LECT

HKU Working papers

Issue 1

香 港 大 學

THE UNIVERSITY OF HONG KONG

Published by:

Department of Linguistics

University of Hong Kong

http://www.linguistics.hku.hk/

Email: lectlang@hku.hk

LECT /lɛkt/

a specific language or language cluster: a language or a dialect; a suffix meaning linguistic 'variety' or 'code' which is ultimately derived from Ancient Greek légō (I speak).

LECT is an online, open-access journal of working papers by students at HKU whose research focuses on linguistics and language sciences. Through a simple peer review process, LECT provides feedback for authors considering further publication. Each submission is reviewed in-house by a scholar from the relevant sub-discipline at HKU. Our editorial board includes experts in phonetics, phonology, syntax, sociolinguistics, as well as hearing and speech sciences. Full-length working papers, squibs (new discoveries or short articles), and reviews are accepted for annual publication.

Arthur Lewis Thompson and Eileen Waegemaekers
Email: lectlang@hku.hk

Table of Contents

Perfective 'already' in Cantonese-English bilingual children: A case of relabeling or grammaticalization?

Pui Yiu Szeto
University of Hong Kong
pyszeto@gmail.com

Abstract

The emergence of perfective 'already' is attested in Cantonese-English bilingual children and a number of Southeast Asian contact languages. Analyzing the use of already in Cantonese-English bilingual children, we find that both the semantic and syntactic properties of the Cantonese [*zo2* ... *laa3*] construction are mapped on *already*. While the emergence of perfective 'already' is usually discussed under the framework of grammaticalization (e.g., Matthews & Yip, 2009), this paper demonstrates that it should also constitute a case of relabeling. Based on our analysis, we will discuss the relationship between relabeling and grammaticalization, and argue that the two research frameworks are complementary instead of mutually exclusive.

Keywords: perfective 'already', Cantonese-English bilinguals, language contact, grammaticalization.

1. Introduction

Relabeling (aka relexification) is a semantically driven mental operation which "consists in assigning a lexical entry a new label derived from a phonetic string drawn from another language" (Lefebvre, 2015, p. 1). It plays a key role in a wide range of language contact situations, from L2 acquisition at the individual level to contact language formation (e.g. creole genesis) at the communal level (Lefebvre, 1998, 2008). The concept of relabeling is theoretically appealing as it

represents a uniformitarian approach to language contact studies which can potentially address the connection between bilingual development and contact language formation, and is consistent with Mufwene (2001)'s important notion that all kinds of linguistic innovations and changes observable at the communal level must logically stem from the idiolectal level. As Lefebvre (2008) illustrates, relabeling is (partly) synonymous with several different terms commonly used in the literature of language contact, such as calquing, transfer, and reanalysis.

Grammaticalization, on the other hand, is considered a separate process (Lefebvre, 2009), despite its relevance to semantic inference. The emergence of perfective aspect in Cantonese-English bilingual children, a phenomenon revealing the intriguing parallels between bilingual development and contact language formation, constitutes the focus of this study. The English utterances of Cantonese-English bilingual children show a number of parallels with Singapore Colloquial English (Singlish) (Yip & Matthews, 2007), one of which is the use of bare verb forms together with already when inflectional morphemes such as the *-ed* and *-en* verbal suffixes are required in Standard English (1-4).

Cantonese-English bilingual children's English (Yip & Matthews, 2007)

(1) You wipe your mouth **already**? (Kasen 3;00)

(2) You swallow the short teeth **already**. (Timmy 3;01)

Singlish (Bao, 2005)

(3) I wash my hand **already**.
'I have washed/washed my hands.' (Bao, 2005, p. 239)

(4) I see the movie **already**.
'I have seen/saw the movie.' (Bao, 2005, p. 238)

Matthews & Yip (2009) argue that the emergence of perfective already presents a case of ordinary contact-induced grammaticalization, and make an important observation that both the [V *zo2* ... *laa3*] and [V *saai3* ... *laa3*] constructions may serve as the models for interlingual identification. Meanwhile, if we adopt the concept of relabeling, we don't have to invoke the theory

of grammaticalization. Instead, "what is being transferred into the creole is a lexical item with all of its functions, thus a multifunctional lexical entry" (Lefebvre, 2004, p. 180). In this case, the emergence of perfective *already* can be viewed as a result of transfer of both the lexical and grammatical functions of some corresponding Cantonese (substrate) lexical entry to the English (superstrate) *already*. Given that Lefebvre (2009) does acknowledge that grammaticalization plays a role in creole genesis, we would like to ask what different roles do relabeling and grammaticalization play in contact situations? Is it necessary (and feasible) to draw a clear distinction between them?

To address these questions and other relevant issues, we first review and discuss issues concerning grammaticalization (especially those occurring in contact situations) (Section 2), thereby outlining a framework for analysis. We then analyze how *already* develops into a perfective marker in Cantonese-English bilingual children, and report findings on the bilingual children's use of *already* in terms of function and placement (Section 3), which lay the foundation for the discussion of issues concerning grammaticalization and relabeling (Section 4).

2. Grammaticalization

Grammaticalization theory is concerned with the emergence and development of grammatical forms and constructions. According to Hopper & Traugott (2003, p. 18), the phenomenon of grammaticalization refers to "the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions". A typical example in English is the grammaticalization of the motion verb go into an auxiliary which expresses immediate futurity, as in "He's going to go to school soon" (Bybee et al., 1994; Hopper & Traugott, 2003). Interestingly, the grammaticalization of a motion verb meaning *go* into a future marker has occurred independently in a wide range of languages, such as English, French, Bari, Sotho, Zulu, Margi, Bassa, Klao, Igbo, Ecuadorian Quechua, Tamil, Basque (Heine & Kuteva, 2002), suggesting that such a phenomenon may be shaped by some universal processes of grammatical change, which may include the inherent conceptual link between go and the sense of futurity.

2.1. Contact-induced grammaticalization: an overview

Until recently, grammaticalization has typically been viewed as a "language-internal" process as there used to be a widespread assumption among linguists that grammatical structure, or syntax, cannot be "borrowed" (Heine & Kuteva, 2010). However, new findings from a number of fields, particularly areal typology (Ansaldo, 2009; Dahl, 2004; Enfield, 2013; Heine & Kuteva, 2003, 2005) and creole studies (Arends et al., 1994), have provided abundant evidence to demonstrate that grammaticalization could result from "external" factors such as geographical clustering and substrate influence. In their seminal work on grammaticalization, Hopper & Traugott (2003, p. 230) also acknowledge the important link between language contact and grammaticalization, concluding that "[c]ontact has been an important factor for most languages, and a strictly monogenetic view of grammaticalization is ultimately inappropriate".

Analyzing data from a wide range of languages, an important observation highlighted by Heine & Kuteva (2005) is that the principles of grammaticalization are the same regardless of whether or not language contact is involved. Similarly, Matthews & Yip (2009) see contact as a catalyst driving change along pathways of grammaticalization, and suggest that "the general principles of grammaticalization are applicable, not only to the substrate language(s) in which grammaticalization originally took place, but also to the contact language affected by it" (Matthews & Yip, 2009, p. 373).

For example, while the use of equivalents of 'already' as a perfective aspect marker (the focus of this study) is observed in contact varieties such as the English-based Singlish (Bao, 2005), the Spanish-based Zamboanga Chabacano (realized as *ya*) (Steinkrüger, 2013) (5), and the Portuguese-based Papi Kristang (realized as *ja*) (Baxter, 2013) (6), it is noteworthy that such a grammaticalization process has also developed 'internally' in languages like Inuit and Buli (Bybee et al., 1994).

(5) *Éle* *ya-matá* *pwérko* *gat* *Alyá* *gránde* *(Steinkrüger, 2013)*
3SG PFV-kill pig really there big

'S/he killed a really big pig there.'[1]

[1]The abbreviations used in this paper are as follows: 1 = 1st person; 2 = 2nd person; 3 = 3rd person; CLF = classifier; NEG = negation; PASS = passive; PERM = permissive; PFV = perfective; PL = plural; PRT = particle; SFP = sentence-final particle; SG = singular.

(6) *Eli ja bai mar onti anoti (Baxter, 2013)*
3SG PFV go sea yesterday night

'S/he went fishing last night.'

Such an observation is consistent with Mufwene (2001, 2008)'s argument against the distinction between "internally motivated change" and "externally motivated change". Contact plays a crucial role in Mufwene's approach to language evolution as he maintains that all kinds of linguistic change are made possible by the interaction between different idiolects, regardless of whether they are artificially labelled as "internally motivated" or "externally motivated". In short, although grammaticalization may be triggered in contact situations, its underlying principles remain unaffected.

2.2. Ordinary contact-induced grammaticalization

Heine & Kuteva (2005, p. 80) distinguish "two main types of contact-induced grammaticalization depending on whether or not there exists already a model source-to-target grammaticalization process to be replicated". If no such model exists, the process is referred to as ordinary contact-induced grammaticalization (whereas the other type is known as replica grammaticalization). The mechanisms involved in the transfer of grammatical concept or structure from the model language (M) to the replica language (R) in ordinary contact-induced grammaticalization are as follows:

(7) Ordinary contact-induced grammaticalization (Heine & Kuteva, 2005, p. 81)

a. Speakers notice that in language M there is a grammatical category Mx.

b. They develop an equivalent category Rx in language R on the basis of the use patterns available in R.

c. To this end, they draw on universal strategies of grammaticalization, using construction Ry in order to develop Rx.

d. They grammaticalize category Ry to Rx.

This process, according to Hickey (2010, p. 155), is "an unconscious one and persists even with speakers who have considerable target language proficiency".

The use of *already* as a perfective aspect marker in Singlish is an example of ordinary contact-

induced grammaticalization, in which a new grammatical category was developed in the replica language (Singlish, a contact variety of English) due to transfer from the model languages (Southern Sinitic and Malay varieties). Compared to Standard English, Singlish seldom employs inflectional morphology. Its aspectual system, as Bao (2005, p. 237) puts it, is "essentially the Chinese system filtered through the morphosyntax of English". For example, already occurs after uninflected verbs to mark perfective aspect without involving any inflectional morpheme. This is consistent with the use of the perfective aspect marker liau51 in Hokkien, one of the most important substrate languages in the formation of Singlish (Ansaldo, 2009):

(8) a. *He* *eat* *already* *[Singlish]*
He eat PFV

'He has eaten'

b. *i*$^{55>33}$ *tsiaʔ*$^{5>32}$ *pŋ*$^{33>21}$ *liau*^{51} *[Hokkien]*
3SG eat rice PFV

'He has eaten'

In this case, the Singlish speakers grammaticalized a lexical item into an aspect marker instead of using inflectional morphology to mark perfective aspect, following the aspect-marking strategy of the substrate languages.

3. Emergence of perfective *already* in Cantonese-English bilingual children

The use of *already* as a perfective marker in Cantonese-English bilingual children was first discussed in Kwan-Terry (1989) and further investigated in Yip & Matthews (2007). According to these studies, the bilingual children's development is parallel to the development of *already* as a perfective aspect marker in Singlish, which is used in post-verbal or clause-final position to express perfective notions such as completion and change of state. Such a use of *already* is observed in all nine children in the Hong Kong Bilingual Child Language Corpus (Yip & Matthews, 2007), regardless of their dominant language (Szeto et al., 2017). The question is – why is *already* (but not any other lexical item) identified as a perfective marker?

3.1. Already and the Cantonese [zo2 ... laa3] construction

The development of *already* into a marker of perfective aspect is attested in a number of languages (see Section 2.1). The existence of such parallel developments is suggestive of some inherent conceptual links between the semantics of *already* and the sense of perfectivity.

The English adverb *already* is often associated with completed events, which makes it a natural perfective aspect marker. For example, in (9), the event of starting is completed when the event of arriving occurs.

(9) The performance had **already** started when we arrived.

However, as Traugott & Waterhouse (1969) and Soh (2009) point out, in addition to a "change of state" interpretation, an important feature associated with the use of already is that the assertion made by the sentence is contrary to what one may expect or assume (referred to as the "contrary to expectation" interpretation by Soh (2009))[2]. For example, the sentence in (9) implies that the performance began earlier than expected. A more noticeable example involving a "contrary to expectation" sense is given in (10), where B uses *already* to correct A's wrong assumption B hasn't finished his/her homework yet.

(10) A: You have to finish your homework first.
B: I've finished my homework **already**!

As Soh (2009) demonstrates, the semantics of *already* is similar to the Mandarin sentence-final particle *le*, which corresponds closely to the Cantonese *laa3* (Matthews & Yip, 2013). The Cantonese *laa3* is a sentence-final particle which functions to express current relevance (Matthews & Yip, 2013) or a change of state (Cheung, 2007). It often co-occurs with the perfective aspect marker *zo2* to describe a completed event which has current relevance, thus entailing a change of state, as in (11), where there must be a prior state in which the speaker has not returned home.

(11)	*Ngo5*	*faan1-zo2*	*uk1kei2*	*laa3*
	1sg	return-PFV	home	SFP

[2] As Soh (2009, p. 624) demonstrates, the "change of state" and "contrary to expectation" interpretations are independent of each other. In some contexts, already may only carry the "contrary to expectation" interpretation without any "change of state" interpretation (but not the other way around).

'I have returned home'

Like *already*, the [*zo2* ... *laa3*] construction may express senses of "change of state" and "contrary to expectation" simultaneously. For example, the utterance of B in (12) is associated with a sense of "contrary to expectation" because the fact that "B has eaten" is in contrast with A's inviting B to join him/her for a meal. At the same time, B's utterance also conveys the meaning that the eating event is completed and s/he has changed from a state of "not having eaten" to "having eaten". (13) is the English equivalent of (12), in which *already* carries the "change of state" and "contrary to expectation" readings.

(12) a.

Heoi3	*m4*	*heoi3*	*sik6*	*faan6*	*aa3?*
go	NEG	go	eat	rice	SFP

'Will you go and eat the meal?'

b.

Ngo5	*sik6-zo2*	*je5*	*laa3*
1SG	eat-PFV	thing	SFP

'I've already eaten'

(13) A: Let's go eat, shall we?

B: No, I've **already** eaten.

Although the [*zo2* ... *laa3*] construction shares some properties with already in the above example, the differences in acceptability between (14) and (15) suggest that they differ in the obligatoriness of the "contrary to expectation" reading when responding to a neutral yes/no question, the use of [*zo2* ... *laa3*] construction is perfectly acceptable, but the use of *already* in (16) sounds unnatural to native speakers. This is because the "contrary to expectation" reading inherently associated with already is incompatible with a neutral context; meanwhile, the use of [*zo2* ... *laa3*] construction in (14) does not pose a problem because it is not obligatorily associated with a "contrary to expectation" reading. See Szeto et al. (2017) for a more detailed analysis of the congruence between *already*, *zo2*, and *laa3*.

(14) a.

Maai5-zo2	*bun2*	*syu1*	*mei6*	*aa3?*
buy-PFV	CLF	book	NEG	SFP?

'Have you bought the book?'

b. *Maai5-zo2* *laa3*
buy-PFV SFP

'Yes, I have'

(15) A: Have you bought the book?
B: Yes, I have **already** bought it.

Based on the above analysis, the high degree of semantic overlap between already and [*zo2* ... *laa3*] construction makes it likely for the bilingual children to identify the former with the latter. If this is the case, two predictions follow. First, *already* will undergo semantic bleaching and no longer be obligatorily associated with a 'contrary to expectation' interpretation. Second, given that Cantonese [*zo2* ... *laa3*] always occurs post-verbally, we predict that *already* will have a strong tendency to occur post-verbally in the bilingual children's utterances, unlike the monolingual use of *already*, which can also occur pre-verbally.

3.2. The use of already in bilingual children

Based on the analysis presented in Section 3.1, the high degree of semantic overlap between *already* and [*zo2* ... *laa3*] construction makes it likely for the bilingual children to identify the former with the latter. If this is the case, we would expect *already* to undergo semantic bleaching and lose the obligatory "contrary to expectation" reading. Consistent with our prediction, there are cases in which the bilingual children's use of *already* is unlikely to be associated with a "contrary to expectation" interpretation. In (15), the child appears to be replying to a neutral question by indicating that the sending action is completed. There appears to be no specific presupposition suggesting a "contrary to expectation" interpretation.

(16) Investigator: Did you send him to the hospital
Child: I send **already** (Timmy 3;01)

Moreover, one child is observed to use *already* when asking yes/no questions with no presupposition about whether the enquired actions should have been done, as illustrated in (16). The data supports our hypothesis that *already* is identified with the [*zo2* ... *laa3*] construction and undergoes semantic bleaching in the bilingual children, losing its obligatory association with the "contrary to

expectation" reading.

(17) Child: You **already** eat?
Investigator: Mm?
Child: Okay, you wipe your mouth **already**? (Kasen 3;00)

Already is expected to have a strong tendency to occur post-verbally in the bilingual children due to influence from the Cantonese post-verbal [*zo2* ... *laa3*] construction. The corpus data confirm this prediction. As shown in Table 1, already predominantly occurs post-verbally (90%) in the bilingual children[3]. Five out of nine children produced *already* only in post-verbal position. Bilingual children's strong tendency to place already in post-verbal position is contrasted with monolingual children's clear preference for preverbal placement of *already*. As shown in Table 2, pre-verbal *already* is used more often than post-verbal *already* (69.9% vs. 30.1%) in the monolingual children. Fisher's (1954) exact test shows that the differences between the two groups are highly significant ($p < .0001$). The results support our hypothesis that the bilingual children have identified *already* with the Cantonese [*zo2* ... *laa3*] construction.

4. Relabeling or grammaticalization?

By now there is ample evidence suggesting that *already* is identified with the [*zo2* ... *laa3*] construction by the bilingual children, as reflected by its semantic and syntactic properties. In this section, we will first discuss whether the emergence of perfective *already* constitutes a case of relabeling in Lefebvre's sense. Then we will come back to the questions raised in Section 1 - what different roles do relabeling and grammaticalization play in contact situations? Is it necessary (and feasible) to draw a clear distinction between them?

[3]The post-verbal category can be further divided into immediately post-verbal *already* and sentence-final *already*. The strong preference for the latter reflects the syntactic properties of the Cantonese sentence-final particle *laa3*. Kathryn appears to be an outlier in this regard – unlike all the other bilingual children, she has slightly more preverbal use, like the monolinguals do. We suspect such as difference may be due to the fact that Kathryns recording started at a much later stage than the other children (see Table 2). At such an age period, the target-like tense/aspect forms may gradually emerge in the bilingual child at the expense of the non-target perfective *already*. See Szeto et al. (2017) for a more detailed discussion.

Name	**Pre-verbal already**	**Post-verbal already**
Timmy	1	12
Sophie	0	20
Alicia	0	4
Llywelyn	0	3
Kathryn	4	3
Charlotte	1	11
Janet	0	24
Kasen	5	18
Darren	0	4
TOTAL	**11** (10%)	**99** (90%)

Table 1: The placement of *already* in the Cantonese-English bilingual children

Name	**Pre-verbal already**	**Post-verbal already**
Adam	14	4
Eve	2	2
Sarah	14	3
Lara	5	5
Thomas	16	8
TOTAL	51 (69.9%)	22 (30.1%)

Table 2: The placement of *already* in the English monolingual children

4.1. Perfective already as a case of relabeling

In relabeling, a lexical entry from the substrate language (18a) is assigned a new label derived from a phonetic string drawn from the superstrate language (18b), and the original phonological representation is eventually removed, yielding (18c). Thus, the resultant lexical entry carries the substrate semantic and syntactic properties with a phonological representation derived from the superstrate language (Lefebvre, 2015, pp. 1-2):

(18)

a.	/phonological representation/$_i$ [semantic features]$_i$ [syntactic features]$_i$
b.	/phonological representation/$_i$ /phonological representation/$_j$ [semantic features]$_i$ [syntactic features]$_i$
c.	/phonological representation/$_j$ [semantic features]$_i$ [syntactic features]$_i$

As discussed in Section 3, the bilingual children's *already* manifests both the semantic and syntactic properties of the Cantonese [*zo2* ... *laa3*] construction. Nevertheless, there is an important difference between this case and cases of relabeling discussed in Lefebvre's work, where the substrate lexical entry involved is typically a polysemous word with both lexical and grammatical functions, such as the Quechua *tiya-ri-*, which means 'to sit', 'to live', 'there is', or serves as a locative marker. By contrast, *zo2* and *laa3* are not independent lexical entries; instead, they must combine with content words to perform grammatical functions. Here, if we extend the applicability of relabeling to constructions, the emergence of perfective already in the Cantonese-English bilingual children arguably constitutes a case of relabeling, where the semantic and syntactic features of the Cantonese [*zolaa3*] construction are mapped onto the English *already*, as exemplified in (19):

(19)

	C-E Bilingual children *already*	Cantonese [*zo2...laa3*]	English *already*
Semantics	[+ completion] [+ change of state] [± contra-exception]	[+ completion] [+ change of state] [± contra-exception]	[+ completion] [+ change of state] [+ contra-exception]
Syntax	[+post-verbal] [-pre-verbal]	[+post-verbal] [-pre-verbal]	[+post-verbal] [+pre-verbal]

In other words, the Cantonese [*zo2* ... *laa3*] construction is assigned a new label derived from the phonetic representation of the English *already*. The high degree of semantic congruence between [*zo2* ... *laa3*] and *already* has triggered the relabeling process.

As illustrated above, whether the emergence of perfective *already* represents a case of relabeling solely depends on whether we will extend the applicability of relabeling beyond individual lexical entries. Regardless of whether a lexical entry or construction is involved, the underlying mechanisms are arguably the same - a new label derived from a superstrate phonetic string is assigned to a substrate lexical entry or construction.

4.2. Two terms for one and the same mental process?

As discussed in Section 2, the emergence of perfective *already* (both in Cantonese-English bilingual children and contact languages) can be studied under the theoretical framework of (contact-induced) grammaticalization. Now that we also consider it a case of relabeling, we may seem to have created a paradox and have to seek ways to reconcile the two theoretical frameworks.

Lefebvre (2009) makes a distinction between relabeling and grammaticalization, and maintains that the former is the main process in the creation of a creole's lexicon, which applies to both lexical and functional items. Given that relabeling is semantically driven, Lefebvre (2009, p. 297) argues that grammaticalization may come into play "when a substrate lexical entry could not have been relabelled because of its lack of semantic content, or because of the absence, in the superstrate language, of an appropriate form to relabel the substrate entry". This statement can be understood as follows in cases where the grammatical items co-exist with their lexical source in the substrate language, the semantic and syntactic properties of this lexical entry (including those of its derived grammatical categories), can be transferred to the emergent creole through relabeling. Meanwhile, in cases where a particular substrate grammatical category lacks a clear lexical source or an exponent which is congruent with the superstrate structure, a new lexical entry manifesting the properties of that grammatical category will be developed in the creole by means of grammaticalization. Notably, such a distinction between relabeling and grammaticalization is strongly reminiscent of that of the two main types of contact-induced grammaticalization, which depends on "whether or not there exists already a model source-to-target grammaticalization process to be replicated" (Heine & Kuteva, 2005, p. 80). The type of contact-induced grammaticalization comparable to relabeling in its narrow sense (which targets lexical entries but not constructions) is termed replica grammaticalization, in which a grammaticalization process is transferred from the model language to the replica language. For example, in Cantonese, the word for 'give' *bei2* is believed to have undergone the grammaticalization pathway 'lexical verb > permissive marker >

passive marker' (Lord et al. 2002). While still retaining the lexical meaning of 'give', as illustrated in (20), *bei2* has also developed into a grammatical marker.

