔顺规则问题讨论

这里我们讨论八个值得重视的有关两岸笔顺规则的问题。

问题一：规则条目太多

内地的笔顺规则有 16 条，台湾的笔顺基本法则有 17 条。数量这么多显然是不易学习和掌握的。况且，这么多规则还不够，还有一些需要补充。这一点将在下一节进一步讨论。

问题二：某些规则欠缺

对于两岸的笔顺规则来说，都存在些笔顺规律没有覆盖到。例如，对于横竖相交的笔画，两套规则中都有"先横后竖" 的条目，对于撇捺相交的笔画，也都有"先撇后捺"的规则。然而，处于同一层次上的横撇相交，涉及的汉字也相当多（例如 "有"、"右"，"在"等字的左上部），而且都遵循"先横后撇"的规律，但这一规律在两岸的笔顺规则中都未能反映出来，应该补上。竖提相交（例如 "辆"、"打"，"特"左偏旁最后两笔）也有类似的问题，两岸也都应该补上"先竖后提"的笔顺规则。

问题三：有些规则内容是多余的。

　　两岸都有一些规则或规则中的部分内容已经蕴含于其他规则之中，不需要写出来。具体情况如下所述。

- 台湾法则七"横画与竖画组成的结构，最底下与竖画相接的横画，通常最后写。"是完全没有必要的。因为这是"先上后下"的一种情形，内容已经涵括于法则二之中。
- 内地的规则11" 右上包围结构，先外后里"、规则12 "左上包围结构，先外后里"、规则15"缺口朝下的，先外后里"。和规则10"点在里面后写点"都为规则5 "先外后里" (基本规则)所覆盖，所以不必列出。
- 内地规则8"点在正上或左上先写点"也为规则3"从上到下"(基本规则) 所包含，不需再单独订立规则。
- 台湾法则十二 "撇在上，或撇与横折钩，横斜钩所成下包结构，通常撇画先写。(例如，千,白,用,凡)"也可省去，因为内容可用法则二" 先上后下"和法则一"自左至右"来解释。
- 台湾法则四"先横后竖：凡横画与竖画相交，或横画与竖画相接在上者，皆先写横画，再写竖画"，其中的"或横画与竖画相接在上者"已被法则二"先上后下"所包含，可省略。
- 法则六"竖画在上或在中而不与其它笔画相交者，先写竖画。"中的"在上或"没有必要，其内容已被法则二"先上后下"所包含，可省略。

问题四：规则之间的优先级关系不明确

　　另一个妨碍笔顺规则应用的原因是：两岸都没有说明规则之间的优先级关系。这样，当多条规则的条件同时满足时，就会出现取舍问题。例如，"米"字右上角的短撇和左下角的长撇。根据其方位关系，如果按照"从上到下"的规则，则应该先写短撇再写长撇；如果使用"从左到右"的规则，则应该先写长撇再写短撇。如果我们根据标准笔顺的一般规律，规定规则 "从上到下"的优先级高于"从左到右"，问题便迎刃而解。道理就像在数学中需要规定"先乘除，后加减"一样。

问题五：某些规则内容不够完整或准确。

　　内地笔顺规则中1至7是比较重要的基本笔顺规则。但是，这七条基本规则除最后一条"先中间后两旁"有大概注明其使用条件是"中间部分长或宽"之外，其余的六条规则都没有说明条件。例如规则1 "先横后竖"没有说明其条件是笔画相交。这方面台湾做得较好，但也存在改善的空间，例如法则一"自左至右：凡左右并排结体的文字，皆先写左边笔画和结构体，再依次写右边笔画和结构体。" 其实，遵循"自左至右"笔顺的除成文字的左右并排结体外，还应该包括不成字的左右并列结构。例如竹字头（䒑），"监"字头不单独成字，但都遵循自左部件至右部件的笔顺原则。

问题六：同一套规则内部出现相互矛盾

　　这往往也是问题五的表现。例如，台湾笔顺法则三"由外而内：凡外包形体，无论两面或三面，皆先写外围，再写里面。"，行文中用了"凡…皆"的句式，意味着没有例外。这与法则十六"凡下托半包的结构，通常先写上面，再写下托半

包的笔画"。和法则十五"凡以辶、廴为偏旁结构的字，通常辶、廴最后写。"的存在是相矛盾的。如果将法则三的"凡"和"皆"两字删去，并将该法则的优先级调到下面两条法则之下，问题就解决了。

问题七：规则的例外字没有指出

　　根据《常用国字标准字体笔顺手册》和《GB13000.1 字符集汉字笔顺规范》中的标准笔顺字表，两地的笔顺规则几乎每一条都有例外的字。例如，"先上后下"的例外字有"心"（标准笔顺中，右边上面两点最后写）；"先左后右"的例外字有"止"（上部笔画先写右边的竖横，再写左边的竖）；"先横后竖"的例外字有"丑"（里面的十字形先竖后横）；"先撇后捺"的例外字有珳、女、戈（其中撇捺／点交叉的笔画后写撇）；"右上包围结构，先外后里"的例外字有"可"（先上面的横，然后是里面的口，最后是外面的竖钩），等等。详见附表中笔顺规则例字后面的例外字。从严谨的科学态度出发，对于各条规则的例外字，如规律性比较弱，数量少，不便使用新的规则加以概括，就应该一一列出注明，每一条有例外的规则都应该列出相应的例外字或部件的标准笔顺。当然，最理想的方法是通过笔顺规范的逐渐完善使所有字的标准笔顺都符合规则。

问题八：两套规则的一致性有待提高

　　为方便海峡两岸的语言教学与应用，两套笔顺规则应该互相取长补短，尽可能达成一致。例如，内地的笔顺规则应该参考台湾的做法，配全每一条规则的应用条件。而台湾特有笔顺法则十一"凡从戈之字，先写横画，最后写点、撇"。与"戈"在内地的标准笔顺（最后写撇、点）不一致。笔者曾在一次中国语文政策国际研讨会上提到这一问题，与会的一位台湾教师说从戈之字他的习惯写法与内地标准一样。从技术上讲，该台湾特殊规则并没有什么优越性，如果又不能很好的放映广大语言工作者的书写习惯，那是完全可以考虑与内地靠拢的。希望两岸的笔顺规范能早日统一。

## 4．结论

　　笔顺规则是掌握笔顺的重要工具，对于中文的学习、书写、检索和计算机输入等都能带来很大的方便。本文对比讨论了海峡两岸的两套较有影响的笔顺规则，指出和分析了八个值得重视的问题，并提出了一些解决方法。

　　笔顺规则的优化需要字形建设的支持，还与笔顺标准字表内部的规律性有关（张小衡，苏永昌，1995）。此外，两岸标准笔顺表之间的一致性也需要加强。据统计，《常用国字标准字体笔顺手册》列明标准笔顺的 4808 字中，与内地字形相同但笔顺不同的占 7.97%(张小衡，张焕淇，2006)，严重影响两地的笔顺应用。有趣的是，两地都认为自己的笔顺有助于把汉字写得最好。这些都是需要进一步研究的相关问题。本文挂一漏万，希望能起到抛砖引玉的作用。

# 参考文献

国家语委，(1997)．现代汉语通用字笔顺规范．北京：语文出版社．

国家语委，(1999)．GB13000.1 字符集汉字笔顺规范．上海：上海教育出版社．

李行健，费锦昌，(主编，2001)．语言文字规范使用指南．上海：上海辞书出版社．

盛九畴，(2003)．商务学生字典．香港：商务印书馆，(第 9 次印刷)．

台湾教育部国语推行委员会，(1996)．常用国字标准字体笔顺手册．台北：教育部.(http://www.edu.tw/EDU_WEB/EDU_MGT/MANDR/EDU6300001/allbook/bishuen/f8.html?open)

许嘉璐，(主编，2000)．现代汉语模范字典．北京：中国社会科学出版社．

张小衡，苏永昌，(2005)．进一步加强汉字规范笔顺的规律性．陈群秀，孙茂松（编），自然语言理解与大规模内容计算．北京：清华大学出版社，pp. 547-552．

张小衡，张焕淇（2006)．两岸汉字规范字形笔顺比较．两岸四地语文政策国际学术研讨会论文．香港理工大学，2006 年 5 月 11-13 日．

# 附表：台湾-内地笔顺规则对比

| 台湾笔顺法则 | 笔顺法则例字 | 内地笔顺规则 | 笔顺规则例字 | 比较与说明 |
|---|---|---|---|---|
| 一．自左至右：<br>凡左右并排结体的文字，皆先写左边笔画和结构体，再依次写右边笔画和结构体。<br>（条件应该包括非字的左右并列结构。） | 川，仁，街，湖<br>（例外：止，小，水，万，甘，阝，爿，樂） | 4. 从左到右<br>（基本规则，缺条件） | 孔，让<br>（例外：止，将，收，北，非，其，耳，兆，才，矛，寸，忙，恭） | ● 完全相同；<br>● 法则的内容较具体； |
| 二．先上后下：<br>凡上下组合结体的文字，皆先写上面笔画和结构体，再依次写下面笔画和结构体。（条件应该包括非字的上下组合结构。） | 三，字，星，意<br>（例外：心，戈，犬） | 3. 从上到下<br>（基本规则，缺条件） | 亏，王<br>（例外：心，戈，犬） | ● 完全相同；<br>● 法则的内容较具体； |
| 三．由外而内：<br>凡外包形体，无论两面或三面，皆先写外围，再写里面。（考虑在前面加"除满足法则14，15，16条件的"。） | 刀，勾，月，問<br>例外：乃，這，區，凶，可，断，乘，爽 | 5. 先外后里<br>（基本规则，应处理好优先级） | 月，问<br>例外：凶，可，断，乘，爽 | ● 完全相同<br>● 法则的内容较具体；； |
| （三） | | 11. 右上包围结构，先外后里<br>（多余规则，含于规则5） | 勾,司<br>例外：或，可。 | ● 部分相同 |
| （三） | | 12. 左上包围结构，先外后里<br>（多余规则，包含于规则5） | 庆，房 | ● 部分相同 |
| （三） | | 15.缺口朝下的，先外后里（多余规则，包含于规则5） | 内，向 | ● 部分相同 |
| 四．先横后竖：<br>凡横画与竖画相交，或横画与竖画相接在上者，皆先写横画，再写竖划。（条件"或横画与竖划相接在上面者，"已被法则二"先上后下"所包含，可省略） | 十，干，士，甘，聿<br>例外：丑，聿，書（1,2画）垂，訊，出 | 1. 先横后竖<br>（基本规则，缺条件） | 十，干<br>（例外：里，黑，垂，丑，贯，耳，子） | ● 完全相同；<br>● 法则的内容较具体； |

| | | | | |
|---|---|---|---|---|
| 五·先撇后捺：<br>凡撇画与捺画相交，或相接者，皆先撇而后捺。 | 交，入，今，長<br>例外：女，戈 | 2．先撇后捺<br>(基本规则，缺条件) | 人，木，八，文<br>(例字"八"、"人"多余，已为"先左后右"和"先上后下"所涵盖。例外：琰，女，戈) | ●完全相同；<br>●法则的内容较具体； |
| 六·竖画在上或在中而不与其它笔画相交者，先写竖画。("竖画"后的"在上或"可省略，已经为"先上后下"所表示。) | 上，小，山，水<br>(例外：川，修) | 7．先中间后两旁(中间部分长或宽)，<br>（基本规则，两旁应对称) | 小，办，<br>(例外：爽，乘，米) | ●部分相同<br>●法则的内容较具体； |
| 十三．横、竖相交，横画左右相称的结构，通常先写横、竖，再写左右相称之笔画。 | 來，垂，喪，乘，插字旁； | (7) | (例外："来"的上两点比竖先写) | ●与规则7部分相同<br>●法则的内容较具体； |
| 十七．凡字的上半或下方，左右夹中，且两旁相称或相同的结构，通常先写中间，再写左右。 | 兜，學，樂，變，嬴 | (7) | (例外：學，輿) | ●与规则7部分相同<br>●法则的内容较具体； |
| 七·横画与竖画组成的结构，最底下与竖画相接的横画，通常最后写。(本法则无必要，含于"先上后下"之中。) | 王，里，告，書 | (3) | (例外：玉) | ●部分相同<br>●法则的内容较具体； |
| 八·横画在中间而地位突出者，最后写。 | 女，丹，母，毋，冊<br>(例外：秉，丑，律，伊) | | | ●台湾独有 |
| 九·四围的结构，先写外围，再写里面，底下封口的横画最后写。 | 日，田，回，國 | 6．先外后里再封口(基本规则) | 日，田<br>例外：母鸟 | ●完全相同；<br>●法则的内容较具体； |
| 十·点在上或上的先写，点在下、在内或右上的，则后写。("点在上或左上的先写"，"点在下，则后写。"的意思已为法则二"先上后下"所包含。) | 卞，爲，叉，犬<br>例外：母，丹，戈，心，必 | 8．点在正上或左上先写点<br>(多余法则，已为规则3"从上到下"所包含。) | 門，斗 | ●部分相同； |
| (十) | | 9．点在右上后写点 | 戈，犬 | ●与法则十部分相同； |

| （十） | | 10．点在里面后写点（多余规则，为规则5"先外后里"所覆盖） | 瓦，叉（例外：母，丹，舟，卵，金，鸟，义，夜） | ●与法则十部分相同； |
|---|---|---|---|---|
| 十一。凡从戈之字，先写横画，最后写点、撇。 | 戍，戒，成，咸 | | | ●台湾特有笔顺 |
| 十二。撇在上，或撇与横折钩，横斜钩所成下包结构，通常撇画先写。(多余法则，可用法则"先上后下"和"先左后右"解释。) | 千，白，用，凡 | （3），（4） | | ●部分相同<br>●法则的内容较具体； |
| 十四。凡竖折、竖曲钩等笔画，与其他笔画相交或相接而后无挡笔者，通常后写。 | 區，臣，也，比，包 | 16．缺口朝右的，先上后里再左下 | 区，匹例外：印，丞 | ●部分相同 |
| 十五．凡以辶、廴为偏旁结构的字，通常辶廴最后写。 | 廷，建，返，迷 | 13．左下包围结构，先里后外 | 近，建(例字还有"断"；例外：飞，心，起) | ●部分相同<br>●法则的内容较准确； |
| 十六．凡下托半包的结构，通常先写上面，再写下托半包的笔画。 | 凶，函，出例外：丞 | 14.缺口朝上的，先里后外 | 击，凶（例外：學，輿） | ●完全相同 |

# Skills Development in a Short-term English Study Abroad Program

Mineharu NAKAYAMA          Noriko YOSHIMURA
The Ohio State University     University of Shizuoka

## 1. Introduction

An increasing number of English study abroad programs have been hosted for Japanese college students in the US, England, Canada, Australia, and New Zealand in recent years. These programs are often existing university affiliated ESL programs with morning classes and afternoon activities. There are various kinds of studies on these study abroad programs (Ellis 1993, Kitao 1993, Suzuki 2001, Tateyama 2002), but most of these studies discuss psychological and socio-cultural aspects of learners' studying such as their motivations and attitudes toward learning English, and their understanding of these English speaking countries' cultures. As noted in Freed (1998), there are rather few articles that discuss the improvements of participants' English language skills in these programs (e.g., Hashimoto 1992, Tanaka & Ellis 2003).[1]

In this paper, therefore, we will discuss what contribution a three-week summer English study abroad program would bring to the development of participants' English proficiency. Specifically, we will discuss the effect of the summer program on the participants' skill improvement based on the results of the Michigan Test of English Language Proficiency of the Institutional Michigan English Language Assessment Battery and the English essay tests (Institutional Test of Written English) given during the summers of 2004-2006. The summer program, specifically tailored for one Japanese university, focused on improving student's oral presentation skills through learning American culture. Because the program was not designed to improve test scores and because the program duration was short, we did not expect to find any significant improvement in the test scores. However, we did find a significant improvement in the scores of the latter test, though not in the former test. The short-term program became effective for a particular skill, namely, writing, even when the program did not focus on it. The oral presentation training served as a practice for formulating ideas quickly and making a coherent and logical argument. Because these skills are also necessary when one writes an essay in a limited time, we believe that their writing improved despite the fact that the curriculum did not teach writing.

The organization of this paper is as follows. We will briefly discuss the previous studies pertinent to the short-term English study abroad in the following section. Then, we will describe our study abroad program in Section 3. Section 4 provides a discussion of the effect of the program referring to the assessment test results. Finally, Section 5 presents a conclusion and future research agenda.

---

[1] These language skills include the knowledge of appropriate contexts for usages and cultural understandings (Hymes 1972) and uses of communication strategies (Canale 1983). As for the improvement of a learner's speaking ability after a study abroad program, it has been mostly discussed based on the results of the ACTFL Oral Proficiency Interviews (e.g., see Freed (1995)).

## 2. Previous Studies

Most of previous studies on short term study abroad programs discussed the program effects on the participants' motivation, their learning strategies, and/or their communication attitudes. They were often based on introspective data collected through questionnaires and interviews. For example, based on two (self-assessment) questionnaires given to 33 freshman students of a women's junior college who participated in the three week summer English abroad program, Kitao (1993) reports that the students felt that their English skills improved and their images of America and Americans also improved. Takeda (1998) investigates the effect of the 3.5 week study abroad program from the perspectives of motivation and strategy of learning English. He gave questionnaires three times (2 weeks before the participation, immediately after the program, and 6 months after the program) to 143 students, but only 63 completed three questionnaires. At the end of the program, it was revealed that the participants' motivation increased, but it was difficult for them to maintain that level of motivation after 6 months. However, it was found that they had more positive attitudes toward learning from their own errors and they were not afraid of making errors while using English.[2] Similarly, Tateyama (2002) reports that a homestay program increased the participants' learning attitude and motivation, based on the results of a questionnaire and an interview of 11 junior college students participating in a three week study abroad program.[3]

Unlike the number of studies that discuss the participants' learning attitudes, there were fewer studies that discuss the improvement of the participants' language skills in study abroad programs whose durations were shorter than one month, as noted in Freed (1998). For instance, Hashimoto (1992) evaluates the participants' improvement of English skills in the 4 week English abroad program held at an American university. He reports that 24 Japanese university students' average Pre-TOEFL scores (70 minutes, maximum 500 points) improved from 406 (before the program) to 422 (after the program), in particular, those who possessed advanced and intermediate levels of English proficiency improved 16 points on average. Tanaka and Ellis (2003) report that 166 Japanese sophomore university students (English majors, ages 19-20) from one university in Tokyo improved 18.55 points on average on TOEFL (from average 426.73 to 445.28) after participating in the 15 week English program at an American university. They (2003: 78) state that "Although this was statistically significant, the gain seems moderate given the length and nature of the learning experience." The "Grammar" section improved 2.48 points (5.88%) while the "listening" section improved only1.15 points (2.65%). Sawasaki and Yoshimura (2005) report that 37 university students in a three week English program at an American university during 2003-2005 self-assessed that their English writing skills improved. Their writing test scores reported in Sawasaki, Yoshimura, Nakayama, and Terao (2006) support their self-assessment.[4] Yoshimura, Nakayama, and Sawasaki

---

[2] Note that study abroad is becoming more popular among junior and senior high schools in Japan as well. For instance, see Brauer (2005) for study abroad in 61 private junior high schools.

[3] In addition to these studies based on the students' self-assessments, Takezawa (1995) reports the six instructors' comments, who taught 39 Japanese female junior college students during the four week spring English program in New Zealand. The major impediments for their learning were their silence and passive attitudes in classroom. They suggested that the participants should be instructed before they join the program that they should express their opinions and raise questions in the classroom.

[4] Based on the 12 Japanese students studying at Monash University, Australia, Marriott (2000) points

(2006) present the analysis of the 2004 and 2005 Institutional Michigan Test of English Language Proficiency and Institutional Test of Written English scores of the same program and report that there was a significant improvement in the 2005 Institutional Test of Written English scores.

## 3.The Current Study

Extending Sawasaki and Yoshimura (2005), Sawasaki, Yoshimura, Nakayama, and Terao (2006), and Yoshimura, Nakayama, and Sawasaki (2006), we will report in this paper the effect of a three-week summer English study abroad program held in the US during the summers of 2004-2006 which focused on oral presentations and discussions. Particularly, we would like to address the following three research questions: a) Can a short-term abroad program make any contribution to the development of participants' English proficiency? b) If yes, then what contribution may be expected? And c) how should such short-term study abroad program be assessed within the Japanese university curriculum?

### 3.1 The Program participants

Thirty-one undergraduate students from one Japanese university participated in a three-week summer English study abroad program held in a large mid-west American university during 2004-2006.[5] The participants were eight sophomores-seniors (mean age 20.2) in 2004, 12 freshmen-sophomores (mean age 19) in 2005, and 11 freshmen–seniors (mean age 19.3). All were Japanese native speakers who participated in the program voluntarily, and a few students each year had been to English speaking countries prior to their program participation (notably, three students in 2005 and one in 2006 spent more than one year in the US prior to entering the university). All but one student with a science major in 2004 were from humanities (e.g., English and Japanese) and social sciences (e.g., international relations and East Asian studies).[6]

### 3.2  The characteristics of the Program

The three-week summer English program was created specifically for one Japanese university in 2003. Originally, the program focused on the improvement of oral communication skills, but it shifted its focus slightly to improve students' academic oral presentation skills specifically. Also the 2003 program had employed different assessment measures. Thus, we will focus our analysis only on the 2004-2006 programs here. At the beginning and the end of the program, each participant took the Michigan Test of English Language Proficiency (MTELP) of the Institutional Michigan English Language Assessment Battery and the Institutional Test of Written English (an essay test; hereafter, ITWE). Although the MTELP taken at the beginning

---

out that a lack of academic writing skill is the biggest problem in order to achieve a success in college.

[5] The participants included one graduate student each year, but we excluded them from the analysis to discuss the improvement of only undergraduate students' English skills. This is to keep the subject group more homogeneous as the participating graduate students were specializing in EFL.

[6] Those with an English major didn't necessarily have better first test scores compared with non-English majors.

The participants included one graduate student each year, but we excluded them from the analysis to discuss the improvement of only undergraduate students' English skills. This is to keep the subject group more homogeneous as the participating graduate students were specializing in EFL.

of the program indicated the level of each participant's English ability, they were not divided into different levels based on their scores as the entire group was already small each year. They were all in one multi-level skill class. The classroom activities were mostly held in the morning. The participants did not use a specific textbook. They read some articles from newspapers, magazines, books, and websites, watched movies, and discussed them in the classroom. They also had special lectures such as American politics and Asian-Americans (different depending on the year) and hosted a Japanese cultural session for American 1-7th graders. Each student set three specific goals at the beginning and reported to what degree they accomplished them by the end of program. Furthermore, each conducted his/her own research project and made a 5 minute presentation on the last day of the program. Although a handout was requested of each presenter, no written paper was required on his/her project. Each participant also had a conversation partner and experienced a weekend homestay.[7] Each conversation partner spent at least 6 hours per week with the participant, and these hours were not necessarily all conversation times (i.e., some may have helped them with their research projects directly, but others may have taken them to parties and other activities as all participants stayed in a dorm together). Accordingly, the program took a broader, holistic approach to improve the students' language skills with a focus on oral presentation skills through learning American culture and working on their own research projects. For the details of the program, see Yoshimura, Sawasaki, Terao, and Nakayama (2006), which also report the results of the 2005 participants' self-assessment.

## 3.3. Assessment measures

In order to measure the effect of the program, all students took both MTELP and ITWE at the beginning and the end of the program each year. The actual contents of those tests were different each time. An analysis discussed in the next section is based on the results of these test scores. The MTELP included Grammar (40 points), Vocabulary (40), and Reading Comprehension (20) questions (maximum 100 points).[8] The ITWE was to write an essay in English on a specific topic within 30 minutes, and the topics were different each time. The evaluation method followed the ETS TWE guidelines (1=low 6=high, with + and -) as stated in (1).

---

[7] It is important to note that conversation partners and weekend homestays were helpful in increasing one's motivation to study English and one's confidence in English ability, and developing a positive attitude toward communicating in English.

[8] Each participant also took a listening comprehension test twice. Although we will not discuss the results in detail, we report here that only the 2006 group improved significantly (t(10), -1.89294, p=0.044). We suspect that this may be related to the fact that the 2006 group had more freshmen, as we will discuss below.

(1) Test of Written English Scoring Guidelines

| score | |
|---|---|
| 6 | Demonstrates clear competence in writing on both the rhetorical and syntactic levels, though it may have occasional errors. |
| 5 | Demonstrates competence in writing on both the rhetorical and syntactic levels, though it will probably have occasional errors. |
| 4 | Demonstrates minimal competence in writing on both the rhetorical and syntactic levels. |
| 3 | Demonstrates some developing competence in writing, but it remains flawed on either the rhetorical or syntactic level, or both. |
| 2 | Suggests incompetence in writing. |
| 1 | Demonstrates incompetence in writing. |

Because the program did not offer a specific curriculum for these tests and because their length of stay in the US was short, no significant improvement was expected to occur in the scores.

## 4. Results

The results of the MTELP and the ITWE will be discussed separately below.

### 4.1 MTELP

Table 1 shows a summary of the MTELP scores during 2004-2006. The lowest and the highest scores in the three years were 48 and 90 in the first (beginning) test of 2005, respectively. The lowest average score (65.8) was reported in the first test of 2005 and the highest average score (73.4) was in the second test of 2006.

Table 1. Summary of the MTELP scores in 2004-6

| Scores/100pts | 2004 Beginning (n=8) | End | 2005 Beginning (n=12) | End | 2006 Beginning (n=11) | End |
|---|---|---|---|---|---|---|
| 41-50 | 0 | 0 | 3 | 1 | 0 | 0 |
| 51-60 | 1 | 0 | 2 | 2 | 1 | 1 |
| 61-70 | 3 | 4 | 3 | 3 | 6 | 2 |
| 71-80 | 4 | 4 | 1 | 4 | 3 | 6 |
| 81-90 | 0 | 0 | 3 | 2 | 1 | 2 |
| Score range | 54~76 | 62~72 | 48~90 | 48~87 | 60~81 | 58~84 |
| Average Score | 67. 6 | 67.4 | 65.8 | 68.5 | 69.6 | 73.4 |
| Over 5 point improvement | 38% | | 67% | | 45% | |
| Standard Dev. | 7. 65 | 4. 24 | 15.1 | 10.98 | 6.25 | 7.43 |

The average score for the 2004 group slightly decreased in the second test. This may be due to the number of the participants (four students = 50%) whose scores went down, although three students increased more than 5 points in the second test. The 2005 group seems to have the largest number of people who gained more than 5 points (eight students), and those who scored low (41-50 points) improved greatly (about 11 points). This group had the lowest scored person – whose major was in science. Also this group had the largest variation of scores as seen in the standard

deviations (Beginning 15.1, End 10.98). The 2006 group scored the highest average points in both first and second tests. The average improvement of the first and second scores was 1.77. No statistically significant improvement was found in the participants' MTELP scores. This is expected because it is difficult to see the actual improvement on one's English ability when the program is shorter than one month (e.g., Day 1987, Milleret 1990). We observed that these participants used the first week to adjust to their lives in a new environment, e.g., culture, daily life style, dorm life, use of facilities such as libraries and computer centers, and jetlag. Realistically speaking, they seemed to start using English more from the second week. Thus, the program duration to improve the language ability is naturally short. Furthermore, the program was not designed to increase test scores. The fact that the scores improved slightly in such a short time might still be worth noting, given the fact that no specific curriculum was implemented to improve the MTELP scores. Furthermore, this was a multi-level class in which specific skill level exercises were not offered. The relatively low increase of the scores appears similar to the results found in Hashimoto (1992) and Tanaka and Ellis (2003), although the scores of the MTELP and TOEFL are not necessarily comparable directly.

## 4.2 ITWE

Table 2 shows a summary of the ITWE scores during 2004-2006. The maximum score of the ITWE was 6 points. The lowest score was 3 in the first test every year and the highest score was 5+ in the second test of 2006. To calculate the average scores, 3- was considered as 2.7 and 3+ was 3.3. The lowest average score (3.6) was reported in the first test of 2006 and the highest average score (4.6) was reported in the second test of 2006. The average score improved after the program every year.

Table 2. Summary of the ITWE scores in 2004-2006

| Scores | 2004 Beginning (n=8) | End | 2005 Beginning (n=12) | End | 2006 Beginning (n=11) | End |
|---|---|---|---|---|---|---|
| 1~1+ | 0 | 0 | 0 | 0 | 0 | 0 |
| 2-~2+ | 0 | 0 | 0 | 0 | 0 | 0 |
| 3-~3+ | 2 | 2 | 2 | 1 | 8 | 1 |
| 4-~4+ | 6 | 5 | 7 | 6 | 2 | 2 |
| 5-~5+ | 0 | 1 | 3 | 5 | 1 | 8 |
| 6-~6 | 0 | 0 | 0 | 0 | 0 | 0 |
| Score range | 3~4+ | 3~5- | 3~5 | 3+~5 | 3~5 | 3+~5+ |
| Average | 3. 9 | 4 | 4.2 | 4.4 | 3.6 | 4.6 |
| More than 2 cat. improved | 38% | | 33% | | 82% | |
| Standard Dev. | 0. 55 | 0. 57 | 0.64 | 0.56 | 0.62 | 0.54 |

In 2004, no one scored 5 in the first test, but there was one student who scored 5- in the second test. No one improved more than 2 categories (e.g., 3+ to 4) and the average score improved only slightly. The standard deviation increased slightly because the number of people who scored 3 did not change, but there was one who improved and scored 5- in the second test. Similarly, two additional students scored 5 in the second test of 2005 and four students (33%) improved more than 2 categories.

The standard deviation became smaller because those who scored low in the first test improved in the second test. A statistically significant improvement was found in the 2005 group (t(11), -2.307, p=0.042). The 2006 group improved the most among the three groups. Nine students (82%) improved more than 2 categories and the standard deviation became smaller. Like 2005, there was a significant improvement in 2006 (t(10), -5.315, p=0.000). The 2005 and 2006 participants included more freshmen compared with the 2004 group. Thus, we conjecture that the younger students learned faster and improved more because they maintained basic foundation skills after their study for the entrance examination. Presumably, English ability would decrease as time goes by after the entrance examination among non-English majors as they lack English contact hours in college.

## 4.3 The characteristics of the 2005 participants' essays

We discuss the 2005 participants' essays as examples. Table 3 shows a summary of the compositions in terms of the average number of words, sentences, and complexity of structures. Because of the daily opportunities to speak in English, we anticipated that these factors would be indicators to measure the improvement of the participants' productivity in English.

Table 3. Summary of the 2005 Essays

|  | Beginning | End |
|---|---|---|
| Word count | 177.00 | 175.00 |
| Number of sentences | 15.00 | 14.50 |
| Word/sentence ratio | 12.16 | 12.36 |
| Simple sentences | 5.58 | 7.58 |
| Infinitives | 0.92 | 1.17 |
| Sentences with（and/but） | 1.00 | 1.42 |
| that-clauses | 2.08 | 0.75 |
| Adjectival clauses | 0.33 | 1.75 |
| Adverbial clauses | 1.92 | 0.17 |
| Wh-questions | 0.58 | 0.83 |
| More than 2 adverbial clauses | 2.58 | 0.83 |

However, a close examination of the 2005 essays revealed that the average number of words per essay, the average number of sentences per essay, and the complexity of sentence structures did not change very much in the first and second compositions. Among the 12 students, the longest sentences in the first and second essays contained about 15.5 words on average. There were three students whose sentences were composed of fewer than 10 words on average (8.27, 8.75, 9.67) in the first essay, but there was only one student whose sentences included under 10 words on average (9.92) in the second essay. This may indicate that they may have become accustomed to producing more words in each sentence after three weeks. Furthermore, the first essays contained ungrammatical incoherent sentences with 2 or 3 adverbial clauses (average 2.58 sentences). For instance, (2) contains three adverbial clauses (two IF-clauses and one EVEN (IF)-clause). Their grammatical knowledge did not seem to have functioned as a monitor.

(2) She doesn't pretend to be happy if she's angry about something, she doesn't pretend to be agree with a certain idea if she doesn't think it's a right thing, even the idea is her friend's.

However, there were very few such ungrammatical sentences observed in the second essays (average 0.83 sentences). Their grammatical knowledge started to function as a monitor.

Next, very few adjectival clauses were used in the first and the second essays (0.33 and 1.75 sentences per essay, respectively). This suggests that it is not difficult for Japanese intermediate and advanced EFL learners to comprehend relative clauses, but it is rather difficult to produce sentences with relative clauses and/or adverbial clauses (cf. Schachter 1974). Examples in (3) show plural, subject-verb agreement, determiner, and categorical errors that are commonly observed among Japanese EFL learners.[9]

(3) a. At last, the qualities that needed for a good parent is how they can tell children to live a real life enjoyable and happiness parent should have both side, gentle and hard(strict) for children.
   b. There parents do not scold their children directly to say "don't do that", instead they say "someone will be angry at you."

Since the program did not provide specific writing lessons, the second essays still contained such errors.[10] However, what changed the most is the organization and coherence of their essays. Let us look at the essays more closely, especially, from a discourse perspective. The following are example essays of one student who improved at least two category scores (i.e., from Score 3 to 4-):

(4) Example
   a. Beginning of the program: <Topic: Ideal parents> (Score 3)

TWE Scoring Guidelines Score 3: Demonstrates some developing competence in writing, but it remains flawed on either the rhetorical or syntactic level, or both.

I think two qualities are necessary for a person to be a good parent. First, we need kindness. My mother is always kind. She is merely angry. She has humor and be laughed at me every time. My father is also kind. I am glad that my parents always think of me.
Second, We need to be strict. My mother is always kind, however, she sometimes scold me. For example, When I want to stay friend's house, she never allow me. She says me, "You should separate good thing to bad thing." It means there're two things that one is I can do, the other thing is I can't do.
If I were mother, I would become to be likes my mother. I think parents should teach their children a social order. Otherwise, children don't know how to grow up.

---

[9] We did not count grammatical errors in the essays because the program did not focus on grammatical writing and we assumed that no significant change would be observed. However, as pointed out by a reviewer, it is important to confirm that this is indeed so.
[10] See Howard (2001) for the effects of study abroad on non-Japanese L2 learners' sentence structures.

If my mother is always kind to me, I would be a bad girl. It is bad to do everything what I want to do.

I can't think of kindness to other person.

So It is necessary to be kind and strict.

I want to be my mother. Sometimes kind, sometimes strict, it's good balance. Parents should believe their children and help for them.


b. End of the program: <Topic: Japan's important event> (Score 4-)

TWE Scoring Guidelines Score 4: Demonstrates minimal competence in writing on both the rhetorical and syntactic levels.


Japanese traditional holiday is New year's day.

We celebrates the new year.

We have a lot of traditional customs in the First January.

First, we go to "jinja" with families or friends. Then we hope our health, steady of home, studying hard, to pass the examination of college or high school, and safety of traffic so on. In jinja, we pull omikuji. Omikuji says what to do in the year in order to come us happiness. Daikichi means the happiest. Daikyo means the worst.

Second, we eat "osechi". It's traditional food in the new year's day. We eat it to spend the year healthy and to take enelgy. Mothers and her daughters cook it. It's also Japanese traditional customs.

Third, we meet families in the new year's day.

Even if my ancle and aunt live so far from my house that I can't meet them freacently, I could meet them in the new year's day. Then we talk about ourselves a lot.

Moreover, children play with "hagoita", "tako" or "karuta". These're traditional toys of new year's day.

New Year's Day is important holiday for Japanese people. I think it's interesting. Because, for a long time, we take over the traditional customs from old generations. We must tell it to next generations. Parhaps, our children also repeat same customs. It's good that all families meet at one place and talk each other.

I'm look forward to meeting them every year. We can meet once a year in the day.

So New Year's day is special day for Japanese people.