(20)

Ngo5	*bei2-zo2*	*bun2*	*syu1*	*keoi3*
1SG	give-PFV	CLF	book	3sg

'I've given him/her a book.'

(21)

Aa3maa1	*m4*	*bei2*	*ngo5*	*tai2*	*din6si6*
mother	NEG	PERM	1SG	watch	television

'My mum doesn't let me watch TV.'

(22)

Ngo5	*bei2*	*keoi5*	*daa2-zo2*	*jat1*	*haa5*
1SG	PASS	3SG	hit-PFV	one	PRT

'I was hit by him/her once.'

In (21), *bei2* functions as a permissive marker, expressing the sense that 'my mum doesn't give me the permission to watch TV'. In (22), *bei2* functions as a passive marker, expressing the sense that the speaker is the patient of the hitting action. Interestingly, Yip & Matthews (2007, p. 243) observe that Cantonese-English bilingual children use give as a grammatical marker in English, i.e. permissive *let* in (23a) and passive marker in (24a), apparently recapitulating the grammatical development pathway of *bei2*:

(23) a. (appearing in swimsuit) Daddy, I **give** you see. (Sophie 3;04)

b.

Baa4baa1	*ngo5*	*bei2*	*nei5*	*tai2*
daddy	1SG	PERM	2SG	see

'Daddy I let you see'

(24) a. (points to scratched leg) Here is **give** Timmy scratch. (Sophie 3;06)

b.

Ni1dou6	*hai6*	*bei2*	*Timmy*	*waa2-can1*	*ge3*
here	be	PASS	Timmy	scratch-PRT	SFP

'Here is scratched by Timmy'

From the perspective of relabeling, the Cantonese equivalents (23b-24b) of the child's utterances illustrate that *give* manifests both the semantic and syntactic features of the Cantonese *bei2*.

Meanwhile, in many cases, it is not possible to "replicate" the grammaticalization pathway of a substrate grammatical item, either because of the absence of a clear lexical source of the grammatical item concerned or that of a corresponding superstrate structure (which often results from a lack of clean typological fit between languages in a given contact situation). Under such circumstances, ordinary contact-induced grammaticalization will be brought into play, through which the substrate grammatical categories will be developed in the superstrate language on the basis of the use patterns available in the latter (see Section 2.2). The emergence of perfective *already* in Cantonese bilingual children, as well as parallel developments in Singlish and other Southeast Asian creoles, represent cases of ordinary contact-induced grammaticalization triggered by the typological differences between the substrate and superstrate languages (Szeto et al., 2017).

The driving force and underlying mechanisms of the two types of contact-induced grammaticalization are essentially identical - speakers of the model (substrate) language would expect the same grammatical distinctions in the replica (superstrate) language when they acquire the latter as a second L1 or an L2. Consequently they would search for equivalents in the superstrate language to categories in the substrate language with which they are already familiar. In both cases, the speakers draw on universal strategies of grammaticalization to develop a new grammatical category in the contact language. The only difference lies in the availability of a salient grammaticalization pathway which can be transferred from the substrate to the superstrate. Therefore, it makes perfect sense to consider ordinary contact-induced grammaticalization and replica grammaticalization to be two subtypes of the same kind of mental operations.

However, if we maintain a rigid distinction between relabeling and grammaticalization by restricting the former to cases involving lexical entries with concrete semantic content, we may essentially be giving two names to a single type of mental process. Moreover, as relabeling in its narrow sense may still involve the emergence and development of grammatical forms from lexical items, there is no good reason to treat relabeling and grammaticalization as two distinct processes in contact scenarios. The roles of relabeling and grammaticalization are schematically illustrated in Table 3.

Contrary to Lefebvre's (2009) position, we see a high degree of overlap between relabeling and grammaticalization. However, we share Lefebvre's view that relabeling plays a major role in creole genesis and other types of contact-induced change. Given that the development of grammatical categories from lexical items can be driven by the conceptual links between them (see Section 2), the

transfer of lexical meaning in contact situations can potentially trigger the subsequent development of grammatical categories. Therefore, the notion of relabeling can provide a coherent framework for

	Relabeling	**Grammaticalization**
Transfer of lexical meaning, e.g. Fongbe hù 'to murder/to mutilate' > Haitian Creole *ansasinen* 'to murder/to mutilate' (Lefebvre, 2008, pp. 92-93)	+	-
"Replication" of grammaticalization pathway, e.g. permissive and passive *give*	+	+
Emergence of new grammatical categories based on substrate constructions, e.g. perfective *already*	±[4]	+
Development of grammatical categories without language contact	-	+

Table 3: Roles of relabeling and grammaticalization

the emergence of new lexical meaning as well as new grammatical categories in contact scenarios (provided we accept that relabeling can also be applied to constructions). On the other hand, the theory of grammaticalization provides a research framework for the emergence and development of grammatical categories, regardless of whether language contact is involved.

5. Conclusions

Unlike its counterpart in Standard English, the *already* in Cantonese-English bilingual children lacks the obligatory "contrary to expectation" reading and has a strong tendency towards post-verbal placement, manifesting both the semantic and syntactic characteristics of the Cantonese [*zo2* ... *laa3*] construction, which arguably constitutes a case of relabeling. Although relabeling is only applicable to lexical entries with semantic content in previous studies, given the similar underlying mechanisms involved, the notion of relabeling should be extended to constructions in order to provide a coherent framework for studying grammatical change in contact scenarios. Further, instead of making a rigid distinction between relabeling and grammaticalization, we see a lot of parallels between these two notions and argue that they can serve as complementary research

[4] In its narrow sense, relabeling only targets individual lexical entries with solid semantic content. In that case, the emergence of perfective *already* does not constitute a case of relabeling. However, as we argue in Section 4.1, the notion of relabeling should be extended to constructions.

paradigms for the study of language change.

References

Ansaldo, U. (2009). *Contact languages: Ecology and evolution in Asia.* Cambridge University Press.

Arends, J., Muysken, P., & Smith, N. (1994). *Pidgins and creoles: An introduction*, volume 15. John Benjamins Publishing.

Bao, Z. (2005). The aspectual system of singapore english and the systemic substratist explanation. *Journal of Linguistics*, *41*(2), 237–267.

Baxter, A. N. (2013). Papiá kristang structure dataset. *Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.*

Bybee, J. L., Perkins, R. D., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world.* University of Chicago Press Chicago.

Cheung, S. H.-n. (2007). *A Grammar of Cantonese as Spoken in Hong Kong.* The Chinese University Press.

Dahl, Ö. (2004). *The growth and maintenance of linguistic complexity*, volume 71. John Benjamins Publishing.

Enfield, N. J. (2013). *Linguistic epidemiology: Semantics and grammar of language contact in mainland Southeast Asia.* Routledge.

Heine, B. & Kuteva, T. (2002). *World lexicon of grammaticalization.* Cambridge University Press.

Heine, B. & Kuteva, T. (2003). On contact-induced grammaticalization. *Studies in Language.*, *27*(3), 529–572.

Heine, B. & Kuteva, T. (2005). *Language contact and grammatical change.* Cambridge University Press.

Heine, B. & Kuteva, T. (2010). Contact and grammaticalization. In R. Hickey (Ed.), *The handbook of language contact* (pp. 86–105). West Sussex: Blackwell Publishing Ltd.

Hickey, R. (2010). Contact and language shift. In R. Hickey (Ed.), *The handbook of language contact* (pp. 151–169). Wiley-Blackwell Oxford, UK.

Hopper, P. J. & Traugott, E. C. (2003). *Grammaticalization.* Cambridge University Press.

Kwan-Terry, A. (1989). The specification of stage by a child learning English and Cantonese simultaneously: A study of acquisition processes. *Interlingual processes*, 33–48.

Lefebvre, C. (1998). *Creole genesis and the acquisition of grammar: The case of Haitian Creole*, volume 88. Cambridge University Press.

Lefebvre, C. (2004). *Issues in the study of pidgin and creole languages*, volume 70. John Benjamins Publishing.

Lefebvre, C. (2008). Relabelling: A major process in language contact. *Journal of Language Contact*, *2*(1), 91–111.

Lefebvre, C. (2009). The contribution of relexification, grammaticalisation, and reanalysis to creole genesis and development. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, *33*(2), 277–311.

Lefebvre, C. (2015). *Relabeling in language genesis.* New York: Oxford University Press.

Matthews, S. & Yip, V. (2009). Contact-induced grammaticalization: Evidence from bilingual acquisition. *Studies in Language. International Journal sponsored by the Foundation Foundations of Language*, *33*(2), 366–395.

Matthews, S. & Yip, V. (2013). *Cantonese: A comprehensive grammar.* Routledge.

Mufwene, S. S. (2001). *The ecology of language evolution.* Cambridge University Press.

Mufwene, S. S. (2008). *Language Evolution: Contact, Competition and Change.* London: Continuum.

Soh, H. L. (2009). Speaker presupposition and Mandarin Chinese sentence-final -le: a unified analysis of the change of state and the contrary to expectation reading. *Natural Language & Linguistic Theory*, *27*(3), 623–657.

Steinkrüger, P. O. (2013). Zamboanga Chabacano structure dataset. *Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.*

Szeto, P. Y., Matthews, S., & Yip, V. (2017). Multiple correspondence and typological convergence in contact-induced grammaticalization. *Journal of Language Contact*, *10*(3), 485–518.

Traugott, E. C. & Waterhouse, J. (1969). alreadyand yet: a suppletive set of aspect-markers? *Journal of Linguistics*, *5*(2), 287–304.

Yip, V. & Matthews, S. (2007). *The bilingual child.* Cambridge University Press.

Testing the Competence of First Language: A Prototype Test for Language Maintenance

Hongdi Ding
University of Hong Kong
hongdiding@gmail.com

Abstract

This research presents the design and application of a prototype criterion-referenced competence test for first language(s) of linguistic minority speakers. It is designed for the linguistic purpose of language maintenance, not for educational or professional purpose. The testing content is the core or implicit knowledge of the language system per se (i.e. phonology, morphology, syntax, lexicon, semantics and pragmatics). The aim is to measure how competent the linguistic minorities are in their first language(s) when they are facing bi- or multi-lingualism. A major function of the prototype test is to identify the imperceptibly incremental change of language competence. This method can be applied to languages without writing systems and illiterate speakers. The testing results yielded from this methodology are presented, with application to bilingual speakers of Nuosu, a Tibeto-Burman language in Sichuan, China, and Mandarin Chinese. Accordingly, first language and bilingual competence, and language shift are also discussed.

Keywords: first language and bilingual competence, language maintenance, prototype, language assessment.

1. Introduction

The purpose of this research is to present the design and application of a prototype criterion-referenced competence test for the first language(s) (L1) of linguistic minority speakers for the

purpose of language maintenance. Many minority languages are said to be under the threat of the culturally dominant languages. It may be true given the domains of language use. However, since many speakers still learn the minority language as L1 and speak it fluently without perceptible difficulty, it is beyond the reach of direct observation to conclude whether there is actual language attrition or shift, manifested by decrease in competence, both individually and communally. Therefore, this research is targeted to answer the question of how competent the linguistic minority speakers are in their first language(s) when they are facing bi- or multilingualism.

If two tests are designed and co-normed for bilingual first language speakers, namely those who learn both the ethnic minority language and the culturally dominant language as L1s, bilingual competence and possible language shift can be identified. The application of the test will be addressed in Section 6 with bilingual speakers of Nuosu, a Tibeto-Burman language, and Mandarin Chinese, from Liangshan Yi Autonomous Prefecture (hereafter known as Liangshan, literarily Cool Mountains), Sichuan, China.

The proposed testing method is a comprehensive, objective and linguistic diagnosis and measurement of the core or implicit grammatical knowledge or the language system per se of the first language(s) of normal speakers. The testing content is phonology, morphology, syntax, lexicon, semantics and pragmatics of the target language. Since grammatical knowledge is widely accepted as the manifestation of competence, the validity of the test is justifiable (see Section 3.4). The first language does not necessarily need a writing system; the L1 speakers can be illiterate.

As a competence test, it is different from the mainstream L2 proficiency tests (e.g. TOEFL, TOEIC) and L1 achievement tests (e.g. final exams, high-stakes entrance exams). Although "proficiency" and "competence" are used interchangeably in common sense, they differ in linguistic implications. Proficiency implies variation and is traditionally related to learning, teaching and testing in second language (e.g., Stern, 1983; Llurda, 2000); but competence implies stability (e.g., Chomsky, 1965; Taylor, 1988). Both proficiency and achievement tests are mainly for the purpose of school education. Since the tasks of these tests integrate language knowledge and skills (e.g. listening, speaking, reading and writing), it involves complex cognitive but non-linguistic abilities, such as the ability of analyzing, evaluating and synthesizing information. These cognitive abilities need explicit or formal school education, and result in large variations among performance of the examinees, whether they are native speakers or not.

Other than these, there is explicit or learned language knowledge (cf. core or implicit language knowledge in Section 3.2). For example, the content of the Chinese subject test of the University Entrance Exam in China includes comprehension of ancient Chinese texts, argumentation writing, etc. The students follow certain curriculum and textbooks, which can be reviewed before the test. Full literacy is required in order to participate. This is far beyond the core language knowledge. People who have not been formally educated are less likely to obtain the learned knowledge. However, any native speaker, without a single day of schooling, should have nearly all the core or implicit knowledge of the target language. Tests for language revitalization (e.g. Southern Min in Taiwan, Mori and Hawaiian) also follow the mainstream integrative approach as proficiency tests. In Section 2, I will compare with the proposed method another three kinds of tests of linguistic purpose: L1 vocabulary test, qualitative L1 competence test, and L1 competence test for therapeutic purpose.

Secondly, a native speaker, who speaks the language as L1 and possesses native competence, should find the test easy. Their competence should be stable (see Section 3.5 and Section 6.1). The instrument can be considered reliable if the test results are also stable under stable sociolinguistic environment. Only those L1 speakers without native competence, probably experiencing language shift, will make a significant number of mistakes. This is also the reason why such a test was not strongly needed several decades ago; by then, many first languages were still stably maintained. However, what is happening now is that an increasing number of speech communities are experiencing "unstable bilingualism or multilingualism" (see Austin & Sallabank, 2011, p. 1).

The methodology of designing the competence test will be presented in Section 3 and 4, such as the testing content and its justification, solution to dialectal difference and the literacy issue. I will illustrate the testing interface and sample questions with Nuosu, a Tibeto-Burman language in Sichuan, China, in Section 5. The application of the test is presented in Section 6, accompanied by comparative data through a vocabulary test. To conclude, I will suggest a research agenda (Section 7).

2. A brief review and comparison with the existing testing tools

Due to their educational or professional purpose, the mainstream L2 proficiency tests and L1 achievement tests are qualitatively different from the proposed method. However, there are three

kinds of tests which also focus substantially on implicit grammatical knowledge, namely L1 vocabulary test, qualitative L1 competence test and L1 competence test for therapeutic purpose. Their purpose is linguistic (e.g. language loss and impairment). But neither can they serve the same purpose of testing language competence of normal speakers.

Vocabulary test is often used for the sake of vocabulary knowledge itself (Read, 2000, see). But it is also frequently used as a quick and convenient instrument to test competence (e.g., Dai, 2011). I will prove, in Section 6.2, that vocabulary test, if used to test competence, can only make distinctions on the condition that the target language competence of the speech community is already low. It will overestimate if the L1 speakers are still fluent in the target language.

The qualitative L1 tests (e.g. grammaticality judgment, spontaneous speech elicitation) is used to understand the existing perceptible phenomena by studying one or more grammatical aspects (e.g., Altenberg, 1991; De Bot et al., 1991). They have taken certain prior value positions. The purpose is not to measure, but to confirm and deepen the understanding of some readily observable phenomena. However, there is no prior value position with the present method. It is the quantitative results from the test, which decide the situation of the target language, such as whether there is language shift and attrition.

The L1 competence tests for people with language impairment, such as Bilingual Aphasia Test (Paradis & Libben, 1987), are mainly for therapeutic purpose. Their results can only conclude whether the person has language-related pathological problems, not whether the person is a competent speaker.

To sum up, the proposed method is different from the existing tools mainly in the following seven features (also see Section 1):

1. *Comprehensive:* the content of the proposed competence test should be as adequate as possible to cover the language system;
2. *Objective:* generalizations are fair and unbiased, without subjective influence (e.g. self-assessment); the test results should not be an underestimation or overestimation;
3. *Linguistic:* the purpose of the test is not to echo pedagogical practice in language education, but for linguistic investigation, such as language maintenance;
4. *Diagnosis and measurement:* the most important functionality of the research is to diagnose

and quantify language competence; no prior value position is needed;

5. *Core or implicit grammatical knowledge of the language system:* the testing content is not the explicit knowledge (including literacy skills) which needs formal learning;

6. *First language:* the target language is not second language or foreign language;

7. *Normal speakers:* the target population should not have language impairment.

Table 1 summarizes the seven aspects of the reviewed tests and the present method:

	Target	**Knowledge**	**Coverage**	**Subjects**	**Test nature**	**If applied to test L1 competence**	**Purpose**
Present Method	L1	Implicit	Comprehensive	Normal	Measurement	Objective	Linguistic
Vocabulary test	L1	Mostly implicit	Fractional	Normal	Measurement & understanding	Overestimating	Linguistic
Qualitative test	L1	Implicit	Fractional	Normal	Understanding	n.a.	Linguistic
Therapeutic test	L1	Implicit & explicit	Comprehensive	Patient	Measurement	n.a.	Linguistic
Achievement test	L1	Explicit	Comprehensive	Normal	Measurement	Underestimating	Educational
L2 test	L2	Explicit	Comprehensive	Normal	Measurement	n.a.	Educational & professional
Revitalization test	L2	Mostly explicit	Comprehensive	Normal	Measurement	n.a.	Educational & professional

Table 1: A comparison between the existing testing tools and the present method

3. The testing content and its justification

For the next two sections, methodology of designing the prototype competence test will be introduced. In this section, I should answer the question: what to be tested. Following the justification for testing grammatical knowledge from Section 3.1 to 3.5. Section 3.6 will address the solution to the dialectal differences between the testing variety and the subjects' first variety.

3.1. The testing content

The testing content is the core or implicit grammatical competence or knowledge. It is the language knowledge possessed by a native speaker, but perhaps without any form of schooling or explicit learning. In order to justify the use of grammatical knowledge as the testing content, there are three sub-questions: 1) what grammatical competence is, 2) why it is the testing content, and 3) whether it is testable. Grammatical competence, according to Chomsky (1965), is the entire essence of the concept of competence. It deals with the knowledge of form and meaning. The three major components are syntax, phonology, and semantics (Chomsky, 1965, p. 16). For other scholars, grammatical competence comprises different components:

Chomsky (1965, p.16)	syntax, phonology, and semantics
Lado (1961, as cited in Purpura, 2004, p.51)	morphosyntactic form
Canale & Swain (1980, p.29)	lexical items, morphology, syntax, sentence-grammar semantics and phonology
Bachman (1990, p.87)	vocabulary, morphology, syntax, and phonology/graphology
Purpura (2004, p.91)	form-meaning correspondence among phonology/graphology, lexicon, morphosyntax, cohesion, information management, and interaction

Table 2: Components of grammatical competence by different scholars

As is indicated by Chelliah & Willem (2010, p. 168), a native speaker should be able to "produce syntactically well-formed utterances that are semantically meaningful, pragmatically adequate, and textually coherent". Larsen-Freeman (1997) characterizes grammatical knowledge along three dimensions: linguistic form, semantic meaning and pragmatic use.

Therefore, the components of grammatical competence for the proposed test are 1) phonology, e.g. place of articulation, 2) morphology, e.g. derivational suffixes, 3) syntax, e.g. NP structure and causative construction, 4) lexicon, e.g. meaning, word collocation, and the syntactic behavior of adpositions and conjunctions, 5) semantics, e.g. meaning of the ideophones, and 5) pragmatics,

e.g. topic-comment construction. Some features can also be at the interface of the components (e.g. morphosyntax). Semantic competence overlaps with lexical competence in that both involve lexical semantic knowledge. If a question targets at merely lexical semantic knowledge of the words, it falls within the domain of semantic competence; if it involves more than lexical semantic knowledge, the feature falls within other domains (e.g. lexicon or pragmatics).

Each domain should have its representations in the test. But the proportion is language specific. For example, some languages are morphologically poor; the domain of morphology may have a small proportion or be absent in the test. Moreover, the features tested should be reflexive of the characteristics of the language. Each feature bears equal weight in the marking.

The next two sections will introduce two important distinctions as the criteria to select candidate testing features.

3.2. Distinction between core knowledge and learned knowledge

There are two levels of grammatical knowledge. The first level is what Lado (1961, p. 20) calls "the central core of language" or the implicit knowledge. Another level is learned or explicit. They can be distinguished by the following features:

1. *Frequency:* the core knowledge is frequently used and easily acquired, while the learned knowledge is rarely used and may not be acquired. An ideal native speaker should have all the core knowledge;
2. *Context:* the core knowledge is mostly contextually general, while the learned knowledge is contextually restricted and often literary;
3. *Variation:* there is usually only one way to express the core knowledge, such as the only form for English comparative construction (e.g. "I am taller than him"). If the core knowledge can be expressed with variations (e.g. causative verbs in Northern Chinese), their frequency is similarly high. However, the learned knowledge can often be expressed alternatively through other means with the core knowledge, such as the learned English noun "endeavour" versus the core one "try";
4. *Inherence:* the core knowledge is inherent in the grammar, newly loaned knowledge is learned. However, if the loaning happened long ago and became an inherent part of the grammar

already, it belongs to the core knowledge, such as the numerous English words of French origin;

5. *Training:* according to Crowley & Lynch (1995, p. 10), L1 or native speakers acquire the core grammatical knowledge subconsciously and effortlessly as they grow up and follow them subconsciously and effortlessly as well. However, formal training may be necessary for the learned knowledge, which, in most cases, is obtained through formal education at school;

6. *Closeness:* the core knowledge is closed, hardly acquiring new members, while learned knowledge is open, such as the continuous borrowing of vocabulary from another language and invention of new expressions. The open lexical categories are open due to the learned knowledge.

3.3. Distinction between normal language change and language errors

Another distinction as the feature selection criterion is between normal language change and language errors. Though both phenomena involve the varied uses of the features, features which undergo normal language change should not be included in the test.

Normal language change is the internal development of the language; it follows universal tendencies that occur independently in unrelated languages (Bybee, 2015); however, language error is language specific. Next, normal language change is often featured with simplification with compensation while language error is reduction without compensation elsewhere (Dressler, 1991). For instance, the case systems of Latin and Old English were simplified with compensation of flourished prepositional constructions, more rigid word order and introduction of obligatory articles. Therefore, it is in no sense wrong for language change (Aitchison, 2001). However, Dressler (1991) points out that language errors, without compensation, make the language dysfunctional.