The first essay was about ideal parents and the second was about Japanese important events. Due to writing on different topics, there may be some differences with respect to whether one is easier to write than the other. For our purposes, we will disregard the possible influence of topic choice. The first essay did not really have an "Introduction". It started with "I think two qualities are necessary for a person to be a good parent" as a kind of "Introduction" without stating what those two qualities are and how the essay would be concluded. It moved to the main "Body" right away and discussed two conditions for ideal parents from the beginning. After the discussion of the second quality, the conclusion is not clearly stated. Paragraphs are not clearly distinguishable in this essay. As the TWE Scoring Guidelines indicates, the essay demonstrates some developing competence in writing, but it remains flawed on either the rhetorical or syntactic level, or both. On the other hand, the second essay had a clearer organization, though paragraphs were still not separated well. An introduction paragraph starts with "Japanese traditional holiday is New year's day" and "We have a lot of traditional customs in the First January", followed by the description of three

activities while celebrating New Year's Day, and the conclusion. In particular, the Body provided examples to explain the traditional customs stated in the introduction. Since the essay was not quite the level of the TWE Scoring Guidelines Score 4: Demonstrates minimal competence in writing on both the rhetorical and syntactic levels, it was scored 4-. Though slight, the second essay shows improvement in the organization of the essay. Note that both essays contain usage errors like "I can't think of kindness to other person" in (4a) and "Because, for a long time, we take over the traditional customs from old generations" in (4b) as well as misspelling errors like "ancle" and "freacently". Even though the score improved from 3 to 4-, there were still errors in the second essay.

In sum, a well organized essay with the introduction-body-conclusion structure was found more often in the second trial. Also, the second essays were written with more consistent themes. This may be due to the emphasis on discussion and oral presentation because coherence is required. This finding indicates that writing skills can improve even through oral presentation training in such a short-term program.[11] Further research is required to investigate whether or not such writing improvement can be achieved in a similar in-home-country program.

## 5. Conclusion

This paper discussed what contribution a three-week summer English study abroad program may have made to the development of participants' English proficiency. In particular, we raised the following three research questions: a) Can a short-term study abroad program make a contribution to the development of participants' EFL proficiency? b) If yes, then what contribution may be expected? And c) how should such short-term study abroad program be assessed within the Japanese university curriculum? Our answers to these questions are as follows: a) Even a three-week program is meaningful for the improvement of English skills. Although the MTELP scores did not significantly improve, the ITWE scores did. Thus, the program was effective even though it was held for such a short-period. b) In particular, it was found that writing improved significantly with respect to the organization and the coherence of essays, even though the program focused on oral presentation. Oral presentation training provided exercises for forming opinions, making coherent stories, and logical arguments. Thus, these exercises also helped their writing. c) How we place this kind of short-term study abroad program within the university curriculum depends on the answer to further research that is required to investigate whether or not such writing improvement can be achieved in an in-home-country program. If so, a study abroad program would be an option rather than a necessity, for students. In such a case, however, we need to assess other values of study abroad programs, e.g., motivation, understanding of English speaking countries' cultures, etc. It is important not to treat a study abroad program as a special activity, but rather incorporate it into the university's English curriculum effectively. After all, as Ellis (1994:166) writes, "...formal instruction helps learners to develop greater L2 proficiency, particularly if it is linked with opportunities for natural exposure." More studies in this area are needed.

---

[11] See Hawke and Yoshimura (2006) for their findings in training academic presentation skills in Japan.

## References

Brauer, H. E. (2005). Short-Term Overseas Study Programs: A Survey of Private Junior High Schools in Tokyo. *JALT Journal* 27.1, 95-118.

Canale, M. (1983). From Communicative Competence to Communicative Language Pedagogy. In Richards, J. and R. Schmidt (eds.) *Language and Communication.* London: Longman.

Day, J. T. (1987). Student Motivation, Academic Validity, and the Summer Language Program Abroad: An Editorial. *Modern Language Journal* 71. 261-266.

Ellis, J. (1993). Japanese Students Abroad: Relating Language Ability in Class and in the Community. *Thought Currents in English Literature* 66, 45-82.

Ellis, R. (1994). *The Study of Second Language Acquisition.* Oxford: Oxford University Press.

Freed, B. F. (1995). *Second Language Acquisition in a Study Abroad Context.* Amsterdam: John Benjamins Publishing Company.

Freed, B. F. (1998). An Overview of Issues and Research in Language Learning in a Study Abroad Setting. *Frontiers: The Interdisciplinary Journal of Study Abroad* 4. 31-60.

Hashimoto, M. (1992). Pre-TOEFL no kekka o tooshite mita Kanagawa daigakusei no tanki eigo kenshuu no seika. [Effects of Kanagawa University's Short-term Study Abroad: A Report Based on the Results of Pre-TOEFL]. *Kokusai Keiei Forum* 4, 114-120.

Hawke, P. and N. Yoshimura (2006). From Everyday Conversations to Scholarly Presentations: Developing Learners' Academic Language Proficiency. *Ars Linguistica* 13, 21-41.

Howard, M. (2001). The Effects of Study Abroad on the L2 Learner's Structural Skills. In S. Foster-Cohen and A. Nizegorodcew (eds.) *EUROSLA Yearbook* 1, 123-141. Amsterdam: John Benjamins Publishing Company.

Hymes, D. H. (1972). On Communicative Competence. In J. B. Pride and J. Holmes (eds.) *Sociolinguistics*, 269-293. London: Penguin.

Kitao, S. K. (1993). Preparation for and Results of a Short-term Overseas Study Program in the United States. *Bulletin of the Institute for Interdisciplinary Studies of Culture* 10, 107-118. Doshisha Women's College.

Marriott, H. E. (2000). Japanese Students' Management Processes and Their Acquisition of English Academic Competence during Study Abroad. *Journal of Asian Pacific Communication* 10.2, 279-296.

Milleret, M. (1990). Evaluation and the Summer Language Program Abroad: A Review Essay. *Modern Language Journal* 74, 483-488.

Sawasaki, K. and N. Yoshimura. (2005). Kaigai tanki eigo kenshuu-no seika-o donoyooni hyooka-suruka [How Do We Evaluate a Short-term English Study Abroad Program?]. A paper presented at the 51st meeting of the Linguistic Circle of Central Japan (December 17, 2005, University of Shizuoka, Japan).

Sawasaki, K., N. Yoshimura, M. Nakayama and Y. Terao. (2006). Tanki kaigai kenshuu program-o hyooka-suru: Eigoryoku-no koojoo-o chuushin-ni [Evaluating the Effects of a Short-term Study Abroad Program on Students' English]. A poster presented at the 8th Annual Conference of the Japanese Society for Language Sciences (June 11, 2006, International Christian University, Japan).

Schachter, J. (1974). An Error in Error Aanalysis. *Language Learning* 24, 205-214.

Suzuki, Y. (2001). Kaigai eigo kenshuu-no igi: Cambridge University-de manabu kyuudaisei [The Meaning of English Study Abroad: Kyuushuu University Students at Cambridge University]. *Gengo Bunkaron Kenkyuu* 13, 163-174. Institute for Language and Culture, Kyuushuu University.

Takazawa, Y. (1995). English Language Study Tour Abroad: Maximizing Its Effectiveness in Helping Students Improve Communicative Competency (Japanese). *Kokusai Kyoyoogaku Ronshuu* 5, 1-27. Chiba Keiai Junior College.

Takeda, A. (1998). Kaigai eigo gakushuu-ni okeru gakushuu dooki-to gakushuu strategy-ni kansuru ichi koosatsu [Remarks on Motivation and Learning Strategy in an English Study Abroad Program]. *Bulletin of Mukogawa Women's University* 46, 11-18.

Tanaka, K. and R. Ellis. (2003). Study-abroad, Language Proficiency, and Learner Beliefs about Language Learning. *JALT Journal* 25. 1, 63-85.

Tateyama, E. (2002).The Response of Japanese Nursing Students to a Vacation English Program Abroad (ERIC Document Reproduction Service No. ED472141).

Yoshimura, N., M. Nakayama and K. Sawasaki. (2006). Tanki kaigaikenshuu program-to eigoryoku-no koojoo [A Short-term Study Abroad Program and the Improvement of English Skills]. *Ars Linguistica* 13, 92-103.

Yoshimura, N., K. Sawasaki, Y. Terao and M. Nakayama. (2006). On demand kaigai eigo kenshuu no kokoromi [Creating an English Study Abroad Program]. *Journal of International Relations and Comparative Culture* 3. 2, 293-308. University of Shizuoka.

# Topic and Focus in Cantonese

Cauvis Suet Man FUNG

The School of Humanities
University of Hong Kong

## Abstract

Topicalization is a common syntactic phenomenon in Cantonese. A topicalized phrase can bear either the TOPIC or the FOCUS function. This paper investigates these two grammaticized discourse functions in Cantonese, based on the definition of topicalization in Lexical-Functional Grammar (LFG) which constrains them to being identified with a missing function in the sentence. The information features, [New] and [Prom] (Choi's 2001), are employed to characterize discourse focus. It is proposed that [+New] is the feature which licenses FOCUS topicalization in Cantonese and FOCUS is defined as a topicalized phrase which contains some [+New] information. An LFG account is presented to show how Functional Uncertainty characterizes TOPIC and FOCUS in Cantonese. The Extended Coherence Condition is revised to accommodate external topic (E-TOP) and it is proposed that E-TOP conforms to the Extended Coherence Condition by having a PRED in its immediate f-structure.

## 1. Introduction

The syntactic phenomenon 'topicalization' has been given various definitions (e.g. Krapova 2004, Lambrecht 1994, Rosén 1998). This study examines the phenomenon based on Rosén's (1998) definition which conforms to the assumption in Lexical-Functional Grammar (LFG). Li and Thompson's (1981) definition of topic in Mandarin will be examined and it will be suggested how the definition can be revised to characterize topic in Cantonese. Based on Aissen's (1992) and King's (1995) study, Cantonese topic will be categorized into internal topic (TOPIC) and external topic (E-TOP). It will be investigated what licenses FOCUS topicalization in order to propose a definition of FOCUS in Cantonese. An LFG account will be presented to show how the functional device – Functional Uncertainty – accounts for Cantonese topicalization and how external topic (E-TOP) can be accommodated in the Extended Coherence Condition.

This paper is organized as follows. After introducing the study in this section, section 2 then presents the definition of topicalization. Section 3 defines topic in Cantonese and differentiates between two types of topics, internal topic and external topic. Section 4 examines what licenses FOCUS topicalization and defines Cantonese FOCUS. In section 5, an LFG account is given to topicalization and E-TOP constructions. Section 6 concludes the paper.

## 2. Topicalization

Rosén's (1998: 184) definition of topicalization is adopted in the present study, which states that '[t]opicalization is a construction in which a leftmost constituent is understood as filling a missing constituent in the sentence'. Sentences (1) and (2)

below are examples of topicalization in English and Cantonese respectively according to this definition[1].

(1)     ***That book,*** *I like.*

(2)     ***nei1 bun2 syu1***     *ngo5*     *zung1ji3*[2]
        DEF CL     book     1.SG     like
        'As for this book, I like it.'     or
        'It is this book that I like.'

The leftmost constituents *that book* and *nei1 bun2 syu1* fill the missing OBJ function of the verb *like* and *zung1ji3* respectively. A topicalized phrase can bear either the TOPIC or the FOCUS function ('TOPIC' and 'FOCUS' are used to refer to *topicalized* discourse topic and *topicalized* discourse focus respectively). In Cantonese, a topicalized phrase followed by no particle can be interpreted as either of the functions depending on the context and the intonational pattern of the utterance. Therefore, with *nei1 bun2 syu1* followed by no particle, the construction has two possible interpretations. Particles in Cantonese specify different speech-acts, evidentiality and emotions of the speaker (Matthews and Yip 1994). Some particles tend to follow a TOPIC and some a FOCUS. In this paper, *aa4* and *ne1* will be used for the TOPIC function and *aa3* for the FOCUS function.

Though a topicalized phrase normally occurs in the leftmost position of a sentence, there are instances where it does not, for example, when there are multiple instances of topicalization in one sentence. Consider (3).

(3)     ***keoi5***     *aa3*     ***bun2 syu1***     *ngo5*     *bei2zo2*
        3.SG     PART CL     book 1.SG     give.PERF
        'It is him/her who I have given the book to.'

In this example, the OBJ₀ *keoi5* and the OBJ *bun2 syu1* are topicalized to become the FOCUS (note the particle *aa3* for the FOCUS function) and the TOPIC (a TOPIC is usually followed by no particle with a FOCUS preceding it) respectively. The FOCUS precedes the TOPIC so that the TOPIC is no longer sentence-initial. *Bun2 syu1* is still regarded as a TOPIC even though it is not in the left periphery since it is extracted from its canonical post-verbal position to the pre-SUBJ position.

---

[1] Symbols and abbreviations used:
1 = First person; 2 = Second person; 3 = Third person; ADJ = Adjunct; AP = Adjective phrase; AdvP = Adverb phrase; CL = Classifier; CLP = Classifier phrase; COMP = Completive aspect; DEF = Definiteness; DEF = Definite determiner; D = Determiner; DP = Determiner phrase; E-TOP = External topic; FOCUS = Topicalized discourse focus; IP = Inflectional phrase; New = Newness; NP = Noun phrase; NUM = Number; OBJ = Primary Object; OBJ₀ = Secondary Object; PART = particle; PERF = Perfective aspect; PERS = Person; PP = Prepositional phrase; PRED = Predicate; Prom = Prominence; QP = Quantifier phrase; SG/SG = Singular; SUBJ = Subject; TOPIC = Topicalized discourse topic/internal topic; VP = Verb phrase
[2] All grammatical Cantonese examples consist of three lines. The first line shows the romanization of the Chinese characters, employing the Jyutping system (Linguistic Society of Hong Kong 2003). The second line provides the word-for-word translation. The third line is the free English translation of the construction. No translation is given for ungrammatical examples.

## 3.  Topic in Cantonese

Topic structures of Chinese, mainly Mandarin, have been widely researched (e.g. Li 2001, Li and Thompson 1981, Lu 2000, Pan and Hu 2003, Xu 1994, Xu and Liu 1998, Zhang 1996). The same structure in Cantonese has however received comparatively less attention. In this section, Li and Thompson's (1981) definition of Mandarin topic is revised to give a definition to Cantonese topic (section 3.1) and two types of topics, internal topic (TOPIC) and external topic (E-TOP), are identified (section 3.2).

### 3.1  Defining topic in Cantonese

According to Li and Thompson's (1981) study on topic in Mandarin, a topic is 'typically a noun phrase (or a verb phrase) that names what the sentence is about, is definite or generic, occurs in the sentence-initial position, and may be followed by a pause or a pause particle' (1981: 87). They adopt Chafe's (1976) definition of Chinese-style topic to explain 'what the sentence is about'. A topic 'limit[s] the applicability of the main predication to a certain restricted domain' and 'sets a spatial, temporal, or individual framework within which the main predication holds' (Chafe 1976: 50). Though Li and Thompson's (1981) definition of topic seems to be compatible with Cantonese topic in many ways, it still requires some revision.

Li and Thompson (1981) mention that a topic is typically a noun phrase or a verb phrase. While noun and verb are also typical categories for topics in Cantonese, topics of other categories can also be commonly found in the language. It will be shown in the following which types of phrase can be a topic in Cantonese.

Example (2) above is an example of a DP being the topic of a sentence. Deleting the D resulting in a CLP topic (*bun2 syu1 ngo5 zung1ji3*), or deleting both the D and the CL resulting in an NP topic (*syu1 ngo5 zung1ji3*) are also possible. A QP can also act as the topic of a sentence as in the following example:

(4)  **saam1  bun2 syu1**  *ne1*  *ngo5*  *zau6*  *mou5*  *laa3*
     three  CL  book  PART  1.SG  zau  no  PART
     'As for three books, I haven't got that many.'

Verbs and adjectives have been proved legitimate to be a topic in Cantonese (Matthews and Yip 1994). Li and Thompson (1981) prove for Mandarin that a topic can be a temporal or locative phrase (see Tsao 1979 and Xu and Liu 1998 for similar arguments). Similar examples can also be found in Cantonese. In (5) below, the topic is the AdvP *kam4jat6* (*yesterday*) referring to a definite time, which sets the temporal framework for the predication *ngo5 heoi3zo2 tai2 hei3* (*I went to see a movie*) to hold:

(5)  **kam4jat6** *ngo5 heoi3zo2*  *tai2 hei3*
     yesterday 1.SG go.PERF  see  movie
     'I went to see a movie yesterday.'

Xu and Liu (1998) suggest that a small clause can also be the topic of a sentence in Mandarin. This is also true in Cantonese. Sentence (6) shows an example where the topic *keoi5 tai2jyun4 bun2 syu1* is an IP.

(6)     **keoi5    tai2jyun4    bun2 syu1**    ngo5 m4  seon3
          3.SG     read.COMP   CL    book    1.SG not believe
          'That s/he has finished reading the book, I don't believe.'

A PP can also occur in the sentence-initial position as the topic. Sentence (7) is an example where the PP *hai2 gung1jyun2* is a topic.

(7)     **hai2 gung1jyun2**  ngo5    gin3dou2    go3  baak3baak3
          in    park    1.SG    see      CL  old man
          'In the park, I saw an old man.'

Based on the above observations, it is concluded that a Cantonese topic can be a DP, a CLP, an NP, a QP, a VP, an AP, an AdvP, an IP or a PP. In other words, phrases of any category can be topics in Cantonese.

Regarding the structural position of a topic, while it often occurs at the left periphery of a sentence, there are exceptions as mentioned in section 2 when, for instance, TOPIC and FOCUS co-exist in a sentence.

Given Li and Thompson's (1981) definition of topic in Mandarin and the above observations, Cantonese topic is concluded as a phrase of any category which 'limit[s] the applicability of the main predication to a certain restricted domain' and 'sets a spatial, temporal, or individual framework within which the main predication holds' (Chafe 1976: 50), *usually* occurs in the sentence-initial position and is optionally followed by a pause or a pause particle.

### 3.2  Internal topic vs. external topic

Consider the definition of Cantonese topic and Rosén's (1998) definition of topicalization. There is one crucial requirement on TOPIC (a topicalized discourse topic) which is not stated in the definition of Cantonese topic – a Cantonese topic is not bound to be identified with a missing constituent. In other words, a Cantonese topic is not necessarily a TOPIC. Based on Aissen's (1992) analysis of the three Mayan languages, Tzotzil, Jakaltek and Tz'utujil, and King's (1995) study on Russian, it is proposed here that there are two types of topics in Cantonese, internal topic and external topic. Consider the following examples where the sentence-initial phrase in (8) is an internal topic and the one in (9) is an external topic:

(8)     **bun2 syu1**    aa4    ngo5    tai2jyun4    laa3
          CL    book    PART   1.SG    read.COMP  PART
          'As for the book, I have finished reading it.'

(9)     **nei1 gin6 si6**   ne1    keoi5    dik1kok3 cung1dung6zo2   di1
          DEF CL  matter PART  3.SG    indeed   impetuous.PERF  a bit
          'As for this matter, it is indeed quite impetuous of him/her (to act like that).

In (8), the sentence-initial CLP *bun2 syu1* fills the missing OBJ function of the following clause and thus conforms to the definition of topicalization. In (9), the PRED *cung1dung6* subcategorizes only for a SUBJ and the construction without the sentence-initial DP *nei1 gin6 si6* is already well-formed, with the DP *keoi5* being the

SUBJ. *Nei1 gin6 si6* cannot be identified with any missing function in the clause and the construction is therefore not an instance of topicalization.

Aissen (1992) identifies two types of topic in the Mayan languages she investigates, which are internal topic[3] and external topic. An internal topic is in fact a TOPIC. It has a closer syntactic connection with the following clause than an external topic. An external topic, or an E-TOP according to King (1995), is 'less integrated into basic clause structure, being essentially prefixed to what is otherwise a fully well-formed clause' (Aissen 1992: 44). Unlike TOPIC, an E-TOP is not required to bind an argument, though they can be coreferent with a pronoun. There can be pragmatic connection but no syntactic connection between the E-TOP and the following clause. It should be noted that, since Cantonese is a pro-drop language (Lee 2003), a 'fully well-formed clause' does not necessarily demand that all arguments be present in the clause. A clause is still well-formed even if some arguments are missing provided that they can be recovered from the context.

Consider (9) again. The sentence-initial constituent *nei1 gin6 si6* does not bind any arguments and has no syntactic relation with the following already well-formed clause. Its relation to the clause is only pragmatic in that it restricts the applicability of the predication *keoi5 dik1kok3 cung1dung6zo2 di1* to the scope of *nei1 gin6 si6*. In other words, it is for this particular matter that the addressee is regarded as being impetuous and this utterance is not commenting on the general personality of the addressee. Given these observations, it is concluded that the sentence-initial DP *nei1 gin6 si6* is an E-TOP.

To summarize this section, a Cantonese topic subsumes TOPIC and E-TOP. A TOPIC is a phrase of any category being identified with a missing function in the clause which 'limit[s] the applicability of the main predication to a certain restricted domain' and 'sets a spatial, temporal, or individual framework within which the main predication holds' (Chafe 1976: 50). An E-TOP is not required to fill a missing constituent.

## 4. FOCUS in Cantonese

In this section, the information features of FOCUS will be discussed and the feature which licenses FOCUS topicalization will be identified in an attempt to define FOCUS in Cantonese.

Languages use different means to encode discourse focus. Green and Jaggar (2003) suggest the following morphosyntactic and prosodic means and observe that languages often make use of more than one of the following options:

---

[3] The internal topic here should be differentiated from the internal topic in Paul (2002, 2005). While Paul regards a preposed object appearing between the SUBJ and the verb as a type of internal topic, the internal topic suggested here occurs in the leftmost position of a sentence.

(10)  a.  Focus *in situ*
      b.  Focus movement[4]
          -  Clause-initial
          -  Pre-verbal
          -  Post-verbal
      (c)  Focus markers
      (d)  Focus stress

In the following, it will be investigated when and how Cantonese makes use of topicalization (clause-initial focus movement in Green and Jaggar's sense) to encode discourse focus.

According to Choi (1997, 1999, 2001), [+New] is the characterizing feature of discourse focus. There are two types of discourse focus including contrastive focus which represents [+New] and [+Prom] information, and completive focus which represents [+New] but [-Prom] information, where [+New] refers to new, informative or unknown information and [+Prom] refers to important, urgent, unexpected, contrastive or surprising information (Choi 2001). Choi (1999) shows that [+Prom] is the motivation for topicalization in English. In order to define FOCUS in Cantonese, it will be investigated in the following what licenses FOCUS topicalization in the language.

The phenomenon of Gap Right Dislocation (Gap RD) in Cantonese, which refers to a kind of leftward movement in which a constituent is moved from the end to the initial position of a sentence (Cheung 1997), has the same sentence structure with certain types of FOCUS topicalization. When the FOCUS is identified with a constituent of which the canonical position is at the end of the sentence, the topicalization construction has the same structure as Gap RD. The Gap RD phenomenon in Cheung (1997) is studied within a derivational approach, contrasting with the LFG approach taken in the present study (see section 5). Regarding information structure, Cheung (1997) proposes that the left-dislocated phrase is the domain of focus where 'the function of focus is to indicate the presence of alternatives to the expression in focus that are under discussion' (Krifka 1997 cited in Cheung 1997: 95). FOCUS is a topicalized phrase designating the discourse focus (Bresnan and Mchombo 1987). The present study however does not employ the concept of 'alternatives' to encode discourse focus since a TOPIC may also involve alternatives as illustrated in the following example:

(11)  *neil bun2*  *syul*  *nel*  *ngo5*  *zau6*  *maai5zo2,*
      DEF CL      book    PART   1.SG    zau     buy.PERF

      *go2 bun2*  *syul*  *nel*  *ngo5*  *zau6*  *mei6*  *maai5*
      DEF CL      book    PART   1.SG    zau     not yet  buy

'As for this book, I have bought it. As for that book, I haven't bought it yet.'

In this example, *neil bun2 syul* and *go2 bun2 syul*, followed by the topic particle *nel*, are TOPICs. At least two alternatives are involved here, *this book* and *that*

---

[4] Within the framework of LFG, this phenomenon is not analyzed as a kind of movement.

*book*. Therefore, 'alternatives' is not an exclusive property of discourse focus. This study adopts Choi's (1999) [+New] feature to characterize discourse focus, employing Lambrecht's (1994) concept of 'new' and Halliday's (1967) idea of 'informative' to interpret '[+New]'.

A phrase carries new or unknown information if 'the denotatum of this phrase stands in a pragmatically construed relation to the proposition such that its addition makes the utterance of the sentence a piece of new information' (Lambrecht 1994: 210). A piece of information is regarded as informative if the speaker intends it 'to be interpreted as informative' (Halliday 1967: 204). Question-answer pairs can be used to identify the discourse focus of a sentence (Rochemont and Culicover 1990):

(12)    Q:    *nei5  maai5zo2  **me1**  aa3?*
               2.SG buy.PERF   what PART
               'What have you bought?'

        A:    **saam1   bun2 zi6din2**    *aa3   ngo5 maai5zo2*
               three      CL    dictionary PART 1.SG buy.PERF
               'It is three dictionaries that I have bought.'

By asking *nei5 maai5zo2 me1 aa3?* (*what have you bought?*), the addresser assumes that the addressee has bought something. *Saam1 bun2 zi6din2* in the answer then represents new information in the sense that its addition to the presupposition *I have bought something* makes the utterance a piece of new information. This piece of information is also presented by the speaker as the most informative part of the whole utterance. Therefore, this phrase bears [+New] information and is thus the discourse focus according to Choi's (1999) classification. A topicalized phrase which designates the discourse focus is a FOCUS.

According to what is defined as 'new', 'unknown' and 'informative' above, an element can still be [+New] even if it has been introduced before (Lambrecht 1994). Consider the following question-answer pair:

(13)    Q:    *nei5  maai5zo2  **zi6din2**    ding6     zaap6zi3  aa3?*
               2.SG buy.PERF   dictionary or       magazine PART
               'Have you bought dictionaries or magazines?'

        A:    **zi6din2**    *aa3   ngo5   maai5zo2*
               dictionary PART 1.SG   buy.PERF
               'It is dictionaries that I have bought.'

Although *zi6din2* is already mentioned in the question, *zi6din2* in the answer should still be regarded as bearing [+New] information since it makes the utterance new by stating that it is dictionaries, instead of magazines, that the speaker has bought.

While the grammatical function assigned to a particular element is fixed, the informational status assigned can differ from speaker to speaker (Choi 2001). For instance, while a speaker regards an element as [+Prom], another speaker under the

same context can regard it as [-Prom]. Consider (12) again. If buying three dictionaries is not regarded as surprising or unexpected, *saam1 bun2 zi6din2* represents [+New] and [-Prom] information and thus designates a completive discourse focus. If buying three dictionaries is considered surprising or unexpected, the phrase bears [+New] and [+Prom] features and therefore designates a contrastive discourse focus. Whether carrying [+Prom] or [-Prom] information, the phrase *saam1 bun2 zi6din2* can still be topicalized. The following example further illustrates that a FOCUS can be either [+Prom] or [-Prom]:

(14)   Q1:  *nei5   maai5zo2   zaap6zi3   aa3?*
             2.SG buy.PERF   magazine   PART
             'You have bought magazines?'

       Q2:  *nei5   maai5zo2   **me1**   aa3?*
             2.SG buy.PERF   what PART
             'What have you bought?'

       A:   ***zi6din2***   *aa3   ngo5   maai5zo2*
             dictionary PART 1.SG   buy.PERF
             'It is dictionaries that I have bought.'

The topicalized phrase in (14A) carries [+New] information with respect to both questions (14Q1) and (14Q2) and is therefore a FOCUS. With question (14Q1), the FOCUS contrasts with *zaap6zi3* in the question and is thus at the same time [+Prom] (i.e., a contrastive focus). With question (14Q2), the FOCUS bears [-Prom] information if it is assumed that buying dictionaries is not a piece of surprising information. Since the topicalization construction (14A) is a legitimate answer to both questions, it is concluded that the [Prom] feature does not affect the acceptability of FOCUS topicalization and it is the [+New] feature which licenses FOCUS topicalization.

Now the [New] feature will be further examined to see how it characterizes FOCUS in Cantonese. In the above examples, the FOCUS designates the discourse focus, that is the [+New] information. A topicalized phrase can include at the same time [+New] and [-New] information as in the following example:

(15)   Q:   *nei5   maai5zo2   saam1 bun2 **me1**   aa3?*
             2.SG buy.PERF   three   CL   what PART
             'What are the three books that you have bought?'

       A:   *saam1   bun2 **zi6din2**   aa3   ngo5 maai5zo2*
             three   CL   dictionary PART 1.SG buy.PERF
             'It is three dictionaries that I have bought.'

Given the question, the presupposition here is that the addressee has bought three books. *Zi6din2* is therefore the [+New] information as its addition to the presupposition makes the utterance a piece of new information. The topicalized phrase contains both [+New] (*zi6din2*) and [-New] (*saam1 bun2*) information. This phrase should still be regarded as a FOCUS since it contains the [+New] information. Therefore, it is proposed that a FOCUS in Cantonese is a topicalized phrase which

*contains* the discourse focus (i.e., [+New] information).

While a FOCUS must contain some [+New] information, it is not constrained to include *all* [+New] information. Consider the following example:

(16)    Q:   *nei5  maai5zo2  **me1**  aa3?*
              2.SG buy.PERF   what PART
              'What have you bought?'

       A:   ***zi6din2***    *aa3  ngo5 maai5zo2  **saam1**   **bun2***
              dictionary PART 1.SG buy.PERF   three     CL
              'I have bought three DICTIONARIES'

The question presupposes that the addressee has bought something. *Saam1 bun2 zi6din2* is what makes the utterance new and is thus the [+New] information. Only part of the [+New] information *zi6din2* is topicalized to be the FOCUS, with *saam1 bun2* in its canonical postverbal OBJ position. This example demonstrates that it is not obligatory for the entire phrase corresponding to the discourse focus to appear inside the FOCUS.

To conclude this section, [+New] is the feature which licenses FOCUS topicalization in Cantonese and Cantonese FOCUS is defined as a topicalized phrase which contains the entire or part of the discourse focus ([+New] information).

## 5.   Characterizing Topicalization in LFG

Most of the studies on topic in Mandarin are descriptive (e.g. Zhang 1996) or done within a derivational approach (e.g. Li 2001, Pan and Hu 2003, Xu 1994). The weaknesses of the movement analysis of the syntactic phenomenon have been discussed (e.g. Lu 2000, Xu and Liu 1998). Lexical-Functional Grammar (LFG) is a framework which is able to provide a nonderivational account of topicalization. Her (1991) and Huang (1992) have analyzed the TOPIC constructions in Mandarin within this framework. Similar research on Cantonese has however not yet been documented. In this section, an LFG account will be presented for topicalization in Cantonese.

In LFG, TOPIC and FOCUS are regarded as grammaticized discourse nonargument functions (Bresnan 2001) which are subject to the Extended Coherence Condition (Bresnan and Mchombo (1987: 746):

(17)    <u>Extended Coherence Condition</u>: [TOPIC and FOCUS must] be linked to the semantic predicate argument structure of the sentence in which they occur, either by functionally or anaphorically binding an argument.

It will be demonstrated in the following how TOPIC and FOCUS are treated in the framework of LFG and how the two functions in Cantonese can be analyzed.

Functional Uncertainty is the formal device in LFG for characterizing topicalization (Kaplan and Zaenan 1989). Consider the following functional uncertainty equation:
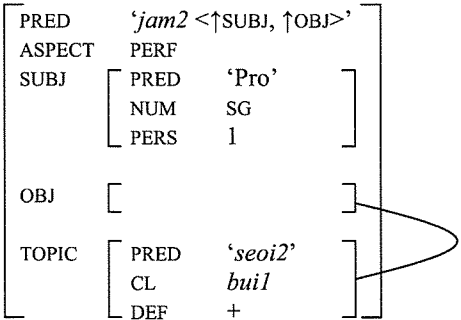
(18)     $(\uparrow\text{TOPIC}) = (\uparrow\text{OBJ})$

The equation states that the TOPIC of the clause is at the same time the OBJ of the clause. Sentence (19) illustrates how this equation applies to Cantonese topicalization constructions. For simplicity reason, I assume here that the topicalized phrase is a TOPIC (in fact it can be interpreted as either a TOPIC or a FOCUS depending on the context and the intonational pattern).

(19)     **bui1   seoi2   ngo5   jam2zo2**
         CL     water   1.SG   drink.PERF
         'As for the cup of water, I have drunk it.'

In this construction, the TOPIC *bui1 seoi2* is at the same time the OBJ of the clause. The following f-structure shows the identification of the two functions:

(20)     F-structure of (19)

$$
\begin{bmatrix}
\text{PRED} & \text{'}jam2 <\uparrow\text{SUBJ}, \uparrow\text{OBJ}>\text{'} \\
\text{ASPECT} & \text{PERF} \\
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'Pro'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 1 \end{bmatrix} \\
\text{OBJ} & \begin{bmatrix} \quad\quad\quad\quad\quad \end{bmatrix} \\
\text{TOPIC} & \begin{bmatrix} \text{PRED} & \text{'}seoi2\text{'} \\ \text{CL} & bui1 \\ \text{DEF} & + \end{bmatrix}
\end{bmatrix}
$$

The curved line connecting the TOPIC and the OBJ represents the unification of the two functions. Therefore, the Completeness condition which states that 'every function designated by a PRED should be present in the f-structure of that PRED' and 'if a designator ($\uparrow$GF) is associated with a semantic role by the PRED, the f-structure element satisfying the designator must itself contain a semantic feature [PRED $v$]' (Bresnan 2001: 63) is still observed though the OBJ function which is subcategorized for by the PRED *jam2* does not have the PRED feature because its PRED feature is provided by the TOPIC. The Extended Coherence Condition is also satisfied with the TOPIC being identified with the OBJ.

In example (16A) above, only part of the OBJ is topicalized to become the FOCUS. The following functional uncertainty equation shows the identification between the FOCUS and the PRED of the OBJ:

(21)     $(\uparrow\text{FOCUS}) = (\uparrow\text{OBJ PRED})$

The f-structure below is proposed to show how this kind of construction can be represented:

(22)    F-structure of (16A)

$$
\begin{bmatrix}
\text{PRED} & \text{'}maai5 <\uparrow\text{SUBJ}, \uparrow\text{OBJ}>\text{'} \\
\text{ASPECT} & \text{PERF} \\[4pt]
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'Pro'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 1 \end{bmatrix} \\[20pt]
\text{OBJ} & \begin{bmatrix} \text{PRED} & \\ \text{CL} & bun2 \\ \text{ADJ} & [\text{PRED} \ \text{'}saam1\text{'}] \end{bmatrix} \\[20pt]
\text{FOCUS} & \begin{bmatrix} \text{PRED} & \text{'}zi6din2\text{'} \end{bmatrix}
\end{bmatrix}
$$
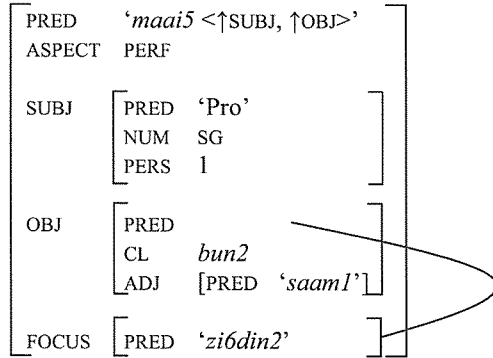
In this f-structure, the FOCUS is linked to the empty PRED of the OBJ, while the CL and the ADJ remain in the f-structure of the OBJ. Compare this account with Kuhn's (1999) account of split NPs in German. Kuhn (1999) links in the f-structure an empty TOPIC to the OBJ. The present study 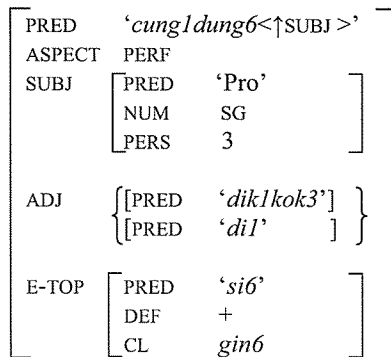assumes that a topicalized constituent appears inside the f-structure of the TOPIC/FOCUS. Therefore, with *zi6din2* being topicalized, it appears inside the f-structure of the FOCUS and *saam1* and *bun2* which remain in the canonical OBJ position appear inside the f-structure of the OBJ.

Now consider example (9) again, repeated here as (23):

(23)    ***nei1 gin6 si6***    *ne1*    *keoi5*    *dik1kok3 cung1dung6zo2*    *di1*
        DEF CL    matter PART    3.SG    indeed    impetuous.PERF    a bit
        'As for this matter, it is indeed quite impetuous of him/her (to act like that).'