Moreover, grammatical change, including the core lexicon, is slow and steady so that it is still possible to recognize the common ancestor of some languages (McMahon, 1994). In contrast, language error can be abrupt, rapid and varied. Finally, media often reinforces the language change, such as loanwords (see Haspelmath & Tadmor, 2009). However, since language errors disrupt satisfactory communication, the changes by them are erroneous and impossible to be widely accepted by the public. Thus they are discarded by the media.

3.4. The reason of testing implicit grammatical knowledge

The reason of using grammatical knowledge as the testing content is due to its prominent and uncontroversial role in any framework of language competence ever since Lado (1961), such as in Hymes (1972), Canale & Swain (1980), and Bachman (1990). For either L1 or L2, grammatical knowledge determines the ability to communicate precisely and effectively. Practically, even for the rating of the integrative language proficiency tests (see Section 1), grammatical accuracy stands out from many considerations (e.g. comprehension and intelligibility) in marking (see McNamara, 1990).

The purpose of the present test is for language maintenance. Its target population are all normal speakers of a speech community, educated or not. The implicit grammatical knowledge is supposed to be learned by all native speakers; however, the explicit or learned grammatical knowledge is not a necessary component of native competence.

3.5. Testability of core competence

Theoretically, Taylor (1988) interprets Chomskys concept of competence as a static mental state with an absolute quality so that it cannot be compared. The interpretation is true in the sense that Chomsky (1965) proposes the concept on the basis of 1) ideal monolingual speakers, and 2) homogeneous speech community. Since the ideal speakers know the language perfectly, their core competence should be constant; no difference can be detected. Practically, to many, test of native core competence seemed unnecessary (Lado, 1961; Valdés & Figueroa, 1994).

Therefore, the core competence of L1 speakers under stable sociolinguistic environment without language loss or shift should be distributed as in Figure 1a. The box represents the core competence and the vertical line stands for the competence level. The ceiling represents the theoretically constant core competence that ideal monolingual speakers should reach. The solid dots stand for the distribution of the actual core competence of L1 speakers. Even though it is hard for them to reach the ideal level, their core competence should be close to the ideal level and similar to each other; the difference from one another should be hard to tell. This is also supported by Ellis (1994) who holds that adult native speakers competence is stable, but L2 learners are not.

Therefore, the distribution of L2 competence should be more widely scattered as is represented

in Figure 1b. It is easier to divide L2 competence into "low", "intermediate" and "advanced" or more detailed classifications if possible. For L2 learners, all the core knowledge of the target language is explicit.

In order to reach similar distinguishability, current L1 tests aim beyond core competence and address the learned or explicit knowledge or even further. Take verbal reasoning from paper-based Graduate Record Examinations (GRE) General Test for example. It measures the ability of analyzing and evaluating written material, synthesizing information, and understanding the relationships among words and concepts. The learned competence of L1 speakers should be similarly varied like the competence distribution of L2 learners (cf. Figure 1b). However, nowadays, due

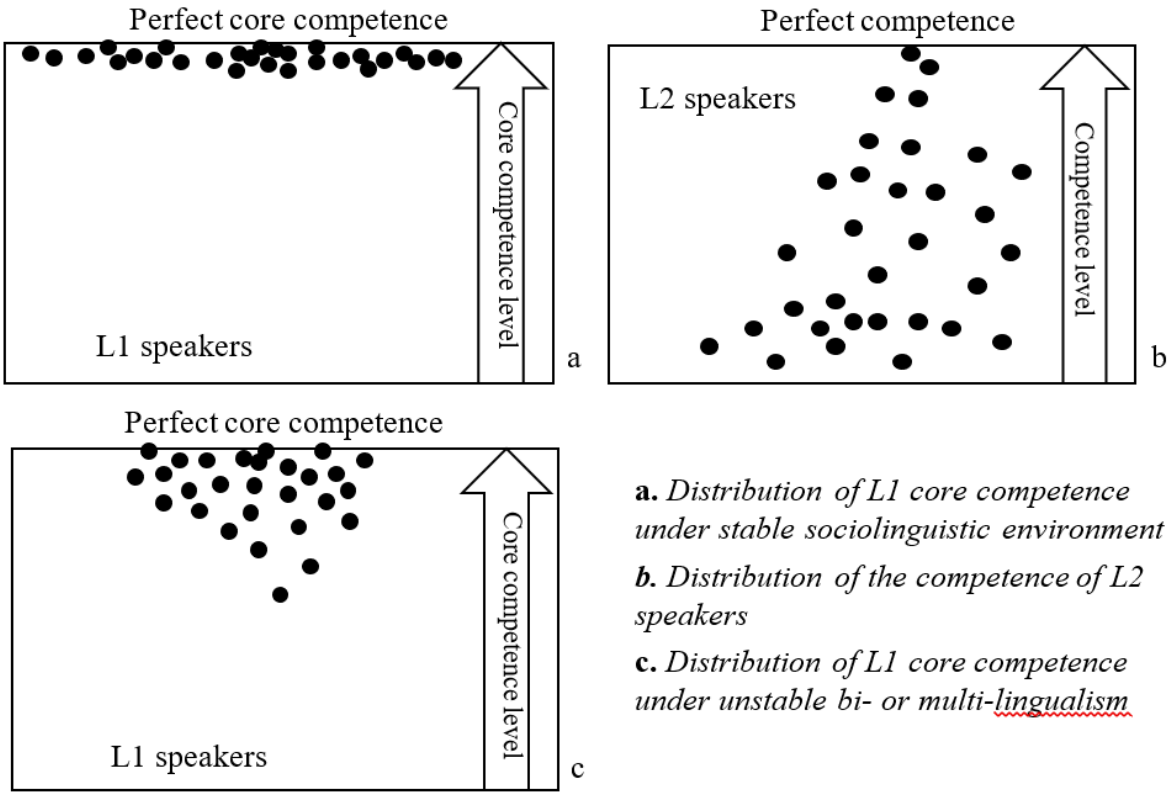


Figure 1: Comparing the distribution of competence

to the learning of the culturally dominant language for better education and career opportunities, the stability of ethnic L1 competence is constantly challenged and even impaired by language shift and language loss. Consequently, the competence of ethnic L1 is becoming less stable and more varied than that in monolingual communities, as is represented in Figure 1c. These communities are often faced with unstable multilingualism. Though the competence variation should not be as

large as among L2 learners of the same language (cf. Figure 1b), it is larger than that of the L1 speakers under stable sociolinguistic variables (cf. Figure 1a). Due to the effect of language shift and language loss, the new state of instability and variation makes the core competence of ethnic L1 distinguishable and testable.

3.6. Solution to the dialectal differences

The variety tested may be identical to the one spoken by the test takers. But the two varieties may also be different, although they can be two closely related dialects of the same language. Therefore, the solution is based on the similarities of the dialects, namely only the shared features between the testing variety and the variety used by the test takers are appropriate candidate testing features; otherwise, the dialectal speakers will be placed in disadvantage due to the dialectal differences.

The dialectal differences can be phonological, lexical and structural. If the mutual intelligibility of the two varieties is low, a separate test should be used for the competence of the target language with the variety of the test takers. For example, in the indigenous language tests in Taiwan for the purpose of revitalization, the test takers can choose which dialect, namely Truku, Toda or Tgdaya (Tsukida, 2005), they prefer to be tested with as a speaker of Seediq, an Austronesian language.

4. The question types and the testing format

So far, I have answered the what-question and provided the justification. In this part, I will continue to answer the question of "how to test the grammatical competence".

4.1. Question types

Discrete-point approach is used in the test. This approach measures one point of the grammar at a time and makes each item independent of the other. It will also make the marking of the test results clearer than integrative approach.

Since the answers need to be controlled, selected-response and limited-production tasks are used in the test. According to Purpura (2004), selected-response tasks measure the recognition of grammatical form and/or meaning, such as multiple-choice question. Limited-production tasks

differ in that the expected answer involves a limited amount of language production, such as gap-filling question. A crucial advantage of the two kinds of task is that they are objective; they, according to Purpura (2004), do not require expert judgment in terms of the marking criteria. Evaluation of extended production tasks requires expert judgment which are often subjective and hard to be controlled. Possible question formats can be: multiple-choice question as selected-response task, and true/false question with correction and sentence-making question as limited-production tasks (see Section 5 for illustration).

4.2. Literacy issue and change of the testing format

About half of the worlds living languages do not have its writing system Lewis et al. (2015). The writing systems of some languages have only historically limited usage, such as the use of the Nuosu writing system mainly for religious purpose. Ordinary speakers are almost illiterate in their ethnic first language. They cannot be tested in the written format.

Literacy ability is independent of ones grammatical competence. It is impossible to refer to ones illiterate monolingual grandparents as non-native speakers of their first language. In order to accommodate languages without writing systems and illiterate speakers, the test should be undertaken in an audible-and-spoken format. It can be realized through a computerized testing interface built with HTML and embedded with a recording module (e.g. Nanogong 5.0) for answering questions (e.g. corrections in the T/F questions).

4.3. Sustainability

Each test can only cover part of the grammatical knowledge. Exhaustive testing is not feasible to both test developers and test takers. What is more important than being exhaustive is the sustainability of the test. There are two ways to achieve sustainability. Firstly, since one test covers a non-exhaustive selection of features, other features can be selected in another test. Secondly, if the same feature is tested again, variation of the input can be realized to test the used features.

5. Testing interface and sample questions

I will illustrate the sample questions with features from Nuosu, a Tibeto-Burman language spoken by over 2 million Nuosu people in Liangshan, Southwest China, which is under the threat of Southwest Mandarin and Standard Chinese as the culturally dominant varieties. I will also integrate the questions into the testing interface. Linguistic description of Nuosu can be found in Ding (2016) and Gerner (2013).

Nuosu has a syllabary writing system. Besides the characters, there is another Romanized phonetic form, called Nuosu pinyin. They are used in the Nuosu-Chinese bilingual education and on the internet. Although the Nuosu writing system reputedly dates back to the 7th century (Wu, 1991), it was historically limited to be used for religious purpose. Ordinary Nuosu speakers, including most Nuosu students in the Nuosu-Chinese bilingual education program, have not mastered it (Ding, 2016). On the testing interface, Nuosu Romanization and Nuosu characters are placed along the questions as visual aids.

5.1. Attitudinal clitics with multiple choice question

Attitudinal clitics of Nuosu occur at the end of the sentence. It belongs to the domain of pragmatics. They do not change the meaning of the utterance, but denote attitudinal information of the speaker, such as the requestive [ha^{33}] to make a polite request or order, the suggestive clitic [ma^{21}] to soften the abruptness of the utterance, the stimulative clitic [lo^{21}] for encouragement or stimulation, and the admonitive clitic [ma^{55}] to signal the warning. The attitudinal clitics attach to the clause, not the verb.

(1) a.

vi^{55}*ga*^{33}	*ga*^{55}=*ta*^{33}	*ts*^{h}*ɯ*^{21}*ʑ*^{33}	*ɣɯ*^{21}	*la*^{33}	*[-----]*
clothes	wear=DUR	flu	get	come	[-----]

'Put the clothes on. Don't catch cold!'

b. ha^{33} lo^{21} ma^{21} ma^{55} (answer)

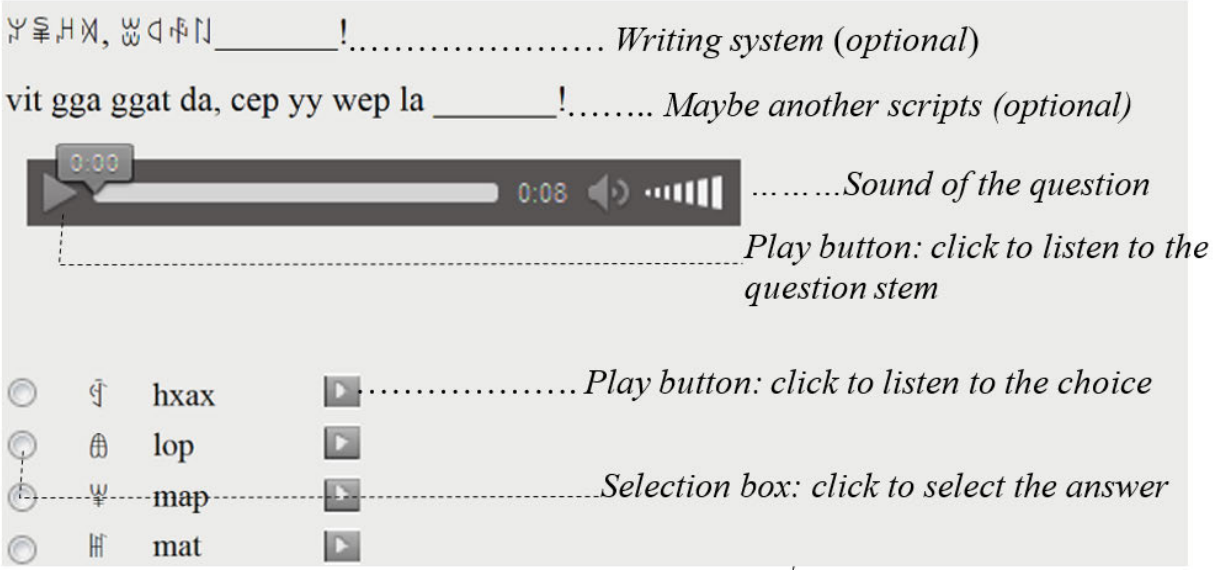


Figure 2: The testing interface of MC question

5.2. *Simplex-causative verb with T/F question*

Nuosu, represented by the Shynra dialect, has a few simplex-causative verb pairs which differ in the voicing of the onset. They are (near) minimal pairs. The simplex variants of such pairs use voiced initial while the causative variants use voiceless onset, such as [gɯ^{33}] listen and [kɯ^{33}] cause to listen, and [ʥɯ^{33}] eat and [tʂa^{33}] feed. It belongs to the morphological domain. Simplex-causative verb is tested with T/F question which requires grammaticality judgment and correction. The mistake is underlined.

(2) *a*^{34}*p*^{h}*u*^{33} *la*^{55}*zɯ*^{33} *tɕo*^{34} *hi*^{21} *ko*^{33} *la*^{55}*zɯ*^{33} *thi*^{55} *la*^{33} *nɯ*^{33}
grandpa Latsse to say COMP Latsse here come 2SG

a^{21}*di*^{21}*hi*^{34} *ŋa*^{33} *si*^{21} *tɕi*^{55} *su*^{33}*ŋo*^{34} *i*^{21}*ɳi*^{21} *=nɯ*^{33} *a*^{34}*p*^{h}*u*^{33} *sʐ̩*^{21}*nda*^{55}
yesterday 1SG take slap since today=FRM grandpa pear

ma^{33} *ka*^{33} *nɯ*^{34} *dzɯ*^{33} *mo*^{33} *di*^{34}
CLF ResV 2SG eat ATT QUOT

Intended meaning: "the grandfather said to Latsse: "Latsse, come here. Since I slapped you yesterday, (as an apology), today (let) grandpa take a pear and feed you."

[dzɯ^{33}] eat should be changed to [tʂa^{33}] feed or make eat. The test takers can record their

answer. They can listen to their response; if they do not feel satisfied, they can rewrite it by recording it again.

Figure 3: The testing interface of T/F question

5.3. Focus-presupposition construction with sentence-making question

Nuosu uses a periphrastic construction, which is a topic-comment articulation, to introduce some constituents as the focus of a sentence. It belongs to the syntactic domain. The construction involves the copula and a nominalized complement: (NP) + (topic clitic) + Focus + -[su^{33}] + [ŋɯ^{33}].

The copula is [ŋɯ^{33}] and the nominalizer is [su^{33}]. Though the construction is similar to the English cleft sentences or Mandarin Chinese pseudo-cleft sentences in terms of discourse function, the single clause is not divided into two parts. Thus the term "cleft" is not adopted. The constituents of the sentence are shuffled and re-arranged in alphabetic order. The test takers are required to assemble them in correct Nuosu sentence.

(3) *tɕ^{h}u^{33}ʂʐ^{33}* *a^{34}ŋi^{33}=mu^{33}* *tshʐ^{33}* *ʑu^{33}=si^{34}* *p^{33}* *i^{34}ko^{33}* *la^{34}*
silver.gold many=ADVL 3SG grab=Res return home come

t^{h}ɯ^{33}=ko^{33}=nɯ^{33} *vu̩55vu^{33}* *ko^{33}* *tɕo^{34}* *n̥a^{33}* *la^{34}=o^{34}* *ts^{h}ʐ^{21}* *tɕo^{34}*
time=C.TOP=FRM elder.brother him to ask come=PFV 3SG to

tɕhu^{33}ʂʐ^{33} _____ *di^{34}*
silver.gold _____ QUOT

"When he took the silver and gold and returned home, the elder brother came to ask him, (he said) to him: _____ the silver and gold?"

The shuffled constituents are given as:

(4) ta^{33} $k^{h}a^{55}$ la^{33} $ŋɯ^{33}$ si^{21} su^{33}
from where come COP take NMLZ

The answer is:

(5) $k^{h}a^{55}$ ta^{33} si^{21} la^{33} su^{33} $ŋɯ^{33}$
where from take come NMLZ COP

'Where do (you) take the (silver and gold) from?'

The topic or the presupposed information, namely [tɕhu^{33}ʂʐ̩^{33}] silver and gold, is already given for the sentence. The test takers should finish assembling the predicate containing the focus domain. Since they may not know how to type the Nuosu characters or the Romanized scripts, they only need to fill in the number to indicate the order.

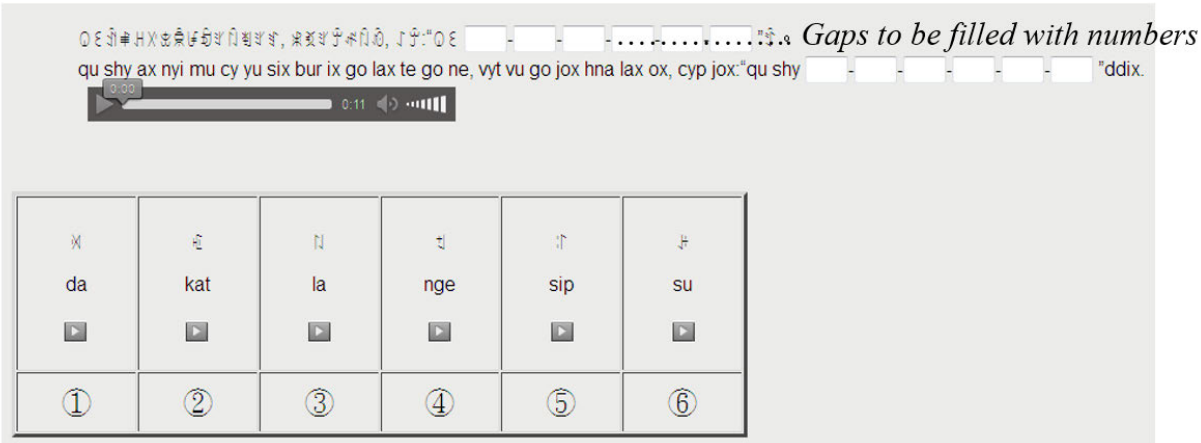


Figure 4: The testing interface of sentence-making question

6. Application of the method: Results and discussion

I have applied a Nuosu test and a Chinese test with similar design on the basis of the proposed method to three generations of Nuosu speakers during a fieldwork in Liangshan. The average age, dialectal background and the number of the participants for each generation are summarized in Table 3. The total number of male participants was 22 and of female participants was 12. The gender imbalance is actually a reflection of the patriarchal society of ethnic Nuosu. In the field,

the females thought that the males should do talking to the strangers. Therefore, they felt rather reluctant to participate. Since the participants were taken care of one after another in the tests, my sample size was not large.

Age label	**Avg. age**	**No.**	**L1**	**Nuosu dialect**		**L2**
Young group	20	12	Nuosu and SWM & SC	Shynra	7	English
				Yynuo	5	
Middle-aged group	40.6	11	mostly Nuosu only	Shynra	10	mostly SWM & SC
				Yynuo	1	
Elder group	66.3	11	Nuosu	Shynra	9	SWM & SC
				Yynuo	2	
Total		34		Shynra	26	
				Yynuo	8	

Table 3: The Nuosu participants (SWM: Southwest Mandarin; SC: Standard Chinese)

There are five dialects of Nuosu: Shynra, Suondi, Adur, Yynuo and Lindimu (or Tianba) (Bradley, 1997; Gerner, 2013). The participants were speakers of Shynra and Yynuo. These two dialects are highly mutually intelligible (Bradley, 2001) and their speakers account for over 70% of the whole population of ethnic Nuosu (Gerner, 2013). The young generation still actively learned Nuosu as their ethnic L1. Except those who have completed shifted to Chinese, such as urban residents in the families of Nuosu cadres, most of the Nuosu young generation, who learned Nuosu as L1, can speak the language, by observation, as fluently as the elder and middle-aged generations without perceptible difficulty. Meanwhile, most of them learned the culturally dominant language, Mandarin Chinese, specifically Southwest Mandarin and Standard Chinese as two dialects, as another L1. However, nearly all of the elder generation and most of the middle-aged generation learned Mandarin Chinese as L2.

The variety used for the Nuosu test was Shynra Nuosu, and the one for the Chinese test was Standard Chinese, which were the varieties the participants were most familiar with under testing scenario. All the subjects participated in the Nuosu test and the young generation took the Chinese test as well. Eight out of the eleven middle-aged Nuosu speakers also took the Chinese test. Only one elder speaker had the confidence to take the Chinese test; the others gave up due to their lack of confidence in their Chinese competence. Additionally, I have invited a control group of 10 young Han Chinese speakers, whose average age was 20.3, to take the Chinese test; they spoke, as well,

Southwest Mandarin and Standard Chinese as L1, with knowledge in English as a foreign language. Han Chinese is the only majority ethnic group in China. The other 55 ethnic groups in China are all officially recognized as minority groups, such as Tibetans, Uyghurs and Mongolians.

Meanwhile, I have also conducted a 40-word Nuosu vocabulary test to the Nuosu participants in order to make a methodological comparison. The 40 words were selected from the 100-word Swadesh list (Swadesh, 1952). They covered words of number, color, body parts, objects, animal, and basic verbs. All are basic so as to measure the core knowledge of the speakers. The subjects were shown the pictures of the words and asked to identify them in the target language.

By the time of the fieldwork, I have not figured out how to treat phonological competence due to the dialectal differences between Shynra and Yynuo. Therefore, the productive competence in both Nuosu and Chinese phonology was not tested. It is admitted that it should be included in future research. But the following results can already prove the feasibility of the proposed prototype competence test. Moreover, since Standard Chinese lacks inflectional morphology, there was no question about Chinese morphology. The content of the Nuosu test and Chinese test are summarized in Appendix I and Appendix II.