It has been argued in section 3 that the leftmost constituent is an E-TOP, instead of a TOPIC. The f-structure of this construction is shown below:

(24)    F-structure of (23)

$$
\begin{bmatrix}
\text{PRED} & \text{'}cung1dung6<\uparrow\text{SUBJ}>\text{'} \\
\text{ASPECT} & \text{PERF} \\
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'Pro'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \end{bmatrix} \\[20pt]
\text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'}dik1kok3\text{'} \end{bmatrix} \\ \begin{bmatrix} \text{PRED} & \text{'}di1\text{'} \end{bmatrix} \right\} \\[16pt]
\text{E-TOP} & \begin{bmatrix} \text{PRED} & \text{'}si6\text{'} \\ \text{DEF} & + \\ \text{CL} & gin6 \end{bmatrix}
\end{bmatrix}
$$

In order to decide whether this is a legitimate f-structure, the requirements on E-TOP have to be investigated. E-TOP, being a grammatical function, should be subject to the Extended Coherence Condition. The Condition however only states the

requirements on argument functions and TOPIC, FOCUS and ADJ (Bresnan 2001), without mentioning how an E-TOP can be integrated into the f-structure. The Extended Coherence Condition has to be revised in order to accommodate the E-TOP function. Since E-TOP is not subcategorized for by the PRED and is not required to bind any argument, it is proposed that an E-TOP, like an ADJ, should be required to have a PRED in its immediate f-structure. The immediate f-structure of the E-TOP in (24) is the matrix f-structure, which contains the PRED *cung1dung6*. Therefore, this is a legitimate f-structure.

While it is possible for a sentence-initial phrase which is not linked to any missing function in the clause to be interpreted as an E-TOP, the semantics of the clause has to be taken into consideration in order to determine the well-formedness of a construction. Consider (25).

(25)    ***bui1 seoi2***    *ngo5 jam2zo2*    *bui1 caa4*
        CL   water    1.SG drink.PERF   CL   tea

Similar to example (23), there is an extra constituent in front of an already well-formed clause. This example is however ungrammatical. In order to decide the grammaticality, it has to be determined whether the sentence-initial phrase conforms to the requirements on a topic. Consider (23). The translation of the construction shows that the sentence-initial phrase *nei1 gin6 si6* can be interpreted as specifying the scope within which the main predication, *keoi5 dik1kok3 cung1dung6zo2 di1*, holds. As mentioned in section 3, the phrase specifies that it is for this particular matter that the addressee is being impetuous. Whether he or she is generally impetuous is not known. Now consider (25). Although *seoi2* (water) and *caa4* (tea) are related in that they are both a kind of drinks, it is impossible to say that *bui1 seoi2* (*the cup of water*) limits the applicability of the main predication *ngo5 jam2zo2 bui1 caa4* (*I have drunk the cup of tea*), as illustrated by the strange translation #*As for the cup of water, I have drunk the cup of tea*. With *water* and *tea* being two parallel items under the family of *drinks*, *the cup of water* fails to set the scope for the action of *drinking the cup of tea*. Therefore, as the leftmost constituent cannot be interpreted as a topic, sentence (25) is ungrammatical. To conclude, a leftmost constituent in front of a SUBJ which cannot be linked with any missing function in the clause is interpreted as an E-TOP if it satisfies the requirement on a topic by setting up a domain for the main predication to hold.

This section has illustrated how topicalization constructions and E-TOP constructions in Cantonese can be characterized in LFG. It has been shown how a construction with part of the OBJ being topicalized should be represented in an f-structure, how an E-TOP can satisfy the Extended Coherence Condition and under what circumstances a sentence-initial constituent should be regarded as an E-TOP.

## 6.   Summary and Conclusions

In this paper, Cantonese topic has been defined based on Li and Thompson's (1981) definition of Mandarin topic as a phrase of any category which 'limit[s] the applicability of the main predication to a certain restricted domain' and 'sets a spatial, temporal, or individual framework within which the main predication holds' (Chafe 1976: 50), *usually* occurs in the sentence-initial position and is optionally followed by

a pause or a pause particle. Two types of topics, TOPIC and E-TOP, have been identified in Cantonese where TOPIC is required to be identified with a missing constituent while E-TOP is not. It has been shown that [+New] is the feature which licenses FOCUS topicalization and FOCUS in Cantonese is defined as a topicalized phrase which contains the entire or part of the discourse focus ([+New] information). An LFG account has been presented and it has been proposed that E-TOP conforms to the Extended Coherence Condition by having a PRED in its immediate f-structure.

It is hoped that this paper will contribute to the area of Cantonese syntax by proposing an LFG account of topicalization. More research is definitely required in order to have a comprehensive account of the syntactic phenomenon.

## Acknowledgements

## References

Aissen, Judith L. (1992). Topic and Focus in Mayan. *Language*, 68(1):43-80.
Bresnan, Joan. (2001). *Lexical-Functional Syntax*. Malden, Mass.: Blackwell Publishers.
Bresnan, Joan and Sam A. Mchombo. (1987). Topic, pronoun, and agreement in Chichewa. *Language*. 63(4), 741–782.
Chafe, Wallace. (1976). Givenness, contrastiveness, definiteness, subjects and topics. In Charles N. Li (ed). *Subject and Topic*. New York: Academic Press.
Cheung, Yam-Leung. (1997). *A Study of Right Dislocation in Cantonese*. Unpublished M. Phil Thesis. Hong Kong: The Chinese University of Hong Kong.
Choi, Hye-Won. (1997). Information structure, phrase structure, and their interface. In Miriam Butt and Tracy Holloway King (eds). *Proceedings of the LFG97 Conference*. Stanford: CSLI Publications.
Choi, Hye-Won. (1999). *Optimizing Structure in Context: Scrambling and Information Structure*. Ph.D. thesis, Stanford University.
Choi, Hye-Won. (2001). Phrase structure, information structure and resolution of mismatch. In Peter Sells (ed). *Formal and Empirical Issues in Optimality Theoretic Syntax*. Stanford, CA.: CSLI Publications
Fung, Cauvis Suet-Man. In preparation. TOPIC and FOCUS in Cantonese: An OT-LFG Account. M.Phil. Thesis. Hong Kong: University of Hong Kong.
Green, Melanie and Philip J. Jaggar. (2003). Ex-situ and in-situ focus in Hausa: syntax, semantics and discourse. In J. Lecarme (ed). *Research in Afrosiatic Grammar II*. Amsterdam: John Benjamins.
Halliday, Michael A. K. (1967). Notes on transitivity and theme in English, part II. *Journal of Linguistics*, 3: 199-244.
Her, One-Soon. (1991). Topic as a grammatical function in Chinese. *Lingua*. 84(1), 1-23.
Huang, Chu-Ren. (1992). Certainty in functional uncertainty. *Journal of Chinese Linguistics*, 20(2): 247-287.

Kaplan, Ronald M. and Annie Zaenan. (1989). Long-distance dependencies, constituent structure, and functional uncertainty. In Mark R. Baltin and Anthony S. Kroch (eds). *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press.

King, Tracy Holloway. (1995). *Configuring Topic and Focus in Russian*. Stanford, California: CSLI Publications.

Krapova, Iliyana. (2004). Word order in Topic-Focus structures in the Balkan languages. *Romània Orientale XVII*, 139-162.

Krifka, M. (1997). Investigations into the Nature of Focus. A series of three lectures. City University of Hong Kong and Chinese University of Hong Kong, May 19-21.

Kuhn, Jonas. (1999). The syntax and semantics of split NPs in LFG. In F. Corbin, C Dobrovie-Sorin and J. M. Marandin (eds). *Empirical Issues in Formal Syntax and Semantics 2, Selected Papers from the Colloque de Syntaxe et Sémantique à Paris (CSSP 1997)*. Thesus, The Haugue, 145-166.

Lambrecht, Knud. (1994). *Information structure and sentence form: topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Law, Ann. (2003). Right dislocation in Cantonese as a focus-marking device. UCL *Working Papers in Linguistics* 15.

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

Li, Daqin. (2001). Guanxihua dui huatihua de yingxiang: hanyu huati jiegou ge'an fenxi [Influence of relativization on topicalization: A case study of topicalization in Chinese]. *Dangdai yuyanxue*, 3(2): 127-131.

Linguistic Society of Hong Kong. (2003). *Jyutjyu pingjam zibiu* [Guide to LSHK Cantonese Romanization of Chinese Characters]. Hong Kong: Linguistic Society of Hong Kong.

Lu, Yingshun. (2000). Guanyu zhuti de laiyuan wanti [A note on the source of the topic in Mandarin Chinese]. *Journal of Fuyang Teachers College (Social Science)*, 78: 6-7, 10.

Lee, On-Man. (2003). *The Subject Function in Cantonese*. M.Phil. Thesis. Hong Kong: University of Hong Kong.

Matthews, Stephen and Virginia Yip. (1994). *Cantonese: A Comprehensive Grammar*. London: Routledge.

Pan, Haihua and Jianhua Hu. (2003). Head noun movement, focus, and topicalization in Mandarin Chinese. *Language Science and Technology Monograph Series No. 1: Chinese Syntax and Semantics*, 119-156.

Paul, Waltraud. (2002). Sentence-internal topic in Mandarin Chinese: the case of object preposing. *Language and Linguistics (Academia Sinica, Taipei)*, 3(4): 695-714.

Paul, Waltraud. (2005). Low IP area and left periphery in Mandarin Chinese. *Recherches linguistiques de Vincennes*, 33: 111-134.

Rochemont, M. and P. Culicover. (1990). *English Focus Constructions and the Theory of Grammar*. Cambridge: Cambridge University Press.

Rosén, Victoria. (1998). *Topics and Empty Pronouns in Vietnamese*. University of Bergen. PhD Dissertation.

Tsao, Feng-fu. (1979). A functional study of topic in Chinese: The first step towards discourse analysis. In Ting-chi Tang (ed). *Monographs on Modern Linguistics*. Taiwan: Student Book Co., Ltd.

Xu, Jinfeng. (1994). "Topicalisation" in Chinese. *Zhengzhou gongxueyuan xuebao (Sheke ban)*, 1: 67-73.

Xu, Liejiong and Danqing Liu. (1998). Huati de Jiegou yu Gongneng [The Structure and Function of Topic]. In *Contemporary Chinese Linguistics*. Shanghai: Shanghai Jiaoyu Chubanshe.

Zhang, Shuzheng. (1996). Guanyu zhuyu ji huati [A note on the subject and topic in Mandarin Chinese]. *Journal of Shandong University (Philosophy and Social Sciences)*, 3:45-50.

# A Preliminary Exploration of Headless Relative Clauses in Chinese

FAN Xiao-qian

Department of Modern Languages and Linguistics
The Chinese University of Hong Kong

## Abstract

This paper investigates the head drop phenomenon in the Chinese headless relatives and seeks a solution to explain why the head of the Chinese relatives can be dropped and how it happens and under what condition it will happen in order to probe for the nature of head drop. From the formation mechanism of the relative clauses to the occurrence of head drop, a detailed elaboration on the nature of head is made in order to find a correct analysis, the topic-drop analysis. It is proposed that head drop is actually a topic-drop phenomenon. The evidence on the impossibilities of the *pro*-drop analysis strengthens this point. Then a close relation between head drop and topic drop is shown up to further support the point. By comparing two kinds of drop, a clear distinction between the topic-drop and *pro*-drop shows up. What can be drawn from the comparison is that the topic-drop analysis well explains the head drop phenomenon in Chinese relatives whereas the *pro*-drop one fails in the impossibility for recovery of a resumptive pronoun in the dropped head position.

Besides, two prerequisites for the occurrence of head drop in Chinese relatives are proposed in this paper. One is discourse setting while the other is focus requirement. Only if a relative clause is set in a discourse the head of the relative is possible to drop. In addition, the head may happen to drop under the circumstance that it originally must be the focus of the discourse and later it changes from new information to old information. While a topic also happens to drop when it changes from new information to old information, which indicates that some relation may exist between the dropped head and the topic. Therefore, it can be concluded that the head drop in headless relatives is topic-drop and the null head is a zero topic. Also a mechanism for head drop is assumed, which is topicalizaiton of the head noun followed by drop of the head noun.

## 1. Introduction

Relative clauses have been at the center of investigation by linguists in the generative grammar tradition since the 1960s. Crosslinguistically, relative clauses can be classified in many different ways (headed, headless, prenominal, postnominal, reduced, etc.), out of which the headed and headless of Chinese relative clauses will be the focus of this paper. The headed relatives in Chinese (prenominal relative clauses) have been studied by linguists over decades (Chao 1968, Tang 1975, Huang 1980, Chiu 1993 & 1995, Ning 1993 & 1996, Simpson 1998a & 2002, Li 2000, Wu

2004). However, Chinese headless relatives are less studied. Unlike English relatives whose head must be overt, Chinese relatives allow its head to be implicit. In this paper, we will analyze the headless relatives from a syntactic point of view and explain why the head of the Chinese relatives can be dropped. Our analysis shows that the head omission in Chinese relative clauses is a topic-drop phenomenon that only occurs in a fully given discourse in Chinese.

The relative clauses discussed here refer to those that have obvious gaps corresponding to the head nouns as in (1). Those that do not have gaps corresponding to their head nouns are Noun-Complement Clause (CNPs) or appositive constructions as in (2). The former with gap can function as transferred-designation while the latter without gap functions as self-designation. [1] The typical Chinese relatives are always associated with an obligatory element '的 DE'.[2]

    (1) [wo zuotian   mai    $e_i$]   de shu$_i$
       *I yesterday   buy   $e_i$   DE book*
       'The book(s) I bought yesterday.'
    (2) [John de muqin qushi] de xiaoxi
       *John DE mother die   DE news*
       'The news that John's mother died.'

The focus of this paper is on the gapped relative clauses in Chinese. Relatives with gap discussed here mainly fall into three types according to the difference of the predicates. All head nouns below are relativized.

    (3) a . Verb Predicate Relatives
       [$e_i$ zai luzi shang <u>zhu</u> zhe]     de   shu$_i$
        *on stove top   boil ASP    DE water*
       'The water that is cooking on a stove.'
       b. Adjective Predicate Relatives
       [$e_i$ <u>hongse</u>] de pingguo$_i$
        *red   DE apple*
       'An apple that is red.'
       c.   Locative Predicate Relatives
       [$e_i$ zhuozi <u>shang</u>] de shu$_i$
        *table    top DE books*
       'Books that are on the table.'

---

[1] See Zhu, Dexi. 1983 for more details.

[2] The cleft sentences marked with '是...的...', '*shi...de...*' (as in Wo shi zuotian   mai de   piao.)

                                        *I   BE yesterday buy DE ticket*

                                       'It is yesterday that I bought the ticket.'

and possessor structures (as in John de shu. ) both of which relate to 'DE' are not considered in this paper.

           *John DE book*

           'John's book'

[e*i* fangzi <u>li</u>]  de ren*i*
*room inside DE people*
'People who are in the room'

All the heads of the above three type of relatives classified in this paper should be potential arguments (either subjects or objects) within the relative clauses when they are not relativized. Such potential arguments observation is very clear in type (3a). And in (4a) it is known that the relation between the head *shui* 'water' and the predicate *zhu* 'boil' is the subject and predicate one. However in type (3b) and (3c), the potential relation (the relation of arguments and predicates) between the head and the predicate in relatives is not so obvious. But the head in type (3b) and (3c) do have relation with the predicate. When we say *pingguo hongse* 'apple is red' or *shu zhuozi shang* ' books are on the table', listeners can still understand the meaning of such utterances because they will unconsciously insert the verb *shi* 'be' and *shi zai* 'be on/ be in' into the type (3b) and (3c) respectively in order to make up Chinese sentences, which is shown in (4b) and (4c).

Below are the counterparts of (3) which the head nouns are not relativized.
(4) a. shui   zai luzi   shang zhuzhe.
       *water on stove   top   cooking*
       'Water is cooking on a stove.'
   b. pingguo shi hongse.
       *apple    is   red*
       'The apple is red.'
   c. shu      shi   zai   zhuozi   shang.
       *books   are   on    table    top*
       'Books are on the table.'
       ren     shi   zai   fangzi   li
       *people are   at    room    inside*
       'people are in the room.'

The gapped relative clauses in Chinese have one phenomenon that the head of the relatives can drop under certain circumstances. In this paper, we centers on such a head drop phenomenon and try to find a correct analysis to explain it. One example of the head drop phenomenon will be given below.
(5) A: wo xihuan CaoYu xie   de   juben.
       *I    like   CaoYu write DE    play*
       'I like plays that Cao Yu writes.'
    A: ni    ne?
       *You   Q*
       'What about you?
    B: wo xihuan Shakespeare xie   de   e.
       *I    like   Shakespeare write DE   e*
       'I like plays that Shakespeare writes.'

In example (5), the head *juben* 'play' drops in a conversation; that is to say, it drops in a given discourse. The key issue is that what motivates the head to drop, and what kind of phenomenon the head drop is, and which empty category the dropped head (zero head) belongs to. We will give an elaboration on those issues in the following context.

## 2. Observations on the head and its left EC inside the relative clauses

In this section the head and its corresponding gap in the relative will be carefully observed. From such observations, we will further probe for the mechanism of the relativization and the head drop phenomenon.

## 2.1 Relativization vs. topicalization

Relativization has some similarities with topicalization. Li and Thompson (1976) argued that Chinese is a topic-prominent language. One major piece of evidence they cited is that a topic in Chinese can be relativized, that is, a topic-comment construction can function as a relative clause with the topic position related to the head noun.

(6) Shu, wo zuotian mai de.
*Books I yesterday buy DE*
'Books, I bought yesterday.'

(7) Wo zuotian mai de shu.
*I yesterday buy DE book*
'Books which I bought yesterday.'

In (6) *shu* 'books' is topicalized to the topic position of the sentence. In (7) *shu* 'books' is relativized. It is inferred that topicalization is a prerequisite of relativization. But such an interpretation from Li and Thompson (1976) does not seem to be reasonable. Firstly, (7) is not a sentence itself. And secondly if (6) and (7) are derivationally related, (6) is not an independent sentence either.

In view of this, Jiang (1991) argues that all relative clauses in Chinese are topic-comment constructions at a certain stage, that is, an NP must be topicalized first in order to be relativized later. He presents two main reasons for his analysis. One is that relativization and topicalization are subject to almost the same structural constraints. The other is that the similarities between the two processes will be easily accounted for if topicalizaiton is assumed to be the first step of relativization. If Jiang's argument is correct, then topic will be the only element that can be relativized.

As pointed out in Ning (1993) and Qu (1994), the main problem for Jiang's analysis is that the similarities between topicalization and relativization can be equally well accounted for by other theory-internal assumptions. For example, if it is assumed that both processes belong to the same type of operation such as so-called A'-movement, the similarities are also expected. The A'-movement analysis has the

additional advantage of allowing certain differences between the two processes.

Jiang's analysis also encounters some empirical problems. Topic is always definite, as pointed out by Li & Thompson (1976) and Tsao (1990), but the NP being relativized does not have to be so. Besides, not all topic-comment constructions have corresponding relative clauses. As pointed out by Chen (1996), the topic in sentence (8a) cannot be relativized, because (8b) is not acceptable.

(8) a. Shuiguo wo zui xihuan chi xiangjiao.
   *Fruit   I   most like   eat banana*
   '(Among) all fruits, I like to eat bananas the most.'
   b. *[IP wo zui xihuan chi xiangjiao] de shuiguo
   *I most like   eat banana   DE fruit*
   '*I like to eat fruits that bananas the most.'

It is very clear that not all topic-comment constructions have corresponding relative clauses, ant thus it is inferred that not all topics can be relativized. But following Jiang it is believed in this paper that topic will be the only element that can be relativized. It is proposed that topics and relativized nouns[3] relation is comprising and comprised relation in terms of their exchanging roles. Such a relation is clearly demonstrated in the pie chart 1-1 below.
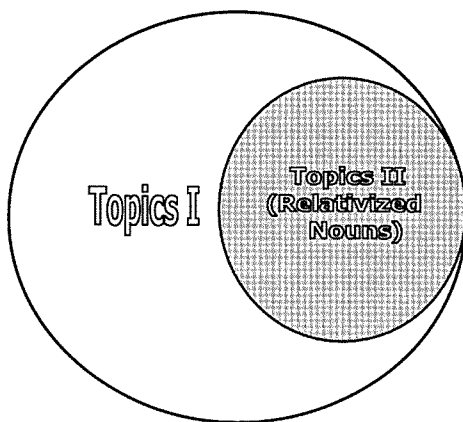


**Chart 1-1**

From chart 1-1, it is shown that the exchanging scope of topics comprises the exchanging scope of the relativized nouns. Under this proposition it is well explained why the topic *shuiguo* 'fruit' can not be relativized. It is because this kind of topic falls into Topics I. In conclusion, not all topics can be relativized (only topics II can be relativized) but all relativized nouns can serve as topics. To such extent, topic, more exactly Topics II, will be the only element that can be relativized.

---

[3] The relativized nouns here refer to the heads of Chinese relative clauses, which is also the key point we discussed in this paper.

All in all, topic is the only element that can be relativized. And it is deduced from Ning (1993) and Qu (1994) that the internal mechanism for relativization will be attributed to the A'-movement analysis. Under this analysis, both subject gap and object gap are A'-bound.

## 2.2 The nature of the gap in relative clauses

This section mainly centers on the nature of the gap inside the relative clauses. We will analyze its nature through the following examples.

(9) Li Xiaojie$_j$ hai zhao-bu-dao [yige [ $e_j$ keyi jia      $e_i$ de] nanren$_i$].

*Li    Miss    still can't-find     one    e can marry $e_i$ DE man*

'Miss Li still cannot find a man who [she] can marry.'

(The agent of the verb *jia* is female)

(10) Li Xiaojie$_j$ hai zhao-bu-dao [yige [ $e_i$    keyi qu      ta$_j$    de] nanren$_i$].

*Li Miss    still    can't-find     one    $e_i$    can marry her DE man*

'Miss Li still cannot find a man who can marry her.'

(The agent of the verb *qu* is male)

(11)  Li  Xiaojie$_j$  hai  zhao-bu-dao  [yige  [ $e_j$  xinzhong  xihuan     $e_i$    de] nanren$_i$].

*Li Miss    still can't-find     one    $e_j$ in-heart     like      $e_i$    DE    man*

'Miss Li still cannot find a man who she loves in her heart.'

NOT 'Miss Li still cannot find a man who loves her in his heart.'

The relativized head noun is the subject of the relative clauses in sentence (10). And the relativized head is the object of the relatives in sentence (9) and (11). ECs in the subject position of the relative clauses in both (9) and (11) are null subjects. Whatever the motivation for relativization is (NP Raising or Null Operator), the trace left by the relativized NP is always an A'-bound element (a variable or a null constant) because movement is involved in the relativization process. And the trace left by an extraction of a target NP to the head position of the relative is coreferential to the noun head and it must get identified through its potential antecedent. The identification requirement of the EMPTY CATEGORY PRINCIPLE (ECP) proposed by Rizzi 1990, which holds for all nonpronominal empty categories, is as follows:

"ECP (Identification):

Empty categories <-pronoun> must be chain-connected to an antecedent where the antecedent can be an A- or A'- or $X^0$-position, depending on the kind of chain." (Rizzi 1994: 160)

According to Rizzi's proposal, the trace left by the relativized NP in relatives must get identified through the noun head in the lexical head position ($X^0$-position, either D or C (see chapter one)).

The fact the *e* can be A'-bound is due to its variable status as a result of the NP relativization. A question might arise here about the nature of the gap: can it be an empty category which is not a variable? The answer to this question seems to be no. It

cannot be PRO given the wide distribution of the gap - it can be found in any governed positions. In (11) the gaps in the relative are subject and object positions, and PRO never occur in governed positions. So the EC in the gaps is not PRO. In addition, the EC in the gaps cannot be an anaphor since its antecedent is never found in its governing category. For instance, in (11) the antecedents of the ECs in object and subject position are all located outside the relative clause that is the local domain of the EC. It cannot be a *pro* since it is located inside a complex NP in general and is available in both object and subject positions. The EC in relatives can either be in subject or object position. When it is in the object position, it is controversial to consider it as *pro*. In this paper, we propose that the null object in Chinese is a variable.

So far it is very clear that the head of a relative is originated from a relativized NP which is previously located in a position of the relative. One may assume that the head may inherit some intrinsic property from the target NP previously seated in the gap. If the target NP is a subject of the relative, then the head may drop like a subject and the null head may be a *pro*. However, this assumption is not on the right track, because the motivation of the head drop comes from the discourse rather than the inside of the relatives. Head drop happens after relativization of the target NP and has nothing to do with the argument structure inside the relative clauses. One cannot analyze the nature of the dropped head regarding its corresponding gap in the relative clause. In the next section, an observation on the nature of the dropped head will be carried on.

## 3.  Head drop of relative clauses in discourse

Headless relative clauses in Chinese result from the drop of the head. The central issue is that which empty category the null head belongs to or what the nature of the head drop is. In this section, the author will analyze the head drop phenomenon in Chinese relative clauses and give an answer to the question.

### 3.1. Reasons for occurrence of head drop in Chinese relative clauses

The occurrence of the head drop in Chinese relatives is decided by the relation between heads and complement clauses; specifically it is attributed to unique properties of the head. Appositive clauses do have no head drop phenomenon because the complement clause of the appositive is a specification about the head and the head does not have the potential subject or object relation with the predicate inside the complement clause, which is also the reason why the appositive clauses are gapless. It seems that the head and the complement clause in the appositive are connected closely and not allowed separation. Therefore head drop is not allowed in the appositive clauses. In contrast to appositive clauses, all relative clauses result from relativization of a target NP that leaves a gap behind it inside the relatives. All target NPs (heads) must have potential subject or object relations with the predicates inside relative clauses. In other words, only if the NPs have potential subject or object

relation with the predicates could they get relativized to become the head of the relative clauses. Therefore, the head NP has unique relation with the predicate in a relative clause and it can be dropped.

> (12) [ta dejiang]       de xiaoxi
> *he win the prize DE news*
> 'The news that he won the prize.'

The sentence like (12) is an appositive clause. Due to lack of the potential relation between the head *xiaoxi* 'news' and the predicate *dejiang* 'win the prize', the head does not drop in this case. It is believed by the author that the relation between the head and the complement clause is closer than that between the head and the predicate in the complement clause in (12). By contrast, the head in the relative clause as in (13) can be dropped because it has closer relation with the predicate inside the relative than the complement clause.

> (13) [wo zuotian mai     $e_i$]   de shu$_i$
> *I yesterday buy   $e_i$   DE book*
> 'The book(s) I bought yesterday.'

In relative clauses the relation between the head and the predicate is closer than that between the head and the complement clause. However, the relation between the head and the complement is closer than that between the head and the predicate in appositive clauses. It is concluded that the intimate relationship of the head and the predicate inside the complement clause decides the occurrence of the head drop in Chinese relative clauses.

## 3.2 Prerequisites for occurrence of head drop in Chinese relative clauses

The first requirement for head drop in Chinese relatives is a fully given discourse, that is to say a relative clause must first appear within a sentence which is set in discourse, and the relative clause as a unit of the sentence must have certain relation with a discourse topic in some way. It is common for the head noun following *de* to be omitted if it is anaphoric and /or can be recovered from the discourse context.

The second requirement for head drop is that the head may happen to drop under the circumstance that it originally must be the focus of the discourse and later it changes from new information to old information. With regard to this, the cleft sentences equivalent to Mandarin '*shi...de...*' will be the best argument in a contrastive illustration.

It is noted that *shi...de* sentences consistently give rise to interpretations similar to English clefts, with the focused element commonly following the copula *shi* and frequently being an adverb or preferring to the time or place where some event has occurred, as for example in (14). Sentence (15) is the alternation found where the object optionally appears positioned after *de*.

(14) wo    shi    zuotian    mai    piao    de.
     *I     BE    yesterday   buy    ticket   DE*
     'It was yesterday that I bought the ticket.'
(15) wo    shi    zuotian    mai    de    piao.
     *I     BE    yesterday  buy    DE    ticket*
     'It was yesterday that I bought the ticket'

Moreover, shi….de sentences can be evolved into a simpler type in which some element like object and *shi* can be omitted as in (16) and (17). The *shi-de* construction commonly encodes a clear focus set off against a strongly presupposed background which often consists of the verb and its object as in examples (14) and (15). As the object is then frequently part of the presupposition and hence old information, there will be a natural tendency for omission of it rather than a repetition of a full descriptive NP form. As Chinese furthermore allows for null objects (*pro* or topic-operator-bound trace as in Huang 1984), *shi-de* sentences then frequently occur without any overt object, and also often without any overt *shi*, as in (16).

(16) (wo)    jintian    mai    de    *e.*
     *(I)      today      buy    DE    e*
     'I bought it *today.*'

It is evident that the null element in (17) is not the dropped head of a relative clause but a null object, the old information in presupposed background. The new information, also the focus, of the sentence (17) is *zuotian* 'yesterday'.

(17) wo    shi    zuotian    mai    de    *e.*
     *I     BE    yesterday  buy    DE    e*
     'It was yesterday that I bought (ticket).'

All facts show that the head of the relative clauses can drop unless it is a focus of its own discourse. The demand of being focus for head drop is reflected in sentence (18). It is known that the head of the relative clause in (18) is the focus of *shi-de* sentence and it can be dropped. And the head of the relative clause in (18) really undergoes drop process when a discourse is given in (19). Consequently two requirements for the occurrence of head drop in Chinese relative clauses are embodied in (18) and (19).

(18) zhe    shi    wo    zuotian    mai    de    piao
     *It     BE    I     yesterday   buy    DE    ticket*
     'It is the ticket that I bought yesterday.'
(19) A: gaosu    wo    yougan    zhezhang    piao    de    shi
        *tell     me    about     this-CL     ticket  DE    matter*
        'tell me things about this ticket.'
     B: zhe    shi    wo    zuotian    mai    de    *e.*
        *I t     BE    I     yesterday   buy    DE    e*
        'It is (the ticket) I bought yesterday.'

It is stated above that the dropped head noun in its own sentence originally must be the focus of the given discourse to some extent it is new information. When the head of a relative clause in discourse changes from new information to old information it may happen to drop. It is the same that a topic also happens to drop when it changes from new information to old information.

(20) wo xihuan shashibiya xie  de juben$_i$, danshi wo geng xihuan laoshe  xie
    *I like Shakespeare write DE play but I more like LaoShe write*

        de     e$_i$.
        *DE ei*
        'I like plays written by Shakespeare but I like plays written by Laoshe more than by Shakespeare.'

In (20) *juben* 'play' is the focus of the discourse, and it is also the head of the relative clause. It is new information when it originally appeared in the first sentence, and then it dropped when it appeared again because it changes into old information in the second sentence. In (20) *juben* 'play' semantically is also a sentence topic in the first sentence, and it happens to drop in the second sentence when it changes into old information. Its drop may be assumed to be topic drop. Arguments illustrated above indicate that some relation may exist between the dropped head and the topic.

## 4. Head drop vs. topic drop

Before going deep into the theme of this paper and probing for the nature of the head drop in Chinese relative clauses, we firstly review the literature on the Chinese drop.

### 4.1. Literature review on the Chinese drop

Chinese is considered as a *pro*-drop language by some scholars, in that the subject of a clause need not to be overt. Thus a Chinese speaker can use either a null subject or an overt pronoun in the subject position of a sentence.

Unlike Italian or Turkish, which also have subject *pro*-drop, Chinese has no inflections to mark subject-verb agreement. Huang (1984 and 1989) argued from a syntactic perspective that for languages like Chinese, a null subject is identified by an NP in a superordinate clause, due to the lack of Agr. Tsao (1979) and Li (1981) observed that subject *pro*-drop in Chinese was actually Topic-NP deletion, which is an optional process, alternating with use of overt pronoun in subject position. Li (1985) and Chen (1986) proposed from a discourse perspective that null subject in Chinese is more likely in cases of topic continuity, in which the information represented by the subject is the component of a series of related actions, events or states. In addition to being a pro-drop language, Chinese is also a discourse-oriented language; a topic is

always present in this kind of language. Discourse-oriented languages are topic-prominent, and topic-comment structures are commonly used as in (21).

> (21) Neichang huo, xingkui xiaofangdui lai de zao.
> *that fire fortunately fire-brigade come DE early*
> 'That fire, fortunately the fire brigade came early.'

In regard with the definition of topics Chao once made a suggestion that 'the grammatical meaning of subject and predicate in a Chinese sentence is topic and comment, rather than actor and action' (1968: 69). According to Shi (2000), the topic-comment construction is a syntactic device employed to fulfill certain discourse functions. It is assumed that a topic has some relation with the comment clause. The topic NP has no independent thematic role of its own and plays no role in syntactic processes that involve specific thematic roles or structural positions, unless the topic is related to the position where a thematic role is assigned. Since topic has no independent thematic role but always depends on an element inside the comment for its thematic role, it has no syntactic function of its own. So far, it can be claimed that Chinese is a *pro*-drop language or a topic-drop language. The two kinds of propositions seem to arbitrarily appear in one language. Topic drop and *pro*-drop can be found in one language. However, the head drop in Chinese relative clauses should belong to one of those two.

## 4.2. The nature of head drop

As previously stated, the motivation of the head drop comes from the discourse. In other words, only if a relative clause is set in a discourse the head is possible to drop.

> (22) A: wo xihuan CaoYu xie de juben.
> *I like CaoYu write DE play*
> 'I like plays that Cao Yu writes.'
> A: ni ne?
> *You Q*
> 'What about you?'
> B: wo xihuan Shakespeare xie de *e.*
> *I like Shakespeare write DE e*
> 'I like plays that Shakespeare writes.'

Sentence (22) shows that the head *juben* 'play' can be dropped unless its reference can be recovered from discourse. *Pro* can behave like an overt pronoun with a deictic reading in discourse while A'-bound elements is linked with discourse through operators. But whether the null head is a *pro* or an A'-bound element is the main concern. The author in this paper believes that the null head must be an A'-bound element and the head drop is in fact the topic drop phenomenon. There are two reasons for the author to hold this point of view.

Firstly, head drop in Chinese relative clauses can be well explained under the topic drop hypothesis. One of two configurations, that is discourse topic configuration,

is stated here.

     (23) Discourse topic:   [$_{TOP}$ Op$_i$  e$_i$…… [$_{IP}$…….t$_i$…….]]    "topic-drop"

Two processes are involved in this topic-drop configuration. Step one is topicalization of a NP inside the comment. Step two is drop of the coreferential NP. According to this analysis, the dropped head can be identified sentence-internally through a null operator, which connects the sentence with the discourse, and recovered from a discourse topic. Thus the dropped head is an A'-bound element. In line with this, the head *juben* 'play' of the relative clause in sentence (22) B should be topicalized first to the topic position of the matrix clause, and then it drops in that position, motivated by the null operator functioning as a link between the discourse and the sentence. That is why head drop of relatives must happen in a discourse. By dropping the topic *juben* 'play' that is originally the head of the relative going through the topicalization process, the null head must be identified through the null operator and recovered its reference from a discourse topic.

Due to topicalization and topic drop involved in the head drop phenomenon of the relatives, the null head is also considered as the zero topic and the head drop is actually the topic drop phenomenon. More exactly, the head drop in relative clauses will be realized only in discourse and such drop phenomenon as matter of fact undergoes two processes that are topicalization and drop. Therefore, head drop is topic drop.

Besides dialogue examples, a textual example is given below in (24) to strengthen the point stated above. In (24) abstract noun *shenghuo* 'life' serves as a head of a relative clause, and it happens to drop in context.

     (24) Meigeren dou xiang zhuiqiu xingfu de shenghuo$_j$, yuanli
        *Everybody all would like to look for happy DE life break away from*

        tongku de e$_j$. Qishi, wulun xingfu de e$_j$, haishi tongku de e$_j$, women
        *Sorrow DE e$_j$. In fact, whatever happy DE e$_j$, or sorrow DE e$_j$ w e*

        bixu yonggan miandui.
        *Must boldly face.*
        'Everybody would like to look for happy life and break away from sorrowful life. In fact, whatever it is happy or sorrowful, we must boldly face it.'

Secondly, the construal of the dropped head as *pro* encounters some empirical problems. Head dorp is not considered as *pro*-drop in this paper because it is impossible for Chinese relative clauses to recover the dropped head with a resumptive pronoun.

*I   and   John yesterday     buy books ASP    but    I    buy    DE    **it***
bei tamen jiezou    le.
*BEI them borrow ASP*
'* I and John bought books yesterday. But **it** which I bought were
borrowed by them.'

It is failed for the pronoun to appear after DE in (25). This is the example in which
DE is followed by the inanimate such as *shu* 'book' in (25). Example (26) is shown
below by contrast, in which an animate thing follows after DE, but it is impossible for
the animate following DE to drop.

(26) Suiran    women    shinian    mei jianguo LiMing    le, danshi women
     yiran
     *Although    we    ten years    not    see    LiMing    ASP    but    we
     still*
     nenggou jiqi,        jinglichongpei   de   ta,   jijixiangshang de   ta,
     *can    remember    vigorous        DE   he   enthusiastic DE   he*
     maomaoshishi de    ta.
     *Boldfaced    DE    he*
     'Although we have not seen him for ten years, we can still remember
     him whatever he is vigorous, enthusiastic or boldfaced.'

(27)* Suiran    women    shinian    mei jianguo LiMing    le, danshi women
     yiran
     *Although    we    ten years    not    see    LiMing    ASP    but    we
     still*
     nenggou   jiqi,        jinglichongpei   de   e,   jijixiangshang de    e,
     *can    remember    vigorous        DE   e   enthusiastic DE    e*
     maomaoshishi de    e.
     *Boldfaced    DE    e*
     '*Although we have not seen him for ten years, we can still remember
     him whatever is vigorous, enthusiastic or boldfaced.'


It is worth thinking that *ta* 'he' after DE can be substituted by the Personal Name
'LiMing' and if 'LiMing' follows after DE, 'LiMing' can still be substituted by the
pronoun 'ta'. Some scholars may think the dropped head can be *pro* and consider (26)
as an counter example to the point that dropped heads in relative clauses can never be
recovered by a resumptive pronoun. As a matter of fact, the pronoun 'he' in (26) can
never drop in that position, which is shown in (27). If the animate cannot happen to
drop after DE, then it does not at all need to be recovered by a resumptive pronoun.
Thus, it is very clear that the dropped heads in relative clauses can never be recovered
by a resumptive pronoun. The null head is least considered as a small *pro* because a
small *pro* can be recovered by a resumptive pronoun.

In this paper it is claimed by the author that the head-drop in relatives is actually
a topic drop phenomenon rather than a *pro*-drop one for EC (Empty Category) in the
head position can never be replaced by a resumptive pronoun, as well head drop

phenomenon in relatives can be explained under topic drop hypothesis. And according to Gu (1998) the trace and the resumptive pronoun can't appear in the same environment, that is to say they display complementary distribution both in Chinese and English, so the EC in the head position is believed to be the trace left by the head, which moves to the topic position, rather than a *pro*. Moreover, only if a relative clause is set in a discourse is the head of the relative possible to drop, which means the reference of the dropped head noun must be recovered from a discourse topic. Besides, the dropped head noun originally must be the focus of the discourse and later it changes from new information to old information, under which circumstance the head of a gapped relative clause may happen to drop. While a topic also happens to drop when it changes from new information to old information, which indicates that some relation may exist between the dropped head and the topic. In short, the null head must be treated as an A'-bound element rather than a *pro*. The head drop in relative clauses will be realized only in discourse and such drop phenomenon as a matter of fact undergoes two processes that are topicalization and drop. Therefore, head drop is topic drop.

## 5. Conclusion

Unlike English, relative clauses whose head must be obligatorily overt, Chinese relatives present the property of head omission once in a given context. In this paper, a large amount of evidence is given to prove that head drop in relative clauses is actually a topic drop phenomenon and the null head is a zero topic. In Chinese, relatives allow head drop because in relative clauses the head nouns are loosely associated with the predicates inside the complement clauses. This makes them differ from appositives that have a tight relationship between the head and the predicate and thus do not allow head drop. Two prerequisites for the occurrence of head drop are that relative clauses must be in a fully given discourse and that the head may happen to drop under the circumstance that it originally must be the focus of the discourse and later it changes from new information to old information. These two requirements indicate that the head of a relative is related to topic to some extent. It is then claimed that the head drop in Chinese relatives is topic-drop. It is assumed that the head of the relative is first topicalized and then dropped. So two processes, topicalization and drop, are involved in this head drop phenomenon. Excluding pro-drop from the analysis on the nature of the head drop proves the impossibility of a resumptive pronoun occupying the head position for recovering the meaning of the null head.

There are also two research questions left here for consideration. What is the relation among head drop, topic drop and topic chain? If head drop is topic drop, what impact will it have on the information structuring of the sentence?

One answer to the first question is that head drop is one kind of topic drop and topic chain is a tool to recover the dropped head from discourse. And the point that head drop is topic drop is proved in this paper will greatly impact the information

structuring of the sentence. Under such a proposition, the basic element for sentence information structuring will be assumed to be topics in Chinese. In regard with the definition of topics Chao once made a suggestion that 'the grammatical meaning of subject and predicate in a Chinese sentence is topic and comment, rather than actor and action' (1968: 69). According to Shi (2000), the topic-comment construction is a syntactic device employed to fulfill certain discourse functions. It is assumed that a topic has some relation with the comment clause. The topic NP has no independent thematic role of its own and plays no role in syntactic processes that involve specific thematic roles or structural positions, unless the topic is related to the position where a thematic role is assigned. Since topic has no independent thematic role but always depends on an element inside the comment for its thematic role, it has no syntactic function of its own. Semantically, topics are subcategorized into superordinate, basic-level and subordiate topics basing on their different linguistic levels and semantic properties. "A superordinate topic, according to Oosten (1986), is a "mental structure" of a schema, a scene, or evaluation. A basic-level topic is a participant in the scene or schema, whereas a subordinate topic is a part or an aspect of a basic-level topic. Whether analyzing from syntax or semantics, topic in Chinese seems to be the basic element of information structure in a sentence. And these two questions are proposed here for the readers' consideration.

## Acknowledgements

## References

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

Deal, Amy Rose. (2005). *Pro-drop, Topic-drop, and the Functional Lexicon: A Constructional Account of Null Arguments*. Honors Thesis, Linguistics Program, Brandeis University.
http://people.umass.edu/amyrose/Deal_thesis_final.pdf

Franks, Steven. (1995). *Parameters of Slavic Morphosyntax*. Oxford: Oxford University Press.

GU, Gang. (1998). *The Topic Structure: More Evidence from English and Chinese*. M. Phil. thesis, The Chinese University of Hong Kong.

Huang, C.-T. James. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15, 4: 531-574.

Kayne, Richard. (1994). *The Antisymmetry of Syntax*, Cambridge, MA: MIT Press.

----- (2000). *Parameters and Universals*. New York: Oxford University Press.

Ning, Chunyan. (1993). *The Overt Syntax of Relativization and Topicalization in Chinese*. Doctoral dissertation, University of California, Irvine.

----- (1996). De as a functional head in Chinese. *Working Papers in Lianguistics*. University of California, Irvine.

Shi, Dingxu. (1989). Topic chain as a syntactic category in Chinese. *Journal of Chinese Linguistics 17*. 223-62.

----- (2000). Topic and topic-comment constructions in mandarin Chinese. *Language 76*,2: 383-408.

Shi, Yuzhi and Charles N. Li. (2001). *A History of Grammaticalization in Chinese —— Motivations and Mechanisms of Evolution of Chinese Morpho-syntax*, Beijing: Peking University Press.

Tsao, Fengfu. (1979). *A Functional Study of Topic in Chinese: The First Step Towards Discourse Analysis*. Taipei: Student Book Co.

----- (1990). *Sentence and Clause Structure in Chinese: A Functional Perspective*. Taipei: Student Book Co.

Xu, Liejiong and D.Terence Langendoen. (1985). Topic structures in Chinese. *Language 61*,1: 1-27.

Xu, Liejiong. (1986). Free empty category. *Linguistic Inquiry 17*,1: 75-93.

Xue, Ping. (1991). *Syntactic Dependencies in Chinese and Their Theoretical Implications*. BC: University of Victoria dissertation, Victoria.

Xiu-Zhi Zoe Wu. (2004). *Grammaticalization and Language Change in Chinese*, Oxfordshire: RoutledgeCurzon.

Zhu, Dexi. (1983). Self-designation and transferred-designation: A study on the syntactic and semantic functions of the nominalizing particles de 的, zhe 者, suo 所 and zhi 之. *Dialect* 1: 16-31.

Zhang Bin, and Qi Huyang. (张斌和齐沪扬). (2000). 《现代汉语短语. 上海：华东师范大学出版社.

# Some Notes on *gam2* and *gam3* in Cantonese[*]

Joanna Ut-Seong SIO and Sze-Wing TANG

The Hong Kong Polytechnic University

## 1. Introduction

This paper discusses the properties of the modifying indexical element *gam* in Cantonese. *Gam* appears in two different forms, with variations in tones: *gam2* and *gam3*. Both elements have the function of 'pointing'. *Gam2* points to a property (eventual, nominal or situational) and *gam3* points to a degree of a scalar property. The usages of *gam2* and *gam3* are illustrated below:

(1) a. **Gam2**   sik6    b. **Gam2** ge3 jan4    c. **Gam2** aa4...    d. **Gam3**      gou1

      GAM     eat        GAM   GE  people    GAM  SFP           GAM           tall
      'eat like that'        'people like that'    'if it is the case...'  'this tall/that tall'

Both *gam2* and *gam3* can point deictically and anaphorically. By deictically, we refer to a situation in which *gam2/3* refers to a reference that is in the immediate non-linguistic surrounding, with the aid of demonstration. By anaphorically, we refer to a situation in which the reference of *gam2/3* is established in the previous discourse. There is yet another use of *gam2/3* in which case *gam2/3* is preceded by a description, as illustrated in (3) and (4). The reference is then the content of the description. We refer to this situation an 'establishing' use of the *gam2/3*.[1]

(3) a.   Maan6-maan6 **gam2**   haang4       b.      Faa1 faa1 fit1 fit1      **gam2**      ge3
    naam4jan2
      slow-slow    GAM    walk              flashy          GAM   GE man
      'slowly walk'                      'flashy type of men'

(4)     Zoeng1        toi2    **gam3** gou1
      CL             table   GAM    tall
      'tall as a table'

The paper is organized as follows. In section 2, we discuss the different usages of *gam2* and show that *gam2* always points to a property. The nature of the property is dependent on what *gam2* is associated with and its position with respect to its associate. In section 3, we illustrate the use of *gam3* in modifying adjectives, showing that it is used

---

[1] We borrow the term 'establishing' from Hawkins (1978, pp.131). He uses the term 'referent-establishing relative clauses' to refer to relative clauses like *the woman he went out with last night*, where a definite referent is established at the point the noun phrase is uttered.

with gradable adjectives and it points to a degree of *adjective*-ness. In section 4, we discuss similarities and differences between *gam2* and *gam3*, as well as some speculations.

## 2. Gam2

### 2.1 Event modification

### 2.1.1 Pre-verbal

When *gam2* is followed by a Verb Phrase (VP), it behaves like a manner adverbial (Sio & Tang 2007). By manner, we refer to both method and style of action (e.g. *killing someone with guns* and *killing someone slowly*). Suppose I see someone swimming at this moment with flippers and I utter (5) with pointing. In such a case, *gam2* is used deictically. Alternatively, if I tell my friend the story a few days later and he utters (5). *Gam2* is then used anaphorically.

(5)    **Gam2** jau4-seoi2    dou1    dak1!
       GAM    swim-water    also    okay
       'How can someone swim like that!'

When *gam2* comes with a preceding description, the description establishes the reference of *gam2*. In (6), *gam2* refers to the manner of having one's mouth wide open.

(6)    Keoi5    maak3-daai6-hau2    **gam2** jau4-seoi2    gaa3
       3SG    make taut-big-mouth GAM    swim-water    SFP
       'S/He swims with her/his mouth wide open.'

Some modifiers can combine with a VP to give rise to a manner reading without *gam2*. For instance, in (6), *gam2* can be absent without affecting the grammaticality.[2] There are also modifiers that need *gam2* to make them into manner adverbials. For instance, without *gam2*, the following sentence would be ungrammatical.

(7)    Peter    seng4 go3    model    *(**gam3**) haang4 lou6    gaa3
       Peter    as-if CL    model    GAM    walk    road    SFP
       'Peter walks as if he is a model.'

This suggests that some modifiers are inherently 'manner' while some are not. For modifiers that are not inherently 'manner' in nature (the precise qualification of which needs further research), adding *gam2* can give it a manner interpretation.

There are, however, a few cases in which *gam2* does not only refer to manner when preceding a VP. It happens to a few complement-taking verbs like *gong2* 'say', *waa6* 'say', *zou6* 'do', *lam2* 'think' etc, . This is illustrated below:

---

[2] Sio & Tang (2007) show that adverbials with-*gam2* and without *gam2* differ in that for the former, the modification is obligatorily restrictive while the latter is optionally restrictive.

(8) A: Ngo5    mou5    jung6   gaa3          B: Nei5 m4     hou2   **gam2**  gong2 laa1
      1SG     NEG     worth   SFP             2SG   NEG    good   GAM     say   SFP
      'I am useless.'                         'Don't say things like that!'


(9)      M4     hou2   juk1-sau2-juk1-goek3   **gam2**  gong2  laa1
         NEG    good   move-hand-move-foot    GAM     say    SFP
         'Don't speak with your limbs moving around'


In (8), *gam2* in B's utterance refers back to the comment made by A (that A being useless), which interpretatively, can be understood as the complement of *gong2* 'say'. In (9), *gam2* refers to the manner in which the action of saying is carried out.

## 2.1.2   Post-verbal

As we have just shown, in the majority of cases, *gam2* gives rise to a manner reading when it precedes a VP. It can in fact also follow a VP, in which case, it is more natural to interpret *gam2* as referring to an evaluation of the event. Compare the two sentences below, pay attention to the position of *gam2* with respect to the VP *coeng2-go1*:

(10) Keoi5  **gam2**    coeng3-go1       ge2
       3SG   GAM      sing-song        SFP
       'How come she sings in such a way!'
       Possible interpretations of *gam2*: with her throat, with two microphones, etc. (manner)

(11) Keoi5    coeng3-go1    (coeng3-sing4)      **gam2**  ge2
       3SG     sing-song     sing-RES          GAM     SFP
       'How come she sings like that!'
       Possible interpretations of *gam2*: too loud, too much of a creaky voice, etc. (evaluation)

The position-interpretation co-relate observed above is also manifested in the question-counterpart of *gam2*, *dim2*. The position and interpretation of *dim2* show the same co-relation. Consider (12) and (13) below:

(12)   Peter   **dim2**   jau4-seoi2      gaa3?
        Peter   DIM      swim-water     QP
        'In what way does Peter swim?'
        Answers: e.g. with his head up, swim naked, etc.


(13)   Peter   jau4-seoi2     (jau4-seng4)   **dim2**  aa3?
        Peter   swim-water     swim-RES      DIM     QP
        'How does Peter swim?'
        Answers: e.g. very good, very fast, etc.

Tsai (1999) makes a similar observation based on the Mandarin counterpart of *dim2*, *zenme yang*, which he glosses as 'how-manner'. Consider the following contrast, taken from Tsai (1999), paying attention to the position of *zenme-yang* with respect to the verb *chang* 'sing':

(14)   Akiu   shang-ci       **zenme-yang**   chang-ge?
       Akiu   last-time       how-manner   sing-song
       'How did Akiu sing last time?'
       Answers: a. you houlong 'with throat' (method)
                     b. hen dasheng 'very loudly'(style of action)

(15)   Ge,   Akiu   shang-ci       chang-de       **zenme-yang?**
       song   Akiu   last-time       sing-RES       how-manner?
       'How did Akiu perform in singing last time?'
       Answer: a. bu zenme-yang 'not so (good)' (resultative)
                    b. tai dasheng 'too loud' (style of resultant state)

Tsai (1999) treats pre-verbal *zenme-yang* 'how-manner' as a pre-verbal modifier and the post-verbal *zenme-yang* as a complement. In particular, he takes pre-verbal *zenme-yang* to be modifying v', giving rise to either method or style of action reading. He postulates a result clause headed by the post-verbal *zenme-yang*, which predicates upon a resultative event/state introduced by the resultative marker *de*. If the subject of the resultative clause is an event, then the resultative reading arises. If the subject of the resultative clause is a state, the style (of state) reading arises. We suggest that the same structure can be applied to the Cantonese *gam2* and *dim2*. Following Tsai (1999), we adopt the following structure for Cantonese:

(16)       vP
        ty
  subj        v'
          ty
  *dim2/gam2*    v'
            ty
        v       VP
          ty
       (Obj)       V'
            ty
          V       RC (Resultative Clause)
              ty
           PRO    *dim2/gam2*

The higher *gam2/dim2* gives rise to a manner reading (method, style of action) and the lower *gam2/dim2* gives rise to an evaluation reading (resultative, style of resultant state).

## 2.2 Nominal modification

### 2.2.1 Pre-nominal

When *gam2* appears within a nominal, it refers to a nominal property. Like most other nominal modifiers, *gam2* has to be followed by a modification marker *ge3*. As noted earlier on, when *gam2* is preceded by a description, it is used 'establishingly'. In this usage, *gam2 ge3* can appear either preceding the demonstrative, as in (17) or between the classifier and the noun, as in (18):

(17)  Wan4-wan4-dei2     **gam2 ge3**     go2     zung2 gam2gok3
      kind-of-dizzy     GAM   GE     that     CL     feeling
      'that kind-of-dizzy-type feeling'


(18)  Go2     zung2 wan4-wan4-dei2     **gam2 ge3**     gam2gok3
      that     CL     kind-of-dizzy     GAM   GE     feeling
      'that kind-of-dizzy-type of feeling'

The [modifier- *gam2ge3*] string can also appear to the left of the numeral 'one', as in (19a), but with other numerals, the resulting phrase sounds very marginal as in (19b). The [modifier- *gam2ge3*] string is focused in (19a).

(19)  a.     Hou2-ci5     gel     gam3 ge3     jat1     gau6     je5
           like     gel     GAM  GE     one     CL     thing
           'a piece of gel-like stuff'

     b.     ??Hou2-ci5     gel     gam3 ge3     saam1 gau6     je5
           like     gel     GAM  GE     three     CL     thing
           Intended reading: 'three pieces of gel-like stuff'

When bare *gam2 ge3* appears in a phrase-initial position or between the classifier and the noun, it is interpreted deictically if it is uttered with demonstration (e.g. pointing to a dog of a peculiar quality in the surrounding) or it is interpreted anaphorically (e.g. referring to some previously mentioned quality of dogs).

(20)  **Gam2 ge3**     gau2     dou1     maai5 ?
      GAM   GE     dog     even     buy
      'You are buying such dogs?'

(21)  Ngo5     soeng2 maai5 jat1     zek3     **gam2 ge3**     gau2
      1SG     want     buy     one     CL     GAM  GE     dog
      'I want to buy a dog of such quality.'

It is however impossible to place *gam2 ge3* to the left of the demonstrative, as shown in (22):

(22) *__Gam2 ge3__   go2   zek3   gau2
     GAM   GE      that   CL     dog

It is possible to put bare *gam2 ge3* to the left of the numeral 'one', as in (23a), but the phrase sounds marginal with other numerals, as in (23b).

(23) a.   __Gam2 ge3__   jat1   go3   naam4jan2
          GAM   GE      one    CL    man
          'a man of such quality'

     b.   ?? __Gam2 ge3__   jat1   baak3      go3   naam4jan2
          GAM   GE         one    hundred    CL    man
          Intended reading: 'a hundred men of such quality'

(24)
A summary of the observations is presented in the following table:

| α= | α-demonstrative | Classifier-α-noun | α-one | α-two, three… |
|---|---|---|---|---|
| [modifier-*gam2 ge3*] | √ | √ | √ (with α in focus) | X |
| [*gam2 ge3*] | X | √ | √ (with α in focus) | X |

When *gam2 ge3* modifies a nominal, it gives rise to a kind reading. As expected, individual-levels modifiers are compatible with *gam2 ge3* as in (25), but stage-level modifiers are not as in (26).

(25)   Seng4-jat6   fai6   gam2   ge3   gau2
       all-day      bark   GAM    GE    dog
       'this/that type of dog(s) that bark(s) all day long'

(26)   *Ji1gaa1   fai6-gan2      gam2   ge3   gau2
       now        bark-PROG      GAM    GE    dog
       Intended reading: 'the now-barking type of dogs'

NPs denote kinds while DPs denote individuals. In order to account for the modifying-kinds reading, we assume that *gam2 ge3* is generated right on top of NP, presumably in the specifier of some projection related to modification. We have observed that for [modifier-*gam2 ge3*] strings, they can appear both to the left of the demonstrative or between the classifier and the noun. We assume that the different positions are derived by movement.

[Modifier-*gam2 ge3*] and [*gam2 ge3*] strings can precede the numeral 'one' and in these cases, the [modifier-*gam2 ge3*] and [*gam2 ge3*] strings are focused. The derivation of the pre-'one' position of [modifier-*gam2 ge3*] and [*gam2 ge3*] strings can be viewed as movement from NP to some focus position in the left periphery of the noun phrase. It is unclear to us as to why the numeral 'one' behaves differently from other numerals in these constructions. It might be the case that in such constructions, 'one' is functioning as an indefinite article rather than an authentic numeral. There is at least one piece of evidence that supports this claim. 'One' can be stressed when the quantity reading has to be emphasized in regular noun phrases, as in (27). However, when 'one' is preceded by [modifier-*gam2 ge3*] and [*gam2 ge3*], 'one' cannot be stressed, as shown in (28a) and (28b). The stressed 'one's are placed within boxes.

(27)       Ngo5   jau5   ⌗jat1⌗   zek3   gau2,   m4   hai6   saam1 zek3   gau2
              1SG    have   one    CL    dog,   NEG   BE    three   CL    dog
              'I have got one dog, not three dogs.'

(28)   a.   *   Hou2-ci5     gel    gam3   ge3   ⌗jat1⌗   gau6   je5
               like        gel    GAM   GE   one     CL     thing
               'A gel-like piece of item'

       b.   *   Gam2   ge3   ⌗jat1⌗   go3   naam4jan2
               GAM    GE   one     CL    man
               'A man of such quality'

Note that without a preceding modifier, *gam2 ge3* cannot appear to the left of a demonstrative, though it can appear to the left of 'one'. In other words, the combination of the absence of a modifier and the presence of a demonstrative make the movement impossible. *Gam2* is an indexical element. It has to get its reference from somewhere. It can be viewed as a variable that needs to be bound either by some contextual operator (when used deictically or anaphorically) or by the preceding modifier. In other words, *gam2* containing an unbound variable cannot move across a demonstrative, the following configuration is ungrammatical:

(29)    *gam2 ge3* [e]$_i$ Dem…t$_i$

### 2.2.2   Post-nominal

*Gam2ge3* cannot appear in front of proper names, showing that in a phrase- initial position, it is obligatorily restrictive. However, it can ascribe a property to a proper name in a predicative sentence:

(30)   a.      ***gam2** ge3    Peter          b.     Peter   hai6    **gam2**   gaa3
                 GAM    GE     Peter                Peter   BE      GAM     SFP
                                              'Peter is like that.'

The grammaticality contrast between (30a) and (30b) suggests that *gam2* gives rise to different interpretations depending on its position. In a phrase-initial position, *gam2* is a pre-nominal modifier and is only restrictive. Thus it can't appear with proper names, which have rigid designations. When *gam2* acts as a predicate, we treat it as the predicate of a Small Clause (SC) with NP/DP as the subject. *Gam2* in this case is attributive in nature, and is thus compatible with proper names. The relevant structural differences are shown in (31) and (32).

(31)       NP/DP          (32)         SC
              ty                          ty
    *gam2ge3*   NP/DP            NP/DP     *gam2ge3*

## 2.3    'Situational' use

When *gam2* is not followed or preceded by anything, it refers to the property of a situation. In (33), *gam2* refers to the situation that A has got no cash. It is interpreted anaphorically. *Gam2* can also be interpreted deictically if A, instead of saying he has no money, shows an empty wallet to B. *Gam2* is not used establishingly when referring to the property of a situation.

(33) A: Ngo5   mou5 daai3   cin2    tim1       B: **Gam2** aa4, ngo5 bei2 sin1 laa1
         1SG   NEG bring     money SFP         GAM  SFP, 1SG   give first SFP
         'I forgot to bring any cash.'          'In that case, I will pay first.'

## *3.  Gam3*

*Gam3* is always followed by a gradable adjective. Let X be the adjective, [*gam3-X*] refers to a degree of X-ness. It combines equally well with open-scale and closed-scale adjectives. It can be used deictically, with demonstration accompanying the utterance as in (34). It can also be used anaphorically as shown in (35). When *gam3* is combined with a non-gradable adjective, it is uninterpretable, as in (36).

(34)   **Gam3** mun5 laa1                              (closed-scale
adjective)
       GAM   full    SFP
       'this full' (with the index finger touching the side of a beer glass)

(35)   A: Tiu4       sing2  jau5   saam1 mai5  **gam3** coeng4 (open-scale adjective)
         CL  rope   have   three   meter GAM   long
         'The rope is three meters long.'

         B: **Gam3** coeng4 dou1 m4    gau3        bo3
           GAM  long  still   NEG   enough     SFP
         'That is still not long enough.'

(36)# Peter **Gam3** daan1san1 gaa3
Peter GAM single SFP
Lit. 'He is that single.'

In addition to a degree reading, *gam3* can have an intensifying meaning on top of the degree reading, meaning 'very', as in (37). It, however, cannot be used alone to mean 'very' without the indexical meaning, as in (38).

(37) Go3 naam4zai2 **gam3** gou1 ge2
CL boy GAM tall SFP
'The boy is so tall.'

(38) *Ngo5 m4 zi1 keoi5 jau5 gei2 gou1,
1SG NEG know 3SG have how tall
daan ngo5 zi keoi5 gam3 gou1
but 1SG know 3SG GAM tall
Intended reading: 'I don't know how tall he is (in measurement), but I know he is very tall.'

When there is a preceding description, only the indexical reading is possible.

(39) Saam1 mai5 **gam3** gou1
three meter GAM tall
'three meters tall'

Note that in (39), *gam3* can be absent. The meaning does not alter much.[3] However, there are degree modifiers that cannot modify the adjective directly without *gam3*, as in (40).

(40) Zoeng1 toi2 **gam3** gou1
CL table GAM tall
'tall like a table'

What this suggests is that, similar to the use of *gam2* in manner modification of events, there are modifiers that are inherently degree-like. For these modifiers, *gam3* is not needed to mediate the relationship between the modifier and the adjective. For those modifiers that are not degree-like, *gam3* can be used to make them into degree-modifiers.

## 4. *Gam2* and *gam3*: Similarities, Differences and Cross-linguistic Relevance

*Gam2* and *gam3* are both indexical elements that can be used deictically, anaphorically and establishingly. Another interesting similarity between *gam2* and *gam3* lies in their establishing use. In manner modification of event, for modifiers that cannot combine with the VP directly, *gam2* makes them into legitimate manner adverbials. Similarly, in degree modification, for modifiers that cannot combine with an adjective

---

[3] The measurement seems to be more precise when *gam3* is absent.

directly to denote a degree, *gam3* makes them into legitimate adjectival degree modifiers. In other words, *gam2* and *gam3* alter the nature of the preceding description in such a way that they can modify an event and the degree of an adjective respectively.

With respect to their differences, *gam2* and *gam3* differ in that *gam2* refers to a property while *gam3* refers to a degree of a property denoted by an adjective. The difference in the reference is not only reflected in their tones. *Gam2* can always be followed by an associative noun *joeng2* 'appearance' while *gam3* cannot (Sio & Tang 2007).

(41)   a.    Gam2 joeng2                   b.    Gam3 (*joeng2)

In other words, the modification relation between *gam2* and the modifiee (e.g. VP, NP, etc.) is not direct. *Gam2* ascribes a property to *joeng2* 'appearance' and *gam2 joeng2* as a whole ascribes a property to the event or the nominal. Furthermore, as is shown, *gam2 (joeng2)* can appear in different environments while *gam3*, on the other hand, always appears with an adjective.

Having one and the same indexical element to refer to nominal/verbal properties and degree is not merely a Cantonese idiosyncrasy. In fact, many languages use identical indexical elements to refer to nominal/verbal properties and degree, for instance, the Czech *tak*, the Dutch *zo*, etc. (see also Landman & Morzycki 2003 for discussion in Polish, Russian and German). Some illustrative Czech examples are provided below:[4]

(42)

| (a) Tak | tančí. | (b) Takový | muž | (c) Tak |
|---|---|---|---|---|
| TAK | dance.3SG.PRES | TAK-M.NOM | man.M.NOM | TAK |
| 'dance in such a way' | | 'such a man' | | 'It is the |

case.'

| (d) Tak (hle) velký | (e) | Tak | velký |
|---|---|---|---|
| TAK(see) big.M.NOM | | TAK | big.M.NOM |
| 'this big' (with demonstration) | | 'so big' (exclamation) | |

Another similar example is the English *such* as in *such is life* and *such a lovely house*, which seems to have both a referential/demonstrative-like usage (referring to a property) and an intensifying usage (Wood 2002 and references therein). This cross-linguistic robustness suggests the usage of the same indexical element in these environments is not a co-incident. The remaining questions are how these different environments are related and what functional category should be assigned to such elements. We leave these questions for further research.

---

[4] We are grateful to Frantisek Kratochvil for the Czech data.

## Acknowledgements

## References

Hawkins, J. (1978). *Definiteness and indefiniteness: a study in reference and grammaticality prediction.* Croom Helm linguistics series. USA: Humanities Press.

Landman, M & M, Morzycki. (2003). Event kinds and the representation of manner. In Antrim, N. M., Goodall, G, Schulte-Nafeh, M and V. Samiian (eds.) *Proceedings of the western conference in linguistics (WECOL), 2002, vol. 11.* California State University, Fresno. (Held at the University of British Columbia).

Sio, U.-S. & S.-W., Tang (2007). The indexical expressions *gam2* and *gam3* in Cantonese. In Sio, U.-S. and S.-W. Tang (eds.) *Studies in Cantonese Linguistics 2.* Hong Kong: Linguistic Society of Hong Kong.

Tsai, W.-T. (1999).The Hows of Why and The Whys of How. In Del Gobbo, F. and H, Hoshi (eds.) *UCI Working Papers in Linguistics 5*, 155-184.

Wood, J. L. (2002). Much about such. *Studia Linguistica*, 56:1, pp. 91-115.

# A study on "*you* (有)+VP" structure in Mandarin Chinese

SUN Jing

College of International Education and Exchange
Tianjin Normal University;

**Abstract**

This paper examines the development of *you* (有) from a lexical verb to a functional category as it is in Cantonese in *you*+VP syntactically and cognitively. Through grammaticalization and language contact, the emergence of *you*+VP construction in Mandarin indicates the possibility of symmetric *you* vs. *meiyou* aspectual system in addition to well-established asymmetric ones such as the *le* vs. *meiyou* systems.

## 1. Introduction

The Chinese word *you* (有) is a polysemous verb that has many meanings and is used widely in various situations. In Mandarin, it mainly expresses 'possession,' e.g. 我有一本书. 'I have a book.'; 'existence', e.g. 墙上有一幅画。'There is a picture on the wall.'; 'comparison, achievement in amount', e.g. 小明都有他爸爸那么高了。'Xiaoming is as tall as his father now.' (Lü, 1999; Liu, 2000; Chao & Han et al, 2005); 'occur, happen', e.g. 他有病了。'He is sick now.' (Liu, 2000; Chao & Han et al, 2005).

Nowadays, we frequently hear *you*+VP constructions among Northern Mandarin speakers and it is becoming popular in Mainland China, especially among the younger generation and the media. It seems that *you* now has one more trendy usage in Mandarin, expressing not only lexical meanings but also temporal notion, as illustrated in (1) and (2):

(1) A: 你　　有　　没　　有　　吃　　饭　　啊？
　　　Ni$^3$　*you*$^3$　mei$^2$　*you*$^3$　chi$^1$　fan$^4$　a?
　　　You　have　NEG　have　eat　dinner　SFP?
　　　'Did you have dinner?"

　　B: 有　　啊。
　　　*you*$^3$　a.
　　　have　SFP.
　　　'Yes, I did.'

(2)　我　　在　　美国　　也　　有　　去　　这样　　的　　超市,
　　Wo$^3$　zai$^4$　mei$^3$guo$^2$　ye$^3$　*you*$^3$　qu$^4$　zhe$^4$yang$^4$　de　chao$^1$ shi$^4$,
　　I　in　USA　too　have　go　this　NOM　super market,

可　人　家　不　让　我　拍！　　　　　　(www.homemsg.focus.cn)
ke³　ren²jia¹　bu²　rang⁴　wo³　pai¹!
but　PRON　not　let　me　photograph!
'I went to this kind of supermarket in the US too, but picture-taking was prohibited.'

This paper will examine the development of *you*(有) from a lexical verb to a functional category as it is in Cantonese and trace the mechanism of development of *you*+VP structure in Mandarin.

## 2. Literature Review

English "have (done)" influence on Cantonese "*yauh* (VP)", which was referred to as "language transfer analysis" (Shi, 2002) couldn't explain "*you*+VP" in Mandarin trendy usage.

In Mandarin, Shi Yuzhi's recent paper concerning Chinese perfective marker (Shi, 2004) pointed out that *le* (了) and *mei(you)* [没(有)] are the affirmative and negative perfective markers in Mandarin Chinese and the pair forms an asymmetric system. Emergence of *you*+VP here gives rise to a possible symmetric system involving a *you* vs. *mei you* contrast.

Heine (1977) depicted possession in some European and African languages from the perspective of cognitive sources and grammaticalization. Heine reports that in English, as well as many other European languages, "constructions using 'have' as their predicate base have given rise to both expressions of verbal possession and of perfect aspect" (p.189). The case of *you* might be the application of this tendency in Mandarin.

Previous studies on word *you* mainly focused on the syntactic scope of *you* constructions (Lü, 1942; Ma, 1983; Huang, 1989, etc.) and post-*you* constituents (Zhu, 1982, etc.). The present study will focus on the possible development of *you* from lexical verb to aspect marker, i.e. from possessive/existential verb *you* to functional *you* in *you*+VP structure in Chinese.

## 3. Lexical Meaning Extensions of *you*

### 3.1 Basic lexical meanings of *you* in Ancient Chinese

In the etymological dictionary entitled *Shuo Wen Jie Zi* (说文解字, Notes on Language and Characters, compiled by Xu Shen (许慎) in the East Han Dynasty around 110-121 AD), the Chinese character *you* (有, 'have') is a compound character formed with the characters 手 'hand' and 肉 'meat' meaning "to *have* (possess), or, the state of *having* (existence) meat in hand". 'Possession' and 'existence' are the two basic meanings of *you* in ancient Chinese (Wang, 2005). *You* in (3) demonstrates the meaning of 'possession' and the sentence here means that 'everyone possesses personal hobbies'.

(3)  民       生       各       有       所       乐       兮，
     min²    sheng¹  ge⁴     *you³*   suo³    le⁴     xi¹,
     people  live    every   have    PRT     hobby   SFP,

     余       独       好       修       以       为       常。
     yu²     du²     hao⁴    xiu¹    yi³     wei²    chang².
     I       only    like    think   to      as      usual.

'Everyone has hobbies. I like doing thinking.'

[*Chuci,Lisaojing* (楚辞·离骚经), in Warring States Period]

In (4), *you* demonstrates the other basic meaning of 'existence', often translatable as English 'there be':

(4)  河       侧       有       两       山            相       对，
     he²     ce⁴     *you³*   liang³  shan¹       xiang¹  dui⁴,
     river   side    have    two     mountain    each    face,

     水       出       其            间。
     shui³   chu¹    qi²          jian¹.
     water   out     PRON         between.

'On both sides of the river, *there are* two mountains facing each other; the water flows out from between them.'

[*Shuijingzhu* (水经注), in Southern and Northern Dynasties]

In Ancient Chinese, *you* is also used with the basic meaning of 'happen' (He, 1988, p.108), 'occur' or 'take place' as in (5). It explains how Zhangheng's seismograph works by depicting a possible situation *happening, occurring, or taking place* in the future. The reading of 'happen', 'occur' or 'take place' here could be cognitively regarded as derived from 'existence' *you* on the condition of situation in context and its possibility of occurrence.

(5)  如       有       地       动，      樽          则       振       龙，
     ru²     *you³*   di⁴     dong⁴,   zun¹       ze²     zhen⁴   long²,
     if      have    earth   quake,   cylinder   CNJ     shake   dragon,

     机       发       吐       丸，      而       蟾蜍          衔       之。
     ji¹     fa¹     tu³     wan²,    er³     chan²chu²   xian²   zhi¹.
     machine launch  vomit   ball,    CNJ     frog        hold    PRON.

'If *there is* an earthquake, a dragon which faces the direction of the earthquake among the eight dragons on the cylinder would shake. The ball in the dragon's mouth would drop down and the frog positioned right below the dragon would hold the ball in its mouth.'