I also want to mention that Question C8 and C14 (see Appendix II) in the Chinese test were intentionally designed to detect dialectal influence from Southwest Mandarin upon Standard Chinese in the bigger project. But the results were re-calculated in the present research. Even though some subjects responded to the two questions with their dialectal knowledge, the answers were accepted as long as they are correct in either Southwest Mandarin or Standard Chinese.

The two tests have different testing content and the internal ratios among the linguistic domains are also different. Since they are different languages with its unique features, it is not necessary to strike an identical internal ratio. The two tests will be independently scored; the absolute scores will be converted to comparable relative percentage, namely correctness rate. The rule to score is to assign each feature with the same weight. The marking scheme is binary: one point as the full mark for a discrete item and zero as non-mastery of a certain feature.

6.1. Instability as the evidence of the gradual decline of Nuosu competence

According to the competence distribution in Figure 5, all the elder and middle-aged Nuosu speakers have reached the correctness rate of 80%. The elder group have stably maintained their performance as high as above or close to 90%. Meanwhile, the Chinese competence of the Chinese control group was also stale, i.e. at or above 85% (see Figure 6).

Regarding the Nuosu competence, the average performance was 93.73% for the elder group, 92.16% for the middle-aged group and 78.74% for the young group. According to the independent samples t-test in SPSS, there was no significant difference (p = .408 > .05) between the performance of the middle-aged group and of the elder group. However, only half of the young speakers reached over 80%. The performance of the young group was significantly lower than that of the middle-aged group (p = .005 ¡ .05) and the elder group (p= .002 ¡ .05). This is also the evidence that the stability of Nuosu competence decreases gradually throughout the three generations. The highest and the lowest correctness rates have been marked in Figure 5 and 6.

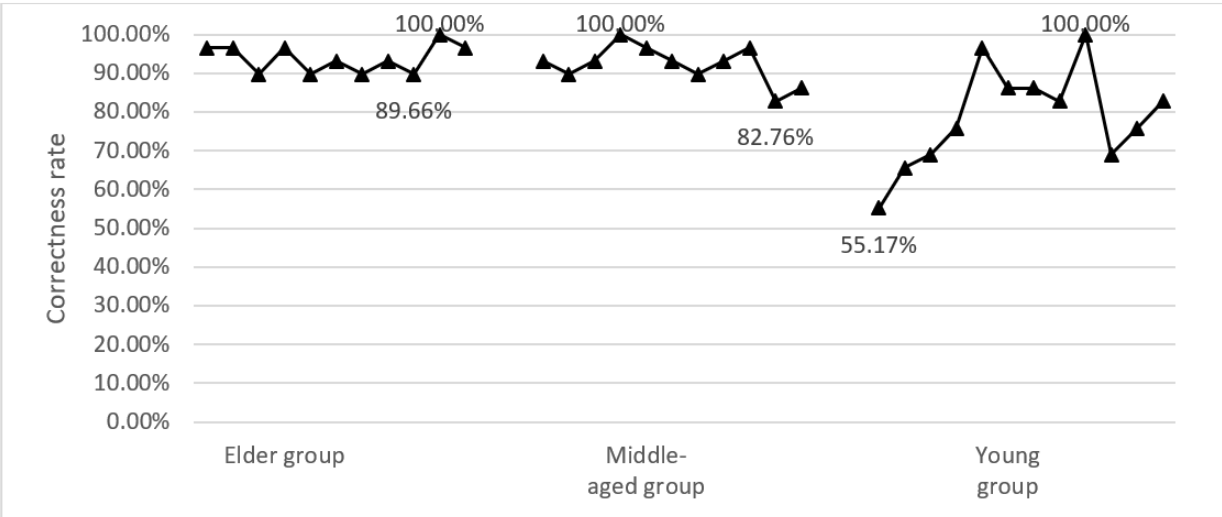


Figure 5: Competence in Nuosu across the three generations

Take the range of difference as the reference, namely highest performance minus lowest performance within a group. The range of difference was 10.34% (i.e. subtracting 100% from 89.66%) for the elder group, 17.24% (i.e. subtracting 100% from 82.76%) for the middle-aged group and 44.83% (i.e. subtracting 100% from 55.17%) for the young group. The larger the range of difference is, the less stable the core competence is.

Secondly, the speed of language loss accelerates more rapidly in the young generation. The reminder from subtracting the range difference was used to indicate the rate of increasing instability. It was found that stability decreased slightly from the elder group to the middle-aged group, namely less than 7% (i.e. subtracting 17.24% from 10.34%), but drastically to the young group, namely 34.49% (i.e. subtracting 44.83% from 10.34%).

Thirdly, some Nuosu young speakers were able to keep similarly high competence like the elder generations while the others were not. This suggests that there is individual variation regarding the paces of adaptation to the changing sociolinguistic environment. Regarding the Chinese competence, although it was more varied among the Nuosu young group, independent samples t-test, namely $p = .365 > .05$, found no significant difference between their Chinese competence and that of the Chinese control group. The average performance was 91.5% for the Chinese control group, 88.75% for the Nuosu young group and 82.5% for the Nuosu middle-aged group. Eight subjects from both the Chinese control group and the Nuosu young group (see Figure 6) reached the correctness rate of 90% in the Chinese test. The Chinese competence of the middle-aged group was more varied than that of the Nuosu young group. The only elder Nuosu speaker who took the Chinese test scored 60%, while his Nuosu competence was 96.55%. His performance confirmed that the elder Nuosu speakers were Nuosu-dominant.

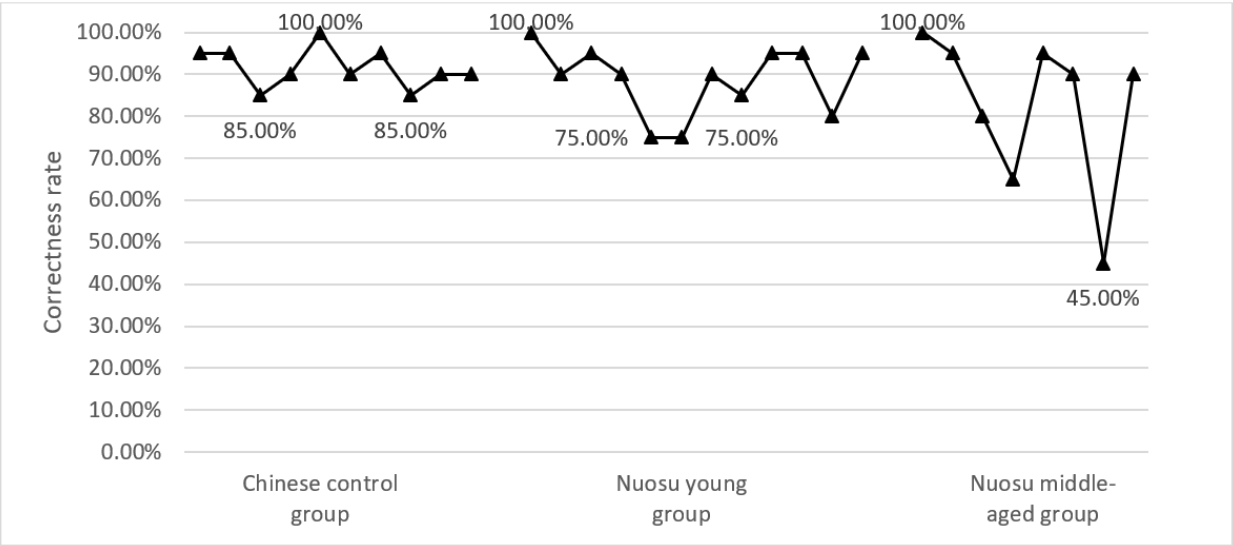


Figure 6: Competence in Chinese across the three groups

6.2. Vocabulary test

However, according to the results of the vocabulary test, all the three Nuosu groups had near-ideal performance, namely over 99% (see Figure 7).

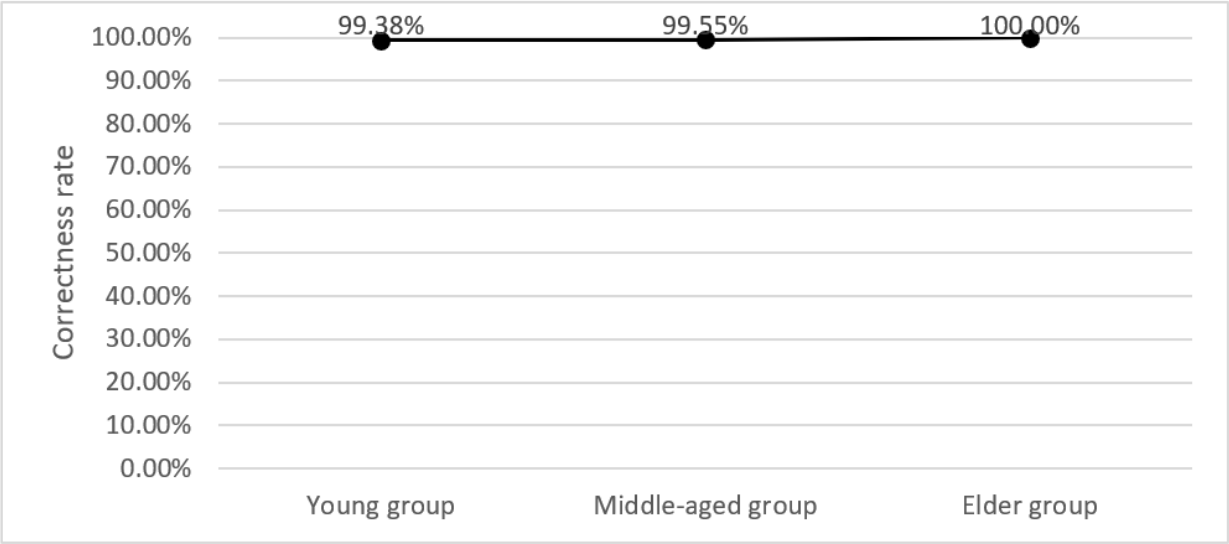


Figure 7: The average competence by Nuosu vocabulary test

This confirms the limited distinguishing ability of vocabulary test. Firstly, those with lower competence by grammatical test also achieved perfect competence by vocabulary test. For example, Subject M10, Y2 and Y10 (see Figure 8 and 9), who had relatively lower Nuosu competence through the grammatical test, also reached 100% through Nuosu vocabulary test. The results by the vocabulary test hide the lack of competence.

However, it is unfair to abandon the use of vocabulary test. Vocabulary test is useful to make distinctions when language shift is almost completed and the competence of the target language is already low. The lowest performance by Subject Y1 (see Figure 9) in Nuosu grammatical test (i.e. 55.17%) and vocabulary test (i.e. 92.5%) was not a coincidence.

This is also the reason why Dai et al. (2011) were able to use vocabulary test to make a 4-level distinction among the Jingpho people whose ethnic L1 is either Jingpho proper or Zaiwa, two Tibeto-Burman languages. Their subjects were urban residents of ethnic Jingpho. Their language shift from the minority language to Chinese was much more substantial and widespread than among the Nuosu people.

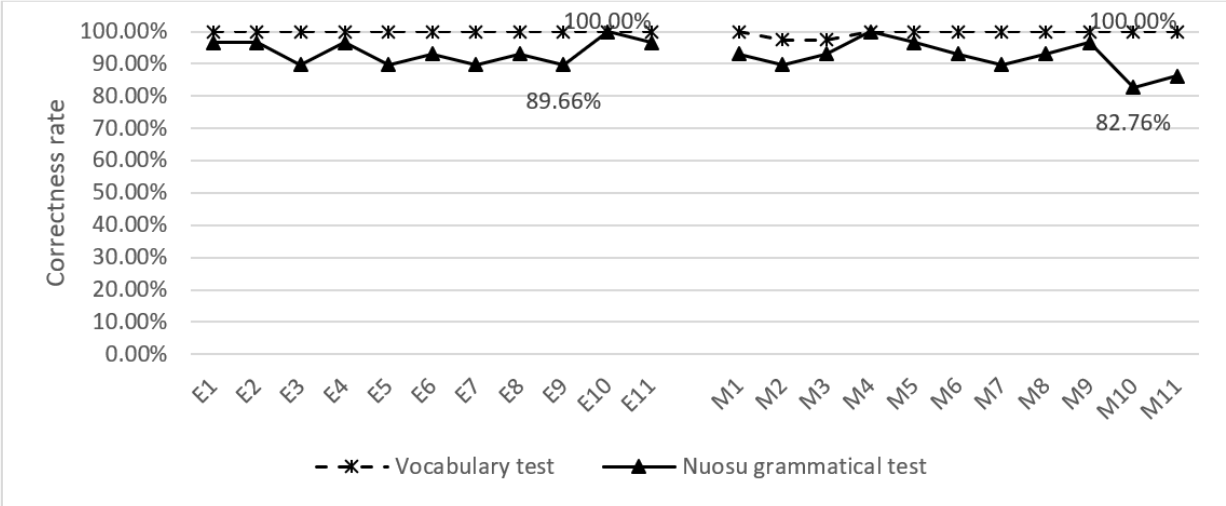


Figure 8: Results through the two methods among elder and middle-aged groups (E: Elder speaker; M: Middle-aged speaker)

6.3. Bilingual competence: Language shift among the Nuosu young generation

The two tests have been co-normed to make further comparison. Since the correctness rates are relative percentages instead of absolute scores, they are comparable between different languages. The column "Chinese better" in Table 4 and 5 indicates the results by subtracting the Nuosu competence from the Chinese competence. The symbol "+" means that the Chinese competence is better than the Nuosu competence and the symbol "-" means vice versa. The last column "difference value" is the subtracted results of the competence of the two languages. Three middle-aged subjects did not participate in the Chinese test, due to their lack of confidence in their Chinese competence.

To determine whether the participants are balanced bilingual speakers or not, confidence interval or CI is calculated to estimate the measurement errors of the criterion-referenced test (Brown, 1996):

$$CI = \sqrt{\frac{M_p(1 - M_p) - S_p^2}{k - 1}}$$

M_p = mean of the proportion scores

S_p = standard deviation of the proportion scores

k = number of test items

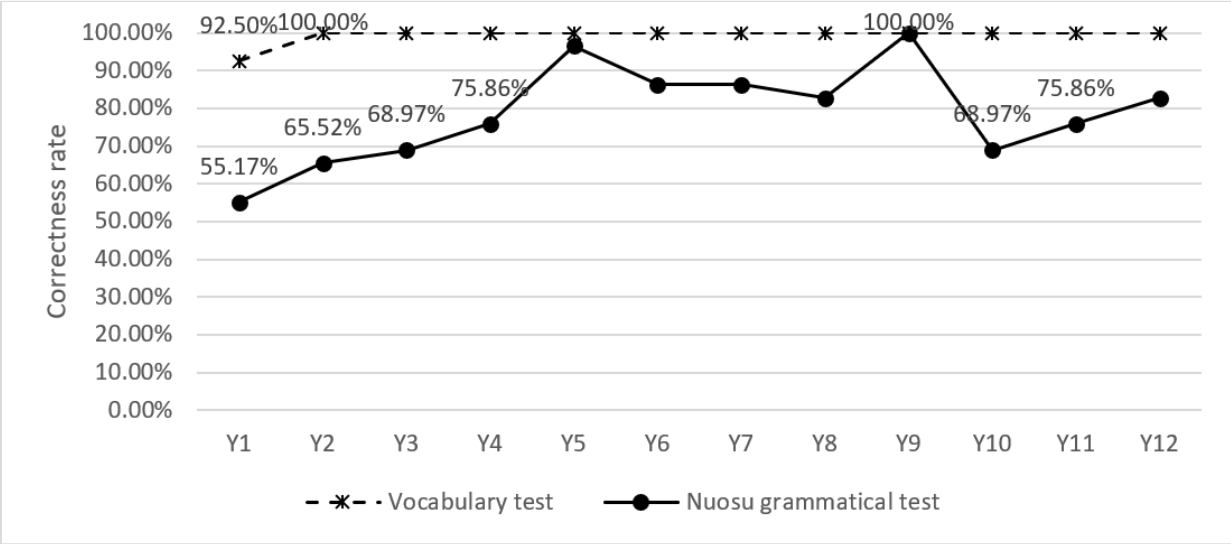


Figure 9: Results through the two methods in the Nuosu young group (Y: young speaker)

Taking the confidence level 68%, the estimated measurement error in the Nuosu test is 5.82%[1]. The figure represents a confidence interval, namely plus or minus 5.82%, around the examinees score within which the test users feel confident that the examinees score will fall 68% of the time if the test is administered to this person many times. For instance, if the correctness rate of the Nuosu test by a test taker is 82.76%, he or she will score as low as 76.94% (i.e. 82.76% - 5.82%) or as high as 88.58% (i.e. 82.76% + 5.82%) if the test is administered to him or her again. The score will fall within this band, namely between 76.94% and 88.58%, 68% of the time if the test is taken repeatedly. Applying the similar calculation to the Chinese test, the estimated measurement error is 7.17%[2].

It was found that most of the Nuosu middle-aged speakers were still more competent in Nuosu than Chinese while most of the Nuosu young speakers were more competent in Chinese than Nuosu (cf. Table 4 and 5).

[1]For the Nuosu test, M_p is 87.93% from all the 34 subjects in the three generational groups. S_p is 10.59%; Brown (1996) specifies that standard deviation of proportion scores should use the N formula, rather than $N-1$. K is 29; though there are 23 test questions of core knowledge, they involve 29 test items; each score is counted as one test item.

[2]For the Chinese test, M_p is 87.10% from all the 31 subjects in the three generational groups of ethnic Nuosu and the Chinese control group. S_p is 12.10% (using the N formula). And k is 20; though there are 18 test questions of core knowledge, they involve 20 test items.

	Nuosu competence	**Chinese competence**	**Chinese better**	**Difference value**
Y1	55.17%	100.00%	+	44.83%
Y2	65.52%	90.00%	+	24.48%
Y3	68.97%	95.00%	+	26.03%
Y4	75.86%	90.00%	+	14.14%
Y5	96.55%	75.00%	-	21.55%
Y6	86.21%	75.00%	-	11.21%
Y7	86.21%	90.00%	+	3.79%
Y8	82.76%	85.00%	+	2.24%
Y9	100.00%	95.00%	-	5.00%
Y10	68.97%	95.00%	+	26.03%
Y11	75.86%	80.00%	+	4.14%
Y12	82.76%	95.00%	+	12.24%

Table 4: Bilingual performance of the Nuosu young speakers (Y: young speaker)

	Nuosu competence	**Chinese competence**	**Chinese better**	**Difference value**
M1	93.1%	n.a.	-	n.a.
M2	89.66%	100.00%	+	10.34%
M3	93.1%	95.00%	+	1.9%
M4	100.00%	n.a.	-	n.a.
M5	96.55%	n.a.	-	n.a.
M6	93.1%	80.00%	-	13.1%
M7	89.66%	65.00%	-	24.66%
M8	93.1%	95.00%	+	1.9%
M9	96.55%	90.00%	-	6.55%
M10	82.76%	45.00%	-	37.76%
M11	86.21%	90.00%	+	3.79%

Table 5: Bilingual performance of the Nuosu middle-aged speakers (M: middle-aged speaker)

Without the measurement errors, the competence gap between the two languages of the test takers may be narrowed. Therefore, with the results of the confidence intervals, the subjects whose difference value is 12.99% (i.e. 5.82% + 7.17%) are considered as balanced bilinguals. The current use of the term "balanced bilingual" is within the literal interpretation since it is possible that the balance exists at a low level of competence (Müller & Pillunat, 2008). Whether it fits the

implicit idea of balanced bilingual who is supposed to possess native level of competence in the two languages (e.g., Bloomfield, 1933) have been investigated in another study of Ding (2016).

Moreover, people whose difference value is >12.99% are categorized into two types. The first type is those who still use Nuosu as the dominant language; their difference value is negative. The second type is those who are experiencing language shift; their difference value is positive and Chinese is their dominant language. All those who gave up the Chinese test were considered as Nuosu-dominant speakers. See Table 6 for a summary.

Typology of bilingualism	**Young group**	**Middle-aged group**	**Elder group**
Nuosu-dominant	1	6	11
Chinese-dominant	5	0	0
Balanced	6	5	0

Table 6: Summary of the subjects with different bilingual status

Most of the middle-aged speakers were still Nuosu-dominant. But balanced Nuosu-Chinese middle-aged bilinguals already existed. They served as a transitional generation of language shift from Nuosu to Chinese among the three generations. Only one Nuosu young subjects were Nuosu-dominant. Half of the young subjects became balanced Nuosu-Chinese bilinguals. There were five young subjects experiencing language shift and became Chinese-dominant. The language shift of some was more drastic than the others (cf. the column of "difference value" in Table 4).

In sum, despite the fact that the change of competence in Nuosu is still imperceptible and can hardly be detected by observation among the young active L1 speakers of Nuosu, the prototype test has successfully identified the competence decline among the Nuosu young generation. Although most of the young Nuosu speakers are still fluent in the minority language, as is indicated by Baker & Jones (1998), simple communication skills may hide someones lack of competence or inexperience in a language. What should be highlighted is that Subject Y1 (see Table 4) with the lowest Nuosu competence (i.e. 55.17%) still learned Nuosu as L1 and spoke it fluently. Simply by observation, I could not tell any difficulty of his use of the language.

It is admitted that competence decline among the young generation of some minority languages may not be new to many field linguists. However, what matters is how the phenomenon is found, namely through measurement with full neutrality, not delayed perception by observation. Even for observable competence decline, this methodology can serve as an objective instrument, unlike

subjective report or qualitative observation, to provide a quantitative linguistic confirmation.

Finally, what is mentionable is that the eight Yynuo speakers were not placed in disadvantage even though the testing variety was Shynra Nuosu. Their average performance (i.e. 88.36%) in Nuosu was slightly higher than that of the Shynra speakers (i.e. 87.8%). More comparison based on dialectal differences were conducted in another study (see Ding, 2016).

7. A general research agenda

The first step for such a research is to describe and understand the target testing varieties and the related dialects. After the shared candidate features are selected, the investigator should conduct a few trial tests to establish validity and reliability.

If necessary, a short background questionnaire can be integrated on the computerized interface to obtain personal and language information. Then a training session should be conducted before the formal test to familiarize the test takers with the characteristics of the test tasks, probably with sample questions. In order that the examinees can perform at their best, there should be no time limit for them to finish the test. The ideal testing time for one language is 20 to 30 minutes.