[*Houhanshu, Zhangheng Zhuan* (后汉书·张衡传), in Southern and Northern Dynasties]

## 3.2 Lexical meanings of *you* in Modern Chinese

(6) 他　有　两　　个　孩子。
ta$^1$ *you$^3$* liang$^3$ ge$^4$ hai$^2$zi.
he have two CL children.
'He has two children.'

(7) 我　有　一　本　书。
wo$^3$ *you$^3$* yi$^4$ ben$^3$ shu$^1$.
I have one CL book.
'I have a book.'

(8) 这　张　　桌子　有　三　　条　腿。
zhe$^4$ zhang$^1$ zhuo$^1$zi *you$^3$* san$^1$ tiao$^2$ tui$^3$.
this CL table have three CL legs.
'This table has three legs.'

(6) and (7) illustrate the meaning of 'to be in possession of', where the possessor is animate, while the possessee can be animate or inanimate. In (8), the possessee is a part of the possessor and both of them can be inanimate.

(9) 香港　　　有　维多利亚　　港。
Xiang$^1$gang$^3$ *you$^3$* wei$^2$duo$^1$li$^4$ya$^4$ gang$^3$.
Hong Kong have Victoria Harbor.
'Hong Kong has the Victoria Harbor.'/'There is the Victoria Harbor in Hong Kong.'

(10) 桌子　　上　　有　一　个　书包。
zhuo$^1$zi shang$^4$ *you$^3$* yi$^2$ ge$^4$ shu$^1$bao$^1$.
table above have one CL bag.
'There is a bag on the table.'

Existential *you* is still a basic meaning in contemporary Mandarin. 'Hong Kong' (*Xiang$^1$gang$^3$*) in (9) can be regarded as the inanimate possessor of the inanimate possessee the Victoria Harbor (*Wei$^2$duo$^1$li$^4$ya$^4$gang$^3$*), or as the place and location where the Victoria Harbor (*Wei$^2$duo$^1$li$^4$ya$^4$gang$^3$*) *exists*. As the term 'existential' suggests, *you* indicates that *there is* something or something *exists* in some place. It can be cognitively regarded as a subset of 'possession' *you* (Zheng, 2004) with the 'possessor' argument being an abstract one, including places and locations such as 'Hong Kong' (*Xiang$^1$gang$^3$*) and 'on the table' (*zhuo$^1$zi shang$^4$*) in (9) and (10) respectively.

(11) 小明　　　已经　有　他　爸爸　那么　高　了。
Xiao$^3$ming$^2$ yi$^3$jing$^4$ *you$^3$* ta$^1$ ba$^4$ba na$^4$me gao$^1$ le.
Xiaoming already have he father that tall SFP.
'Xiaoming is as tall as his father now.'

(12) 小刚　　　有　五　尺　高　了。
　　　Xiao³gang¹　*you³*　wu³　chi³　gao¹　le.
　　　Xiaogang　　have　five　feet　high　SFP.
　　　'Xiaogang is five feet tall now.'

Possession verb *you* in earlier examples takes two arguments, both of which can be animate or inanimate. Both of (11) and (12) have a comparative construction viewed as an abstract possessee argument: In (11), Xiaoming *has* something, that something being the same height as his father's (lit. such height as his father's). Xiaoming's height is compared to his father's to show how tall he is now. It can be interpreted as 'Xiaoming *has* the height of his father now'. In (12), the height Xiaogang *has* is expressed by saying 'He *achieves* the height of five feet'. It conveys abstract 'possession' with an abstract 'possessee' argument. I assume the 'comparison or achievement in amount' reading in Modern Chinese (Lü, 1999; Liu, 2000; Chao & Han et al, 2005) is derived from 'possession'.

(13) 他　有　病　　　了。
　　　ta¹　*you³*　bing⁴　le.
　　　ta　have　sickness　SFP.
　　　'He is ill.'

(14) 这　孩子　嘴　里　　有　牙　了。　　　(*Zhao*, 1979)
　　　zhe⁴　hai²zi　zui³　li³　*you³*　ya²　le.
　　　this　kid　mouth　inside　have　teeth　SFP.
　　　'Teeth are growing the kid's mouth.'

In Modern Chinese, 'happen', 'occur' or 'take place' *you* (Liu, 2000; Chao & Han et al, 2005) also usually expresses that something happens to someone or occurs/takes place somewhere as shown in (13) and (14) which can be alternatively interpreted as 'There is sickness in his body now' and 'There are teeth in the kid's mouth now' respectively, indicating both the result of *occurrence* and the state of *existence*. In such circumstances, the 'happen', 'occur' or 'take place' reading of *you* in is conditionally regarded as derivative of 'existence' *you*. Available examples of 'occur' *you* this paper dealing with in context types are providing "bridging context", not yet evolved to a "switch context" (Heine, 2002) in its lexical grammaticalization. Figure 1 shows its lexical extension.

Verb　*you* 'have'
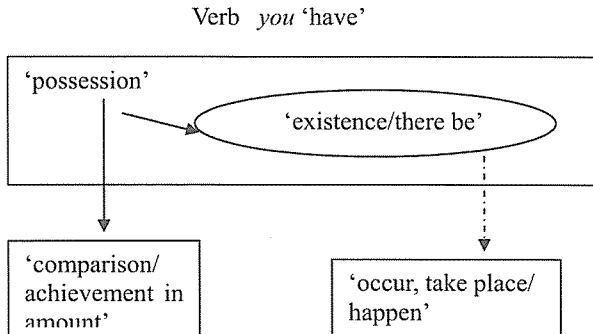


Fig. 1

From the lexical meanings analysis of verb *you*, we see a progression of its meaning from concrete to abstract. In addition, *you* begins to develop from a lexical verb in the contexts to a functional category syntactically and cognitively, which will be examined in the following sections.

## 4. Another Look at *you*+VP in Cantonese

Like in Mandarin, 有 (*you*) is also a verb in Cantonese *yauh* (in Cantonese Romanization). Besides being used as a verb, *yauh* is categorized as a primary auxiliary in Cantonese (Luke & Nancarrow, 2004), which can be placed before verbs in *yauh*+VP in Cantonese.

Cantonese grammar books identify two main functions for *yauh*+VP structure: As an auxiliary, 有/*yauh5* and 冇/*mou5* (有/*you* and 没有/*meiyou* in Mandarin) before a verb is used to indicate past events, but is highly emphatic (Matthew & Yip, 1994; Luke & Nancarrow, 2004) as shown in (15) and (16):

(15) A: 你　　有　　冇　　　食　晒　0的　藥　　　啊?
　　　　Leih *yauh mouh* sihk saai di yeuhk a?
　　　　You have not-have eat all CL medicine PRT?
　　　　'Have you taken all the medicine?'

　　B: 當然　　　有　食　晒　啦。
　　　　Ganghaih *yauh* sihk saai la.
　　　　surely　　have eat all PRT.
　　　　'Of course, I have taken all (of them).'　　　　(Matthew & Yip, 1994: 281)

(16) A: 個　次　　我 0的　去 0左　淺水灣。
　　　　Go chi ngohdeih heui-jo Chekchyuh waan.
　　　　That time we　　go-PFT Stanley.
　　　　'That time we went to Stanley.'

　　B: 系　啊，我　都　有　去。
　　　　Haih a, ngoh dou *yauh* heui.
　　　　Is PRT, I also have go.
　　　　'That's right, I went too!'　　　　(Matthew & Yip, 1994: 281)

*Yauh*+VP in Cantonese is also used to describe the present with a habitual meaning, often conveyed or reinforced through the use of adverbs of frequency (Matthew & Yip, 1994), as shown in Example (17):

(17) 報紙　　　日日　　都　有　講　股市。
　　　bouji yaht-yaht dou *yauh* gong gusih.
　　　newspaper day-day all have talk stock-market.
　　　"The paper talks about the stock market every day." (Matthew & Yip, 1994:281)

This kind of sentence with *yauh*+VP is quite typical and commonly used in Cantonese. Cantonese speakers retain this grammatical structure to express the

function of (emphatic) past when they speak Mandarin. For example in (18), the Mandarin-speaking host asked Cantonese-speaking Director why he slowed down script writing when he was directing. In the director's utterance, *yauh*+VP structure in Cantonese was phonetically converted into *you*+VP in Mandarin. Such utterance indicates him trying to affirm and emphasize his action of 'providing the actor a finished script' with *you (yauh)*. Same situation also found in actor's response in (19) where *you*[3] jiang[3] (有讲) is marked with past tense 'told a story'.

(18)—主持人：我们所知道的是，八十年代写过十三部作品，据说速度挺快的，几个月，就可以完成一部戏的写作。你当导演的时候，为什么就这么慢了呢？你替人家想过吗？

'—Host: As far as I know, you had written 13 scripts in the 80's. It's quite fast to finish one script in a few months. Why did you slow down script writing after being a director? Have you ever been concerned about the actors?'

—王家卫：   我  有  给  剧本  给  他  的，  也  写得  很  快。
—Wang[2] Jia[1]wei[4]: wo[3] *you*[3] gei[3] ju[4] ben[3] gei[3] ta[1] de, ye[3] xie[3] de[2] hen[3] kuai[4].
—Director Wang: I have give script give him PRT, also write PRT very fast.
'—Director Wang: I did give a script to him. I wrote very fast, too.'

(Excerpt from 《艺术人生》，CCTV)

(19)—主持人：第一次王导来找你的时候，也没有剧本吗？

'—Host: Didn't you have a script when you first joined Director Wang's movie?'

—梁朝伟：    那  一  次，他  有  讲  故事  给  我  听，我觉得很新颖，所以才有第一次的合作。
—Liang[2] Chao[2]wei[3]: na[4] yi[2] ci[4], ta[1] *you*[3] jiang[3] gu[4]shi gei[3] wo[3] ting[1], wo[3] jue[3] de[3] hen[3] xin[1] ying[3], suo[3] yi[3] cai[2] you[3] di[4] yi[1] ci[4] de he[2] zuo[4].
—Actor Liang: that one time,   he have tell story   give me listen, I feel very interesting, so have first cooperation.
'—Actor Liang: That day, he told me the story. I thought it was very interesting. Then we cooperated."

(CCTV, 《艺术人生》之《重温你的"花样人生"》, 2004 年 9 月 29 日)

"It can be expected that different kinds of archaic and medieval features be potentially preserved in certain of the more conservative dialect groups of Sinitic" (Chappell, 2004). Southern Chinese dialects such as Min, Kejia and Yue (Cantonese) retain more Old Chinese grammar than Modern Mandarin does (Yuan, 1983; Norman, 1995; Shi, 2002). Given considerable dialectal contact, it is directly related to and may be the synchronic resource of *you*+VP pattern in Mandarin.

## 5. *you* Evolves to a Functional Category in Mandarin

*You*+VP structures in conversations and in declarative sentences are mainly of two types in Mandarin.

## 5.1 *you*+VP in conversations

*You*+VP in conversations usually appears in response to the *youmeiyou*+VP form in questions, which often invites an affirmative answer with *you*+VP. This perfective response token morphologically and syntactically conforms to its negative and interrogative counterparts, *meiyou*+VP and *youmeiyou*+VP respectively.

In Mandarin, the negative and interrogative form of possession verb *you* before noun phrases are *meiyou* (没有/not have) and *youmeiyou* (有没有/have not have) respectively:

| | NP |
|---|---|
| Affirmative | *you*+NP<br>你 有 这 本 书。<br>ni$^3$ *you$^3$* zhe$^4$ ben$^3$ shu$^1$.<br>you have this CL book.<br>'You have this book.' |
| Negative | *meiyou*+NP<br>你 没 有 这 本 书。<br>ni$^3$ *mei$^2$ you$^3$* zhe$^4$ ben$^3$ shu$^1$.<br>you not have this CL book.<br>'You don't have this book.' |
| Interrogative | *youmeiyou*+NP<br>你 有 没 有 这 本 书?<br>ni$^3$ *you$^3$ mei$^2$ you$^3$* zhe$^4$ ben$^3$ shu$^1$?<br>you have not have this CL book<br>'Do you have this book?' |

*Youmeiyou*+VP is a "lately emerged" interrogative form in Mandarin (Ding, 1961). The development of such an interrogative form is from *meiyou*, which modifies a VP (Ota, 1958). The following chart highlights the comparison between *you*+NP construction and *you*+VP construction.

| | NP | VP |
|---|---|---|
| Interrogative | *Youmeiyou*+NP<br>你 有 没 有 这 本 书?<br>Ni$^3$*you$^3$mei$^2$you$^3$* zhe$^4$ben$^3$shu$^1$?<br>You havenothave this CL book<br>'Do you have this book?' | *Youmeiyou*+VP<br>你 有 没 有 吃 饭?<br>Ni$^3$ *you$^3$ mei$^2$ you$^3$* chi$^1$ fan$^4$?<br>You have not have eat dinner<br>'Have you eaten dinner?' |
| Affirmative | *You*+NP<br>你 有 这 本 书。<br>Ni$^3$ *you$^3$* zhe$^4$ ben$^3$ shu$^1$.<br>you have this CL book<br>'You have this book.' | *You*+ VP<br>我 有 吃 饭。<br>Wo$^3$ *you$^3$* chi$^1$ fan$^4$.<br>I have eat dinner<br>'I have eaten dinner.' |
| Negative | *Meiyou*+NP<br>你 没 有 这 本 书。 | *Meiyou*+VP<br>你 没 有 吃 饭。 |

| | Ni³ mei² you³ zhe⁴ ben³ shu¹. | Ni³ mei² you³ chi¹ fan⁴. |
|---|---|---|
| | you not have this CL book | You not have eat dinner |
| | 'You don't have this book.' | 'You haven't eaten dinner.' |

*Le* (了) and *meiyou* (没有) are the standard pair of affirmative and negative perfective markers in Mandarin (Shi, 2004). In this conversation, the yes answer should be utterance (20):

(20)　我　　吃　　饭　　了。
Wo³　chi¹　fan⁴　le.
I　　eat　dinner　PFT.
'I have eaten dinner.'

The answer to the question *youmeiyou*+VP should be VP+*le* or *meiyou*+VP in Mandarin, depending on whether affirmative or negative responses are intended. This perfective marker system involving *le/meiyou*+VP is therefore asymmetric.

*Meiyou* ('not have') before VP is an auxiliary for negating completion of the action (Zhao, 1979). In the affirmative response of the conversation, by getting rid of the negator *mei* in negative answers, *you* might be the affirmative perfective marker which functions equally as *le₁* marking perfect aspect [indicating "perfection in the sense of completion of the relevant act" (Radford, 1999)] as shown in sentence (21B). It is possible for Mandarin speakers to add *you* before VP since there exists a natural gap before VP. *You*+VP and *meiyou*+VP are possibly to form a symmetric system of perfective markers.

(21) A: 你　有　没　有　吃　饭？
ni³　you³　mei²　you³　chi¹　fan⁴?
you　have　not　have　eat　dinner?
'Have you eaten dinner?'

B: 我　有　吃　饭。
wo³　you³　chi¹　fan⁴.
I　have　eat　dinner
'I have eaten dinner.'

### 5.2 *you*+VP in declarative sentences

*You*+VP structure is also observed in declarative sentences with some verbal phrases such as 提高 'progress', 增加 'increase' and 变化 'change', all of which are also abstract NPs:

(22) a 我　学习　有　进步。
wo³　xue²xi²　you³　jin⁴bu⁴.
I　study　have　progress.
'My study shows progress.'

b 我　　学习　　进步　　了。
wo$^3$　xue$^2$xi$^2$　jin$^4$bu$^4$　*le.*
I　　study　　progress　PFT
'My study has progressed.' ('My study shows progress.')

(23) a 今　年　　的 产量　　　有　增加。
jin$^1$　nian$^2$　de chan$^3$liang$^4$　*you$^3$*　zeng$^1$jia$^1$.
this　year　de　crop　　　have　increment.
'This year's crop shows an increment.'

b 今　年　　的 产量　　　增加　　了。
jin$^1$　nian$^2$　de chan$^3$liang$^4$　zeng$^1$jia$^1$　*le.*
this　year　de　crop　　　increase　PFT
'This year's crop has increased.'('This year's crop shows an increment.')

(24) a 情况　　　　有　　变化。
qing$^2$kuang$^4$　*you$^3$*　bian$^4$hua$^4$.
situation　　　have　change.
'There is some change in the situation.'

b 情况　　　　变化　　　了。
qing$^2$kuang$^4$　bian$^4$hua$^4$　*le.*
situation　　　change　　PFT.
'The situation has changed.' ('There is some change in the situation.')

Examples (22a), (23a) and (24a) are definitely good in accordance with *you*+NP in Chinese grammar. 提高 'progress', 增加 'increase' and 变化 'change' are construable as either abstract nouns or verbs. Since Chinese lacks derivational morphemes, these words (namely 'progress', 'increase' and 'change') are morphologically the same when used as noun or verb. Their usage paves the way for the rise of *you*+VP construction. They provide evidence of a 'bridging context' (Evans & Wilkins 2000; Heine 2002) for the possibility of *you* being aspect marker in *you*+VP.

The lack of morphological marking between noun and verb in Chinese is an important feature facilitating this development. Interestingly though, it is language contact with southern dialects that sparked *you*+VP in Mandarin.

Another usage of *you*+VP in declarative sentences is illustrated in (25) to (27) below, where 给 'give', 去 'go' and 看 'watch' are pure verbs without ambiguity. It appears to be influenced by similar usage in Cantonese as demonstrated in section 4. General and emphatic indication of past event is shown in (25) and (26). Expression of habitual meaning is reinforced by frequency adverbs, as shown in (27).

(25) 那 天　生日，　　我　有　给 他 礼物。
na$^3$ tian$^1$　sheng$^1$ri$^4$, wo$^3$　*you$^3$*　gei$^3$ ta$^1$ li$^3$wu$^4$.
That day　birthday, I　have　give he present.
'I gave him a present on his birthday.'

(26) 我　有　去　参加　你的　婚礼，　恭　喜　你　啦。
　　　wo$^3$ *you$^3$* qu$^4$ can$^1$jia$^1$ ni$^3$de hun$^1$li$^3$, gong$^1$ xi$^3$ ni$^3$ la.
　　　I　have　go　attend　your　wedding, congratulations.
　　　"I did go to your wedding. Congratulations!"

(27) 我　天天　都　有　看　你的　影集，非常　　的　不错。
　　　wo$^3$ tian$^1$tian$^1$ dou$^1$ *you$^3$* kan$^4$ ni$^3$de ying$^3$ji$^2$, fei$^1$chang$^2$ de bu$^2$cuo$^4$.
　　　I　everyday　all　have see　your　album, very　　　good
　　　"I have been looking at your album everyday. It's really very nice." [新路网-留言板-李述钦留言本，2002. (cf. Yang & Dong, 2003)]

In Mandarin, 了 *le* and 没有 *meiyou* form an asymmetric aspect system. That is, *le* is attached postverbally, and preverbal *meiyou* has no etymological relation with affirmative *le*.

In this section, *you* is not a lexical verb in context but a functional category. There is just a negative morpheme *mei* before *you* to be removed for affirmation. In this case, *you* and *meiyou* are both preverbal for perfect aspect. A symmetric perfective marker system in Mandarin provides syntactic and cognitive foundation for *you* acquiring its functional features.

Bybee et al. (1994) presented the path of development of 'be/'have' leading to perfective and simple past as follows: 'be/'have'> resultative > anterior > perfective/simple past (p105, Fig. 3.1). By contrasting between perfective and past, Bybee et al. (1994) assumed that "simple pasts are more grammaticized than perfectives, and in some cases, are further developments from perfectives" (p.92). *You* in (23) and (24) in Mandarin and (17) in Cantonese (presented in section 4) indicates that *you* somehow is even undergoing extension to simple past context.

From earlier analysis of *you*+VP in Contemporary Mandarin, ancient Chinese and Cantonese, it is worth noting that *you* in *you*+VP structure in Mandarin is a functional category being a perfective marker which is complementary with *le$_1$* and in "bridging contexts" a simple past marker. However, in contrast to *le*, *you* has its specialness and can not completely be a substitute of *le*.

## 6. The Specialness of *you* in *you*+VP

Tense and aspect are common in languages in the world, yet every language may have its own logic and system of tense and aspect. Chinese is believed to be a tenseless language in which no overt temporal morpheme like *-ed* in English is attached to the predicates. But Chinese is not an aspectless language.

As a grammatical notion, aspect deals with how the event described by a verb is viewed. There are two main categories of aspect, namely, of imperfective and perfective (Comrie, 1976). According to Yang and Bateman (2002), "(in) the Chinese aspect system, the combination of individual aspects is much more semantically restricted and much less grammatically regular than that in English."

In Mandarin, *you*+VP structure usually occurs in response to *youmeiyou*+VP

questions, which leaves a natural gap before VP as highlighted in (28) 吃饭 'eat dinner'. If the question involves postverbal perfective marker *le₁* and sentence final interrogative particle 吗/*ma* (VP+*le*+*ma*) instead of preverbal *youmeiyou*, Mandarin speakers would refer to use perfective marker *le* as in (29). The postverbal and preverbal position of *le₁* and *you* is a major difference, especially in discourse.

(28) A: 你　有　没　有　　吃　饭　?
　　　Ni³ *you³ mei²* *you³* chi¹ fan⁴?
　　　You have not have　eat dinner?
　　　'Have you eaten dinner?'

　　B: 我　有　吃。
　　　wo³ *you³* chi¹.
　　　I　have eat.
　　　'I have eaten.'

(29) A: 你　吃　饭　了　吗?
　　　Ni³ chi¹ fan⁴ *le ma¹*?
　　　You eat dinner PFT *ma*?
　　　'Have you eaten dinner?'

　　B: 我　吃　了。
　　　wo³ chi¹ *le*.
　　　I　eat PFT.
　　　'I have eaten.'

Cantonese perfective marker "咗*zo²*" (Zhang, 1972; Yuan, 1989; Yue, 2004, etc.) can not appear together with "有*yauh*" (Zhang,1972), neither can *you* with *le* co-occur in Mandarin.

(30) *有　買　咗　　這　本　書。
　　　yauh maaih zo　ze bun syu.
　　　have buy　PFT　this CL book
　　　'I have bought this book.'

(31) *有　买　了　这　本　书。
　　　*you³* mai³ *le* zhe⁴ ben³ shu¹.
　　　have buy PFT this CL book.
　　　'I have bought this book.'

In practical speech, *you*+VP usually emphasizes a fact insisted by the speaker. In "我吃饭了/wo³ chi¹ fan⁴ *le*", *le₁₊₂* *here* refers to the completion of eating and it is relevant to the current state that "I'm not hungry now." In contrast, by saying "我有吃饭/wo³ *you³* chi¹ fan⁴", the speaker intends to emotionally affirm the fact of "I have eaten". This strong affirmation or assertion alludes to a (potential) contrastive view, perhaps one which the speaker suspects the hearer subscribes to, and which the speaker needs to counteract.

*You* also expresses habitual meanings when referring to the present, as discussed in section 5.2 and shown in Example (27), reproduced here as (32). This is another difference between you and *le* and *you*.

(32) 我 天天　都 有 看 你的 影集，非常　 的 不错。
　　 wo³ tian¹tian¹ dou¹ you³ kan⁴ ni³de ying³ji², fei¹chang² de bu²cuo⁴.
　　 I everyday all have see your album, very　　good
　　 "I have been looking at your album everyday. It's really very nice." [新路网-留
　　 言板-李述钦留言本，2002. (cf. Yang & Dong, 2003)]

## 7.　More Evidence of *you*+VP in Pre-Qin Works

(33) 皆 古　 圣 人 也，吾 未 能 有 行 焉；
　　 jie¹ gu³　 sheng⁴ ren² ye³, wu² wei⁴ neng² you³ xing² yan¹;
　　 all ancient saint people SFP, I NEG can have do SFP;

　　 乃 所 愿，则 学 孔子 也。
　　 nai³ suo³ yuan⁴, ze² xue² kong³zi³ ye³.
　　 I PRT wish, PRT learn Confucius SFP.

'All the three are ancient scholars; I could not have done as they did. What I wish is to learn (of the ways of) Confucius.'

[*Mengzi* (孟子), Warring States Period]

Evidences more than (33) in Pre-Qin works show that *you* had once been used as a perfective marker indicating completion of an action or event (Liu, 1999). Southern Chinese dialects such as Min, Kejia and Yue retain more Old Chinese grammar than Modern Mandarin does (Yuan Jiahua, 1983; Jerry Norman, 1995). *You* has aspect marker function after its grammaticalization in Kejia Dialect (Zheng, 2005); *yauh*+VP structure is still being used as aspect auxiliary in Cantonese further shows that the functional usage of *you* is preserved in the Chinese language. The trendy use of *you* as a functional category in Mandarin appears to have been triggered after language contact with southern dialects. This provides fertile soil for *you*+VP to be reacquired in Northern Mandarin. Heine and Kuteva (2005, p.14) pointed out that in the 'contact-induced' language change, "…the same grammatical category may re-emerge in a given language, and that this persistence or 'diachronic stability' is a phenomenon entirely language-internally conditioned."

## 8.　From Possession to Perfective Marker—Cognitive Development of *you*

Heine (1977): in English, as well as many other European languages, "constructions using 'have' as their predicate base have given rise to both expressions of verbal possession and of perfect aspect" (p.187). The case of *you* demonstrates the application of this tendency in Mandarin. Consider the examples below:

| | | |
|---|---|---|
| English: | I have a house. | I have eaten. |
| German: | Ich habe ein Haus. | Ich habe gegessen. |
| French: | J'ai une maison. | J'ai mangé. | (Heine, 1977) |

Heine (1977) noted that "the development from possession to perfect aspect has

been accounted for with reference to a kind of metaphorical process, whereby the possessor of an object is used as a vehicle to express the agent of an action … this process was triggered by specifying possession" (p.192). The stages of the metaphorical process are sketched as steps from A to E:

    A. He has a letter
    B. He has a letter # (a) written (one)
    C. He has written # a letter
    D. He has written
    E. He has gone            (Heine, 1977)

| A. | 我　有　一　个　通知。<br>$Wo^3$ *you*$^3$ $yi^2$ $ge^4$ $tong^1zhi^1$.<br>I　*have*　one　CL　notice<br>'I have a piece of notice.' | possessor has possessee→ |
|---|---|---|
| B. | 我　有　一　个　通知　（要）告诉　你。<br>$Wo^3$*you*$^3yi^2$ $ge^4$ $tong^1zhi^1$ $(yao^4)$ $gao^4su^4$ $ni^3$.<br>I　*have*　one　CL　notice　(to)　tell　you.<br>'I have a piece of notice to tell you.' | possessor has event / action → <br><br>agent has action→ |
| C. | 我　有　告诉　　你。<br>$Wo^3$ *you*$^3$ $gao^4su^4$　$ni^3$.<br>I　*have*　tell　　you<br>'I have notified you.' | agent has done something→ |
| D. | 我　有　看　这　部　电影。<br>$Wo^3$ *you*$^3$　$kan^4$　$zhe^4$ $bu^4$　$dian^4ying^3$.<br>I　have　watch　this CL　movie<br>'I have watched this movie." | agent has done something |

This metaphorical process also cognitively accounts for the grammaticalization of *you* in Mandarin.

At stage A, possession verb *you* is in a transitive structure where "I" am the possessor of the possessee "a piece of notice". At stage B, the possessee is modified by "to tell you" and "I" am the possessor of the action or event of "telling you (the piece of notice)" as well. Meanwhile, then, the possessor of the action of "telling you (the piece of notice)" can be viewed as the agent of the action. At stage C, "tell" is used as the main verb, the possessor is reinterpreted as the agent. The interpretation focuses on the agent 'having an action done (by the agent himself/herself). In effect, this is construable as 'the agent has done something'. Finally in stage D, the main verb might be preceded by the perfective marker *you*.

## 9. Conclusion

Chinese is believed to be a tenseless language, but rich in aspect. Through cross-dialectal influence in late 20th century, *you*+VP structure has become popular in Mandarin speaking areas. In Ancient Chinese, *you* was an auxiliary indicating perfect aspect before verbal phrases. *you*+VP template is thus available for Mandarin even though it was not recounted until recently. This paper tries to trace the possible development of *you* from a lexical verb to a functional category by comparing its

usage in Cantonese; and suggests that *you*+VP structure might not just be due to language contact, but there must be cognitive motivation and variable and structural facilitation within the language for its occurrence. The path of development of *you* proposed in this study is outlined in figure 2.

Analyses of *you*+VP in conversations (led by the *youmeiyou*+VP question) and in declarative sentences reveal that *you* has undergone its evolvement to be an aspect marker, which has some specialness that *le* has not. The Mandarin perfective marker system has the possibility of acquiring a symmetric system with *you* besides the asymmetric systems with *le*. The cognitive motivation of this new usage of *you* in Mandarin conforms to a commonly observed development of possession verbs evolving into aspect markers in many other languages.
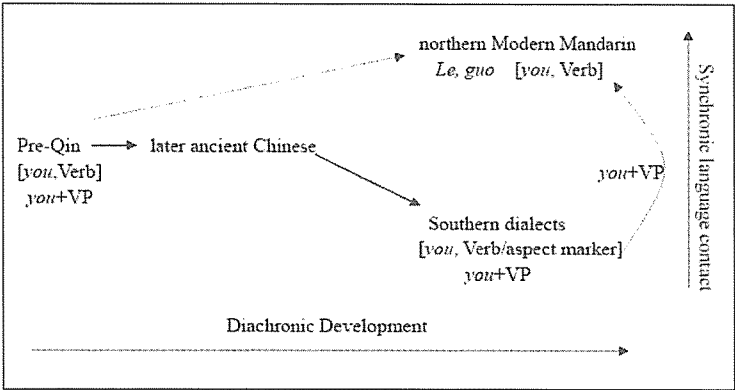


Fig. 2

## References

Bybee, Joan L. et al. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
Chao, Jizhou & Han Jingti et al (晁继周、韩敬体等). (2005). *Modern Chinese Dictionary: 5th Edition* (现代汉语词典: 第五版). Beijing: Commercial Press (北京: 商务印书馆).
Chappell, Hilary. (2004). *Chinese Grammar: Synchronic and Diachronic Perspectives*. Oxford: Oxford University Press.
Comrie, Bernard. (1976), *Aspect: an introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
Ding, Shengshu (丁声树). (1961). *An Introduction to Modern Chinese Grammar* (现代汉语语法讲话). Beijing: Commercial Press (北京: 商务印书馆).

Evans, Nicholas and Wilkins, David. (2000). In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76: 546‒592. Baltimore: Linguistic Society of America.

He, Jinsong (何劲松). (1988). The semantic and origin of the "*Zhi*" between subject and predicate (主语谓语之间 "之" 的词义及其源流). *Journal of China Middle-Southern Ethnic College (中南民族大学学报)*. V.6 106-117

Heine, Bernd. (1977). *Possession: Cognitive Sources, Forces and Grammaticalization.* Cambridge: Cambridge University Press.

----- (2002). *World Lexicon of Grammaticalization.* Cambridge; New York: Cambridge University Press.

Heine, Bernd and Tania Kuteva. (2005). *Language Contact and Grammatical Change.* Cambridge University Press.

Hopper, Paul and Elizabeth Traugott. (2003). *Grammaticalization.* Cambridge: Cambridge University Press.

Hou, Xuechao (侯学超). (1998). *Xiandai Hanyu Xuci Cidian* (现代汉语虚词词典). Beijing: Beijing University Press (北京: 北京大学出版社).

Huang, C.-T. James (黄正德). (1989). On 'be' and 'have' in Chinese. (说 "是" 和 "有"). *Bulletin of the Institute of History and Philology.* Vol.59. 43-64. Taiwan.

Luke, K. K. and O. T. Nancarrow. (2004). Auxiliary Verbs in Cantonese. *Studies in Cantonese Linguistics.* Hong Kong: Linguistic Society of Hong Kong.

Liu, Lianli (刘镰力). (2000). *A Dictionary of Chinese Usage: 8000 Words* (汉语 8000 词词典). Beijing: Beijing Language and Culture University Press (北京: 北京语言大学出版社).

Lü, Shuxiang (吕叔湘). (1942). Chinese Language-Grammar (中国文法要略). Beijing: Commercial Press (北京: 商务印书馆).

----- (1999). *Eight Hundred Words of Modern Chinese* (现代汉语八百词－增订本). Beijing: Commercial Press (北京: 商务印书馆).

Ma, Jianzhong (马建忠). (1983). *Ma's Chinese Grammar* (马氏文通). Beijing: Commercial Press (北京: 商务印书馆).

Matthews, Stephen and Virginia Yip. (1994). *Cantonese: A Comprehensive Grammar.* Routledge.

Norman, Jerry (罗杰瑞著, 张慧英译). (1995). *Chinese* (汉语概说). Beijing: Yuwen Chubanshe (北京: 语文出版社).

Ota, Tatsuo. (1958) (Translated by Jiang Shaoyu and Xu Changhua). *The Historical Grammar of Chinese.* Beijing: Peking University Press.

Radford, Andrew et al. (1999). *Linguistics: An Introduction.* Cambridge: Cambridge University Press.

Shi, Dingxu (石定栩). (2002). Syntactic changes induced by language contact: the case of Hong Kong written Chinese (香港书面汉语句法变异: 粤语的移用、文言的保留及其他). *Applied Linguistics (语言文字应用).* 3. Beijing: Language Press (北京: 语文出版社).

Shi, Yuzhi (石毓智). (2004). Possession verb and perfective marker in Chinese (汉语领有动词与完成体的表达). *Studies in Language and Linguistics (语言研究).* 24, 2. Beijing: Science Press (北京: 科学出版社).

Shi, Yuzhi and Li Na (石毓智, 李讷). (2000). Syntactic Change around the 15th Century and Formation of Chinese Negative Marker—The Grammaticalization of mei(you) (十五世纪前后的句法变化与现代汉语否定标记系统的形成—否定标记 "没（有）" 产生的句法背景及其语法化过程). *Linguistic Researches* (语言研究). *2*. Beijing: Science Press (北京: 科学出版社).

----- (2001). Chinese: motivations and mechanisms of evolution of Chinese morpho-syntax (汉语语法化的历程—形态句法发展的动因和机制). Beijing: Beijing University Press (北京: 北京大学出版社).

Wang, Li et al (王力等). (2005). *Gudai Hanyu Changyongci Cidian*: 4th Edition (古汉语常用词词典:第四版). Beijing: Commercial Press (北京: 商务印书馆).

Yang, Guowen and John A. Bateman. (2002). The Chinese Aspect System and its Semantic Interpretation. Paper submitted to the 19th International Conference on Computational Linguistics. Taiwan.

Yang, Wenquan and Dong Yuwen (杨文全, 董于雯). (2003). Language Variation: *you*+VP in Mandarin (语言变异: 汉语 "有+VP" 句简析). *Yuwen Jianshe Tongxun (语文建设通讯) 75*. Hong Kong: The Chinese Language Society of Hong Kong (香港: 香港中国语文学会).

Yuan, Jiahua (袁家骅). (1989). *Outline of Chinese Dialects* (汉语方言概要). Beijing: Wenzi Gaige Press. (北京: 文字改革出版社).

Yue, Anna. (2004). Materials for the diachronic study of the Yue Dialects. *The Joy of Research: A Festschrift in Honor of Professor William S-Y. Wang on His Seventieth Birthday*, ed. by Feng Shi and Zhongwei Shen. p 246-271. Tianjin: Nankai University Press.

Zhang, Hongnian (张洪年). (1972). *Cantonese as Spoken in Hong Kong* (香港粤语语法的研究). Hong Kong: the Chinese University of Hong Kong (香港: 香港中文大学).

Zhao, Yuanren (赵元任). (1979). *A Grammar of Spoken Chinese* (汉语口语语法). Beijing: Commercial Press (北京: 商务印书馆).

Zheng, Ying. (2004). Grammaticalization of "you" in Chinese (汉语 "有" 的语法化). *Language Journal (语文学报). 11*. Taiwan.

----- (2005). Grammatical feature and grammaticalization of "you" in Kejia Dialect of Taiwan (台湾客语动词 "有" 的语法特点与语法化). *Taiwan Language and language Education (台湾语言与语文教育). 6*. Taiwan.

Zhu, Dexi (朱德熙). (1982). *Lecture Notes on Grammar* (语法讲义). Beijing: Commercial Press (北京: 商务印书馆).