Appendix I: The content of the Nuosu test

Order	Features tested	Characteristics	Domain tested
		MC questions with one best answer	
N1	verbs of animal call	with strict collocation	semantics
N2	sibling terms	distinguished by the relative age, the addressees gender, and the speakers gender	semantics
N3	logophor	as source-tracing pronoun in reported speech	morphology
N4	NP structure	head noun + Adj. + DEM + NUM + CLF	syntax
N5	derivational suffixes	to derive nouns or nominals from verbs	morphology
N6	attitudinal clitics	to denote attitudinal information, such as the requestive, the assentive and the expectative	pragmatics
N7	ideophones	the same monosyllabic root can be compounded with different reduplicated syllables to express different vivid impressions	semantics
		True or false questions	
N8	number-sensitive classifier	form of the general classifier changes if the number is three or more	morphology
N9	simplex-causative verbs	voiced initial for simplex variant and voiceless onset for causative variant	morphology
N10	double object construction	agent + theme + (V1) + recipient + V2	syntax
N11	oblique complement with deictic adverbs of space	with obligatory deictic adverb	syntax
N12	structure with resultative clitic	V + resultative clitic + resultative complement	syntax
N13	resultative construction	V + resultative clitic + resultative complement or V V + comitative + resultative complement	syntax
N14	position of negative clitic	negative clitics occupy the penultimate position in the clitic group	morphology
N15	comparative construction	morphology of the adjectives/descriptive verbs determines the structures for comparison	morphosyntax
		Gap-filling questions	
N16	grammatical tone	tones of the verbs determine the thematic roles of the preceding argument	morphology
N17	aspect	some of the temporal references are encoded through aspectual enclitics	lexicon
N18	alternative question	realized through the copying of the final syllable of the predicate	syntax
N19	conjunction	conjunctions in Nuosu is mainly clause-final	lexicon
		Sentence-making question	
N20	focus-presupposition construction	(NP) + (topic clitic) + focus + -NMLZ + copula	syntax
		MC questions with more than one answer	
N21	verb of borrowing	distinguished by whether the object returned is the original one or not	semantics
N22	Epistemic and deontic modals	distinguished as the abilitive, the skillitive and the permissive	semantics
N23	Causative construction	The structures are determined by the controllability of the causer over the causee, namely make versus let	syntax

Table 7: The content of the Nuosu test

Appendix II: The content of the Chinese test

Order	Features tested	Characteristics	Domain tested
		MC questions with one best answer	
C1	preposition	distinguished by goal and source	semantics
C2	verb-copying	V_i + object + V_i + resultative clitic + resultative clause	syntax
C3	ideophones	the same monosyllabic root can be compounded with different reduplicated syllables to express different vivid impressions	semantics
C4	sentence-final clitics	to express the pragmatic functions of interrogation and attitude	pragmatics
		True or false questions	
C5	verb 'eat' and 'drink'	with strict collocation according to the solidness and fluidness of the food	semantics
C6	idiomatic expression	with strict collocation	semantics
C7	structure of delimitative aspect	patterns determined by the number of the syllables of the predicate	syntax
C8	double object construction	agent + V + recipient + theme	syntax
C9	negation	distinguished as the general negator, the perfective negator and the deontic negator	semantics
C10	epistemic and deontic modals	distinguished as the abilitive, the skillitive and the permissive	semantics
		Gap-filling questions	
C11	durative aspect	to mark ongoing situation as an enclitic	lexicon
C12	V-*de* construction	*dé* appears after the main verb as the modal auxiliary	syntax
C13	pseudo-cleft sentence	copula + focus + -NMLZ	syntax
		Sentence-making question	
C14	adverbial word order	the order is determined by the types of the adverbials, namely the predicational adverbials, functional adverbials and PPs/DPs	syntax
C15	topic-comment construction	NP as topic + (topic clitics) + clause as comment	pragmatics
C16	comparative construction	comparee NP + comparative preposition+ standard NP + verb + resultative clitic + adverb + (degree expression) or comparee NP + verb + resultative clitic + comparative preposition + standard NP + adverb + (degree expression)	syntax
		MC questions with more than one answer	
C17	causative construction	causer + causative verb + causee + main verb	syntax
C18	passive construction	undergoer / patient + prepositional passive marker + (actor / agent) + predicate	syntax

Table 8: The content of the Chinese test

References

Aitchison, J. (2001). *Language change: Progress or decay?* Cambridge University Press.

Altenberg, E. P. (1991). Assessing first language vulnerability to attrition. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition: Structural and theoretical perspectives* (pp. 189–207). Cambridge University Press.

Austin, P. K. & Sallabank, J. (2011). *The Cambridge handbook of endangered languages.* Cambridge University Press.

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford University Press.

Baker, C. & Jones, S. P. (1998). *Encyclopedia of bilingualism and bilingual education.* Multilingual Matters.

Bloomfield, L. (1933). *Language.* University of Chicago Press.

Bradley, D. (1997). Tibeto-burman languages and classification. *Pacific Linguistics. Series A. Occasional Papers*, *86*, 1–72.

Bradley, D. (2001). Counting the family: family group classifiers in Yi (tibeto-burman) languages. *Anthropological Linguistics*, 1–17.

Brown, J. D. (1996). *Testing in Language Programs.* Upper Saddle River, N.J.: Prentice Hall Regents.

Bybee, J. (2015). *Language change.* Cambridge University Press.

Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied linguistics*, *1*(1), 1–47.

Chelliah, S. L. & Willem, J. (2010). *Handbook of descriptive linguistic fieldwork.* Springer.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, Mass: MIT Press.

Crowley, T. & Lynch, J. D. (1995). *The design of language: An introduction to descriptive linguistics.* Longman Paul.

Dai, Q. (2011). *Language use and change of Jingpho at Dehong, Yunnan.* Beijing: The Commercial Press.

De Bot, K., Gommans, P., & Rossing, C. (1991). L1 loss in an L2 environment: Dutch immigrants in France. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition: Structural and theoretical perspectives* (pp. 87–98). Cambridge University Press Cambridge.

Ding, H. (2016). *Testing the competence of first language(s) : a cross-generational study of ethnic Nuosu in Liangshan, Sichuan. (Thesis).* University of Hong Kong.

Dressler, W. (1991). The sociolinguistic and patholinguistic attrition of breton phonology, morphology, and morphonology. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition: Structural and theoretical perspectives* (pp. 87–98). Cambridge University Press Cambridge.

Ellis, R. (1994). *The study of second language acquisition.* Oxford University.

Gerner, M. (2013). *A grammar of Nuosu*, volume 64. Walter de Gruyter.

Haspelmath, M. & Tadmor, U. (2009). *Loanwords in the world's languages: a comparative handbook.* Walter de Gruyter.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). London: Penguin.

Lado, R. (1961). *Language Testing.* New York: McGraw-Hill.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied linguistics*, *18*(2), 141–165.

Lewis, M. P., Simons, G. F., Fennig, C. D., et al. (2015). *Ethnologue: Languages of the world*, volume 16. SIL international Dallas, Texas.

Llurda, E. (2000). On competence, proficiency, and communicative language ability. *International Journal of Applied Linguistics*, *10*(1), 85–96.

McMahon, A. M. S. (1994). *Understanding language change.* Cambridge University Press.

McNamara, T. F. (1990). Item response theory and the validation of an esp test for health professionals. *Language Testing*, *7*(1), 52–76.

Müller, N. & Pillunat, A. (2008). Balanced bilingual children with two weak languages. In G.-F. P., M. P. Larraaga, & J. Clibbens (Eds.), *First language acquisition of morphology and syntax: Perspectives across languages and learners*, volume 45 (pp. 269–294). John Benjamins Publishing.

Paradis, M. & Libben, G. (1987). *The assessment of bilingual aphasia.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Purpura, J. E. (2004). *Assessing grammar.* Cambridge: Cambridge University Press.

Read, J. (2000). *Assessing vocabulary.* Cambridge university press Cambridge.

Stern, H. H. (1983). *Fundamental concepts of language teaching: Historical and interdisciplinary perspectives on applied linguistic research.* Oxford University Press.

Swadesh, M. (1952). Lexico-statistic dating of prehistoric ethnic contacts: with special reference to North American Indians and Eskimos. *Proceedings of the American philosophical society*, *96*(4), 452–463.

Taylor, D. S. (1988). The meaning and use of the term competencein linguistics and applied linguistics. *Applied linguistics*, *9*(2), 148–168.

Tsukida, N. (2005). Seediq. In K. A. Adelaar & N. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 291–325). London: Routledge.

Valdés, G. & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias.* Ablex Publishing.

Wu, Z. (1991). Traditional Yi script. In D. Wang (Ed.), *The written scripts of Chinese ethnic minorities* (pp. 12–33). Beijing: China Tibetology Press.

Swallow alignment

Jonathan Yip[1], Yadong Liu[1], Diana Archangeli[1,2]
University of Hong Kong[1], University of Arizona[2]
yipjonat@hku.hk, yadong@alumni.ubc.ca, dba@email.arizona.edu

1. The Problem

One of the challenges in collecting articulation data using ultrasound is how to address the fact that both the probe and the head are moveable objects, yet quality data requires that the two be aligned with each other in the same way throughout the data collection session. Without some means of ensuring comparable head-probe relations across a recording, data may give the appearance of variation where none exists, and so obscure instances where there truly is variation, such as the difference between bunched and retroflex /ɹ/ (Mielke et al. 2016) or between alveolar and postalveolar stops (Archangeli et al. 2016).

There are two basic ways in which a reliable head-probe relation achieved, either by fixing the relationship throughout the data collection session (Stone & Davis 1995; Whalen et al. 2005; McLeod & Wrench 2008; Derrick et al. 2015; Sigona et al. 2013) or by post-hoc adjustment of the images (Mielke et al. 2005). Fixed-relationship methods require specialized equipment which can be costly, bulky, and/or difficult to adjust, and even the Mielke et al. (2005) system requires specialized equipment for data collection (to mix video camera images with ultrasound images).

There are some research situations where such equipment is difficult or inappropriate to use, for example in the field (Gick 2002; Bird 2011; Whalen & McDonough 2015) as well as with special populations such as children. Beyond linguistic research, the calibration of head position relative to the ultrasonic probe is critical in the areas of clinical speech research and second-language pronunciation training, in which the assessment of individual-talker patterns must be compared against well-established standards of linguistic performance. However, fixed-head methods of head-probe stabilization are not always feasible in a clinical or educational environment. This has led

to research exploring whether certain types of tongue positions are more or less sensitive to the probe/head relation (Ménard et al. 2012; Zharkova et al. 2015).

In our own work, due to field conditions, we often collect ultrasound data with the head and probe loosely fixed using three mechanical arms developed to hold camera equipment. Two are positioned as a head rest for the participant's forehead, and a third holds the probe solidly beneath the chin. Participants are asked to maintain their forehead in contact with the arms throughout data collection. Because head and probe are fixed with respect to each other only so far as the speaker can maintain the fixed head position, there is always potential for some spurious variation to appear.

In this paper, we present a type of post-hoc adjustment in contour data that can be used to improve the quality of data collected without a fixed head/probe relation, using low cost, portable equipment. We call the method "swallow alignment" since it relies on wet and/or dry swallows throughout the ultrasound recording. Swallow alignment described here reduces the degree of spurious variation. (As with Mielke et al. (2005), this system can deal only with translation, not with pitch or roll.)

2. Swallow alignment

To realize post-hoc positional adjustment, the tongue contours have to be anchored to a relatively fixed reference, and contour image adjustment can be made on the basis of the movement of the reference. The palate is considered an excellent reference since it retains its shape during production and its position remains constant in relation to the position of the head. Importantly, too, it is visible in ultrasound imaging: the location of the palate can be identified when a speaker dry-swallows or swallows a water bolus ("wet swallow").

2.1. Method

Speakers were asked to sip a small amount of water at the beginning and end of a data collection session. Video recordings of every speaker were examined carefully to locate both wet and dry swallows. The last frame of every swallow process was considered to be the optimal palate image since the soft palate can be pushed upward during the swallow or by the water bolus; taking the

last frame gives consistency across swallows. For each identified palate frame, a smooth spline was fit to the lower edge of the visual palate contour in the frame using EdgeTrak (Li et al. 2005) and hand-corrected.

Each speaker's palate traces were plotted onto a single Euclidean coordinate space; the aggregated contours were inspected for uniformity in position. See examples in Figure 1. In some cases, palates aligned well with each other regardless of their temporal position in the recording. In other cases, misalignment was visibly obvious, indicating head movements during the recording session. Figure 1 (left) shows variability and Figure 1 (right) shows consistency.

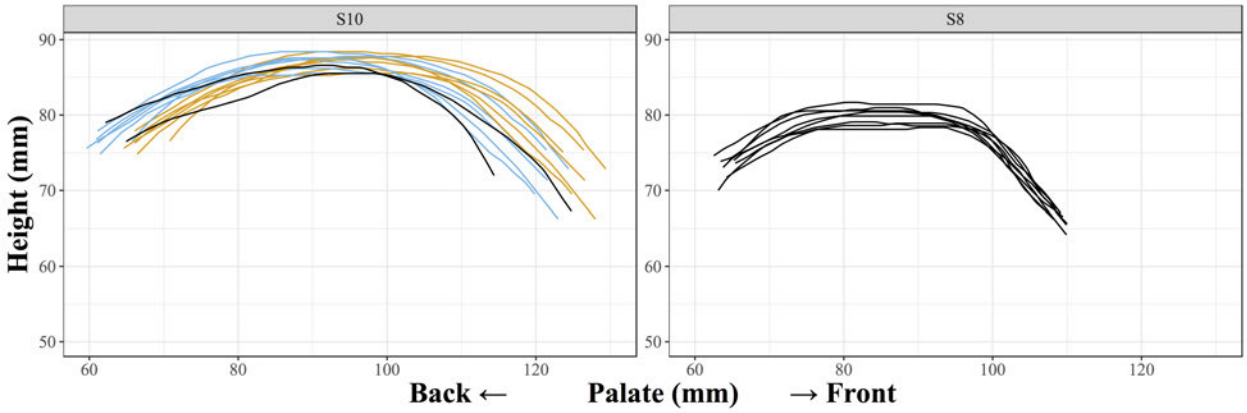


Figure 1: Variable palate position shown with speaker S10 (left); consistent position shown with speaker S8 (right). Variability correlates to changes in the head-probe relation, and adjustments for S10 are shown by contours from before alignment (orange) versus after alignment (blue). Unadjusted palate contours for this speaker are shown by black lines.

Misalignment, such as that illustrated in Figure 1 (left), can be codified as differences in x- and y-coordinates at one or more specific time points during the recording. To correct the misalignment, adjustments are made to the x- and y-coordinates of tongue-trace points occurring after the head movement occurred. These adjustments were achieved by identifying the x/y-coordinate position of each successive palate trace with the smallest sum of distances to corresponding points along the previous palate trace, yielding the optimized readjustment of the palate trace to correct for head movement. We assume that all subsequent frames recorded before the time of the next palate trace are subject to the same extent of movement and so receive the same conversion via translation within the Euclidiean space.

This adjustment improves the uniformity of ultrasound data by reducing the variation in the

spatial positions of data points for the talkers with large variability in palate traces, as illustrated in Figure 2.

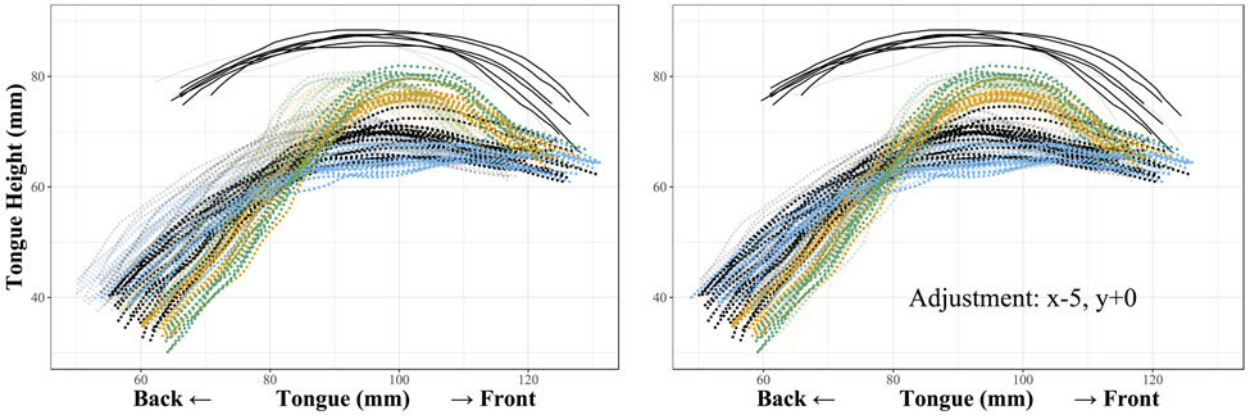


Figure 2: Tongue contours (colored dotted lines) and palate contours (black solid lines) for participant S10 before adjustment (left) and after adjustment (right). Pale traces in both images indicate contours that were not adjusted. Contours of each color are more uniform after adjustment.

2.2. Evidence that it works

Archangeli et al. (2016) uses a root-mean square (RMS) calculation of the distances between compared contours within a polar-coordinate grid, depicted in Figure 3, in order to quantify the similarity between lingual articulations. In principle, other measures of similarity/difference between articulatory contours could be taken, such as the mean-distance calculation used in Zharkova & Hewlett (2009). See Archangeli et al. (2016) for some discussion of the relative merits of different methods.

To examine whether our adjustments, as discussed in Section 2.1, actually resulted in more stabilized data, we compare statistics such as the mean and standard deviation of RMSD values in the adjusted data. If positional adjustments to tongue contours effectively corrected for movements of the head during data collection, then we expect the adjusted data to more closely resemble what they would be had the head not moved at all. This means that tongue contours for identical sounds, e.g. [tɕ] in two iterations of the Sasak word *cemplung* [tɕəmpluŋ] 'dive', should have decreased physical distance measured between them. In terms of data reported in Archangeli et al. (2016), RMSD values between tongue contours obtained from the same sound should be lower as a result of reduced noise cause by head movement in the data.

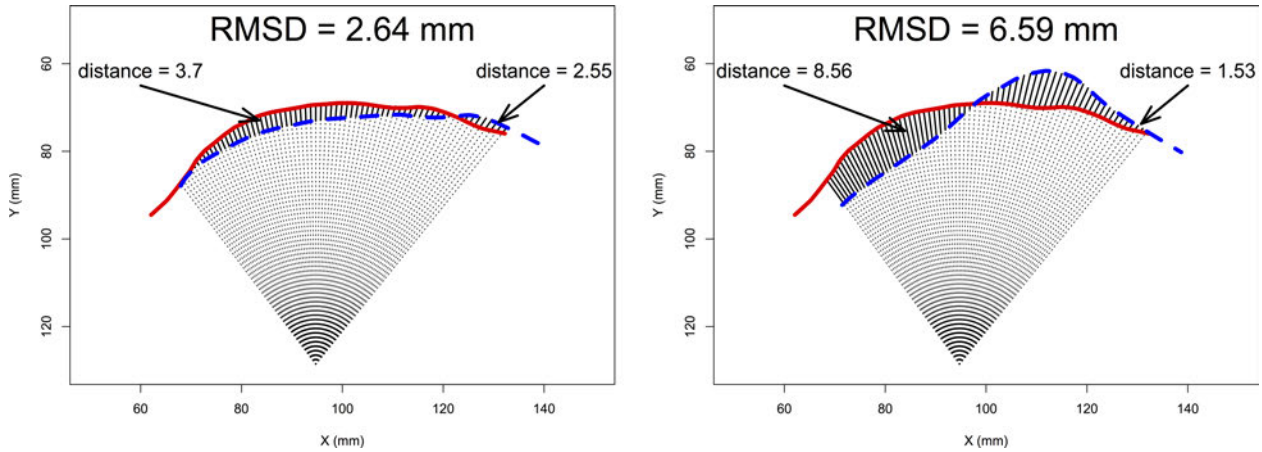


Figure 3: Example illustrations of the RMS calculations of radial distances between pairs of tongue contours sharing the same articulation (left; two iterations of [t]) and contrasting in place (right; solid red=[t], dashed blue=[tɕ]). Individual radial distances are represented by solid lines connecting points from each contour at each angle (in integer degrees).

This is indeed the result, as shown for talkers S9, S10, J6, and J7 in Figure 4 (subjects from Archangeli et al. 2016). For same-sound comparisons for each of these talkers, both mean RMSD and standard deviation values decrease as a result of contour adjustment. Not only do these data become more consistent with our understanding of contour-to-contour similarity and RMSD value (i.e. a lower value indicates greater similarity), but the same-sound RMSD values for such talkers become more tightly distributed, with fewer large RMSD values skewing the data toward the right. This result suggests that our adjustments are successful, with noise due to movement of the head during speech production being, at least partially, eliminated from the RMSD data.

3. Discussion: Potential improvements

Having established that the swallow alignment method of adjusting tongue contour positions introduced in Section 2 is to some extent successful, we now discuss further ways in which this kind of data can be improved.

One straightforward approach to reducing noise in the ultrasonic data due to sporadic movements of the head relative to the ultrasonic probe (transducer) during collection is to simply examine

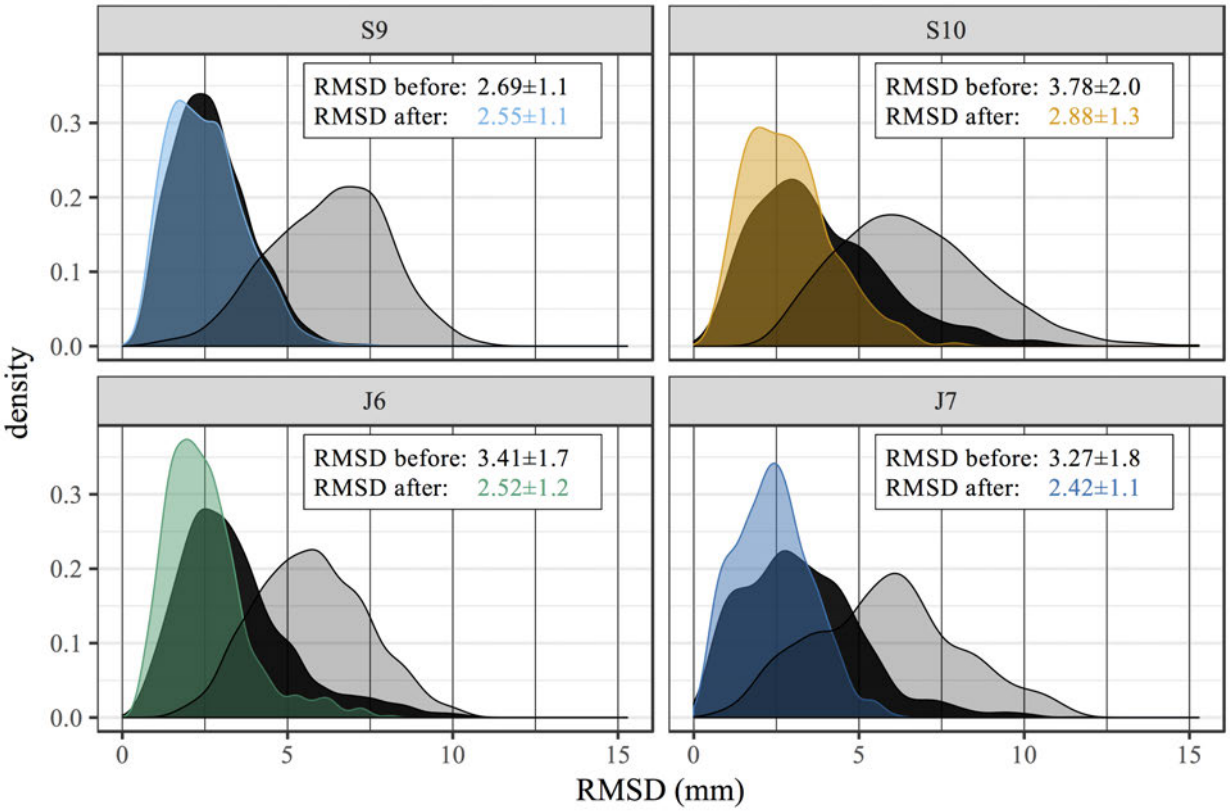


Figure 4: Comparison of histograms for RMSD values calculated from tongue contours taken from before adjustment (black shapes) and after adjustment (colored shapes), for articulations of the same sound produced in the same word. Slight improvement is shown for S9; data from the other talkers show substantial improvement. For reference, distributions of RMSD values from comparisons involving different/heterorganic sounds, which are relatively large, are shown in gray.

the data collected from each talker and omit data from the talkers exhibiting obvious head movement. Identifying such talkers is not a simple matter and would minimally require an inspection of all extracted tongue-contour splines overlaid on top of each other in order to examine any inconsistencies in the position of the tongue for the articulation of target sounds. Ideally, during the production of the same sound, articulatory positions of the tongue are nearly identical, and identical tongue contour shapes for the same sound without matching spatial positions would strongly indicate a change in head-to-transducer positioning.