# Aspectual Contributions of Verbal Particles: Spatial, Temporal and Functional

Colleen WONG and Patricia MAN

Department of English
The Hong Kong Polytechnic University

## 1. Introduction

Verbal particles keep receiving interests from linguists of Germanic languages (e.g. Dehe et al. 2002). Apart from the special syntactic links between the particle and its host verb, the apparently unpredictable semantic contributions of the particle to the resulting construction in many cases arouse questions such as the complications involved for learners to acquire such knowledge. This paper looks at the acquisition of English verbal particles by Cantonese-speaking secondary school students in Hong Kong.

We study the operation of 'aspectuality marking' by means of 'particles'. In English the post-verbal elements are analysed as 'particles' since they form lexicalized combinations with the preceding verbs. We call these verb-particle constructions VPCs for short.

We devise two levels of aspectuality marking using particles. One is marking at semantic level, and this has two sub-types. The other is marking at functional level (cf. Hampe 2002). The picture looks like this:

> ### Marking at semantic level:
> *Type 1a: V + telic particle → bounded VPC*
> *Type 1b: V+ atelic particle → unbounded VPC*
>
> ### Marking at functional level:
> *Type 2: telic V + telic particle → bounded VPC (strongest event interpretation)*

Unlike Brinton (1988) who treats the telic particles as marking 'aktionart' meaning and the atelic particles as marking 'aspectual' meaning, we do not differentiate the semantic level on which these two groups of particles operate since they both mark 'aspect' in the broader sense of the term. The difference between the two particle groups lies in boundedness: telic particles make bounded VPCs and atelic particles make unbounded VPCs.

We find in our Cantonese learner corpus of written English that most particles are telic. The telic particles (Type 1a) largely retain their literal spatial meanings in the resultant VPCs. At the same time, a smaller number of VPCs are formed with telic particles used in their extended or metaphorical sense. Less than 10% of all VPC types involve particles used in their temporal, non-spatial sense which meet Jackendoff's (2002) description of aspectual particles. Some of these aspectual particles fall into our Type 1a, and some others, Type 1b. Only around 10% of all VPC types we have found are formed with atelic particles (Type 1b). And less than 10% of all VPC types involve type 2 marking.

This reveals a picture of the learners' acquisition of aspectuality. Students' use of verbal particles shows that aspectuality is expressed largely through retaining the spatial meaning of verbal particles rather than idiomatic meaning.

## 2. The Meanings of Verbal Particles

In this paper, we will follow the treatment of Lindner (1981) and use 'particle' as a term to refer to an adverb when the construction it enters into is shown to be VPC. VPCs are distinguished from verb-preposition constructions in the relative position between the particle or preposition with the noun object. This is shown in (1)-(4) below:

(1a) He sped *up* the process.
(1b) He sped u̲p̲ the pole.

(2a) The man reeled *in* the line.
(2b) The man reeled i̲n̲ the street.

(3a) He sped the process *up*.
(3b) *He sped the pole u̲p̲.

(4a) The man reeled the line *in*.
(4b)*The man reeled the street i̲n̲.

While a preposition can only appear immediately after the verb, therefore right before the object noun phrase, a particle can either precede or follow it. So in (1-4) above, only the (a) sentences contain particles (as italicized). The (b) sentences contain prepositions (as underlined).

Lindner (1981) takes both idiomatic and non-idiomatic constructions as VPCs. This differs from the more traditional studies such as Fraser (1976) which treat only idiomatic expressions as VPCs. The consequence of such difference is revealed in their analysis of the meaning of the particle: Whereas Fraser argues that the semantic contribution of the particle in a VPC is unsystematic and not independent, Lindner argues that the particle contributes systematically to the meaning, including literal and extended meanings, of the VPC.

### 2.1 Systematic vs. unsystematic semantic contribution

Fraser (1976:5) points out that there are VPCs of which 'the particle appears to have retained an adverbial force.' For example:

(5) bolt down, drink down, gulp down, swallow down
(6) hang up, nail up, paste up, screw up

The particle in another group of VPCs such as *faded out* 'appears to modify the meaning of the verb, giving it a completive sense' (p.6). These two groups of VPCs Fraser calls systematic combinations since 'the particle appears to cause a consistent change in meaning of the verb stem.' (p.5) But he takes this to be a small group as compared to the majority of English VPCs which involve figurative combinations.

Fraser attempts to define English VPCs as a particular type of lexical idiom. He argues that like idioms in general, the semantic interpretation of the VPC is independent of its component parts and so the particle should not be seen as contributing to the meaning of the VPC systematically.

Adopting the cognitive approach of Langacker's[1], Lindner (1981) treats both idiomatic and non-idiomatic expressions as proper members of VPCs by showing the semantic connection among the apparently diverse meanings of the same particle in various contexts. She offers explanations for the meanings of a particle in combination with different verbs in a schematic way. Taking *out* and *up* as examples, she shows how these adverbial particles can be interpreted through a network of semantic extensions using subschemas of 'removal', 'expansion' and 'departure' for *out*, and those of 'vertically higher' and 'approach' for *up*.

Among the three subschemas of *out*, 'expansion' (as in *flatten out*) and 'departure' (as in *set out*) are treated as limiting cases of 'removal'(p.139). Extensions of the subschema 'removal' bring the meanings of 'distinguishing, choosing, rejecting and change from hiddenness to accessibility' (as in *find out*) (p.103). Extensions of the subschema 'expansion' bring the meanings of 'clarification' (as in *lay out*) and 'distribution' (as in *give out*) (p.128-9).

As compared with *out*, the meaning of *up* is more abstract. And between the two subschemas, 'approach' is more common than 'vertically higher'. The completive sense is one of the extensions of the 'approach' subschema (p.147).

Later in this paper we will show that in our data, many particles retain spatial meanings.

## 2.2 Aspectual contributions of particles: semantic level

Now that we have moved to semantic analysis of particles, we will show Brinton's analysis of the aspectual meaning of particles. She argues that the most common particles like *up, down, off* and so on, instead expressing resultative meaning, mark telicity, which contributes to the aktionsart meaning. This corresponds to what we call Type 1a marking. She also suggests that a few other particles like *on* and *along* mark iterative or continuative aspect and they correspond to our Type 1b marking.

*Marking at semantic level:*
*Type 1a: V + telic marker -> bounded VPC*
*Type 1b: V + atelic marker -> unbounded VPC*

In Brinton's analysis, six particles (*up, down, out, off, through* and *over*) show clear telicity marking (our Type 1a marking). Three other particles (*on, along* and *away*) mark continuative or iterative aspects (our Type 1b marking).

The group of particles Brinton (1988) analyzes as marking telicity includes the more commonly used *up, down, out, off,* and the less frequent *through, over* and

---

[1] We refer to Langacker's versions of Space Grammar, the latest one being 1999.

*away* (in the sense of 'to the end', 'completely', 'until it is finished' or 'all of it/them' (1988:169).

(7) The children are *eating up* the candy.
(8) The management decided to *close down* the plant.
(9) The lights are *fading out*.
(10) You should *shut off* the electricity.
(11) Have you *thought through* the problem?
(12) We have *read over* the documents.
(13) She is *throwing away* her money.

She points out that the telic particles show their nature not so much through vague paraphrases of 'completely' or 'to an end', but through a series of formal tests which set apart telic from atelic expressions (Brinton 1988:171-3).

We agree with Brinton that the particles in (7-13) above are telic. But we think the meaning of *away* in (13) cannot be taken to be parallel to those other particles in the sentences cited here. We think *away* retains a strong spatial sense here. Thus we will treat this version of *away* as a telic particle not because it expresses any aspectual meaning (e.g. 'completion', 'continuation') through non-spatial denotation but through the spatial boundary denoted.

Telic particles are incompatible with stative verbs.[2] They can occur with activity verbs, which are by definition unbounded. But unexpectedly, they may also occur with bounded aspectual classes like accomplishments and achievements. Although not every achievement verb can take these telic particles (14-15), some are fine with them (16-18) (Brinton 1988:173-4):

(14) *notice up/out
(15) *realize out/through

(16) won out
(17) found out
(18 met up

Accomplishment verbs are in general possible with telic particles:

(19) heal/clean up
(20) flatten out

To explain this semantic redundancy of telic particles in VPCs involving accomplishment and achievement verbs, Brinton adopts Traugott's (1982) comment that the addition of a telic particle to an already telic verb 'serves to make a covert endpoint overt' and Lindner's analysis that 'the particle serves to "profile" the goal' (Brinton 1988: 175).

---

[2] Fraser (1976, p.11), for example, also observes that particles in general cannot combine with stative verbs.

Another group of particles which also contribute to the aspectual meaning of the VPC are *on, along* and *away.* Brinton analyzes them as marking continuative/iterative aspect (in the sense of 'continue' or 'keep (on)') as in the following sentences (Brinton 1988:175-6):

(21) He worked *away* at the problem for hours
(22) You should carry *on* with your work.
(23) You shouldn't string him *along*.

For activity verbs, which by definition refer to unbounded situations, it is natural to conceive that the particle delimits the situation denoted by the VPC by adding a boundary to it. However, Brinton's explanation of this semantic redundancy of telic particles going with telic verbs (of accomplishment and achievement) as to help 'make a covert endpoint overt' leaves room for more elaborate explanations. Accounts established on functional level and cognitive level may render the picture somewhat more convincing and comprehensive.

## 2.3 Aspectual contributions of particles: functional level

### 2.3.1 Hampe (2002): strongest event interpretation

Hampe (2002) argues that on functional level, semantically redundant particles are indices of the speaker's emotional involvement. She claims that the redundant VPCs represent the 'superlative' versions of their simple-verb counterparts in the sense that they are formally and lexically marked and in that they are used to refer to the 'strongest' event interpretation possible (the one referring to its entirety)' (p.153). Support for this argument has been found in corpus-based investigations.[3] It is observed that most redundant VPCs (e.g. *cool down, queue up, cover up*) are significantly less frequent than the corresponding simple verbs (eg. *cool, queue, cover*). This supports the hypothesis that the redundant VPC is the formally marked member since it meets the requirement of involving material surplus (the fact that a 'redundant' particle is added to the simple verb counterpart), and the requirement of significantly lower frequency.[4]

In terms of text types, redundant VPCs occur more frequently in spoken and less formal discourse (e.g. telephone conversations) and in those with 'evaluative' and 'persuasive' speaker goals (e.g. commentaries, reviews and editorials).

Hampe does not exclude the possibility that the semantically redundant part icles still contribute to the aspectuality of the VPC and express the speaker's involvement at the same time. The particle may be enhancing the lexical meaning of the VPC (p.147) at the same time.

On the cognitive level, a semantically redundant particle repeats a substructure of the conceptual base of the verb to which it is added. This results in a conceptual overlap that becomes a salient part of the conceptualizations. Such profiling function

---

[3] The corpora investigated include the British corpus collection which consists of the MCA, the LOB, the LLC and the SEC, and the Collins Online, an extract of the "Bank of English".

[4] Cf. Dik, Simon (1989), *The theory of functional grammar*. Dordrecht: Foris Publications. P. 41.

(cf. Lindner 1981) of the particles is congruent with their function to express the speaker's emotional involvement (Hampe 2002, p.181).

In the following we will introduce Jackendoff (2002)'s analysis of verbal particles to lend support for Hampe's 'strongest event interpretation' account.

### 2.3.2 Jackendoff (2002): excessiveness

Although Jackendoff (2002)'s basic concern is the syntax of VPCs, his analysis of both idiomatic and productive VPCs reveals that there is a correlation between 'excessiveness' and certain types of VPC. Before going into his analysis of VPCs which involve excessiveness, let us take a look at his semantic categorization.

Jackendoff divides VPCs into six groups according to their meanings. The first group are verb-particle idioms such as *throw up* (=vomit), or *bring up* (a child) and semi-productive expressions such as *pass out, black out, knock (NP) out, burn (NP) out* and so on (in the sense of 'go into an unusual mental state').

The second group involves directional particles such as *take up* as in sentences like '*Beth took the food up*'. We will see that this is the largest group of VPCs found in our data (section 3 Tables 7 and 8).

The third group involves aspectual particles including *up* in the sense of 'complete', *away* and *on* in the sense of 'keep on V-ing', *through* in the sense of 'from beginning to end', and *over* in the sense of 're-V'. We will come back to this group in section 3.2.3.

The fourth group is the *'time-away' construction* as in sentences like '*Bill slept the afternoon away*'. The subject of the sentence is in some sense heedlessly using the time up (p.83).

The fifth group is the totally productive *'V/ N-d out' construction* as in sentences like '*I'm (all) coffeed out.*' The subject of the sentence is 'worn out from too much V-ing or too much N' (p.85).

The last group contains the idiomatic constructions in the *'his heart out' family* as in sentences like '*Richard ran his head/butt off*' and '*Harold sang his heart out*'. The adverbial meaning of these constructions could be roughly paraphrased as 'to excess' (p.87).

Among the six groups of VPCs, four involve the meaning of excessiveness. They include (1) the semi-productive V+*out* , a sub-group of group one listed above, (eg. *pass out, black out, knock NP out, burn NP out, stress NP out*), which has the general meaning of 'go into an unusual mental state'; (2) the 'time-*away*' construction, which has the basic meaning of 'heedlessly using the time up'; (3) the totally productive 'V/N-d *out* ' construction, which means to be 'worn out from too much V-ing/too much N', and (4) idiomatic constructions in the family of *his heart out* (e.g. *Richard ran his head/butt off* or *Harold sang his heart out.*) which roughly means doing something 'to excess'.

Although the four groups of VPCs which express excessive meanings as analyzed in Jackendoff (2002) are not redundant VPCs, there is a link between 'excessiveness' and Hampe (2002)'s 'strongest event interpretation'. Both point to a more than expected completion of an event. The event is not only brought to an end by reaching its presumed endpoint. It has gone beyond this, making the conventional event boundary a measure of the extravagance of the whole situation.

Although in our corpus we do not find VPCs expressing excessive meanings as described in Jackendoff (2002), we do find a number of redundant VPCs as listed in Hampe (2002).

## 3. Data and Analysis

### 3.1 Data

In this section, we will present the VPCs found in our corpus. We have consulted O'Dowd (1998)'s list of twenty alternating P-forms (ie. forms that can be either preposition or adverb).[5] We have searched through our corpus for these items and found thirteen particles which form VPCs. They are *up, out, down, off, through, over, away, back, in, on along, around* and *by*.

As shown in Table 1, we have put the thirteen particles into two groups ie. telic and atelic, corresponding to our Type 1a and Type 1b marking respectively. Nine particles are telic and seven of them are mentioned in Brinton (1988), namely, *up, down, out, off, through, over* and *away*.[6] In addition, we have also found *in* and *back*. Four atelic particles are found in our corpus, namely *on, along, around,* and *by*. Among these, *on* and *along* are mentioned in Brinton (1988).

We identified constructions as VPCs only if the particle (ie. any of the P-forms) is not followed by any object NP, or the object NP following the particle can be replaced by a pronoun that can occur before the particle (ie. occurring in the VNP sequence). We have found 259 VPC types (not tokens) according to this criterion, of which 237 are formed with telic particles and 22 with atelic particles.

Table 1   telic and atelic particles (Type 1 marking)

|  | Telic particles (Type 1a marking) | Atelic particles (Type 1b marking) |
|---|---|---|
| Particles | *Up, down, out, off, through, over, away, in, back* | *On, along, around, by* |
| No. of VPC types formed | 237 | 22 |

Concerning our Type 2 marking which involve redundant particles, we found less than 10% (ie 20 out of 259) of the VPC types in our corpus that match Hampe (2002)'s list of redundant VPCs. A few other VPCs in our data may also be treated as redundant, though they are not found in Hampe's list. They include:

---

[5] 'away' is not on O'Dowd's list because it is not an alternating P-form, but we have included it in our list as it is treated as a telic particle in Brinton (1988) and meets the requirement of a directional particle, a group of verb particles identified in studies like Jackendoff (2002).
[6] 'away' is put under the telic group in our findings because no occurrence of VPCs with this particle involves the continuative aspect. All occurrences involve the directional meaning of the particle.

*retreat back, return back, raise up, sit down, stand up*

In Table 2, we have listed the particles and their host verbs of the 20 redundant VPCs found.

**Table 2  redundant VPCs (type 2 marking)**      *(total:20)*

| *particle* | *host verb* |
|------------|-------------|
| Up | queue, rise, break, cheer, cover, fix, light, open, pack, save, shut, smash, tidy, tie, wake |
| Out | cry, check |
| Down | calm, cool |
| Away | Hide |

In Tables 3-5, we have arranged the verb-particle combinations according to the aspectual class (ie. stative, activity, accomplishment or achievement) of their host verb.  Most VPC types (ie.153 out of 259) are formed with activity verbs (Table 3).  94 VPC types are found to form with accomplishment verbs (Table 4).  There are only 8 VPC types formed with achievement verbs (Table 5) and no VPC types are formed with stative verbs.  Our findings thus support the observation made in studies like Fraser (1976) and  Brinton (1988) that stative verbs do not occur with particles.

**Table 3  particles with activity verbs as host**      *(total VPC types formed: 153)*

| *particle* | *host verb* |
|------------|-------------|
| Up | blow, bring, call, climb , eat , be fed , grow , hold , jump, look , pick , pull , queue, raise, rise , sit, stand, stay, take, turn , use, walk, wash |
| Out | bring, call, carry, clamber , crash , cry, drop , fly, hold, jump, lay, leap, look, march, move, pick, pull, read, run, rush, scream, shoot, shout, sing, 'slap' (=slip), speak, stay, step, take, teach, tell, throw, turn, walk, wipe, work |
| Down | bent, calm, climb, cool ,count, drop, fall, follow, jump, knock, look, move,  push, run, sit, turn, walk, write |
| Off | bomb, fall, show, take, tear |
| Through | bring,   drive, jump, look, march,   read, 'rumble'(=rummage),   rush, spread, squeeze, sweep, think, walk |
| Over | fall, jump, run, talk, trip, turn |
| Away | bring, draw, fly, hide, keep,  move, pass,  run, show, stay, take, throw, walk |
| Back | bring, bully, call, drive, fight, look, move, phone, pull, run, take, teach, think, travel,  turn, wear |
| In | 'boom'(= bomb), join, move, rush, shoot, step, storm,   swim, take, walk |
| On | drive, keep,  pick, turn, walk |
| Along | walk |
| Around | crawl, look,  move, run, turn,  wander, watch |
| By | Pass |

**Table 4  particles with accomplishment verbs as host**      *(total VPC types formed: 94)*

| *particle* | *host verb* |
|------------|-------------|
| Up | break ,_build, catch, cheer, come, cover , fix, give, get, go, lessen, light, loosen , make, open, pack, put, save, send, set , shut, sign, smash, sum, tidy,  tie, wake |
| Out | break, burst , check, come, give, get , go, lock, put, sell, send, set, sort, wear |
| Down | break, come, cut, faint, flush, get, go, mark, put, settle, slip |

| Off | cut, get, go, pay, put, set |
|---|---|
| Through | burst, come,  go, get, pass |
| Over | faint, go, put |
| Away | get,  give, go, send, speed |
| Back | answer, come, change, get,  give, go,  retreat, return, send |
| In | check, come, fill, go, get, put |
| On | clip, come, get, go, put, switch |
| Around | go |
| By | go |

**Table 5   particles with achievement verbs as host**          *(total VPC types formed: 8)*

| Particle | host verb |
|---|---|
| In | fit |
| Out | find, leave |
| Off | die |
| Over | leave |
| Back | find,  see , welcome |

4 VPC types are formed with nouns or adverbs (*dress up, gangs up, be brimmed over,* and *double back*)

## 3.2. Analysis

### 3.2.1   Are particles mostly telicity markers?

As shown in Table 1, most VPCs found are formed with telic particles.  While 237 VPC types are telic, (Type 1a) only 22 VPC types are atelic (Type 1b).  It is true that more than half of the VPC types are formed with activity verbs and all of them are turned into accomplishments after combining with the telic particles except in one case (*take up*) (see Table 3).  But a considerable amount of VPCs are formed with already bounded aspectual classes (ie. accomplishment and achievement).  Recall that there are 94 VPCs formed with accomplishment verbs (Table 4) and 8 with achievement verbs (Table 5).  We do not think that the particles are added to mark a boundary, something which is already inherent in the host verb.  As mentioned in section 2.2, Brinton (1988) points out that the telic particles added to accomplishment or achievement verbs may help 'make a covert endpoint overt'.  But we think that there are two other good reasons to add a telic particle to an accomplishment or achievement verb: (1) to express spatial meaning; and (2) to express the speaker's involvement, as suggested in Hampe (2002).  We will first deal with spatial meaning.

Particles are analyzed as expressing spatial meaning when they are used in the literal directional sense.  As shown in Tables 6 and 7 below, among the 13 particles found in our data, 9 (except *up, over, on* and *by*) are found to be used in the literal spatial sense in at least half of the VPCs formed with them.

**Table 6          telic particles used in the literal spatial meaning**

| Telic particles | No. of VPC types with the literal spatial meaning of the particle | Total no. of VPC types formed with the particle |
|---|---|---|
| Up | 15 | 52 |
| Out | 27 | 52 |
| Down | 15 | 29 |
| Off | 6 | 12 |

| Through | 15 | 18 |
| --- | --- | --- |
| Over | 2 | 11 |
| Away | 17 | 18 |
| Back | 14 | 28 |
| In | 12 | 17 |
| (total) | (123) | (237) |

**Table 7**          **atelic particles used in the literal spatial meaning**

| Atelic particles | No. of VPC types with the literal spatial meaning of the particle | Total no. of VPC types formed with the particle |
| --- | --- | --- |
| On | 3 | 11 |
| Along | 1 | 1 |
| Around | 8 | 8 |
| By | 0 | 2 |
| (total) | (12) | (22) |

Given the high frequency of particles used in the spatial sense, we do not want to say that particles are telicity markers that express non-spatial meanings. Rather, we want to retain Brinton (1988)'s view by saying that the particles are mostly telicity markers but telicity is constructed through not only non-spatial (ie. temporal) but also spatial meanings expressed by the particles.

Lindner (1981)'s analysis provides a more comprehensive account of our data since it not only directly explains the large group of VPCs with particles retaining the spatial meaning, but also the extended uses of the particles.[7] More importantly, her analysis would also allow us to see why a number of VPCs are not used in their extended sense if the spatial meaning is the basic one. The following VPCs are those with particles which could have contributed to an idiomatic meaning but do not:

*call up (= wake up, =/= phone up), hold up (=/= rob), look up (=/=search for), pick up (=/=learn , improve etc.), run out (=/= use up), turn up (=turn louder, =/= show up)*

### 3.2.2 To what extent the particles contribute to the temporal (ie.non-spatial) meaning to the VPCs?

In section 2.3, we have introduced Jackendoff (2002)'s semantic categorization of verbal particles. He has put particles under six groups, one of which is the aspectual particles. He identifies the following five particles in the specified use as aspectual particles:

| | |
| --- | --- |
| (i) up (in the sense of "completely") | (our type 1a), |
| (ii) away (in the sense of "keep on V-ing") | (our type 1b) |
| (iii) on (in the sense of "keep on V-ing") | (our type 1b) |
| (iv) through (in the sense of "from beginning to end") | (our type 1a) |
| (v) over (in the sense of "re-V") | (our type 1b) |

---

[7] Although Lindner did not offer explanation for particles other than *up* and *out*, we could follow her approach in extending the different versions of a particle from its basic spatial meaning.

In our data, we found altogether 110 VPCs formed with *up, away, on, over* and *through* . Less than a quarter (ie. 23) of these VPCs involves the aspectual sense as specified in Jackendoff (2002). These VPCs are listed as below:

(i)15 (out of 52) VPCs with *up*, which include:

*blow up (a house), break up, bring up (a child), , cover up, dress up, be fed up (with), fix up, eat up, light up, loosen up, open up (one's mind), pack up, save up, shut up, smash up, sum up, tidy up, tie up, use up, wash up*

(ii) zero (out of 18) VPCs with *away*

(iii) 4 (out of 11) VPCs with *on*, which include:
*drive on, go on, keep on, walk on*

(iv) zero (out of 11) VPCs with *over*
(but we found 2 VPCs with *over* in other aspectual senses:
"completely": *be brimmed over*
"until it is finished": *go over*)

(v) 4  (out of 18) VPCs with *through*, which include:
*think through, look through, read through, go through (=experience)*

### 3.2.3 How many redundant VPCs' are found in our corpus? (Type 2 marking)

Apart from expressing spatial meanings, sometimes a telic particle is added to an accomplishment or achievement for functional reason. As shown in Table 2, there are 20 redundant VPC types found in our data. As mentioned in section 2.3, redundant VPCs occur more frequently in spoken and less formal discourse, it is unsurprising to find a low frequency of redundant VPCs (less than 10% out of all VPCs) in our corpus  which involves all written compositions. But it is interesting to see some redundant VPCs listed in Hampe (2002) (the ones put in brackets below) have the same host verb with those found in our data (shown below), only that they are formed with different particle so that the meanings of the particle and the host verb do not overlap:

(i) sell out (cf. sell up/off)
(ii) mark down (cf. mark out/off)
(iii settle down (cf. settle up)
(iv) write down (cf. write up)
(v) fill in (cf. fill up)

The above 5 VPCs found in the data all involve particles in the literal or extended spatial senses.

### 4. Conclusion

To the extent that Type 1a marking (i.e. the particles forming VPCs are mostly telic) overwhelmingly outnumbers type 1b marking, and that most of the host verbs in the telic VPCs are activity verbs (i.e. they are atelic before combing with the particles),

we conclude that particles found in our learner corpus contribute primarily to telicity. This is in line with Brinton's (1988) view that particles are basically telicity markers. But our study shows that we have to allow such aspectual feature (i.e. telicity) to be constructed not only by temporal (i.e. non-spatial) meaning, but also spatial meaning. Aspectuality of VPCs is often composed by spatial meanings, sometimes by temporal meanings and at other times, functional meanings.

## References

Binnick, Robert. (1993). *Time and the Verb: A Guide to Tense and Aspect.* New York and Oxford: Oxford University Press.

Brinton, Laurel. (1988). *The Development of English Aspectual System.* Cambridge: Cambridge University Press

Dehe, N. et al. (eds.) (2002). *Verb-particle Explorations.* Berlin, New York: Gruyter

Fraser, Bruce. (1976). *The Verb-particle Combination in English.* New York: Academic Press

Hampe, Beate. (2002). *Superlative Verbs: A Corpus-based Study of Semantics Redundancy in English Verb-particle Constructions.* Tubingen: Gunter Narr Verlag Tubingen.

Jackendoff, Ray. (2002). English particle constructions, the lexicon, and the autonomy of syntax. In N. Dehe et al. (eds.) *Verb-particle explorations.* Berlin, New York: Gruyter

Langacker, Ronald. (1999). *Grammar and Conceptualization.* Berlin, New York: Gruyter

Lindner, Susan J. (1981). *A Lexico-semantic Analysis of English Verb Particle Constructions with OUT and UP.* Ph.D. dissertation, University of California, San Diego.

O'Dowd, Elizabeth. (1998). *Prepositions and Particles in English: A Discourse-Functional Account.* New York: Oxford University Press

Quirk, R, Greenbaum, S., Leech, G. and Svartvik, J. (1985). *A Comprehensive Grammar of the English Language.* London: Longman.

Traugott, Elizabeth. (1982). From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmmann and Yakov Malkiel (eds.) *Perspectives on Historical Linguistics*, 245-271. Amsterdam and Philadelphia: John Benjamins.

# 普通话中情态词和情态层级的儿童习得

## FAN Li

## College of Foreign Language
## Beijing Forestry University

**摘要**

根据对3名北京儿童143次访问记录转写的语料，文章首先讨论了儿童早期习得主要情态词的情况。从语义发展的角度，研究发现动态性和道义性情态意义的获得要早于认知性的。这个习得顺序与许多跨语言研究的发现是一致的，从而证明存在一个普遍的习得情态意义的顺序。为了探究情态层级的习得情况，在60名儿童（04;00- 08;00）和15名成人中进行了语言测试。测试结果表明，4-5岁儿童已经开始意识到情态词肯定式间的语义差异，7-8岁儿童的相关知识已较成熟。与此相对照，当判断双重否定式间语义强度差异时，7-8岁儿童仍表现出与成人的显著差异。根源性和认知性情态词的差异不会影响儿童习得情态层级。

## 1．0　已有研究

对儿童习得情态词和情态层级的研究可以帮助我们观察到语言发展和认知发展之间的联系，因此是一个非常重要的语言习得和语义学的课题。

关于儿童情态词汇的发展，跨语言的文献非常丰富。关于英语的研究，如 Kuczaj (1977) 、Wells (1979，1985)、Pea 等（1982）、Shepherd (1981, 1982)、Perkins (1983)、Gee 和 Savasir (1985)、Shatz 和 Wilcox (1991)；关于德语的，如 Stephany (1993)；关于法语的，如 Bassano (1996)；关于韩语的，如 Choi（1995）。关于汉语的，如 Guo (1994, 1995)。更详细的情况，可参见 Papafragou（1997; 2000; 2001）对此类研究所作的回顾。以上研究均发现：在儿童语言发展过程中，与根源性的(Root)情态词相比，认知性的(Epistemic)情态词出现时间晚，而且在语言发展早期使用频率低。

关于情态词语义层级的习得，Hirst 和 Weil (1982)曾就情态词的语义强度差异对 54 名 03;00-06;06 儿童进行过测试。根据他们的测试结果，儿童掌握含有认知性情态词的命题的语义强度的时间约在 05;00，而掌握含有道义性(Deontic)情态词的命题的语义强度的时间则大约晚一年，约在 06;00。

对于以上两个课题，汉语的现有研究存在以下不足：

第一，从情态词词汇习得的角度来考察，缺乏长期跟踪所得的自然语料的支持，所得的研究结论不够充实；对 3 岁以前儿童的语言发展关注不够，难以把握相关语言知识和认知能力发展的脉络；考察的情态词有限，无法完整勾勒出儿童早期情态词知识的概况。

第二，从情态层级的习得研究来审查，根据所掌握的文献，没有发现综合利用长期观察的语料和语言试验的结果来系统讨论普通话儿童如何习得情态层级的。

## 2.0 研究方法

本研究采用跟踪取样和语言测试相结合的方法来获取语料。

## 2.1 跟踪取样

### 2.1.1 取样信息

关于儿童对情态词的习得情况，本研究运用了对 3 名生长在北京的儿童进行家庭访问所得的 143 次转写记录。语料收集工作是从他们 1 岁左右开始。下表含他们的简要信息：

| 儿童 | 性别 | 出生日期 | 访问时段 |
|------|------|-----------|-----------|
| CY | 女 | 2001/12/30 | 00;10;05-02;04;31 |
| SJQ | 女 | 2001/10/16 | 01;02;06-01;11;29 |
| ZTX | 男 | 2001/12/20 | 00;11;18-02;06;02 |

语料收集采用家庭访问的形式。每次访问的时间原则上是 1 个小时，但是由于一些客观原因，极少数的访问不足 1 个小时。除了非典流行和其他一些特殊原因以外，访问在 2 岁以前，1 周 1 次；2 岁以后，2 周 1 次。在访问过程中，我们对儿童与周边人群的自然语言交流进行录音和录像，接着对音像资料进行数字化处理，而后根据CHILDES[1]对儿童语料的转写格式人工进行文字转写，最后再由一名非转写人员[2]进行核查。

### 2.1.2 语料分析

在研究情态词的语言发展时，本文关注儿童对认知性、道义性和动态性情态词的使用情况。在筛选儿童语料时，遵循以下原则：

    1. 不讨论发音不清晰，而且转写和核查人员都无法辨别内容的词语。
    2. 不讨论语音形式相同，但不是情态词用法的词语。
    3. 不讨论无法判断情态词确切用法的情况；
    4. 不讨论儿童在歌谣和诗歌中使用情态词的情况。
    5. 考察情态词出现在肯定句、否定句和疑问句中的所有情况。

## 2.2 语言测试

关于儿童对情态层级的习得情况，采用语言测试的方法来获取语料。

---

[1] 关于CHILDES的有关信息，可参见http://childes.psv.cmu.edu。
[2] 所有的转写和核查人员均说标准和流利的普通话。

## 2．2．1 被试信息

语言测试共涉及 75 名被试，被试的来源信息见下表：

| 测试的学校 | | 测试的时间 | 测试的年龄组 | 测试的人数 |
|---|---|---|---|---|
| 儿童组 | 解放军总后勤部五一幼儿园 | 2004/11/8-2004/12/3 | 4-5 岁组 | 15 人 |
| | | | 5-6 岁组 | 15 人 |
| | | | 6-7 岁组 | 2 人 |
| | 前八家小学 | 2004/11/29-2004/12/17 | 6-7 岁组 | 9 人 |
| | | | 7-8 岁组 | 10 人 |
| | 海淀区实验三小 | 2004/12/20-2004/12/23 | 6-7 岁组 | 4 人 |
| | | | 7-8 岁组 | 5 人 |
| 成人组 | 北京林业大学外语学院 | 2004/11 | | 15 人 |

各组儿童的年龄信息：

| 测试的年龄组 | 起止年龄 | 平均年龄 |
|---|---|---|
| 4-5 岁组 | 04;02-04;11 | 04;05 |
| 5-6 岁组 | 05;03-05;11 | 05;05 |
| 6-7 岁组 | 06;00-06;09 | 06;06 |
| 7-8 岁组 | 07;00-08;01 | 07;07 |

成人组的被试是英语专业的学生，他们没有相关的语言学背景，能说标准流利的普通话。


## 2．2．2 测试方法

测试的方式是被试看图象、听故事，而后回答测试问题。测试故事是由北京林业大学 4 名普通话标准和流利的大学生帮助拍摄的。测试问题的录音者是北京语言大学 1 名从事对外汉语教学的女教师。

在进入测试故事前，都会对每个被试进行 1 个预备项目，以保证被试能集中精力进行随后的测试，并帮助他们熟悉测试进行的方法。

成人组的测试是在一间安静的多媒体教室内集体进行的。在测试过程中，要求被试通过大屏幕看故事，之后根据故事内容回答录音机播放的测试问题，并将答案写在测试卷上，无需就答案进行说明。

儿童的测试是以主试和被试一对一的形式在一间安静的教室内进行的。测试故事是通过手提电脑播放，测试问题仍由录音机播放。在测试进行中，主试会要求儿童对他们的回答进行简短说明，以确保他们的回答是有根据的，并能专心进行测试。在测试过程中，对每个儿童的测试都进行了录音。测试结束后，根据录音资料将儿童对问题的答案和说明进行转写。

## 2．2．3 测试内容

本文考察了认知性和道义性的情态层级的习得情况。对这两种情态层级的测试是通过情态词的肯定式和双重否定式来完成的。语言测试涉及的层级关系包括：

| 层级的属性 | 肯定式 | 双重否定式 |
|---|---|---|
| 认知性的情态层级 | "一定"和"应该" | "不可能不 VP"和"不应该不 VP" |
| | "应该"和"可能" | "不应该不 VP"和"不一定不 VP" |
| | "一定"和"可能" | |
| 道义性的情态层级 | "必须"和"应该" | "不可以不 VP"和"不应该不 VP" |
| | "应该"和"可以" | |
| | "必须"和"可以" | |

预备项目和测试故事详见附录一。

## 2．2．4 统计方法

在统计时，只有答案具备下面两个条件，我们才能确保被试儿童能区别故事中涉及到的两个情态词在语义强度上的差异，并且能根据故事提供的情节来进行判断：

第一，选择的答案与所有或绝大多数被试成人一致；

第二，在对答案进行说明时，能重复测试故事中的情态词或使用同义词或近乎同义的词。

以测试故事 6 为例。成人组 100%（15/15）的人次选择"小狗"，表明"一定"表达确定的语气，在语义强度上要强于"可能"。根据成人的判断和以上两个判定儿童回答的标准，我们认为以下 6-7 岁组的 儿童都是把握了"一定"和"可能"的层级关系的：

(1)　YTM：小狗。小狗说我一定去。　　（以重复的方式来解释答案）
　　　　YTX：小狗。小狗说它肯定去。　　（以使用同义词来解释答案）
　　　　XRX：小狗。它说它必须去。　　　　（以近乎同义的词来解释答案）

XRX 答案中的"必须"和故事中的"一定"的核心意义都是表达必然性，只不过"必须"表达的是道义逻辑上的必然，而"一定"表达的是认知逻辑上的必然。

## 3．0　情态词的习得

要讨论情态层级的习得问题，就必须首先了解情态词的习得问题。分析长期的语料，范莉（2006）详细报告了普通话儿童在词汇和语义两方面习得主要情态动词的情况。由于篇幅所限，在此只能总结与本文相关的主要发现。表5是主要的情态动词在 3 名儿童的语料中始现的时间：

3

---

3 由于篇幅要求，无法在此文提供详细语料。

| 情态动词 | 语义概念 | 分类 | CY | SJQ | ZTX |
|---|---|---|---|---|---|
| 会 | 表能力 | 动态性（根源性） | CY01;04;00 | SJQ01;04;14 | ZTX01;07;12 |
| | 表估计 | 认知性 | | SJQ01;09;04 | ZTX01;10;20 |
| 能 | 表能力 | 动态性（根源性） | CY01;08;16 | SJQ01;10;15 | ZTX01;07;00 |
| | 表客观条件许可 | 动态性（根源性） | CY01;04;00 | SJQ01;08;13 | ZTX01;03;17 |
| | 表准允 | 道义性（根源性） | CY01;10;18 | SJQ01;09;04 | ZTX01;11;03 |
| | 表估计 | 认知性 | CY02;00;13 | | ZTX02;03;22 |
| 敢 | 表有胆量 | 动态性（根源性） | CY02;01;10 | SJQ01;11;08 | |
| 可能 | 表估计 | 认知性 | | SJQ01;10;02 | ZTX02;03;08 |
| 可以 | 表能力 | 动态性（根源性） | CY02;01;10 | SJQ01;11;00 | |
| | 表准允 | 道义性（根源性） | CY01;11;09 | SJQ01;08;21 | ZTX01;08;11 |
| | 表客观条件许可 | 动态性（根源性） | CY02;01;10 | SJQ01;11;29 | |
| 应该 | 表情理 | 道义性（根源性） | | | ZTX01;09;01 |
| | 表估计 | 认知性 | | | ZTX02;02;09 |

第一，从情态动词作为一类词始现的时间来说，它们在普通话的儿童语言中出现的时间是在 01;04 左右。

第二，从个体情态动词的始现顺序来说，"会"和"能"是儿童最早获得的词，"敢"、"可能"和"应该"是较晚获得的词。成人语言中的情态动词并没有全部在所掌握的 02;05 左右的儿童语料中出现，例如表意愿的"肯"。与本文相关的情态词和义项的习得的情况是：表估计的"可能"在 SJQ 和 ZTX 的语料中出现的时间分别是 01;10;02 和 02;03;08，而在 CY 的 02;04;31 之前的语料中未出现；表准允的"可以"出现是在 01;08-01;11 左右；表情理和估计的"应该"只在 ZTX 的语料中出现。除了表准允的"可以"外，"应该"和"可能"的始现时间较晚或在个别儿童语料中根本没有出现。

第三，习得顺序优先的词也是在早期儿童语言中使用频率最高的。从各个词的使用频率来看，"能"（共 191 次）和"会"（共 181 次）最高，"可以"（共 48 次）其次，"敢"、"可能"和"应该"很低。与本文相关的情态词，除了表准允的"可以"使用次数较多外，"可能"和"应该"在 02;05 左右以前的儿童语料中的使用频率也是很低的。

| 情态动词 | 语义概念 | 分类 | CY 的语料 | SJQ 的语料 | ZTX 的语料 | 总次数 |
|---|---|---|---|---|---|---|
| 会 | 表能力 | 动态性（根源性） | 40 | 33 | 97 | 170 |
| | 表估计 | 认知性 | 0 | 3 | 8 | 11 |
| 能 | 表能力 | 动态性（根源性） | 20 | 9 | 50 | 79 |
| | 表客观条件许可 | 动态性（根源性） | 12 | 3 | 67 | 82 |
| | 表准允 | 道义性（根源性） | 6 | 3 | 18 | 27 |
| | 表估计 | 认知性 | 1 | 0 | 2 | 3 |
| 敢 | 表有胆量 | 动态性（根源性） | 1 | 1 | 0 | 2 |
| 可能 | 表估计 | 认知性 | 0 | 3 | 2 | 5 |
| 可以 | 表能力 | 动态性（根源性） | 1 | 1 | 0 | 2 |
| | 表准允 | 道义性（根源性） | 2 | 15 | 25 | 42 |
| | 表客观条件许可 | 动态性（根源性） | 3 | 1 | 0 | 4 |
| 应该 | 表情理 | 道义性（根源性） | 0 | 0 | 3 | 3 |
| | 表估计 | 认知性 | 0 | 0 | 2 | 2 |

第四、从语义发展的角度来看，普通话的儿童对不同情态意义获得的时间不同：首先是表"客观条件许可"和"能力"的，而后是表"准允"的，最后是表估计的。习得顺序优先的意义也是儿童在早期使用语言表达最多的。

第五、从情态词内部类别来看，首先是在 01;04 左右出现动态性的情态词（表"客观条件许可"和"能力"），而后是在 01;08 左右出现道义性的情态词（表"准允"），最后是在 01;09 之后才出现认知性的情态词（表估计）。在一个情态动词可表根源性和认知性情态意义的情况下，前者的获得要早于后者。无论从多义词的不同情态意义的使用情况来看，还是从不同情态意义在已出现的语料中的总体分布情况来考察，儿童早期最主要是使用动态性的情态动词，道义性的情态动词占少量比率，认知性的情态动词则较少使用。

在 143 次访问记录中，搜寻 3 名儿童对情态副词"一定"和"必须"的使用情况，发现：

第六、在所掌握的 02;05 之前的语料中，没有 3 名儿童使用"一定"的记录；只有 ZTX 在 02;02;23 一次"必须"的记录：

(2) *FAT: 再来一次.

　　*CHI: 我必须#饿了吃这西红柿.　　　　　　ZTX02;02;23

　　根据所掌握的语料，除了表准允的"可以"之外，其它与下面层级研究相关的情态词，"一定"、"必须"、"应该"和"可能"，被儿童熟练掌握至少应是在 02;05 之后。


## 4．0　情态层级的习得

### 4．1　肯定式的情态层级

　　在可能逻辑中谈情态层级，是说表达"可能性"和"必然性"的情态词可根据所表达的"可能性"的大小构成一个语义层级。与本文相关的情态词以下列方式位于情态层级上：
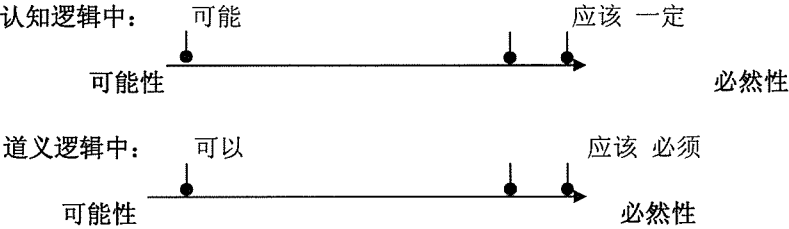


### 4．1．1　认知性的情态层级

　　下表是肯定式中认知性情态层级的测试情况：

| 组别 | "一定"强于"应该" | | "应该"强于"可能" | | "一定"强于"可能" | |
|---|---|---|---|---|---|---|
| | 人数 | 所占比例 | 人数 | 所占比例 | 人数 | 所占比例 |
| 4-5 岁组 | 4 | 26.67% | 2 | 13.33% | 4 | 26.67% |
| 5-6 岁组 | 11 | 73.33% | 11 | 73.33% | 13 | 86.67% |
| 6-7 岁组 | 11 | 73.33% | 10 | 66.67% | 11 | 73.33% |
| 7-8 岁组 | 14 | 93.33% | 11 | 73.33% | 15 | 100% |
| 成人组 | 15 | 100% | 15 | 100% | 15 | 100% |

从测试的结果来看，

　　第一，在 4-5 岁组，已经有少数儿童开始意识到认知性情态词在表达可能性时的差异；在 5-6 岁组，绝大多数儿童已经基本掌握了"一定"、"应该"和"可能"的语义强度的差异。

　　第二，对于儿童来说，"应该"与"可能"间的差异不如"一定"与"应该"、"一定"与"可能"间的差异显著。在对"应该"和"可能"层级差异的判断时，7-8 岁的儿童仍然与成人存在显著差异（$z=2.15, p<.05$）。不同的是，对"一定"和"可能"进行

判断，7-8 岁的儿童与成人的判断已经完全一样；对"一定"和"应该"进行判断，7-8 岁的儿童与成人也不存在显著性差异（z=1.02, p>.05）。产生这种差别的原因是"应该"和"可能"是"极大可能性"和"可能性"间的程度区别；"一定"与"应该"、"一定"和"可能"是"必然性"和"可能性"间的质的区别。

## 4．1．2 道义性的情态层级

关于道义性层级的测试结果和发现如下：

| 组别 | "必须"强于"应该" | | "应该"强于"可以" | | "必须"强于"可以" | |
|---|---|---|---|---|---|---|
| | 人数 | 所占比例 | 人数 | 所占比例 | 人数 | 所占比例 |
| 4-5 岁组 | 3 | 20% | 1 | 6.67% | 5 | 33.33% |
| 5-6 岁组 | 13 | 86.67% | 10 | 66.67% | 11 | 73.33% |
| 6-7 岁组 | 12 | 80% | 10 | 66.67% | 12 | 80% |
| 7-8 岁组 | 15 | 100% | 11 | 73.33% | 15 | 100% |
| 成人组 | 15 | 100% | 15 | 100% | 15 | 100% |

第一，4-5 岁时，少数儿童开始意识到道义性情态词的层级关系。到 5-6 岁，大部分儿童已经掌握了"必须"、"应该"和"可以"之间在意义上的区别。

第二，7-8 岁的儿童对"必须"与"应该"，"必须"与"可以"的判断与成人已经一样，而对"应该"与"可以"的判断与成人的存在显著性差异（z=2.15, p<.05），这个现象和导致的原因与认知性情态层级测试中发现的一样。

## 4．1．3 讨论

测试结果反映出，在 4-5 岁阶段，已经有少数儿童开始意识到情态词在表达认知上和道义上的可能性的差异；5-6 岁和 6-7 岁阶段，是这方面的知识飞速发展的阶段；到了 7-8 岁阶段，儿童这方面的知识已经比较成熟，并且他们能用语言较好地说明情态词在语义强度上的区别，例如部分儿童对测试故事 1 和 2 中问题的回答：

> (3) 测试故事 1
>     7-8 岁组
>     WZY：小熊。小熊说了"一定"，就是它不能不去。
>     YZZ：小熊。小狗说我应该，不是肯定去；小熊说我肯定去。
> (4) 测试故事 2
>     7-8 岁组
>     WQF：小兔子。它说它应该会。"应该"比"可能"可能性要大一些。
>     WZY：小兔子。它说应该。"应该"比"可能"确定一点。

情态词意义的差别越大，这种差别就越早被儿童掌握。这与 Hirst 和 Weil (1982)的发现是一致的。认知上的"一定"与"应该"和"一定"与"可能"，道义上的"必须"与"应该"和"必须"与"可以"间的差异对儿童来说比较明显，因为它们间的区别是本质上的"必然性"和"可能性"的差异。"应该"与"可能"、"应该"与"可以"间的差别是程度上的"极大可能性"和"可能性"的差异，所以早期儿童的掌握情况就稍差些。

情态词的认知性和非认知性属性的差异不会影响儿童对情态层级的习得。虽然长期跟踪语料表明出儿童在获得这两种不同的情态词时有先后次序，但测试结果却并没有反映出，儿童掌握道义性情态层级的情况明显比掌握认知性情态层级的情况好。将相关情态词的测试结果进行对比：
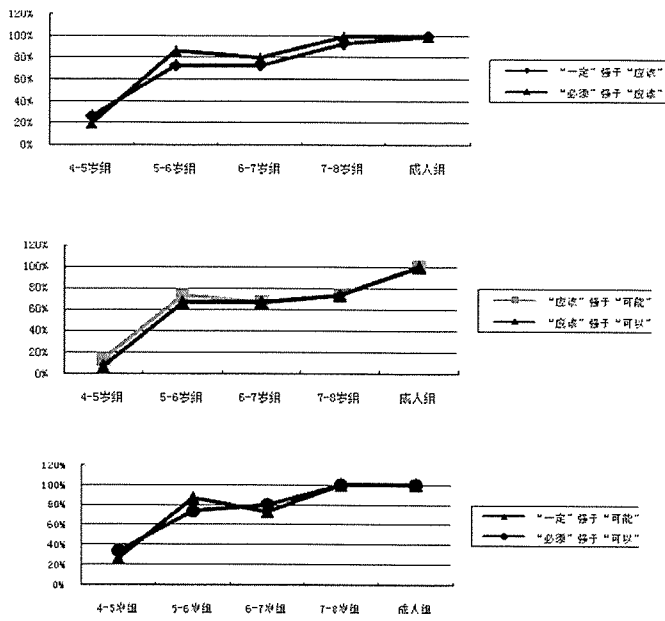






一方面，本文的测试结果不支持这样的观点：道义性的情态意义比认知性的情态意义容易习得，所以相对于认知性的情态层级，儿童更易掌握道义性的情态层级。我们的测试结果证明，儿童一旦从词汇和语义的角度掌握了这些词语，就不会在情态层级的判断上再受到概念复杂性的困扰。

另一方面，本文的测试结果也不支持 Hirst 和 Weil (1982)的发现。

Hirst 和 Weil (1982)的测试结果反映出，儿童掌握认知性情态词语义强度的时间要早于道义性情态词语义强度的时间。他们(1982: 666)认为这是由于道义性的情态要牵涉到一些复杂的社会力量(Social dynamics)，而认知性的情态只是简单的猜测游戏。在他们进行道义性层级的测试中，有两个老师分别给儿童下指令，而后看儿童如何反映。其中一个老师先说：

(5) 老师 A："You $X$ go to the green(red) room." "你 $X$ 去绿色（红色）的房间里。"

而后另一个老师说：

(6) 老师 B："You $Y$ go to the red (green) room." "你 $Y$ 去红色（绿色）的房间里。"

$X$ 和 $Y$ 会用情态词对 must/should、must/ may 和 should/may 分别加以替换。这样的测试方法会产生一些非语言的问题来干扰被试，比如一个被试听从了老师 A 的指令，可能仅仅是因为老师 A 在他的心目中比老师 B 更具有权威性。对于类似的问题，Hirst 和 Weil (1982)也应该是意识到的，如他们(1982: 666)所说的"…两个对立的权威所带来的复杂的社会力量将孩子们弄糊涂了…("…, there was the complicated social dynamics involved in two contradictory authority figures to confuse the children;… .")"。

在我们的测试中，通过动物来讲述测试内容，强调被试要根据故事情节来判断，要求儿童对自己的答案进行说明。因此，Hirst 和 Weil (1982)测试中非语言的干扰因素应该是得到了很好地控制。从本文的测试结果来看，儿童对两种情态层级掌握的情况相似:汉语儿童在5-6岁已经表现出对情态意义差别的极大敏感，7-8 岁时这方面的知识已经基本成熟。

## 4．2 双重否定式的情态层级

总结汉语中必然性情态算子、可能性情态算子分别与否定算子的演算规则：

（不一定不）　¬□¬　　　　　　　　　　　　　　　¬□　（不一定）

（可能、能、会）◇　　　　　　　　　　　　　　　◇¬　（可能不）

（不可能不、不能不、不会不）¬◇¬　　　　　　　　　¬◇　（不可能）

（一定）　□　　　　　　　　　　　　　　　　　　□¬　（一定不）
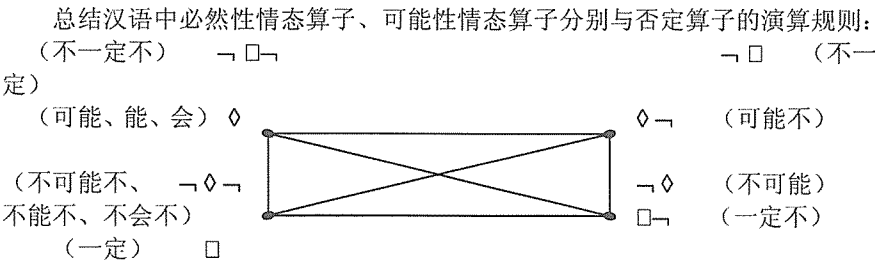
汉语情态词进入双重否定式后，语义强度上会有以下三种变化：

第一，否定加否定弱于肯定，如"不一定不 VP"；
第二，否定加否定强于肯定，如"不可能不 VP"；
第三，否定加否定等于肯定，如"不应该不 VP"。

## 4．2．1 认知性的情态层级

为了考察双重否定对儿童习得层级的影响，将双重否定式间的层级关系和在意义上与之对应的肯定式间的层级关系进行对照。表 9 是"极大可能性"和"可能性"层级关系的测试结果：

| 组别 | "不应该不 VP"强于"不一定不 VP" | | "应该"强于"可能" | |
|------|------|------|------|------|
| | 人数 | 所占比例 | 人数 | 所占比例 |
| 4-5 岁组 | 1 | 6.67% | 2 | 13.33% |
| 5-6 岁组 | 8 | 53.33% | 11 | 73.33% |
| 6-7 岁组 | 8 | 53.33% | 10 | 66.67% |
| 7-8 岁组 | 11 | 73.33% | 11 | 73.33% |
| 成人组 | 15 | 100% | 15 | 100% |

从语义差异度对习得的影响来看，上文已谈到，由于"应该"与"可能"间的语义差异是"量"上的不同，所以儿童差异的把握有些延缓。在判断双重否定式和肯定式时，7-8 岁的儿童与成人仍存在显著性差异（z=2.15, p<.05）。

再对比双重否定式和肯定式的测试结果，5-6 岁的儿童绝大部分已经了解"应该"和"可能"之间在意义上的区别，但是对"不应该不 VP"和"不一定不 VP"间语义强度差别的了解到了 7-8 岁阶段才到达一个比较好的水平。显然，这是双重否定复杂的计算过程造成的。要分辨"不应该不 VP"和"不一定不 VP"间的语义强度就必须先进行计算。测试结果表明，绝大部分儿童到了 6-7 岁，甚至 7-8 岁阶段才能熟练进行双重否定式的计算。

双重否定式给儿童所造成的困难不仅体现在他们较晚才能理解它所表达的意义，而且也反映在他们口头重复双重否定式时，仍有可能错误地将第二个否定词漏掉，比如测试故事 7 中 7-8 岁组的 ZJY 的回答：

(7)　ZJY：小兔子。天线宝宝说小熊你不应该穿的裙子是红色的，说小兔子你穿的裙子一定是红色的。

双重否定式中的第二个"不"，对于没有能力处理双重否定的儿童来说，它是没有语义内容，甚至没有语音内容的。

下面是"必然性"和"极大可能性"的测试结果和发现：

| 组别 | "不可能不 VP"强于"不应该不 VP" | | "一定"强于"应该" | |
|---|---|---|---|---|
| | 人数 | 所占比例 | 人数 | 所占比例 |
| 4-5 岁组 | 0 | 0% | 4 | 26.67% |
| 5-6 岁组 | 3 | 20% | 11 | 73.33% |
| 6-7 岁组 | 3 | 20% | 11 | 73.33% |
| 7-8 岁组 | 8 | 53.33% | 14 | 93.33% |
| 成人组 | 14 | 93.33% | 15 | 100% |

　　绝大部分儿童在 5-6 岁已经能意识到"一定"和"应该"间在语义强度上的差别，7-8 岁儿童的判断与成人已不存在显著性差异（$z=1.02$, $p>.05$）。与此相对照，5-6 岁和 6-7 岁的儿童中只有个别能区别"不可能不 VP"和"不应该不 VP"，而 7-8 岁的儿童中也只有 53.33%的能区别这对双重否定式间的强度差异，他们的判断与成人的仍存在显著性差异（$z=2.48$, $p<.05$）。

## 4．2．2　道义性的情态层级

| 组别 | "不可以不 VP"强于"不应该不 VP" | | "必须"强于"应该" | |
|---|---|---|---|---|
| | 人数 | 所占比例 | 人数 | 所占比例 |
| 4-5 岁组 | 0 | 0% | 3 | 20% |
| 5-6 岁组 | 7 | 46.67% | 13 | 86.67% |
| 6-7 岁组 | 5 | 33.33% | 12 | 80% |
| 7-8 岁组 | 8 | 53.33% | 15 | 100% |
| 成人组 | 13 | 86.67% | 15 | 100% |

　　在处理"必须"和"应该"时，5-6 岁和 6-7 岁儿童的判断与成人的已不存在显著性差异（$z=p>.05$），7-8 岁儿童的判断与成人的一致。但是，双重否定给分辨"不可以不 VP"和"不应该不 VP"带来的困难让 7-8 岁的儿童仍不能都克服掉，他们的判断与成人还存在显著性差异（$z=1.996$, $p<.05$）。

　　除此之外，双重否定给分析带来的困难也反映在成人组。根据逻辑上的分析，"不可以$_2$不VP"等于"必须VP"，"不应该$_2$不VP"等于"应该$_2$VP"。成人 100%（15/15）判定"必须"强于"应该"，而在判断相对应的双重否定结构时比例却是 86.67%(13/15)。13.33%(2/15)的成人被试是确实认为"不应该$_2$不VP"要强于"不可以$_2$不VP"呢？还是他们也误用了计算双重否定的策略？

## 4.2.3 讨论

将肯定式和双重否定式的结果进行对比：
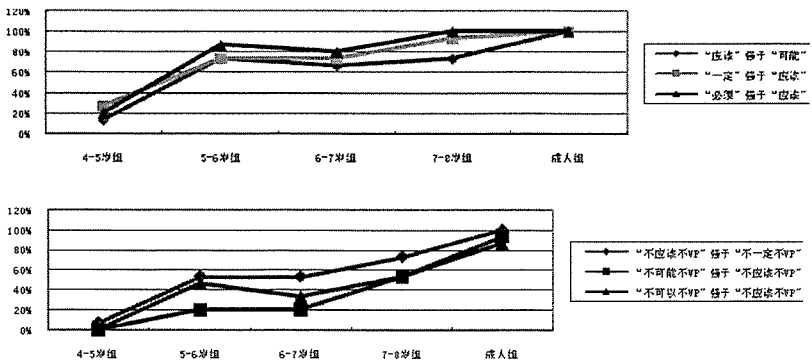




可以清晰地看出，在 5-6 和 6-7 岁，儿童对肯定式情态词间语义差别的了解已经能达到一个比较好的水平；到了 7-8 岁，儿童对相对应的双重否定式间语义差别的判断与成人的仍存在差距。导致这种情况的症结在于：一方面是双重否定使命题计算更加复杂；另一方面双重否定使相关的情态算子在语义强度上出现了三种不同的情况，加强、减弱和不变。

在测试中，还发现在处理双重否定式时，儿童会错误地使用两种计算策略：

第一，省略策略： 不 Mod[4] ~~不~~ VP

将第二个否定算子省略掉，因为儿童没有完全具备计算这样一个复合命题的能力。

如有些被试对测试故事 9 中问题的回答：

(8) 5-6 岁组

    FJH：小猴。"不可能不在后面"就是"不在"。

    ZYH：小猴。小猴"不可能不在后面"是它不在后面。小兔子"不应该不在后面"就是它也不应该在后面。

第二，相消策略： ~~不~~ Mod ~~不~~ VP

在另一些儿童对测试故事 9 中问题的回答中，发现他们在理解双重否定式时，只是单纯地将两个否定词"不"一并消去，没有意识到双重否定给情态词原本的意义所带来的变化，如：

(9) 7-8 岁组

    WZY：小兔子。"不可能不在后面"就是不太确定；"不应该不在后面"就是绝对这样。

相消策略要优于省略策略，应该是在省略策略基础上发展的结果，因为它还是可以处理像"不应该不 VP"这样的结构。但是，相消策略并不能用来处理所有情态词的双重否定结构。只有当儿童认识到双重否定给情态词语义带来的不完全是

---

[4] Mod表示情态词。

"负负得正"的事实后，他们才必然会摒弃相消策略。

在测试中，我们也发现不少 7-8 岁组的儿童不仅理解双重否定式间语义强度的差异，而且能用语言非常好地分析这种差异。如：
部分儿童是这样来回答测试故事 9 的问题：
(10) 7-8 岁组
　　ZSX：小猴。天线宝宝说小猴不可能不在后面。"不可能不在后面"就是"一定在后面"。
　　WQF：小猴。它说小猴不可能不在后面，就是小猴在后面；"不应该不在后面"就是可能在后面。
　　YZZ：小猴。小猴不可能不在后面。"不可能不"就是肯定在后面；"不应该不"就是可能在后面，但不肯定。

更多关于情态词的否定式和双重否定式的儿童习得情况，可参见范莉（2005）。

## 5．0　总结

长期跟踪的语料表明，普通话儿童情态动词的词汇发展开始在 01;04 左右，到 02;05 之前使用的词汇中仍缺少成人常用的一些情态词；相比认知性的情态词，根源性的在语料中始现的时间早，而且在早期使用的频率高。认知性和根源性的差异不会影响儿童对情态层级的习得。从 4-5 岁组到 7-8 岁组的测试结果没有反映出儿童掌握道义性情态层级的情况要比掌握认知性情态层级的情况好。

对肯定式间的情态差异进行判断时，在 4-5 岁阶段，已经有少数儿童开始意识到情态词在表达认知逻辑上和道义逻辑上的可能性的差异；到了 5-6 岁和 6-7 岁阶段，这方面的知识飞速发展；7-8 岁儿童相关知识已经比较成熟。

对情态词双重否定式知识的了解，4-5 岁是个起步的阶段。到了 6-7 岁和 7-8 岁，绝大多数儿童才掌握了相关的知识。很多儿童早期在处理双重否定式时误用了省略策略和相消策略，即使在 7-8 岁的儿童中也发现使用省略策略的现象。双重否定给儿童在判断情态层级时带来了困难，从而导致很多儿童到了 7-8 岁与成人的判断仍有明显差异。

# 参考文献

Bassano, D. (1996). Functional and formal constraints on the emergence of epistemic modality: A longitudinal study on French. *First Language* 16: 77-113.

Choi, S. (1995). The development of epistemic sentence-ending modal forms and functions in Korean children. In J. Bybee and S. Fleischman (eds.), *Modality in Grammar and Discourse.* Amsterdam: Benjamins, 165-204.

Coates, J. (1983). *The Semantics of the Modal Auxiliaries.* London: Croon Helm.

Gee, J. and I. Savasir. (1985). On the Use of Will and Gonna: Toward a Description of Activity-Types for Child Language. *Discourse Processes* 8: 143-175.

Guo, Jiansheng. (1994). *Social Interaction, Meaning, and Grammatical Form: Children's Development and Use of Modal Auxiliaries in Mandarin Chinese.* Ph. D. dissertation. University of California at Berkeley.

----- (1995). The interactional basis of the Mandarin modal *néng* 'can'. In J. Bybee and S. Fleischman (eds.), *Modality in Grammar and Discourse.* Amsterdam: Benjamins, 205-238.

Hirst, W. and J. Weil. (1982). Acquisition of epistemic and deontic meaning of modals. *Journal of Child Language* 9:659-666.

Kuczaj, S. (1977). Old and new forms, old and new meanings: The form-function hypothesis revisited. Paper presented at the Meeting of the Society for Research on Child Development, New Orleans.

Papafragou, A. (1997). Modality in language development: a reconsideration of the evidence. *UCL Working Papers in Linguistics* 9.

----- (2000). *Modality: Issues in the Semantics-Pragmatics Interface.* Oxford: Elsevier Science Ltd.

----- (2001). Linking early linguistic and conceptual capacities: The role of theory of mind. In A. Cienki, B. Luka and M. Smith (eds.), *Conceptual and Discourse Factors in Linguistic Structure.* Stanford: CSLI Publications, 169-184.

Pea, R., R. Mawby and S. MacCain. (1982). World-making and world-revealing: semantics and pragmatics of modal auxiliary verbs during the third year of life. Paper presented at the 7th Annual Boston Conference on Child Language Development, October.

Perkins, M. (1983). *Modal Expressions in English.* London: Frances Pinter.

Shatz, M. and S. Wilcox. (1991). Constraints on the acquisition of English modals. In S. Gelman and J. Bymes (eds.), *Perspectives on Language and Thought.* London: Cambridge University Press, 319-353.

Shepherd, S. (1981). *Modals in Antiguan Creole, child language acquisition, and history.* PhD dissertation. Stanford University.

----- (1982). From deontic to Epistemic: An analysis of modals in the history of English, creoles, and language acquisition. In A. Ahlqvist (ed.), *Papers from the 5th International Conference on Historical Linguistics.* Amsterdam: Benjamins, 316-723.

Stephany, U. (1993). Modality in first language acquisition: The state of the art. In N. Dittmar and A. Reich (eds.), *Modality in Language Acquisition*. Berlin: de Gruyter, 133-144.

Wells, G. (1979). Learning and using the auxiliary verb in English. In V. Lee (ed.), *Cognitive Development: Language and Thinking from Birth to Adolescence*. London: Croom Helm, 250-270.

----- (1985). *Language Development in the Pre-School Years*. Cambridge: Cambridge University Press.

丁声树等，1961，《现代汉语语法讲话》，北京：商务印书馆。

范莉，2005，儿童和成人语法中的否定和否定辖域，未发表，博士论文，北京语言大学。

----- 2006，普通话中情态动词的儿童早期习得，待发表。

方立、范莉，2005，"应该"及其否定句式，《现代外语》，第 4 期， 331-340。

吕叔湘，1982（1942），《中国文法要略》，北京：商务印书馆。

马建忠，1983（1898-1900），《马氏文通》，北京：商务印书馆。

王力，1954，《中国语法理论》，北京：中华书局。

附录一　　关于情态词层级的语言测试

**测试者信息**

姓名：　　　　　　　　　年龄：　　　　　　　　　性别：

学校：　　　　　　　　　测试时间：　　　　　　　测试地点：

**预备项目**

小兔子：　哎哟，哎哟，打不开。

测试问题：小兔子能打开书吗？

**测试项目**

1 小青蛙：我家有好多玩具，你们去我家玩吧。



小熊：我一定去。

小狗：我应该去。

测试问题：如果只有一个小动物去了小青蛙家，你猜它是谁？

2 小猴：你们会踢球吗？



小兔子：我想我应该会。

小熊：我想我可能会。

测试问题：如果只有一个小动物会踢球，你猜它是谁？

3（画外音：小青蛙要选一个动物和他去比赛。）

（指着小熊。）



小青蛙：小熊，你应该去。

（指着小兔子。）

小青蛙：小兔子，你必须去。

测试问题：如果小青蛙只选了一个小动物，你猜它是谁？

4 青蛙老师：小兔子生病了，她在医院。你们谁去看她？

（小熊和小狗齐声说："我去，我去。"）

青蛙老师：小熊，你应该去。

（停顿。）



青蛙老师：小狗，你可以去。

测试问题：如果青蛙老师只选了一个小动物去看小兔子，你猜它是谁？

5 小猴：我们玩玩具吧。

小熊：那谁去拿玩具呢？

小猴：我选一个小动物去拿玩具。

（指着小兔子。）

小猴：小兔子，你可以去。

（停顿。）



（指着小熊。）

小猴：小熊，你必须去。

测试问题：如果小猴子只选一个小动物去拿玩具，你猜它是谁？
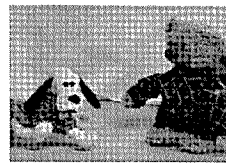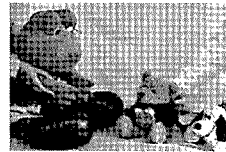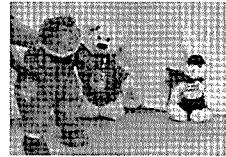
6（画外音：谁去公园？）

小狗：我一定去。

小熊：我可能去。

测试问题：如果只有一个小动物去公园，你猜它是谁？

7 天线宝宝：啊，小熊和小兔子今天都穿了裙子。是什么颜色的呢？



天线宝宝：是红色的吗？我最喜欢红色了。

天线宝宝：让我来猜一猜。

天线宝宝：小熊的裙子不一定不是红色的。

（停顿。）

天线宝宝：小兔子的裙子不应该不是红色的。

测试问题：如果只有一个小动物的裙子是红色的，你猜它是谁？

8 小兔子：哎哟哎哟，推不动。

天线宝宝：我选一个小动物帮助小兔子。



（对着小猴。）

天线宝宝：小猴，你不可以不帮小兔子。

（对着小熊。）

天线宝宝：小熊，你不应该不帮小兔子。

测试问题：如果天线宝宝只选一个小动物帮助小兔子，你猜它是谁？



9 天线宝宝：后面有小动物。是谁呢？让我猜一猜。

天线宝宝：小猴子不可能不在后面。

天线宝宝：小兔子不应该不在后面。

测试问题：如果只有一个小动物在后面，你猜它是谁？
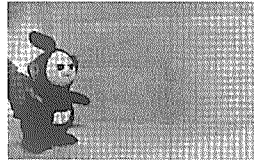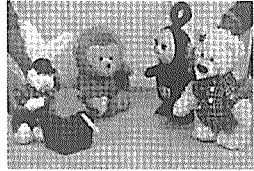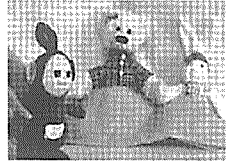
# Null vs. Overt Subjects and Pre- vs. Post-verbal Modals in English-Cantonese Interlanguage: Testing the Full Transfer/Full Access Hypothesis

John WAKEFIELD

The Hong Kong Polytechnic University

## 1. Introduction

Linguists who study second language acquisition (SLA) within a Universal Grammar (UG) framework focus largely on first language (L1) transfer and/or UG access (see White, 2003 for background and details). One of the more prominent theories addressing both L1 transfer and UG access is Schwartz and Sprouse's (1994; 1996) Full Transfer/Full Access Hypothesis (FT/FA), which argues that 'all the principles and parameter values as instantiated in the L1 grammar immediately carry over as the initial state of a new grammatical system on first exposure to input from the target language' (Schwartz and Sprouse, 1996: 41). In addition to complete transfer of the L1's parameter values as the initial state, FT/FA also argues that full access to UG restructures the interlangauge, causing its UG parameter settings to be reset from the settings of the L1 to the settings of the L2. This is conditional, however, on the learner being exposed to a sufficient quantity of positive and unambiguous L2 input data.

The study reported in this paper was a test of the full transfer component of the FT/FA hypothesis. It was a follow-up study of Wakefield (2006), which looked at the acquisition of copula-drop in addition to all of the grammatical features examined in the present study: null subjects; overt subjects; the preverbal modal *ho2ji5*; and the post-verbal modal *dak1*. The informants in both studies were English-speaking learners of Cantonese (ELCs). The results of Wakefield (2006) were consistent with the predictions of FT/FA regarding full access to UG for all the grammatical structures tested, because the scores of the advanced-learner informants on the language test were virtually the same as those of the native-speaker controls for every grammatical structure tested. However, except for the copula-drop sentences, the results were not entirely consistent with FT/FA's predictions regarding full transfer of the L1. Based on speculations about the reasons for this, the language task was modified for the present study to see if, in addition to support for full UG access, support for full L1 transfer could also be obtained for each of the structures examined (see §4.2 for details). The results were that both the full transfer and full access components of FT/FA did receive research support in all the structures tested. However, the evidence of full transfer regarding null vs. overt subjects was not as strong as expected; the results appear to show that ELCs acquire null subjects at a surprisingly early stage.

## 2. Previous Tests of FT/FA

An investigation of FT/FA must look for one, or both, of the following: 1) evidence that the initial-state parameter settings are either the same as, or different from, the parameter settings of the L1, thus providing evidence either for or against full L1 transfer, respectively; and 2) evidence of either the success or failure of

parameter resetting, thus providing evidence for or against full UG access, respectively. Among the studies that have claimed to find evidence in support of FT/FA, some found evidence supporting full L1 transfer (e.g., Slabakova, 2000; Grüter and Conradie, 2006). Others found evidence supporting full UG access (e.g., White et al, 2004; Slabakova, 2005). And several found evidence of both full transfer *and* full access (e.g., Craats, Corver and Hout, 2000; Montrul, 2001; Marsden, 2003; Augusto, 2003).

In contrast to the above findings, other studies have claimed to find evidence against FT/FA. Several found no evidence of L1 transfer (e.g., Ionin and Wexler, 2001; Yuan, 2003; Sopata, 2005; Cabrera and Zubizarreta, 2005). I am aware of only one study that claimed to find positive evidence of *no* L1 transfer, and therefore to have falsified FT/FA (Håkansson, Pienemann and Sayehli, 2002). And I am only aware of one FT/FA-related study that claimed to find evidence against parameter resetting (Kong, 2005). No study I know about found counter evidence to *both* the full transfer *and* full access components of FT/FA.