Under circumstances in which the omission of data is not desired, the primary strategy for making use of ultrasonic data that contains ample head movement is to ensure that there is greater

certainty as to where fixed points on the head are located in the ultrasonic imaging signal at all times during the recording. With the swallow method used in Archangeli et al. (2016) and described here, only a few instances of palate-contour traces were collected in each talker's recording. To increase the certainty about whether each target production in a recording involved repositioning the head relative to the ultrasonic transducer, it would be optimal to instruct the talker to either press their tongue against their palate or simply perform a dry swallow frequently during scanning, such as before every 5–10 target utterances. These techniques for producing traces of the palate traces in the recording are discussed in the "Palate Trace Protocol" appendix found in McLeod & Wrench (2008). However, having talkers perform a dry swallow or press their tongue onto their palate constantly during the session may be infeasible if there are prohibitive time constraints during data collection or if the excessive swallowing makes the subject's mouth too dry in a longer recording session.

Another improvement to the adjustment method described above is to allow for rotation about a point of origin (representing the location of the ultrasonic transducer) in the adjustment of palate and tongue contours. In our method, palate contours were adjusted via an approximate method of translational correction (in x and y within a Euclidean plane) between ostensibly mismatched contours. Incorporating rotational adjustments would address changes in pitch angle of the head relative to the transducer during data collection, which is quite possible when the head is not entirely fixated relative to the transducer. Additionally, the right type of positional adjustment between palate contours could be determined by using an optimization algorithm that minimizes the overall distance between palate-contour pairs or sets.

4. Conclusion

The swallow alignment contour-correction method discussed in Section 2 describes only a rough method of identifying participants whose head moved to a substantial degree during speech production, and a precise, quantitative way of categorizing participants who produced "clean" versus "messy" data has yet to be determined. However, we set forth our method as technique to make improvements on previously-recorded ultrasonic imaging data where there would otherwise be no contour adjustment to correct for physical head movement during collection.

References

Archangeli, D., Yip, J., Qin, L., & Lee, A. (2016). Phonological and phonetic properties of nasal substitution in Sasak and Javanese. University of Arizona and University of Hong Kong.

Bird, S. (17–21 August 2011). Phonetic fieldwork in the Pacific Northwest. In The Hong Kong Consortium for ICPHS 2011 (Ed.), *Proceedings of the 17th International Conference of Phonetic Sciences. Hong Kong*, (pp. 76–79).

Derrick, D., Best, C. T., & Fiasson, R. (10–14 August 2015). Non-metallic ultrasound probe holder for co-collection and co-registration with EMA. In The Scottish Consortium for ICPHS 2015 (Ed.), *Proceedings of the 18th International Conference of Phonetic Sciences. Glasgow, UK*, (pp. 119.1–5).

Gick, B. (2002). The use of ultrasound for linguistic phonetic fieldwork. *Journal of the International Phonetic Association*, *32*, 113–121.

Li, M., Kambhamettu, C., & Stone, M. (2005). Automatic contour tracking in ultrasound images. *Clinical Linguistics & Phonetics*, *19*(6/7), 545–554.

McLeod, S. & Wrench, A. (2008). Protocol for restricting head movement when recording ultrasound images of speech. *Asia Pacific Journal of Speech, Language, and Hearing*, *11*(1), 23–29.

Ménard, L., Aubin, J., Thibeault, M., & Richard, G. (2012). Measuring tongue shapes and positions with ultrasound imaging: A validation experiment using an articulatory model. *Folia Phoniatrica et Logopaedica*, *64*, 64–72.

Mielke, J., Archangeli, D., & Baker, A. (2016). Individual-level contact reduces phonological complexity: evidence from bunched and retroflex /ɹ/. *Language*, *92*, 101–140.

Mielke, J., Baker, A., Archangeli, D., & Racy, S. (2005). Palatron: a technique for aligning ultrasound images of the tongue and palate. In S. Jackson & D. Siddiqi (Eds.), *Coyote Papers*, volume 14 (pp. 96107). Tucson AZ: University of Arizona, Department of Linguistics.

Sigona, F., Stella, A., Gili Fivela, B., Montagna, F., Maffezzoli, A., Wrench, A., & Grimaldi, M. (6–8 November 2013). A new head-probe stabilization device for synchronized ultrasound and electromagnetic articulography recordings. In Queen Margaret University, Clinical Audiology,

Speech and Language (CASL) Research Centre (Ed.), *Proceedings of the Ultrafest VI. Edinburgh, UK*, (pp. 1–4).

Stone, M. & Davis, E. P. (1995). Support system for ultrasound imaging. *J. Acoust. Soc. Am.*, *98*, 3107–12.

Whalen, D. H., Iskarous, K., Tiede, M. K., Ostry, D. J., Lehnert-LeHouillier, H., Vatikiotis-Bateson, E., & Hailey, D. S. (2005). The Haskins Optically Corrected Ultrasound System (HOCUS). *Journal of Speech, Language, and Hearing Research*, *48*, 543–553.

Whalen, D. H. & McDonough, J. (2015). Taking the laboratory into the field. *Annual Review of Linguistics*, *1*, 395–415.

Zharkova, N., Gibbon, F. E., & Hardcastle, W. J. (2015). Quantifying lingual coarticulation using ultrasound imaging data collected with and without head stabilisation. *Clinical Linguistics & Phonetics*, *29*(4), 249–265.

Zharkova, N. & Hewlett, N. (2009). Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /ɑ/. *Journal of Phonetics*, *37*(2), 248–256.

A Pilot Study on Voice Onset Time of English Words in Cantonese-English Code-Switching in Hong Kong

Jeremy Tzi Dong Ng
University of Hong Kong
jntd@hku.hk

Abstract

Code-switching offers an interesting angle to explore structural outcomes between two languages. Research on bilingual phonetics conveys that the phonetic elements of L1 and L2, the base and guest language in code-switching per se, exert influence on each other (Flege, 1995), thus phonetic outcomes from such interactional phenomena as code-switching merit investigation. Cantonese-English code-switching is a prevalent linguistic phenomenon in Hong Kong, whereby prior studies were conducted largely from sociolinguistic and educational perspectives (Li, 1999; Lin, 1996; Luke, 1998). In the literature, speakers with higher English proficiency were shown to demonstrate flexibility in switching between the productions of phonological features in code-switching and those in the monolingual English environment (Ng & Chen, 2016). To gain a rudimentary insight into the phonetic outcomes of Cantonese-English code-switching and to take advantage of the objectivity of acoustic analysis and gradient nature of phonetic data, this pilot study examines the VOT productions of three voiceless aspirated stops and a voiced stop across contexts. Findings revealed an absence of differences in such phonetic productions when comparing the code-switching and monolingual English environments, suggesting that code-switching might not be impactful on the phonetic productions of the guest language. However, the voiceless aspirated velar stop in monolingual English exhibits a significantly higher, more English-like VOT in code-switching than in monolingual Cantonese, as observed in these L2-dominant speakers. Also, the voiced bilabial predominantly experienced devoicing in Cantonese-English code-switching. Implications regarding bilingual language production theories are discussed. This study invites follow-up research on

other features (e.g., vowel quality) for a better picture of the phonetics of Cantonese-English code-switching in HK.

Keywords: code-switching, VOT, Cantonese-English bilinguals, bilingual language production.

1. Code-switching

Code-switching (CS) is a language contact phenomenon broadly defined as the alternation between at least two languages, involving readily identifiable constituents from both languages. It is thus often a topic of interest in studies of bilingualism, and less commonly, in bilingual language production studies. Different types of CS have been observed among bilinguals, including inter-sentential and intra-sentential CS. Inter-sentential CS means the switching from one language to another language at clause boundaries, usually between the first language (L1) and second language (L2) of a speaker. Intra-sentential CS typically takes place within a sentence. According to a typology of CS patterns (Muysken, 2000), a strategy employed by bilingual speakers is *insertion*, involving the embedding of a constituent (e.g., a word, a phrase) in a "nested" A-B-A structure (Muysken, 2000). With reference to the Matrix Language Frame (MLF) model proposed by Myers-Scotton (2001), A and B are the matrix language (ML) and the embedded language (EL) respectively, while A is referred to as the "base language" and B as the "guest language" in the phonetics and phonology of CS (e.g., Bullock, 2009). However, the base language might not necessarily be the phonetic equivalent of the morpho-syntactic matrix language. Sociolinguists have been interested in speakers' attitudes towards CS (e.g., Dewaele & Wei, 2014) while others are fond of the influence of CS on syntactic outcomes (e.g., Cacoullos & Travis, 2016; Yeung et al., 2012). Such a contact phenomenon as CS always yields investigation into the structural outcomes that arise. The morpho-syntax of CS was widely studied, while limited efforts have been dedicated purely to the sound structure of CS.

Hong Kong is arguably a diglossic city where a significant portion of speakers are bilinguals speaking both Cantonese and English (Census and Statistics Department of HKSAR, 2011). Hong Kong has been experiencing a profound history of language contact with English since the 17th century (Luke, 1998). Inevitably, Cantonese and English interact with each other morphologically, phonologically, syntactically, and even sociolinguistically. CS, one of the most common contact phenomena in Hong Kong, offers an interesting angle on the structural outcomes of language

contact in Hong Kong, one of the motives being the typological distinction between Cantonese and English.

CS in Hong Kong is generally defined as a Cantonese utterance or sentence interspersed with an English lexical item or phrase (Li, 2000; Hyland, 1997; Hoet al. , 2007), thus being equivalent to intra-sentential CS. An example of CS in Hong Kong is shown in the following, adapted from Li (2002).

(1) 我 想 *book* 三點 一號場
I want book 3=o'clock no.1=court

"I want to book court no. 1 at 3 o'clock."

The English word "book" was selected instead of its Chinese equivalent as it helped to communicate the intended meaning of 'to make a reservation' more unambiguously. Pioneering studies on CS in Hong Kong (e.g., Gibbons, 1987) showed that this kind of CS was commonplace among Cantonese-speaking Hongkongers, whereas The prevalence of CS among these speakers can be confirmed by evidence ranging from pioneering studies (e.g., Gibbons, 1987) to recent endeavours (Li, 2017). Particularly, it has been found that speakers are often not even aware that they are performing CS (e.g., Li & Tse, 2002).

According to a comprehensive review of major studies of CS in Hong Kong (Li, 2000), research on CS in Hong Kong were predominantly from the sociolinguistic or educational perspectives (e.g., Li, 1999; Lin, 1996; Luke, 1998). CS research between this Y2K review and the recent period has seen a similar trend in topics being investigated (e.g., attitudes towards CS) and the approaches adopted (e.g., discourse analysis), though a more rigorous review is needed to arrive at a more precise conclusion. Still, studies on CS that looked into the structural outcomes or adopted the structural approach were relatively rare,and largely employed the morpho-syntactic level of analysis (Leung, 1988; Chan, 1998, 2009). Insufficient attention is given to the phonetic or phonological level. On another note, bilingual phonetics itself has remained a much alive area among studies of bilingualism. In particular, the interaction and segregation of two languages on the phonetic level has been scarcely investigated (Fabiano-Smith & Barlow, 2010), not to mention such a typologically distant pair as Cantonese and English.

2. Phonetics of Code-Switching

The majority of prior studies on the phonetics of CS used voice onset time (VOT) as a measure. The VOT values of bilinguals tend to 'drift' in the speech context surrounding it, e.g., in CS (Sancier & Fowler, 1997; Bullock et al., 2006). One of the first attempts of investigating the phonetics of CS was Grosjean & Miller (1994). They set out to explore whether the "phonetic momentum" of the *base language* would be carried over into the *guest language* in CS. French-English bilinguals were experimented with three sets of stimuli, including monolingual French and monolingual English as the control stimuli, and the experimental ones were French (base language) utterances with inserted English (guest language) tokens (i.e., French-English CS). The VOT durations of the onset consonant in the code-switch itself were measured. These measurements in the three conditions were compared against one another. Results showed that the phonetics of the base language exerted no impact on the inserted guest-language word in CS production, concluding that bilinguals are both "flexible" and "precise" when entering and exiting CS. With reference to the Matrix Language Framework, this conclusion also implies that the base language did not serve as the phonetic matrix language.

Bullock et al. (2005) later studied Spanish-English bilinguals and specified that the speakers possessed advanced proficiency in both languages. They reported that the VOT value of the guest language in a code-switch was significantly different from that in a monolingual context. This contrasted with the results in Grosjean & Miller (1994) who supported the idea of bilingual flexibility. This also means that the base language could be inferred as the phonetic equivalent of the matrix language in CS. Later, Bullock (2009) demonstrated that degree of difference potentially exist between the sound structure of two participating languages in CS, where this justifies the need to perform acoustic analysis of sounds, heavily centered around VOT productions. Among the few but scrutinized studies, Bullock & Toribio (2009b) studied whether VOT variations as the evidence of cross-linguistic influence of "a subtle nature" would emerge in CS. They found that a significant inter-lingual effect of CS on phonetic production was observed. Antoniou et al. (2011) investigated the phonetic productions of Greek-English CS and demonstrated a "pervasive" influence of L1 on the guest-language phonetic production in CS among L2-dominant bilinguals.

Regarding bilingual language production, the classical Inhibitory Control Model (Green, 1986) hypothesized that in Language A-Language B codeswitching, if the phonetic realization of B is

selected by the speaker, then A would be inhibited or suppressed. While Flege (1995) working on bilingual speech production conveyed that phonetic elements of L1 and L2 must exert influence on one another, the case of phonetic productions from interaction between these two languages in Hong Kong merits an exploration through CS, one of the most common interactional phenomena. Despite the lack of studies on the sound structure of CS in HK, Ng & Chen (2016) nonetheless offered a novice account. The study attempted to identify how Hong Kong ESL learners demonstrated phonological changes in English (guest language) words in Cantonese-English CS, and also their attitudes towards CS. There are however several methodological flaws in their study. First and foremost, the study focused largely on identifying discrete phonological features (e.g., presence or absence of voicing in word-initial consonant, etc.) instead of the phonetic features. Structurally speaking, phonetics possesses the gradient nature that could help to explore bilinguals' production of CS along an acoustic dimension, whereas phonological features are regarded as categorical. Since code-switches are expected to "manifest only marginal cross-linguistic assimilation or, ideally, none at all" (Bullock, 2009, p. 165), the interaction between CS and sound structure should better be explored by examining the discrete phonetic properties in an utterance. Also, their study employed auditory judgment, which could lead to inevitable biases. This would otherwise have been at least mitigated, if not avoided through adopting acoustic analysis, which could give more objective evidence for the potential phonetic outcomes of CS. Since "accent", as studied in Ng & Chen (2016), advances to the paradigms in sociolinguistics, this study reserves to the layer of phonetics and thus aims at yielding implications on the phonetic flexibility with regard to CS.

Previous studies (e.g., Chen, 2013; Ng & Chen, 2016) show that voiced consonants such as the voiced alveolar fricative /z/ experienced *devoicing* to become its allophone /s/. These consonants, however, do not possess a VOT value as they are not stops. The question as to whether voiced stop consonants (e.g., /b/, /d/, /g/) would undergo devoicing, in the Cantonese-English CS context, have been left unanswered. Bullock & Toribio (2009b) once addressed the need for researching the phonetics of code-switching between Chinese and English in the sense that these two languages have different "implementations" of voicing contrast. Particularly, Ng & Chen (2016) included the token 'book' as one of their stimulus words, but not much about its phonetic outcomes of this word in code-switching is known. It was shown that bilabial stops (e.g., /b/ in 'book') were, despite rarely, still prone to being devoiced as a consequence of its following certain vowels (Pape et al., 2003).

In view of these gaps, the present pilot study hopes to gain a rudimentary insight on the VOT as

one of the phonetic outcomes of Cantonese-English intra-sentential code-switching in Hong Kong. The following research questions are established:

1. Are the VOT values of word-initial voiceless aspirated stop consonants of an English word in Cantonese-English code-switching different from their counterparts in monolingual English?
2. Does a word-initial voiced bilabial stop consonant of an English word get devoiced in Cantonese-English code-switching?

Following the conventions of prior studies on CS phonetics, this study focuses on the VOT values of embedded guest-language (i.e., English) words in CS productions. Extant literature has shown that Cantonese as a first language (L1) in Hong Kong interferes with the phonological features of English as a second language (L2) (e.g., Deterding et al., 2008), and this interference was also demonstrated in the phonological features of inserted English words in CS (Ng & Chen, 2016). Supplemented with the findings in Bullock & Toribio (2009a) and Bullock & Toribio (2009b), it is hypothesized that the VOT values of word-initial voiceless stops in English words in CS will be significantly different from their counterparts in the monolingual English context. Chen (2013) and Ng & Chen (2016) showed that consonants of English words in CS were devoiced. Moving on from the previous hypothesis, this study also hypothesizes that word-initial voiced bilabials will be devoiced in Cantonese-English CS. Ultimately, it is hypothesized that Cantonese would be the phonetic matrix language in CS.

3. Methodology

3.1. Participants

A total of 10 Cantonese-English bilinguals were recruited to participate in this study through convenience sampling. All participants grew up in Hong Kong, speaking Cantonese as their native language and received formal school-based language education, particularly learning English since kindergarten. None reported any speech or hearing disorders or difficulty. Local university students in Hong Kong were shown to be speakers with frequent CS (Chen & Carper, 2005). Participants were therefore recruited at a large comprehensive university in Hong Kong and they participated in this study voluntarily without compensation. They were aged between 20 and 24 years old

(mean = 20.7), where 6 of them were female and 4 were male. Ethical consents were obtained for their participation. They all achieved Band 7.0 or above in their overall score of the International English Language Testing System (IELTS) examinations, or achieved Level 5* or above in the English Language subject of the Hong Kong Diploma of Secondary Education (HKDSE), or Grade B or above in the Use of English subject of the Hong Kong Advanced Levels Examination (HKALE). The performance levels equated across IELTS and the two public examinations in Hong Kong were provided by the Hong Kong Examinations and Assessment Authority (HKEAA, 2004). Participants' performance in such examinations instead of their self-rated competence in the target language was used for determining their proficiency, as self-ratings are more prone to inevitable biases.

3.2. Instruments and Procedures

A short CS monologue with 376 words (Chinese/Cantonese and English, Appendix I) was first written by the researcher based on his personal experience. Monologues are advantageous over conversations in that CS in conversations would derive sociolinguistic consideration, out of the scope of this study. It was written from the perspective of a local undergraduate student in Hong Kong so that participants, who were also undergraduates, would find the content relatable to their own background. Along with the informal style, the monologue contains colloquial Cantonese expressions (e.g., "勁神秘" meaning "super mysterious") to offer a sense of locality. This design was to make the monologue appear less fabricated, hoping to model a casual monologue despite the unavailability to collect naturalistic CS data owing to the costs and ethical concerns involved.

The monologue contained either pure Cantonese clauses or such clauses inserted with English words (the CS clauses). These code-switches (N = 23) included single-words that were either monosyllabic (e.g., *friend*) or bi-syllabic (e.g., *report*) and two instances of two-word phrases ("*small potatoes*", "*print notes*"). Minimal pairs of stop-voicing (e.g., /pə/ vs /bə/) or stop-aspiration sharing the same place of articulation, were not arranged consecutively in order to minimize the effects of hyper-articulation. For the first research question, 3 code-switches were chosen as stimulus words ("paper", "taste", "cake"), each beginning with a voiceless aspirated stop consonant (i.e., /p/, /t/, /k/) followed by the same diphthong /eɪ/ for the sake of similar phonological environments. Following are the excerpts from the monologue, each containing a stimulus word:

(2) 自己 做 *paper* 嘅 速度
oneself do paper POSS speed

"One's own speed of working on a paper"

(3) 俾 姊妹 大讚 有 *taste*
PASS sisters praise have taste

"praised to have a good taste by sisters"

(4) 買 返 件 *cake* 食
buy back piece cake eat

"to buy a piece of cake to eat"

For the second research question, there was one stimulus word ("book"), beginning with a voiced bilabial stop consonant. The corresponding excerpt is shown below:

(5) 未 *book* 場 打波
Not=yet book court playing=basketball

"(I) have not yet booked a basketball court."

Other code-switches acted as the fillers in this pilot study. The monologue was translated to a monolingual English (guest language in CS) version (Appendix II), where all the stimulus words appearing in the original CS monologue were maintained. An English major student who is also a Cantonese-English bilingual and did not participate in the study was involved in checking the fluency, grammatical accuracy and choice of diction of the monologue before using it on participants. As opposed to prior studies offering several settings including CS, monolingual base language, and monolingual guest language, a pure monolingual Cantonese (i.e., base language) version of the monologue was not prepared because CS is a highly common and natural linguistic phenomenon in Hong Kong. An entire monologue in Cantonese would make participants feel unnatural (e.g., Li & Tse, 2002).

The 23 code-switches were also extracted into a word list containing the 25 English words, similar to a list of vocabulary item used in ordinary English classrooms in Hong Kong. This word list served as the monolingual setting for the first research question of this study owing to an accidental omission of the token "paper" in the English version of the monologue, while the English

version of the monologue served as the monolingual setting for the second research question. The production of a token in isolation might be able to offset the effects of the preceding sound on a target stop, which might lead to a larger variance in the resultant VOT values.

Each participant first filled in a short questionnaire on their personal and background information, including their gender, age, and their scores in IELTS, HKDSE or HKALE. They were then asked by the Cantonese-speaking researcher, in Cantonese, to read aloud the CS monologue, its English version and the word list (Appendix I, II and III) separately in the form of three reading tasks. To ensure that participants demonstrated their most comfortable pace and volume, no specific instructions were given before participants did the reading. All the reading was audio-recorded in a quiet environment on campus. An optional de-briefing session was conducted after the reading tasks if participants would like to know more about the study.

3.3. Data Analysis

A recording was segregated into chunks on both the sentential level and the word level, for acoustic analysis using Praat. The researcher carried out the data pre-processing, inspection of both the waveforms and spectrograms, and measurements (e.g., VOT measurements) for all productions to ensure consistency. 2.7% of stops were discarded from the dataset as they were inaudible due to unclear pronunciation or skipping of a sound. To answer the first research question, VOT values of each stimulus word of each participant in the two target contexts, namely CS and monolingual English, were measures. For the second research question, the VOT of the voiced stop consonant in the token was measured, also in the two target contexts. At the same time, the spectrogram of the token for the second research question was also inspected, to explore whether its onset stop consonant would be "devoiced". These quantitative data were later analyzed using SPSS. Statistical significance was taken and indicated at the levels of $p < .05$* and $p < .01$**.

4. Results

4.1. VOT of Word-initial Voiceless Aspirated Stops

Among the ten participants of this study, /p/ in CS entailed a larger range of VOT values than its counterpart in monolingual English. A similar phenomenon on the range difference also applies to

the VOT of /k/ in CS and in English. On another note, the VOT measurements of /k/ in both CS and monolingual settings gave rise to the highest SD among those of the three stops. Also, the maximum values of VOT across the three stops in both settings also belonged to /k/.

Figure 1 displays the mean VOT values of the word-initial stop consonants in both the settings of CS and in the monolingual setting. A paired-samples t-test was conducted for comparing the VOT values of each stop in the two settings, yielding no statistically significant difference ($p > .05$).

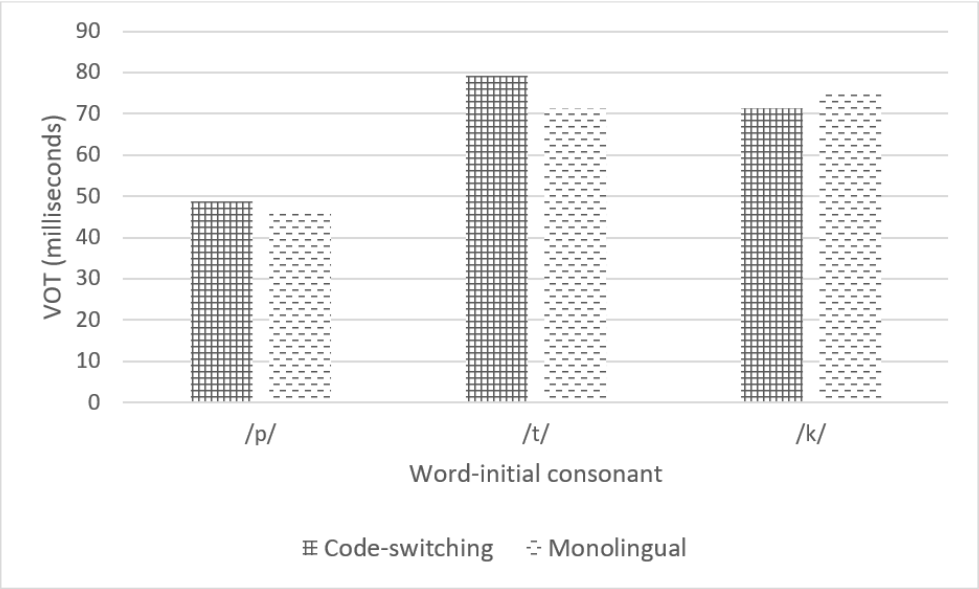


Figure 1: Mean VOT of English words with word-initial voiceless stops in two settings of HP participants

Later, a revisit of the CS monologue revealed that it contains the word '其' (/kei4/) that could be taken into account. It appeared in one of the Cantonese sentences in the CS monologue. Since it does not precede or follow any English elements, it is regarded as a word in the monolingual Cantonese context.

(6) 佢 其實 都 係 搵錢
He actually just is making=a=living

"He's just making a living."

The voiceless aspirated velar stop /k/ and its following diphthong /eɪ/ imply a similar phonological environment with the stimulus word 'cake' (/keɪk/). This enables comparison among the

instances of this stop in not only the monolingual English and CS but also the monolingual Cantonese context. Subsequently, the VOT values of the onset consonant of '其' were measured. In the monolingual Cantonese setting, participants demonstrated a range of VOT values from 27.17 ms to 75.26 ms (SD = 15.19 ms). Modelling the pioneering research design of Grosjean & Miller (1994), the VOT values of /k/ across the three settings were compared. The SD of the VOT of /k/ in Cantonese was actually the lowest among that in the three settings. Figure 8 shows the mean VOT across the three settings. Results were 48.86 ms in Cantonese, 74.95 ms in English, and 74.40 ms in CS. A one-way ANOVA test was conducted on the VOT values of /k/ across the three settings. A strong statistical significance at $p = .005^{**}$ was yielded. Post-hoc tests for pairwise differences showed that the VOT values of /k/ in Cantonese and in English are significantly different ($p = .003^{**}$), while the difference between that in Cantonese and CS is also significant ($p = .007^{**}$). However, the VOT of /k/ in CS and that in English are shown to be statistically indifferent ($p = .750$).

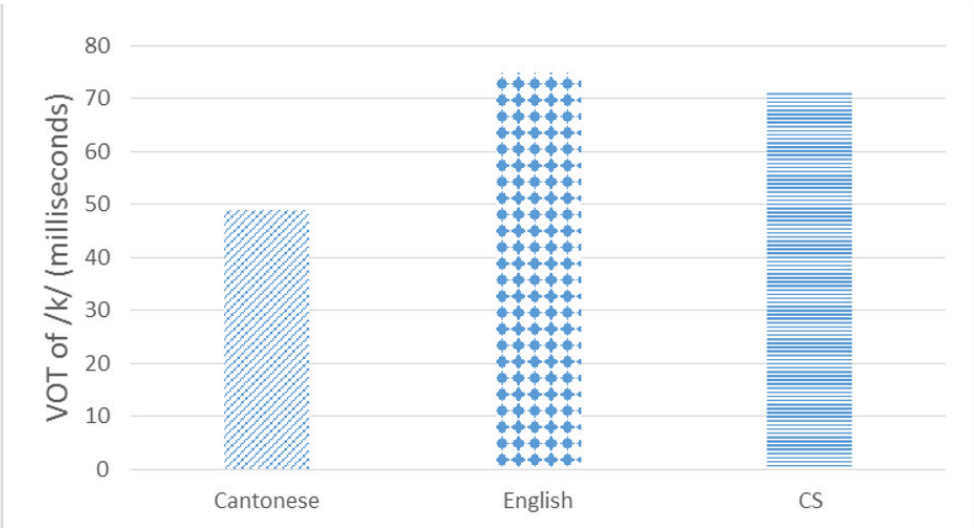


Figure 2: VOT of /k/ under different settings

4.2. Devoicing of Voiced Stop

All instances of /b/ in the token 'book' in CS were devoiced, while in the monolingual English setting, 88.9% of the instances were devoiced. Figure 3 displays the acoustic representation of /b/ in English without devoicing and Figure 4 shows a devoiced counterpart with the lowest positive VOT value among participants. VOT values of the devoiced /b/ were considered for analysis.

Participants demonstrated a larger range of VOT for /b/ in English than that in CS. The VOT values of /b/ in CS nonetheless had the smallest SD and also the smallest value (5 ms) among all observations in this study. An independent samples t-test showed that the per-participant average VOT of /b/ in CS (8.529 ms) and that in monolingual English (14.196 ms) were significantly different ($p = .045$*). For data from each of the ten individual participants, please refer to Appendix IV.

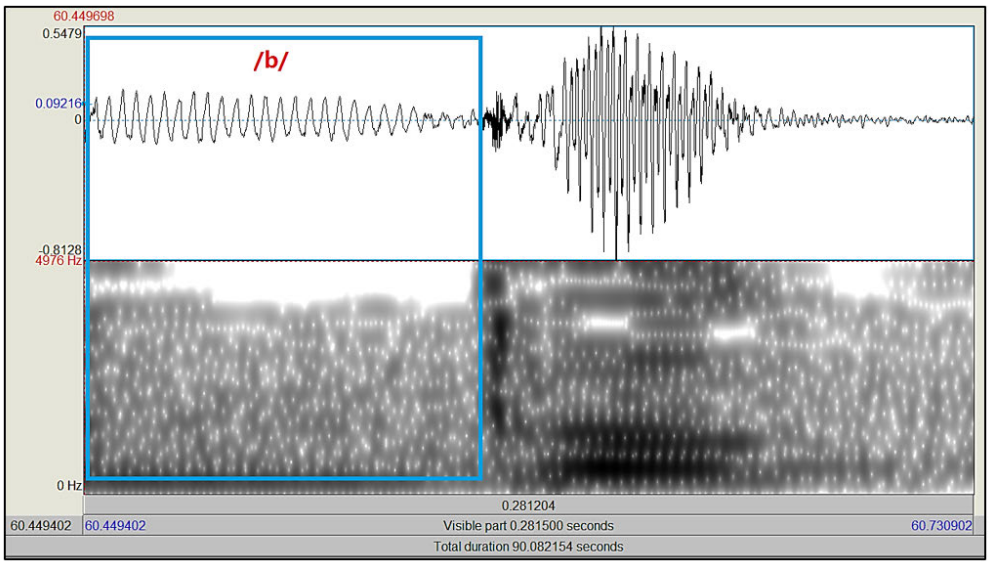


Figure 3: Voiced /b/ in "book" produced by Participant 3 in monolingual English

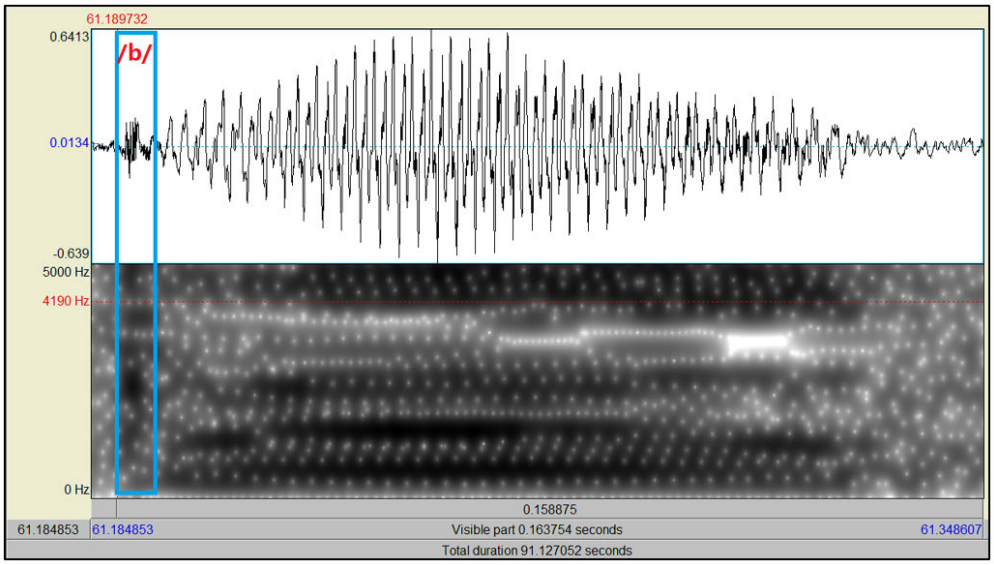


Figure 4: Devoiced /b/ in "book" produced by Participant 5 in monolingual English

5. Discussion

The VOT values of the voiceless aspirated bilabial /p/ and voiceless aspirated velar /k/ in the CS context were diverse. This indicates that these two specific stops were not phonetically stable in CS. This is in alignment with previous findings that the VOT production of a stop produced by bilinguals were not static owing to the speech environment (Sancier & Fowler, 1997; Bullock et al., 2006). For /p/ in the token "paper" in CS, auditory inspection yielded that some speakers produced the token as /ˈpeɪˌpɑː/ in CS and /ˈpeɪˌpə/ in monolingual English. Though there is only a difference in the vowel quality of the second syllable, this vowel lengthening in CS can be seen as an influence from Cantonese since various English loanwords in Cantonese have a similar phonetic adaptation before transforming to a Cantonese word (Bauer & Wong, 2008). Phonetic realization of the entire token (i.e., the code-switch) with the influence from the base language would therefore lead to variance in its onset consonant. For /k/ in "cake" in CS, another round of auditory inspection suggested that speakers produced its following diphthong /eɪ/ in at least two identifiable ways. It was realized as either /eɪ/ or a phonetic variant between /eɪ/ and /ɪ/. The variance in the vowel following a stop therefore gave rise to the phonetic instability of the stop, since the VOT of a stop is shown to be influenced by its following vowel (Pape et al., 2003). The Cantonese context was included at a later stage of the study. Interestingly, a low dispersity was found in /k/ in the Cantonese context for participants. This depicts that /k/ was phonetically stable when realized as a Cantonese consonant, regardless of the dominance of the guest language. Another eye-opening finding was that the maximum VOT value of /p/ in CS was even higher than the its maximum value in English. These speakers proficient in English might have acquired their production capability in the aspirated /p/ in English by deliberately extending their aspiration in /p/ in the Cantonese phonetic inventory. These speakers being English majors might have been meta-linguistically aware of their pronunciation and exhibited over-controlling on their phonetic production. This phenomenon was also observed in Bullock et al. (2006), and was interpreted as hyper-correction in Bullock & Toribio (2009b). The present study thus shows that this phenomenon on L2-dominant individuals also occurred in Cantonese-English CS in Hong Kong.

The mean VOT values of word-initial voiceless aspirated stops were non-significant across contexts. This demonstrated that this phonetic property of the guest language (English) in CS was no different from that in the monolingual context, rejecting that the hypothesis to the first research question. However, the conclusion as to whether CS is impactful on the guest-language VOT cannot

be drawn at this point, since it is not known whether the phonetic productions of English words, before taking CS into consideration, were already influenced by the base language (Cantonese). Results were unable to show whether the language switch from the base language (Cantonese) to the guest language (English) was completed. Prior studies showed that the VOT value of the guest language in CS differed significantly from that in the monolingual environment (e.g., Bullock, 2009; Bullock et al., 2005) while the present study showed otherwise. It was also suggested that CS exerted an impact on the phonetic productions of the guest language. Nonetheless, they studied pairs of languages closer in the family of languages, such as French- English (Grosjean & Miller, 1994), and Spanish-English (Bullock et al., 2005), while the present study explored CS of two typologically distinct languages, having phonologies with a number of contrastive differences (Chan & Li, 2000). A preliminary implication is that the phonetic outcomes of L2 on its own and those of L2 as the guest language in CS are not significantly different from each other if the L2 (English, in the Hong Kong case) is typologically distant from the base language that is also the L1 (Cantonese).

There are two ways of looking into the sound structure of CS. The objective acoustic analysis can derive phonetic outcomes of CS while auditory judgment helps identify its phonological features (e.g., Chen, 2013; Ng & Chen, 2016). Still, the gradient nature of phonetic features might generate more useful implications on bilingual speech production, when compared with the discrete categories of phonological features. While it is evident that the *phonological* features of English in Cantonese-English CS are different from those of English on its own (Ng & Chen, 2016), this study nonetheless shows that such *phonetic* properties as VOT in the two contexts might be very similar. This is a first attempt highlighting the importance of subtle but significant structural changes reflected from CS (Bullock & Toribio, 2009b) as in the context of Hong Kong. For instance,Ng & Chen (2016) showed the deletion of an onset or word-final consonant in several tokens (e.g., "present" as a verb) of the guest language in CS, though consonants might not necessarily get completely dropped and there could be a continuum between the "presence" of a consonant and its "absence", like the variation of VOT which helps to infer the degree of aspiration of a stop.

An unexpected discovery emerged during the process of investigating preliminary findings for the first research question. The context of monolingual Cantonese (i.e., the base language) was also involved for the voiceless aspirated velar stop /k/. Results showed that its mean VOT in CS was no different from that in a monolingual English environment. However, its mean VOT value in Cantonese (48.86 ms) was significantly lower than that in English (74.95 ms, $p = .003^{**}$) and in CS (71.40 ms, $p = .007^{**}$) while VOT of /k/ in English and CS were not significantly different (p

$> .05$). This implies that the phonetic momentum of the base language (Grosjean & Miller, 1994) failed to be carried over to the guest language in CS, at least for the said onset consonant. The language switch from Cantonese as the base language to English as the guest language was therefore completed at the onset of the code-switch. Countering the argument by Bullock et al. (2006), results of this study implies the speakers did not interpret such embedded words as borrowings as they were not integrated into the base-language phonology.

Nonetheless, this at least depicts that when speakers with a higher L2 proficiency code-switch, their phonetic outcomes of CS tend to be productions in the guest language and away from the base language. Participants were English majors who were immersed in an English-dominant environment, thus producing English-like VOT in CS. These speakers could therefore be inferred as not suffering from phonetic interference from L1 in the CS context, at least at the onset. They managed to suppress the impact of L1 on L2 in at least this particular phonetic production at the code-switch boundary and prevented "phonetic bleeding". This aligned with the Inhibitory Control Model (Green, 1986) that hypothesized the inhibition or suppression of Language A (i.e., Cantonese in this study) if the phonetic realization of Language B (i.e., English) is selected in A-B codeswitching. Following how Bullock & Toribio (2009b) framed the phonetics of CS with the MLF model, Cantonese is therefore not the *phonetic matrix language* in Cantonese-English CS in Hong Kong. This consideration is at least valid for the speakers with a higher proficiency in the embedded language and further in-depth investigation is necessary to validate the relationships between the base language and the matrix language. Also, findings at this point could be coupled with an argument posited by Olson (2012) who studied bilingual language switching at the phonetic level. Assumed that these L2-dominant bilingual speakers were able to differentiate between phonetic norms, phonetic activation of English did occur after lexical selection in English, i.e., when switching from Cantonese into an English token.

This study presented acoustic evidence for more objective findings. In addition to the voiced fricative /z/ of an English word in CS (Ng & Chen, 2016), results for the second research question suggested that voiced bilabial stops /b/ also largely experienced devoicing in CS, where only a small portion of instances were not devoiced in the monolingual English context. This validates the hypothesis to the second research question. The inherent phonological differences between Cantonese and English can account for the predominance of devoicing. For instance, voicing contrast exists in English but not Cantonese. Voiced stops such as the bilabial /b/, the alveolar /d/, and the velar /g/, and unvoiced ones (i.e., /p/, /t/, /k/; each can be further categorized into aspirated

and unaspirated) contribute to minimal pairs, while voicing distinction of stops is phonetically not meaningful in Cantonese. This could explain why even these speakers with higher proficiency in English devoiced bilabials. These findings thus demonstrated the opposite to earlier findings on voiceless stops, and those from prior studies (e.g., Flege et al., 2002) where bilinguals dominant in their L2 (a.k.a. proficient in L2) were immune to L1-L2 interference. Nevertheless, findings still strengthen the empirical support towards the theoretical existence of "language-specific" phonetic categories (Antoniou et al., 2011) for voiced bilabials. In the case of Hong Kong ESL speakers in this study, an interlanguage /b/ category houses the voiced and unvoiced /b/ as two language-specific phonetic variants, in which the unvoiced /b/ seems more dominant in this phonetic inventory. Moreover, since the VOT values of unaspirated stops range from 0 to 25 ms (Lisker & Abramson, 1964), most of the participants can be inferred as realizing /b/ as a Cantonese unaspirated stop.

6. Limitations

This study has several limitations. Firstly, a limited number of participants were recruited. Findings are not intended for generalization. Regarding the research design, read-aloud tasks from prescribed materials might not yield the most naturalistic spoken data. This is however an inevitable option due to high costs and ethical concerns involved in obtaining first-hand naturalistic CS data. Notably, the inherent variance of VOT across individuals was also a methodological shortcoming. For the research instruments, the number of stimulus words might be insufficient to reflect a big picture of the phonetic outcomes of CS. Also, the setting of a base language was absent for the voiceless aspirated alveolar stop, the voiceless aspirated velar stop, and the voiced bilabial stop. Further implications on the bi-directionality notion were therefore not drawn. Therefore, caution has to be taken in interpreting preliminary findings of this pilot study.

7. Conclusion and future work

The sound structure of code-switching in Hong Kong is unique in that it involves the interaction of two typologically distinct languages. In an attempt to gain a rudimentary insight into the phonetic outcomes of Cantonese-English code-switching, the present study examined the voice onset time (VOT) durations of three English words each with a voiceless aspirated stop, in the context of Cantonese-English code-switching against the context of monolingual English. Findings

revealed that this phonetic property presented no significant differences between the two contexts, while in the literature, phonological features exhibited differences across contexts (Ng & Chen, 2016). A preliminary implication is that these bilinguals demonstrated phonetic inflexibility between code-switching and the monolingual guest-language environment. This study also explored whether devoicing occurred in the voiced bilabial stop of an English word in CS. Findings showed that most Hong Kong bilinguals in this study demonstrated devoicing likely due to the inherent difference in terms of voicing contrast between the two participating languages. This also led to the interpretation that an interlanguage category for this voiced bilabial contained a voiced and a unvoiced phoneme as two language-specific phonetic variants. Unforeseen findings emerged later in this study after considering the Cantonese context for the voiceless aspirated velar shop. Hong Kong bilinguals with a higher proficiency in their L2 (English) demonstrated similarities in its VOT between code-switching and English, and significant differences between code-switching and Cantonese. Code-switching was therefore inferred as posing no impact on the phonetic production of this stop in English. Suppression of the base language in the phonetic realization of this stop in the guest language served as an explanation.

Since this study focused entirely on the observations in the word-initial position, future research can inspect word-medial stops of code-switches, as studied in Antoniou et al. (2011) in other languages. To come up with a more comprehensive picture of the phonetics of Cantonese-English code-switching, other phonetic features including but not limited to the vowel quality in code-switches also merit further exploration. Suprasegmental features, as investigated in Olson (2012), also deserves examination in the unique context of Cantonese-English code-switching in Hong Kong. Finally, efforts can be devoted to investigate the phonetic interaction between other pairs of typologically distinct language. Forthcoming studies will consider attitudinal data from bilinguals and analyze the relationships between their perceptions of code-switching and their phonetic production in code-switching, with the aim to derive sociophonetic implications.

Appendix I: The monologue with Cantonese-English code-switching
(The actual material was without formatting. Tokens have been italicized here for easier reference.)

我係個大學生，又係個 small potato，中學嗰陣好多 friend 架，有啲而家喺外國讀書，好 happy，有啲考完 oral 之後都無見過，勁神秘。但不得不提嘅，就係嗰啲你明明無佢電話 number 但佢又搵到你嘅「舊同學」。無錯，佢地會無啦啦打俾你，原來係要 sell 你買咩最新防水、防火、防塵、防臭嘅高科技 jacket，仲話男嘅著咗會好 man，女嘅就會俾姊妹大讚有 *taste*。我心諗，嘩，明星咩，著件咁嘅嘢，就會有好多 fans？喺個時候，佢地會問埋你著開咩 size，跟住，你會突然醒覺，佢其實都係搵錢姐。但係，你望住眼前嗰張補習社老細尋日開俾你嘅 check，再凝望住部電腦，嘆息自己做 *paper* 既速度仲慢過舊壞咗嘅 RAM，諗吓自己其實係咪應該努力讀書。無奈嘅係，當你醒起自己未 book 場打波，又想落大中買返件 *cake* 食，又要今晚十二點前覆 email，又要去 print notes，又俾班上面最摺嘅阿水 fake 咗，原來聽日已經要 present 唔知邊科 report，仲記起尋日 claim 少咗上莊啲宣傳費……唉，都係抖多陣，起身叫個 pizza 再算。

Appendix II: The translated English version of the monologue
(The actual material was without formatting. Tokens have been italicized here for easier reference.)