The only way to justify further FT/FA research is to assume that none of the above studies has unequivocally falsified FT/FA. There is good reason to take such a stance. Ionin and Wexler (2001), for example, claimed to find no evidence of L1 transfer, but for some reason they did not discuss the performance of the beginner-level informants in their study. Similar to Ionin and Wexler, Montrul (2001) also found no evidence of L1 transfer in the interlanguage of her intermediate informants regarding one of the grammatical structures she examined. Unlike Ionin and Wexler, however, Montrul did not conclude that L1 transfer did not occur. She said that she did not know how her intermediate-level informants would have performed had they been tested at an earlier stage; it is possible that L1 transfer existed in the early stages of their interlanguage, and that it disappeared as they progressed beyond that stage. In order for a study to test the full transfer portion of FT/FA, beginner-level informants must be included. Since these two studies did not discuss the performance of beginner-level informants, it is reasonable to assume that they did not falsify FT/FA, though they did raise enough doubt in FT/FA to justify more research.

The strongest claim of falsifying FT/FA came from Håkansson, Pienemann and Sayehli (2002), whose informants had not successfully acquired the verb second (V2) word order in German, even though V2 is also a feature of their L1 Swedish. A potential reason for this was that German was the informants' L3; they had all been learning L2 English for a year longer than they had been learning L3 German. This could therefore have been a case of L2 transfer, and the only way to determine this for sure would be to replicate the study using Swedish-speaking informants who have no L2 English.

Regarding access to UG, Kong (2005) concluded that his informants, who were Mandarin-speaking learners of English, did not have full UG access because they did not successfully acquire obligatory overt subjects. It could be argued, however, that his informants had not yet reset the parameters relevant to overt subjects because their entire language-learning experience was classroom based. It was also not clear as to whether or not the informants language instructors were native-English speakers. Perhaps the informants had not yet been exposed to a sufficient quantity of positive and unambiguous L2 input data. I stated above that beginner-level informants must be

used in research studies that test for full L1 transfer, and by the same logic, it is reasonable to propose that research studies that test for full UG access should include advanced-level informants who have been acquiring their L2 in an L2 environment.

In short, I assume that FT/FA has not been unequivocally proven false, and that the overall results of FT/FA-related research are inconclusive. Further research is therefore needed. White (2003: 17) said that linguistic competence can only be tested indirectly, but '[w]hen results from different tasks and different groups of learners show the same trends, this suggests that we are indeed gaining insight (indirectly) into the nature of the underlying linguistic competence.' In this light the present study is justified by the fact that it contributes to the collection of FT/FA-related findings using different types of grammatical constructions within a different type of interlangauge.

## 3. Linguistic Background and Predictions

To test FT/FA, the present study looked at two pairs of syntactic structures, each of which was comprised of one structure that was L1-like and one structure that was not L1-like. On the surface, the two members of each pair of structures were virtually identical in terms of their number and type of overt constituents. One pair compared the acquisition of a subject-verb (SV) construction to that of a topic-(null subject)-verb (T(S)V) construction. The other pair compared the acquisition of a preverbal modal (*ho2ji5*) to that of a post-verbal modal (*dak1*).

### 3.1: SV vs. T(S)V

In the discussion that follows, some of the citations refer to Mandarin, but the relevant facts are the same for both Mandarin and Cantonese. Matthews and Yip (1994: 73) said that moving the object of the verb to the front of a Cantonese sentence is the 'most common and straightforward form of topicalisation... [and it is a form that] is also widely used without a subject, where the implicit subject is unknown or generic.' Sentence (1) is an example of this:

(1)     Go2-gaa3   ce1  zing2-zo2.
        that-CL    car  repair-PERF
        'That car, (someone) repaired it.'

Sentence (1) is an NP-VP construction, i.e., the NP *go2-gaa3 ce1* is followed by the VP *zing2-zo2*. It is a topic-comment construction derived through left-dislocation of the object to the topic position. Li and Thompson (1981: 160) said that '[w]hen the direct object is the topic, the subject may be unexpressed if it is unimportant, unknown, or understood.' To show that the NP in (1) is the object, and that the sentence has an unexpressed subject, the NP can be placed back into object position and a subject can be added before the verb:

(2)     Keoi5 zing2-zo2    go2-gaa3 ce1.
        S/he  repair-PERF  that-CL  car
        'S/he repaired that car.'

This can then be transformed into a topic-comment construction by topicalizing the object through movement:

(3)　　[Go2-gaa3 ce1]ᵢ, keoi5 zing2-zo2 tᵢ
　　　　　that-CL car　s/he　repair-PERF
　　　　　'That car (topic), s/he repaired.'

The subject can then be made null, resulting in the equivalent of (1):

(4)　　[Go2-gaa3 ce1]ᵢ Ø zing2-zo2　tᵢ
　　　　　that-CL car　　repair-PERF
　　　　　'That car, (someone) repaired it.'

Under this analysis, (1) and (4) are construed in this way:

(5)　　[$_{Topic}$ Go2-gaa3 ce1ᵢ [$_{Subject}$ Ø [$_{VP}$ zing2-zo2　tᵢ ]]]
　　　　　　　　that-CL　car　　　　　　　　repair-PERF
　　　　　'That car, (someone) repaired it.'

The above demonstrates that sentence (1) is a derived topic-comment construction with a null subject. Li and Thompson (1981: 89) said that sentences like (1) are 'topic-comment constructions in which the subject of the verb is not present,' which is also the conclusion of other linguists (e.g., Yip and Rimmington, 2004; Li, 2005).

Yip (1995: 89) said that 'English topicalization structures… are essentially restricted to the colloquial left-dislocation of subjects and objects.' The Chinese word order in sentence (3) is just such a structure and could actually be used in English, especially in a context of contrast:

(6)　　This car *(s/he) repaired, that car *(s/he) didn't.

Topicalized sentences like (6) are only acceptable in English, however, if the subjects are overt. The difference between English and Cantonese regarding this structure is therefore *not* that this form of topicalization is grammatical in Cantonese but not in English, but rather that subject nullification is grammatical (and common) in Cantonese but not in English.[1] It is worth noting, however, that although this structure is grammatical in both languages when it has an overt subject, it is not commonly used in English.

The Cantonese construction shown in (1) only works with transitive verbs, of course, because the derived topic is actually the object, and without a transitive verb there can be no object. Cantonese therefore has two types of NP-VP sentences that share the same construction on the surface, but have very different underlying structures. When the verb is intransitive, the NP is not a topicalized object; it is a subject:

---

[1] It is of course well known that English frequently drops the subject in second-person imperatives, i.e., 'Go home', 'Sit down', etc. However, the English counterparts to the structures being discussed here cannot have null subjects.

(7)     Go2-bun2 syu1   dit3-zo2
        that-CL  book    fall-PERF
        'That book fell.'

Here, the verb "fall" indicates a change of location and does not require an agent. An inanimate entity can function as the grammatical subject in such a case. Even though sentences (1) and (7) are both NP-VP constructions, their underlying structures differ: (1) is a topic-verb construction with a null subject (T(S)V), while (7) is a subject-verb (SV) construction.

This study looked at ELCs' acceptance of sentences such as (1), a T(S)V construction, and compared that to their acceptance of sentences such as (7), an SV construction. If the arguments of FT/FA are correct, then sentences such as (7) should be easily acceptable due to L1 transfer, while sentences such as (1) should be unacceptable at the initial state, but gradually become acceptable later on as the interlanguage parameter affecting null subjects is reset. The parameter involved, according to Huang (1982/1998), is the feature [-Agr]. Huang argued that a language licenses null subjects that are *pro* if 1) it has both the features [+Agr] and [+strong Infl] (e.g., Italian), or 2) it has the feature [-Agr] (e.g., Chinese). English does not allow *pro-drop* because it has the features [+Agr] and [-strong Infl]. According to this analysis, if ELCs reject null subjects in Cantonese at first, but then acquire them later on, this would indicate that their interlanguage resets the L1 parameter setting of [+Agr] to the L2 parameter setting of [-Agr].

### 3.2: *ho2ji5* vs. *dak1*

In Cantonese there are two modals, *ho2ji5* and *dak1*, that have similar meanings to each other (i.e., 'can/able to' or 'may/permitted to') but have different feature-related properties, resulting in different word-order constructions; *ho2ji5* occurs pre-verbally and *dak1* occurs post-verbally. All Cantonese linguists consider *ho2ji5* to be a modal auxiliary. Wong (1998) and Kwok (1971, cited in Luke and Nancarrow, 1998) said that *dak1* is also a modal auxiliary, but not all linguists agree. According to Luke and Nancarrow (1998: 92), for example, *dak1* 'should not be described as an auxiliary' because it follows verbs and does not precede VPs.

Regardless of whether or not *dak1* is an auxiliary, however, it is unquestionably a modal that shares certain properties with *ho2ji5*. Matthews and Yip (1994) said that *ho2ji5* is a modal auxiliary that 'means "can" or "may" in the sense of something being possible or permissible' (p. 231), and they referred to *dak1* as a verb that can express modality indicating 'potential, including both possibility and permission' (p. 242). Cheng and Sybesma (2004) referred to *dak1* as a modal element that denotes permission or potential, and they included two sentences in their study that were only different in the respect that one used *dak1* and the other used *ho2ji5*. They said the two sentences were 'very close in meaning' (p. 421). There is no complete agreement about the exact categorization of *dak1*, but there is agreement that 1) both *ho2ji5* and *dak1* are able to express the modal notions of possibility 'can' or permission 'may'; 2) *dak1* is used post-verbally; and 3) *ho2ji5* is used pre-verbally.

The language task used in this study only included sentences whose meanings could be expressed similarly using either *ho2ji5* or *dak1*. Some sentences expressed permission:

(8)   Lei5   ji4gaa1 ho2ji5   co5.
      you    now    can    sit
      'You can sit now.'

(9)   Lei5   ji4gaa1 co5   dak1.
      you    now    sit    can
      'You can sit now.'

Others expressed ability:

(10)   Li1-zek3   maa5 ji4gaa1 ho2ji5 paau2.
       this-CL    horse now    can    run
       'This horse can run now.'

(11)   Li1-zek3 maa5   ji4gaa1 paau2 dak1.
       this-CL    horse   now    run    can
       'This horse can run now.'

Each sentence included an ungrammatical counterpart. The ungrammatical counterparts of (8) and (9), for example, are (12) and (13):

(12)   *Lei5 ji4gaa1 co5 ho2ji5
        you   now   sit   can

(13)   *Lei5 ji4gaa1 dak1 co5
        you   now   can   sit

The adverbial *ji4gaa1* 'now' was included in all of the *ho2ji5* and *dak1* sentences to provide a temporal contrast.

Since the word order for *ho2ji5* is the same as it is for English modals, it should be readily accepted from the beginning if the L1 fully transfers. The word order used with *dak1*, however, would require parameter resetting and would only be acquired successfully with UG access after sufficient exposure to the right kinds of L2 input data. Assuming the analysis of Cheng and Sybesma (2004), the functional features affecting the word order of *dak1* involve two separate modal positions (Mod1 and Mod2), one preverbal and one post-verbal. The *dak1* that expresses permission occupies the preverbal Mod1 slot, but unlike all other modals in Cantonese, it is unable to check for features of [AGR] with regard to $I^0$, which is empty. Therefore the verb must raise, forming a complex modal head in the Mod1 position that is able to check for features of [AGR] with $I^0$. After the verb raises, it is attached to the front of *dak1* in the Mod1 position, making *dak1* post-verbal. The *dak1* that expresses 'potential' (similar to ability) is aspectual in nature, and therefore occupies the post-verbal Mod2 position and takes an AspP as its complement. It needs to check features against $V^0$, but since it is adjacent to $V^0$, it can remain in its post-verbal position. In either case, therefore, *dak1* ends up in the post-verbal position.

## 4. The study

### 4.1 The informants

The informants were 41 ELCs, aged 19 to 26. All of them were missionaries that had received 3 months of full-time (i.e., 9 hours/day), intensive Cantonese training before arriving in Hong Kong. In Hong Kong they spoke Cantonese with locals on a daily basis. They were divided into two groups based on the length of time they had been acquiring Cantonese. Group one was comprised of 23 informants who had been learning Cantonese from 3 to 7 months, and group two was comprised of 18 informants who had been learning Cantonese from 8 to 15 months. In addition to the ELC informants, 10 native-speaker controls were also tested.

### 4.2 The experiment

A grammaticality judgment task was used. In creating the judgment task, I took into account the five things that Cowan and Hatasa (1994) concluded are necessary for making grammaticality judgments reliable and valid: 1) use common language that is not too context specific; 2) test for at least three samples of each sentence type; 3) use short sentences because long, grammatically complex sentences do not get consistent judgments even when tested on native speakers; 4) use sentences without any potentially ambiguous stimuli; and 5) use a native-speaker control group.

The results from the native-speaker controls were used to select the final list of sentences for the language task. The only sentences used were those that all the native speakers agreed were either definitely acceptable or definitely *not* acceptable. As a result, 2 of the 34 non-spoiler sentences that were tested on the native speakers were not included on the language task for the ELCs. The resulting language task included 6 tokens each of T(S)V and SV sentences (i.e., sentences like (1) and (7), respectively), as well as 10 tokens each of *ho2ji5* and *dak1* sentences, half pre-verbal and half post-verbal (i.e., sentences such as (8), (9), (12), and (13)). In addition, there were 12 spoiler sentences, making a total of 44 sentences in all. For the ELC informants, the sentences were written using the Yale system of Cantonese Romanization because that was what they were familiar with. For the native-speaker controls, the sentences were written in traditional Chinese characters.

The sentences in this study were simplified modifications of those used in Wakefield (2006). Based on the results of that study, as well as on follow up interviews, it was concluded that the L1-like sentences contained non-L1-like features that interfered with the informants' judgments. For example, the SV and T(S)V sentences began with bare classifiers (i.e., *Bun2 syu1 dit3 zo2*). To rectify this, demonstratives (such as *go2*) were added, making the sentences like (7).

During the language task, a timed audio recording was played for the informants to follow. This reduced the likelihood that they could rely on their metalinguistic knowledge. On the audio each sentence was spoken aloud twice by a native speaker, with a two second pause in between. After that there was a twenty second pause, during which the informants judged and translated the sentence. They were instructed to work as quickly as they could and to wait until the recording moved on to the next sentence before continuing. To judge acceptability, the informants marked one of four

statements with a tick. The scoring system was similar to the one used in Hawkins and Chan (1997). For acceptable sentences, the informants were given scores as follows according to which statement they chose:

1. Definitely acceptable                       3 points
2. Not sure, but it seems acceptable       2 points
3. Not sure, but it does not seem acceptable   1 point
4. Definitely not acceptable              0 points

For unacceptable sentences, they were given scores as follows:

1. Definitely acceptable                       0 points
2. Not sure, but it seems acceptable       1 point
3. Not sure, but it does not seem acceptable   2 points
4. Definitely not acceptable              3 points

If they judged a sentence to be either definitely or seemingly acceptable, they translated it into English. If they judged a sentence to be either seemingly or definitely not acceptable, they wrote: "I think this sentence is unacceptable." This prevented informants from unintentionally noticing which informants were translating and which were not. They were therefore unable to know whether or not their fellow informants had judged a given sentence as acceptable or unacceptable.

## 5. The Null Hypothesis

Because the findings of Wakefield (2006) supported the full access portion of FT/FA, the null hypothesis of this follow up study, as stated in (14), (15), and (16), addresses only the full transfer portion of FT/FA:

(14)      For all comparable structure pairs (i.e., *ho2ji5* vs. *dak1* and SV vs. T(S)V), group one will not perform better on the L1-like structure than it will on the non-L1-like structure.

(15)      Groups one and two will not belong to the same statistical population with
         regard to their performance on each of the L1-like structures.

(16)      Groups one and two will belong to the same statistical population with regard to their performance on each of the non-L1-like structures.

If group one performs better on each L1-like structure than it does on that structure's non-L1-like counterpart, then (14) can be rejected (i.e., if group one's performance on each type of *ho2ji5* sentence is better than its performance on each type of *dak1* sentence, and if its performance on the SV sentences is better than its performance on the T(S)V sentences). It will be made clear below as to how parts (15) and (16) of the null hypothesis can be rejected.
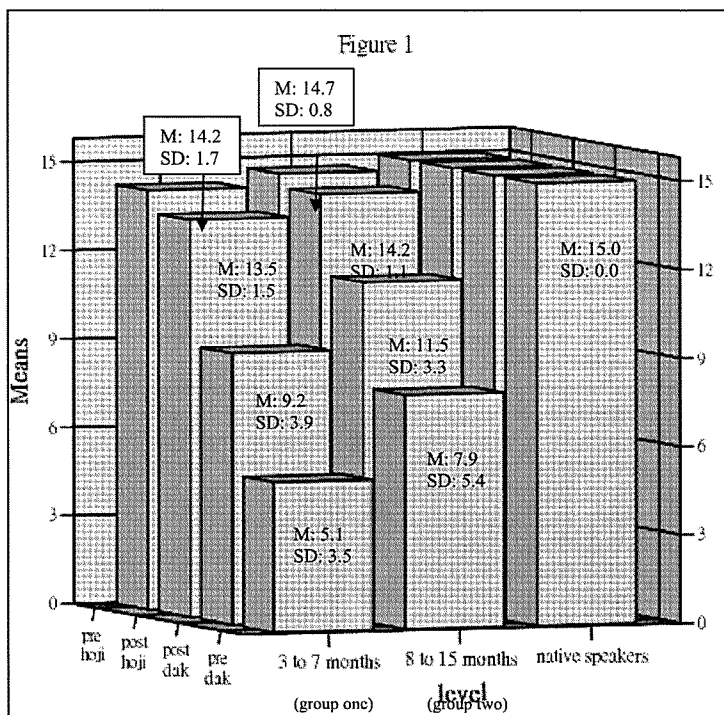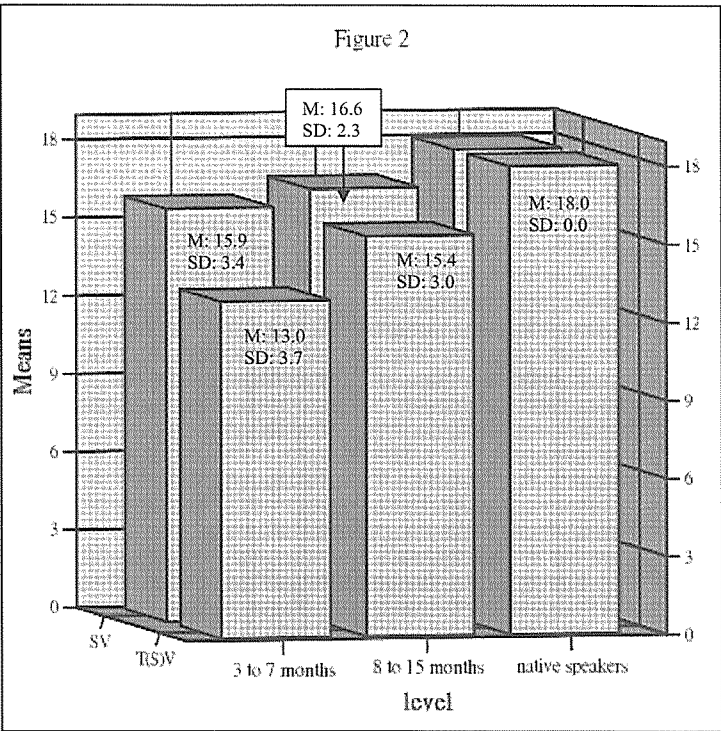
Figure 1

**6. The Results and Discussion**

The graph in figure 1 shows the means and standard deviations for each of the four types of modal sentences for group one (3 to 7 months), group two (8 to 15 months), and the native-speaker controls. There were 5 tokens of each type of sentence, each having a maximum score of 3. Therefore, the maximum possible score for each of the four types of modal sentences was 15. The graph in figure 2 shows the means and standard deviations of all three groups for both the SV and T(S)V sentences. There were 6 tokens of each type of sentence, each having a maximum score of 3. Therefore, the maximum possible score for each of the two types of sentences was 18. The native-speaker controls obtained the maximum possible score for every sentence on the language task.

Figure 1 shows that group one accepted a much higher number of preverbal modal sentences than post-verbal modal sentences. This translated into significantly better mean scores for the sentences using L1-like *ho2ji5* than for the sentences using non-L1-like *dak1*. More specifically, the graph in figure 1 shows that the mean score for either type of *ho2ji5* sentence is higher than the mean score for either type of *dak1* sentence.



Figure 2

Correspondingly, figure 2 shows that group one also performed better on the L1-like SV sentences than on the non-L1-like T(S)V sentences (though not to the degree that was expected). Based on these combined results, (14) can be rejected.

I will now consider (15) and (16). A One-Way ANOVA was used to compare the means of the two groups for each of the six types of sentences included on the language task. In this way, the two variables 'group one' and 'group two' were treated as two categories of a categorical variable in the same way that, for example, male and female would be treated as categories in order to test whether, on average, belonging to one or the other category affects performance on a given task. Each of the six One-Way ANOVAs used one of the six sentence types as a dependent variable. The purpose of this was to see whether or not groups one and two belonged to the same statistical population with regard to each sentence type. It was assumed that if L1 transfer existed in the ELC informants' interlanguage, this would help group one's performance on L1-like structures and hinder its performance on non-L1-like structures. The result would be that group one would tend to belong to the same statistical population that group two does with regard to L1-like structures, but not with regard to non-L1-like structures. This is what one would predict, for example, for the categories 'male' and 'female' with regard to a comparison of means for 'IQ' as

opposed to a comparison of means for 'height', where 'IQ' for our purposes here is analogous to L1-like and 'height' is analogous to non-L1-like.

Looking first at the non-L1-like sentences, there was a statistically significant difference between groups one and two for preverbal *dak1* ($F = 4.1$; $p = 0.05$; $df = 1$), for post-verbal *dak1* ($F = 4$; $p = 0.05$; $df = 1$), and for the T(S)V sentences ($F = 5.1$; $p < 0.05$; $df = 1$). Now looking at the L1-like sentences, there was not a statistically significant difference between the two groups for post-verbal *ho2ji5* ($F = 3$; $p = 0.09$; $df = 1$), for preverbal *ho2ji5* ($F = 1.1$; $p = 0.3$; $df = 1$) or for the SV sentences ($F = 0.5$; $p = 0.47$; $df = 1$). This is not surprising if we assume full transfer. With regard to interlanguage features that are properties of both the L1 and the L2, there should be a tendency for early L2 learners who are helped by transfer to perform in a way that is *not* significantly differently from more advanced learners. At the same time, there should be a significant difference between early and more advanced L2 learners regarding structures that are different from those in the L1. Based on the results, both (15) and (16) of the null hypothesis can be rejected.

The study reported in this paper provided research support for FT/FA, but the results were not exactly as expected because the informants in group one performed better than expected on the T(S)V sentences. One possibility is that the beginner-level informants judged the T(S)V sentences to be like English passives with a dropped auxiliary, or to be a subject-predicate construction with a dropped copula. Many of the informants translated these sentences as, for example, 'The clothes were washed' rather than 'Someone washed the clothes.' In this case, the contrast between the two sentences 'The book fell' and 'The car (was) fixed' would not be one of topicalization and *pro-drop*, but rather one of copula- or auxiliary-drop. Even if this was the case, however, it is still surprising that the informants were able to acquire the T(S)V constructions so early on, and that there was not a bigger difference between their performance on the T(S)V constructions and the SV constructions. This is especially true in light of the fact that the results of Wakefield (2006) indicated that the majority of ELCs take more than 2 years to acquire copula-drop in sentences such as *Go2-zi1 bat1 (hai6) ngo5-ge3* 'That pen *(is) mine.'

If one were to adopt an empiricist (i.e., non-UG) perspective, the most likely prediction would be that ELCs could process, and would therefore accept, a copula-drop sentence whose word-for-word gloss is (17) just as readily as (or more readily than) a T(S)V sentence whose word-for-word gloss is (18).

(17) that pen mine

(18) that car fixed

However, since there appears to be such a large difference in acquisition time between these two structures (only a few months for T(S)V compared to a few years for copula-drop), the implication is that something is influencing the ELCs' language acquisition process instead of (or in addition to) learning mechanisms such as analogy, inductive generalization, puzzle solving, etc. If FT/FA is correct, the results can be explained in terms of parameter resetting. Specifically, if the feature [+AGR] changes to [-AGR] in ELCs' interlanguage sooner than the feature [+TNS] changes to [-TNS], this could explain why ELCs acquire *pro-drop* before copula-drop or auxiliary drop.

Assuming Huang's (1982/1998) analysis along with FT/FA, resetting the [+AGR] parameter to [-AGR] should cause language learners to acquire *pro-drop*. And since topicalization of the object is acceptable in English, ELCs should theoretically acquire T(S)V constructions at roughly the same time that they acquire the feature [-AGR]. If [-AGR] is acquired early, but [-TNS] is not, then beginner ELCs probably do not consider T(S)V sentences to be cases of copula- or auxiliary-drop, because copulas and auxiliaries both carry [TNS] in English. This analysis provides a possible explanation as to why ELCs take such a long time to acquire copula-drop sentences, while they acquire T(S)V sentences very early on.

A problem still remains, however, because FT/FA would predict that L1 transfer would cause ELCs to reject T(S)V sentences at the initial state. To address this we could say that perhaps there was a larger gap between the acquisition of the two contrasting constructions SV and T(S)V at an earlier stage of the ELCs' interlanguage due to L1 transfer, and that for the reasons just explained, the informants were able to acquire *pro-drop* very quickly as a result of the [AGR] parameter being quickly reset. Further research to investigate this would be helpful, but it would be very difficult to test for judgments of these sentence structures on informants who had been learning Cantonese for less than 3 months. It would perhaps be more effective to investigate the acquisition of other *pro-drop* constructions in beginner and intermediate ELCs (or L2-Mandarin learners) to see if *pro-drop* actually is acquired at such an early stage in L2-Chinese interlanguage. To test whether the above speculations are on the right track, further linguistic research could also design language tasks that attempt to isolate the features [-AGR] and [-TNS] in order to examine when each feature is acquired in L2-Chinese interlanguage.

## 7. Conclusion

The aim of the present study was to test for L1 transfer in the acquisition of null vs. overt subjects and pre- vs. post-verbal modals in the interlanguage of ELCs. The results indicated that the informants' performance on the L1-like *ho2ji5* and SV sentences was aided by L1 transfer, and that their performance on the non-L1-like *dak1* and T(S)V sentences was hindered by L1 transfer. Based on the results of the experiment, all three parts of the null hypothesis, as stated in (14), (15) and (16) were rejected, and it was claimed that, combined with the findings of Wakefield (2006), both components of FT/FA received empirical support. The informants' performance on the T(S)V sentences was much better than expected; it was assumed that there would be a larger gap between the scores of the SV and T(S)V sentences within group one. This result indicates that ELCs acquire T(S)V constructions surprisingly early on. One possibility is that the informants judged the T(S)V sentences as passive sentences with a dropped auxiliary, or as subject-predicate constructions with a dropped copula. However, this would conflict with the results of Wakefield (2006), which concluded that ELCs take a very long time to acquire copula-drop. It was concluded that the results can be explained better by assuming that ELCs reset the feature [+AGR] to [-AGR] much sooner than they reset the feature [+TNS] to [-TNS]. Suggestions were given for further research, specifically to see if English-speaking learners of Chinese actually acquire *pro-drop* as soon as was indicated by the study reported here, and to see if the feature [-AGR] is acquired sooner than the feature [-TNS].

The research study reported here provided more evidence in favor of a UG-based approach to language acquisition and language teaching. While lending further support to the idea of a natural, communicative style of language acquisition, its support of FT/FA also suggests that knowledge of the parametric differences between L1 and L2 grammars can be beneficial to teachers and curriculum writers. As linguists continue to learn more and more about which parameters must be reset with regard to specific L1s and L2s, curriculum writers should be more able to aim accurately towards the right kinds of language input data that trigger the resetting of parameters. This is especially relevant to the beginning stages of acquisition, but is also relevant to later stages, since some parameter settings appear to take much longer to reset than others.

## Acknowledgements

## References

Augusto, M. (2003). Recognizing 'then' as an Adverb: An experiment with Brazilian learners of English. In Juana M. Liceras et al (eds), *Proceedings of the 6th GenerativeApproaches to Second Language Acquisition Conference (GASLA 2002)*. Somerville, MA: Cascadilla Proceedings Project.

Cabrera, M. and Zubizarreta, M. L. (2005). Are all L1 grammatical properties simultaneously transferred? Lexical causatives in L2 Spanish and L2 English. In Laurent Dekydtspotter et al (eds), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)*. Somerville, MA: Cascadilla Proceedings Project.

Cheng, L. and Sybesma, R. (2004). Postverbal 'can' in Cantonese (and Hakka) and agree. *Lingua, 114(4)*, 419-445.

Cowan, R. and Hatasa, Y. (1994). Investigating the validity and reliability of native speaker and second-Language learner judgments about sentences. In Tarone, E., Gass, S. and Cohen, A. (eds), *Research methodology in second language acquisition*. Hillsdale, New Jersey: Lawrence Erlbaum.

Craats, I. van de, Corver, N., and Hout, R. van (2000). Conservation of grammatical knowledge: On the acquisition of possessive noun phrases by Turkish and Moroccan Learners of Dutch. *Linguistics, 38(2)*, 221-314.

Grüter and Conradie, S. (2006). Investigating the L2 initial state: Additional evidence from the production and comprehension of Afrikaans-speaking learners of German. In Slabakova, R., Montrul, S., and Prévost, P. (eds), *Inquiries in Linguistic Development: In Honor of Lydia White*. Amsterdam: John Benjamins.

Håkansson, G., Pienemann, M., and Sayehli, S. (2002). Transfer and typological proximity in the context of second language processing. *Second Language Research, 18/3*, 250-273.

Hawkins, R. and Chan, C. Y. (1997). The partial availability of Universal Grammar in Second language acquisition: The 'failed functional features hypothesis.' *Second LanguageResearch, 13,3,* 187-226.

Huang, C.-T. (1982/1998). *Logical Relations in Chinese and the Theory of Grammar,* NewYork & London: Garland.

Ionin, T. and Wexler, K. (2001). Optional Verb Raising in L2 Acquisition: Evidence from L1-Russian Learners of English. In João Costa and Maria João Freitas (eds), *Proceedings of the GALA 2001 Conference on Language Acquisition.* Lisboa, Associação Portuguesa de Linguistica.

Kong, S. (2005). The partial access of Universal Grammar in second language acquisition: An investigation of the acquisition of English subjects by L1 Chinese speakers. *Journal of East Asian Linguistics, 14,* 227-265.

Li, C. and Thompson, S. (1981). *Mandarin Chinese: a functional reference grammar,* Berkeley: University of California Press.

Li, W. (2005). *Topic chains in Chinese: A discourse analysis and applications in language teaching,* Muenchen: LINCOM Studies in Asian Linguistics.

Luke, K.K. and Nancarrow, O. (1998). Auxiliary verbs in Cantonese. In Matthews, S. (Ed.), *Studies in Cantonese Linguistics.* Hong Kong: Linguistics Society of Hong Kong.

Marsden, H. (2003). Inverse scope in L2 Japanese. In Beachley, B., Brown, A.,and Conlin, F. (eds), *BUCLD 27 Proceedings.* Somerville, Massachusetts: Cascadilla Press.

Matthews, S. and Yip, V. (1994). *Cantonese: A Comprehensive Grammar,* London:Routledge.

Montrul, S. (2001). First-language-constrained variability in the second-language acquisition of argument-structure-changing morphology with causative verbs. *Second Language Research, 17,2,* 144-194.

Schwartz, B. D., and Sprouse, R. (1994).Word order and nominative case in nonnative language acquisition. In T. Hoekstra and B. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar.* Amsterdam: Benjamins.

Schwartz, B. D., and Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research, 12,* 40–72.

Slabakova, R. (2000). L1 transfer revisited: the L2 acquisition of telicity marking in English by Spanish and Bulgarian native speakers. *Linguistics, 38,4,* 739-770.

Slabakova, R. (2005). What is so difficult about telicity marking in L2 Russian? *Bilingualism: Language and Cognition, 8,1,* 63-77.

Sopata, A. (2005). Optionality in non-native grammars: L2 acquisition of German constructions with absent expletives. *Bilingualism: Language and Cognition, 8,3,* 177-193.

Wakefield, J. (2006). Acquisition of Topics, Auxiliaries, and Non-verbal Predications in English-Cantonese Interlanguage: A Test of the Full Transfer/Full Access Hypothesis. Unpublished master's thesis, University of New England: Australia, 2006.

White, L. (2003). *Second Language Acquisition and Universal Grammar,* Cambridge: Cambridge University Press.

White, L., Valenzuela, E., Kozlowska-Macgregor, M., and Leung, Y-K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholoinguistics, 25,* 105-133.

Wong, C. (1998). First auxiliaries and modality in child Cantonese. In Matthews, S. (ed.), *Studies in Cantonese Linguistics*. Hong Kong: Linguistics Society of Hong Kong.

Yip, P-C. and Rimmington, D. (2004). *Chinese: A Comprehensive Grammar,* Oxon: Routledge.

Yip, V. (1995). *Interlanguage and Learnability: From Chinese to English*, Amsterdam: John Benjamins.

Yuan, B. (2003). The syntax of clausal negation in French and English Speakers' L2 Chinese.   In Juana M. Liceras et al (eds), *Proceedings of the 6th GenerativeApproaches to Second Language Acquisition Conference (GASLA 2002)*.Somerville, MA: Cascadilla Proceedings Project.

# Notes on Contributors

**FAN Li**'s research interests are in semantics and child language acquisition. She teaches at the College of Foreign Language, Beijing Forestry University.

**FAN Xiaoqian** is interested in linguistics and is doing linguistics at the postgraduate level in the Department of Modern Languages and Linguistics at the Chinese University of Hong Kong.

**Cauvis FUNG** is an M.Phil. candidate in the School of Humanities, The University of Hong Kong. Her research interests include Optimality Theory, Lexical-Functional Grammar and Cantonese syntax. She is completing a thesis on topicalization in Cantonese.

**David C.S. LI** is an associate professor at City University of Hong Kong. He has published mainly in four areas: Cantonese-English code-switching, Hong Kong Chinese EFL learners' learning difficulties and correction strategies, English in Hong Kong and global English.

**GE Lingling** is currently a PhD student at the Translation Department, the Chinese University of Hong Kong. As a professor of Nanhua University, she has been doing research on pragmatics and translation and has published several articles in different journals.

**Patricia MAN** studied linguistics and has been pursuing a career in music.

**Mineharu NAKAYAMA** teaches and does research in the Ohio State University, USA.

**Joanna Ut-Seong SIO** received her PhD from Leiden University Centre for Linguistics in 2006. She completed her work as a post-doctoral fellow in the Hong Kong Polytechnic University at the end of summer 2007.

**SUN Jing** studied linguistics in the Chinese University of Hong Kong and is now a lecturer of Teaching Chinese as a Foreign Language (TCFL) at the College of International Education and Exchange, Tianjin Normal University. Formerly Sun held positions as Lecturer TCFL (05-06), Confucius Institute at the University of Nairobi, Kenya and Lecturer TCFL (06-07), Confucius Institute at Bansomdejchaopraya Rajabhat University, Thailand.

# Notes on Contributors

**Sze-Wing TANG** received his PhD from University of California at Irvine in 1998. He holds the position of assistant professor in the Hong Kong Polytechnic University.

**John WAKEFIELD** has an MA in applied linguistics from the University of New England, Australia. He is currently a PhD student researching L2 Chinese in the department of Chinese and Bilingual Studies at the Hong Kong Polytechnic University.

**Colleen WONG** was President of the LSHK (2006-2007) and associate professor in the Hong Kong Polytechnic University, and is studying theology and spirituality in Dublin, Ireland.

**WONG Tak-Sum** is currently an MPhil student in the Division of Humanities of the HKUST. He received his BEng degree from the same institution in 2004. His research interest includes Cantonese linguistics, Chinese historical phonology, Chinese dialectology and palæography.

**Noriko YOSHIMURA** teaches and does research in the University of Shizuoka, Japan.

**ZHANG Xiaoheng** is an assistant professor in computational linguistics in the Department of Chinese and Bilingual Studies of the Hong Kong Polytechnic University. His research interests include Chinese information processing, computer assisted language learning and Modern Chinese characters.