I'm a university student and also a **small potato**. But I've got many **friends** in secondary school. Some of them are now studying abroad. How **happy**! Some mysteriously disappeared after taking their **oral** exams. And there are always some former classmates who you've never got their phone **numbers**. Still, they can find you out of nowhere. Guess what they want? To **sell** a hi-tech **jacket** to you. They would say something like 'you'll be more like a **man**' if you're a boy, or 'your sisters will praise your ***taste'*** if you're a girl. Wow. If I can get **fans** after wearing this thing, then I'm already a pop star. They'll then ask which **size** you prefer. And you'll suddenly realize they just want to earn some money. At this moment, you're looking at the **check** your boss gave you yesterday. You stare at your computer, feeling that you work more slowly than the **RAM** of a defected computer. You are thinking whether you should work harder. But life isn't fair. Because you forgot to **book** a basketball court. Then you're hungry and you want a ***cake***. But you need to reply to an **email** before mid-night and to **print notes**. But J said something **fake** to trick you. Oh, you realize you have to **present** a **report** tomorrow. Well, you also forgot to **claim** the expenses for the students' association. Sigh! You should sleep for a little longer and order **pizza** later.

Appendix III: The list of tokens in separation

(The actual material was presented in the form of a vertical list of words.)

1. small	2. potato	3. friend	4. happy	5. oral
6. number	7. sell	8. jacket	9. man	10. taste
11. fans	12. size	13. check	14. paper	15. RAM
16. books	17. cake	18. email	19. print	20. notes
21. fake	22. present	23. report	24. claim	25. pizza

Appendix IV: VOT values (in milliseconds) of each stop in different contexts in the HP group

	/p/		/t/		/k/			/b/	
	CS	English	CS	English	CS	English	Cantonese	CS	English
HP1	27.09	23.807	78.323	61.724	56.877	58.741	38.244	7.263	24.708
HP2	44.265	51.972	109.19	56.845	132.448	82.771	52	10	16.86
HP3	49.209	45.698	78.929	90.318	52.037	90.105	75.258	5	-124.49
HP4	68.316	76.241	88.308	89.918	50.216	82.109	27.165	11.75	dropped
HP5	49.161	55.391	75.249	76.451	73.792	59.897	dropped	6	8.625
HP6	50.105	50.506	53.533	69.905	82.146	54.84	46.183	7	10.901
HP7	59.925	21.213	88.273	91.626	95.756	61.817	46.737	9	10.492
HP8	78.474	27.709	81.882	78.661	53.162	77.921	dropped	9.824	23.176
HP9	13	41.23	44.824	67.584	38.269	123.682	64.547	8	11.837
HP10	48.77	72.131	92.579	31.113	79.267	57.623	40.709	11.455	10.868
Max	78.474	76.241	109.19	91.626	132.448	123.682	75.258	11.75	24.708
Min[1]	13	21.213	44.824	31.113	38.269	54.84	27.165	5	8.625
Mean[2]	48.8315	46.5898	79.109	71.4145	71.397	74.9506	48.8554	8.5292	12.6498
SD	18.7401	18.8631	18.5824	18.6591	27.825	21.3233	15.1859	2.26248	46.7513

[1]The minimum VOT value of /b/ was taken among the *devoiced* instances.

[2]The mean VOT value of /b/ was taken at the mean value of all *devoiced* instances.

References

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics*, *39*(4), 558–570.

Bauer, R. & Wong, C. (2008). Database of English loanwords in Cantonese. *URL: http://funstuff.engl.polyu.edu.hk/loanwords/*.

Bullock, B. E. (2009). Phonetic reflexes of code-switching. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-switching*. Cambridge University Press.

Bullock, B. E. & Toribio, A. J. (2009a). *The Cambridge handbook of linguistic code-switching*. Cambridge University Press.

Bullock, B. E. & Toribio, A. J. (2009b). Trying to hit a moving target: On the sociophonetics of code-switching. *Multidisciplinary approaches to code switching*, *41*, 189–206.

Bullock, B. E., Toribio, A. J., Botero, C. G., & Davis, K. A. (2005). Preseverative phonetic effects in bilingual code-switching. In *Theoretical and laboratory approaches to Romance linguistics* (pp. 206–291). John Benjamins Publishing Amsterdam, The Netherlands.

Bullock, B. E., Toribio, A. J., González, V., & Dalola, A. (2006). Language dominance and performance outcomes in bilingual pronunciation. In *Proceedings of the 8th generative approaches to second language acquisition conference*, (pp. 9–16). Cascadilla Proceedings Project Somerville, MA.

Cacoullos, R. T. & Travis, C. E. (2016). Two languages, one effect: Structural priming in spontaneous code-switching. *Bilingualism: Language and Cognition*, *19*(4), 733–753.

Census and Statistics Department of HKSAR (2011). Hong Kong Special Administrative Region (2012, february). 2011 Population census summary results.

Chan, A. Y. & Li, D. C. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language Culture and Curriculum*, *13*(1), 67–85.

Chan, B. H. S. (1998). How does Cantonese-English code-mixing work. *Language in Hong Kong at century's end*, 191–216.

Chan, B. H. S. (2009). *Code-switching between typologically distinct languages.* Cambridge University Press.

Chen, K. & Carper, G. (2005). Multilingual Hong Kong: Present 一個 project (video).

Chen, R. H.-C. (2013). Phonological changes in code-mixing and the effects on foreign accents: A comparative study. *English Teaching & Learning*, *37*(4), 55–89.

Deterding, D., Wong, J., & Kirkpatrick, A. (2008). The pronunciation of Hong Kong English. *English World-Wide*, *29*(2), 148–175.

Dewaele, J.-M. & Wei, L. (2014). Attitudes towards code-switching among adult mono-and multilingual language users. *Journal of Multilingual and Multicultural Development*, *35*(3), 235–251.

Fabiano-Smith, L. & Barlow, J. A. (2010). Interaction in bilingual phonological acquisition: Evidence from phonetic inventories. *International journal of bilingual education and bilingualism*, *13*(1), 81–97.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience*, *92*, 233–277.

Gibbons, J. (1987). *Code-mixing and code choice: A Hong Kong case study*, volume 27. Multilingual Matters Clevedon.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and language*, *27*(2), 210–223.

Grosjean, F. & Miller, J. L. (1994). Going in and out of languages: An example of bilingual flexibility. *Psychological science*, *5*(4), 201–206.

Ho, J. W. Y. et al. (2007). Code-mixing: Linguistic form and socio-cultural meaning. *The International Journal of Language Society and Culture*, *21*, 1–8.

Hyland, K. (1997). Language attitudes at the handover: Communication and identity in 1997 Hong Kong. *English World-Wide*, *18*(2), 191–210.

Leung, Y.-B. (1988). Constraints on intrasentential code-mixing in Cantonese and English. *Hong Kong Papers in Linguistics and Language Teaching: Special Issue*, 23–40.

Li, D. C. (1999). Linguistic convergence: Impact of English on Hong Kong Cantonese. *Asian Englishes*, *2*(1), 5–36.

Li, D. C. (2000). Cantonese-English code-switching research in Hong Kong: A Y2K review. *World Englishes*, *19*(3), 305–322.

Li, D. C. (2002). Cantonese-English code-switching research in Hong Kong: A survey of recent research. *Hong Kong English: Autonomy and Creativity*, 79–99.

Li, D. C. (2017). *Multilingual Hong Kong: Languages, Literacies and Identities.* Springer.

Li, D. C. & Tse, E. (2002). One day in the life of a "purist". *International Journal of Bilingualism*, *6*(2), 147–203.

Lin, A. M. (1996). Bilingualism or linguistic segregation? Symbolic domination, resistance and code switching in Hong Kong schools. *Linguistics and education*, *8*(1), 49–84.

Lisker, L. & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384–422.

Luke, K.-K. (1998). Why two languages might be better than one: Motivations of language mixing in Hong Kong. *Language in Hong Kong at century's end*, 145–159.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*, volume 11. Cambridge University Press.

Myers-Scotton, C. (2001). The matrix language frame model: Development and responses. *Trends in Linguistics Studies and Monographs*, *126*, 23–58.

Ng, T. D. & Chen, H. C. (2016). Phonological changes in Cantonese-English code-mixing for ESL learners in Hong Kong and their attitudes toward code-mixing. *The Journal of AsiaTEFL*, *13*(3), 162–185.

Olson, D. (2012). The phonetics of insertional code-switching: Suprasegmental analysis and a case for hyper-articulation. *Linguistic approaches to bilingualism*, *2*(4), 439–457.

Pape, D., Mooshammer, C., Hoole, P., & Fuchs, S. (2003). Devoicing of word-initial stops: A consequence of the following vowel. In *Proceedings of the 6th International Seminar on Speech Production*, (pp. 207–212).

Sancier, M. L. & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of phonetics*, *25*(4), 421–436.

Yeung, G., Wu, J., Ng, J., & Lau, E. (2012). Intra-sentential code-switching of Cantonese-English bilinguals in Hong Kong: An experimental approach to the evaluation of the Matrix Language Frame (MLF) model. In *Annual Research Forum 2016 of the Linguistics Society of Hong Kong*, Aspen, Colorado.

Factors affecting object preposing in Cantonese complex predicates

Ryan Ka Yau Lai
University of Hong Kong
kayaulai@hku.hk

Abstract

This paper examines the factors affecting the choice between two word orders in Cantonese complex predicate constructions: the object-medial [V_1 Obj *m* V_2] and the object-final [V_1 *m* V_2 Obj]. Data were collected using Google Search and the probability of choosing the object-medial construction under different situations was estimated using nonparametric statistical methods. It is discovered that the choice depends on many factors, including the object's position on the referential hierarchy, verbal semantics, and the degree of lexicalisation of the complex predicate. Potential synchronic and diachronic explanations of these effects are then proposed.

Keywords: complex predicate constructions, Cantonese, object preposing.

1. The explicandum

In modern Cantonese, there exist complex predicates where the object may be sandwiched *between the* two verbs only when second verb, traditionally known as the *complement* (*buyu*) in the Chinese literature, is negated, i.e. preceded by the negation marker *m4* (or *dak1*: see discussion in section 5):[1][2]

[1]There exist complex predicates where the second verb simply cannot be negated regardless of object position, e.g. **giu3 m4 co3 keoi5*/**giu3 keoi5 m4 co3* (call no wrong 3sg/call 3sg no wrong, 'cannot call him by his wrong name'), corresponding to the affirmative construction *giu3 co3 keoi5* (call wrong 3sg, 'call (him/her) by the wrong name'); such predicates will not be discussed.

[2]We shall adopt the Jyutping system of Cantonese romanisation (Tang et al., 2002).

(1) a. *saat3* *sei2* *keoi5*
kill die him

'to kill him'

b. *saat3 keoi5 sei2

c. *saat3* *m4* *sei2* *keoi5*
kill no die him

'to be unable to kill him'

(The subject attempted to kill him, but to no avail.)

d. saat3 keoi5 m4 sei2

Such constructions are historically derived from serial verb constructions that exhibit asymmetry between the verbs - the so-called *asymmetric* SVCs (Matthews, 2006). In this paper, we shall refer to the dominant verb as V_1 and the restricted verb as V_2. We are therefore concerned with [V_1 Obj *m* V_2] / [V_1 *m* V_2 Obj] alternations.

As has been previously noted Li (2013), similar alternations exist in other Chinese varieties, such as Xiang, Hakka, Min and Wu. For instance, the Yiyang variety of Xiang exhibits the following alternation (Xu, 2001):[3]

(2) a. *wei* *bu* *da* *ta*
feed NEG big 3sg

'to not be able to raise (a child, a pet, etc.) successfully'

b. wei ta bu da

c. wei bu ta da

(3) a. *da* *bu* *dao* *che*
ride NEG DAO vehicle

'to not succeed at getting on a vehicle'[4]

b. da che bu dao

c. da bu che dao

From Xu's data, all three construction types seem fully productive in Yiyang. This is not

[3]Xu does not provide phonetic transcriptions for each word; as such, cognate-by-cognate Mandarin romanisations using *pinyin* were used here.

[4]The semantics of the *dao* particle (*dou2* in Cantonese) will be discussed below.

the case in Cantonese. If (2-3) are translated morpheme-by-morpheme to Cantonese, the forms in (2c), (3b) and (3c) will be ungrammatical. The ungrammaticality of the object-medial (c) constructions and grammaticality of the object-final (a) constructions are categorical in Cantonese (with exceptions to be discussed below); however, less is known about the situations under which the object-medial (b) patterns are acceptable:

(4) a. *joeng5* *m4* *daai6* *keoi5*
feed NEG big 3sg

'to not raise (a child, a pet, etc.) successfully'

b. joeng5 keoi5 m4 daai6

c. *joeng5 m4 keoi5 daai6

(5) a. *daap3* *m4* *dou2* *ce1*
ride NEG DOU vehicle

'to not succeed at riding on a vehicle'

b. *daap3 ce1 m4 dou2

c. *daap3 m4 ce1 dou2

This paper focuses on this issue: What factors explain the variability in acceptability of the object-medial order [V_1 Obj *m* V_2] in Cantonese?

We assume that [V_1 Obj *m* V_2] is a marked word order in Cantonese. This is because object-medial order has a more restricted distribution. The object-final order can usually be used in place of the object-medial order (4ab), but the reverse is not always true (5ab); it is rare to find cases where the object-medial order is acceptable but the object-final one is not. Thus the object-medial order is more marked according to one of the criteria detailed in Comrie (1976). Nonetheless, local markedness reversals do exist, and will be touched on below.

With reference to previous studies of constructional competition (*inter alia* Gries, 2006; Bresnan et al., 2007; Arppe, 2009), we hypothesise that an account of the object position alternation in Cantonese may consider the contribution of at least seven groups of factors, not all of which may turn out to be useful:

- Main effect of V_1: phonological weight, TAM value, verbal semantics

- Main effect of V_2: phonological weight, level of grammaticalisation, verbal/adjectival semantics
- Main effect of Obj: phonological weight, topicality, animacy, definiteness
- Interaction effect between V_1 and V_2: level of lexicalisation of the V_1-V_2 pair
- Interaction effect between V_1 and Obj: semantic 'closeness' between V_1 and Obj, level of lexicalisation of the V_1-Obj pair
- Interaction effect between V_2 and Obj: whether Obj can serve as the subject or object of V_2
- Interaction effect among all three: construction-specific idiosyncracies

2. Methodology

2.1. Data collection

This short paper cannot explore all the effects outlined above; instead, we will examine them by presenting data from three classes of Cantonese complex predicates: those with the highly grammaticalised *dou2* as V_2, the resultative construction, and constructions with a highly co-lexicalised V_1/V_2 pair. In each class, we will perform exploratory data analysis based on the frequency counts obtained.

We looked for objects according to a simplified version of the Silverstein hierarchy (Silverstein, 1987) that excludes plural referents:

(6) First-person singular pronoun (*ngo5*) > Second-person singular pronoun (*nei5*) > Third-person singular pronoun (*keoi5*) > Singular kinship terms/Proper nouns (*aa3* X) > Singular definite NPs (*go3* X) > Singular indefinite NPs (*jat1 go3* X)

All data came from Google search and were based on the number of search results obtained. For example, to investigate speakers' preference for the object-medial *gin3 ngo5 m4 dou2* viz-à-viz the object-final *gin3 m4 dou2 ngo5*, the search string "見我唔到" or "見唔到我" was used, and each token of each construction type was counted.

For full NPs, the wildcard character * was used to aid in searching. For kinship terms and proper nouns, we used the morpheme *aa3* along with the wildcard character, yielding kinship terms and terms of endearment like *Aa3 Ming4* for someone called Ming; thus "見阿*唔到" or "見唔到阿*" was used to look for tokens of *gin3 aa3 X m4 dou2* (see AA * NEG DOU) and its object-final coutnerpart. For definite NPs, we used the bare generic classifier *go3* with the wildcard character, which generally results in definite NPs (Cheng & Sybesma, 2005). For indefinite NPs, we used the generic classifier *go3* with the numeral *jat1* ('one').

Several kinds of tokens were not included in our counts. Firstly, in the object-final case, if the object in the search string is not the true object of the sentence (but, say, the possessor of some head noun or the first element of a compound), then the token is discounted. For example, *gin3 m4 dou2 ngo5 aa3maa1* (see NEG DOU 1sg mother, 'to not be able to see my mother') and *gin3 m4 dou2 ngo5 sik6 hoeng1ziu1* (see NEG DOU 1sg eat banana, 'to not be able to see me eat bananas') are not counted as instances of *gin3 m4 dou2 ngo5* (see NEG DOU 1sg, 'to not be able to see me').

Secondly, when V_1 takes more than one argument (7a-b), or when there is an additional 'complement' verb V_3 which carries less semantic weight than V_1 (7c), the object-medial order seems categorically ungrammatical:

(7) a. **gaau3 keoi5 m4 dou2 pou2tung1waa2*
teach 3sg NEG DOU Putonghua

'to not be able to teach him Mandarin'

b. **gaau3 keoi5 m4 dou2 jung6 Excel*
teach 3sg NEG DOU use Excel

'to not be able to teach him to use Excel'

c. **bong1 keoi5 m4 dou2 sau2*
help 3sg NEG DOU hand

'to not be able to lend him a helping hand'

As such, when these constructions appear in their object-final form, the tokens are not counted.

The only significant exception to this general rule is *gu2* + Obj + *m4 dou2* (guess Obj NEG DOU, 'to not be able to guess Obj'), which is acceptable even when the verb subcategorises for an additional complement clause.[5] This is likely because of a constructional idiosyncrasy. *Gu2*

[5]We did find one single example outside of *gu2*, with the *bong1 ... dou2* (help ... DOU) predicate, but this very

was found to highly prefer the object-medial order: the object-final *gu2 m4 dou2* + Obj is simply unattested when the object is pronominal and no additional arguments or verbs are present (8a). When an additional complement clause is present, both orders are possible (8c-d), though the object-final construction is then indistinguishable from a sentence with a single complement clause with an overt subject (8d):

(8) a. **Gu2 m4 dou2 nei5.*
guess NEG DOU 2sg

'to not expect someone (to do something)'

b. *Gu2 nei5 m4 dou2.*
guess 2sg NEG DOU

c. *Gu2 nei5 m4 dou2 gam3 hou2waan2.*
guess 2sg NEG DOU so fun

'to not expect you to be so much fun'[6]

d. *Gu2 m4 dou2 nei5 gam3 hou2waan2*
guess NEG DOU 2sg so fun

For consistency, however, we still discounted *all* tokens with additional arguments.

Note, however, that V_3 is only incompatible with the object-medial order when it is less prominent than the V_1. This does not exclude the possibility of including a third serial verb which is comparable in semantic importance to the first verb (and thus does not form part of the complex predicate). This is fairly common in the data:

(9) a. *ding2 keoi5 m4 seon6 m4 lou1*
endure 3sg NEG along.the.flow.of NEG work

'(Carmen Cheung) could not stand her and stopped working for her.'[7]

2.2. Data analysis

After collecting frequency data, we examined the proportion of object-medial tokens for each group (each group generally comprises all tokens with the same V_1, V_2 and Obj), which is used to estimate

rare.

[6]Taken from https://www.youtube.com/watch?v=5q8b2zhkHJY, accessed 4/27/2017.

[7]Taken from https://www.youtube.com/watch?v=nG_9ZkEyfbk, accessed 4/27/2017.

the probability that a speaker will choose the object-medial order:

(10) a. Parameter: P(Object-medial—V_1, V_2, Obj)

b. Estimator: $n_{medial}/(n_{final} + n_{medial})$ (number of object-medial tokens divided by the total number of tokens per construction)

To estimate the probabilities with greater certainty, exact Clopper-Pearson confidence intervals were used because the counts of some treatment groups were insufficient for the Central Limit Theorem to justify standard normal confidence intervals.[8] We also corrected for multiple testing within each group of constructions. Since the sample proportions are assumed independent, all confidence intervals have been corrected by Šidàk's method - i.e. when there are g intervals, all intervals are given at the $100\%(0.9^{1/g})$ confidence level - yielding slightly sharper intervals than Bonferroni's method at a family confidence level of 90%.

In addition, when we examined predicates with the V_2 *dou2*, we used Fisher's exact test with the Bonferroni correction, a more conservative method, along with bootstrap resampling, a less conservative method, for pairwise comparison between different V_1s; the reason for doing these comparisons will shortly be clear.

All the statistical results are presented in the appendix; we shall describe the more important results in the next section.

3. Three sample construction types

3.1. Highly grammaticalised V_2: dou2

Some V_2s are so frequent and grammaticalised that they act like particles in English phrasal verbs, like *up* in *write it up / write up the paper*. These are known as satellites (Talmy, 2007). One such verb in Cantonese, *dou2* (lit 'arrive', glossed DOU), denotes success, similar to the French periphrastic *arriver à* + V construction, though more grammaticalised. In fact, it is so semantically bleached that it probably cannot be considered a verb any more, but is more like a

[8] One additional advantage of this is that when only one order is observed (i.e. all observed tokens are object-medial or object-final), we can still have an interval estimate.

particle (Matthews, personal communication); however, we shall continue to refer to it as a V_2 for convenience's sake.[9]

Here, we investigate the tendency of speakers to produce the [V_1 Obj *m dou*] construction under eight values of V_1: *bong1* ('help'), *gaau3* ('teach'), *gau3* ('rescue'), *gin3* ('see'), *gu2* ('guess'), *jing2* ('photograph'), *wan2* ('look for') and *zuk1* ('catch').

When the object is a personal pronoun, both orders were widely attested. Confidence intervals for the proportions can be found in (23) of the appendix. Overall, the object-medial order seems to be the most common when the object is in the second person. The value of the point estimate $\hat{p}$ is higher than 0.5 for all verbs but *gaau3* ('teach'). The lower bounds of these verbs' confidence intervals lie above 0.25, and two of the verbs (*gu2* and *gau3*) have lower bounds above 0.5. By contrast, the third-person pronoun is least likely to occur in the medial position: apart from *gu2* ('guess'), all of the proportions seem to lie near 0.5. (As we have mentioned above, *gu2* seems to categorically prefer the object-medial order when the object is pronominal.)

Perhaps the most interesting case is the first-person pronoun. The most variable V_1s seem to be verbs taking benefactive objects (*bong1* 'help', *gau3* 'rescue', henceforth *benefactive verbs*) and those related to capturing (*jing2* 'photograph', *wan2* 'look for' and *zuk1* 'catch', henceforth *capturing verbs*). When the object is the second-person pronoun, the proportions of [V_1 Obj *m dou*] are high for both groups of verbs; when the object is the third-person pronoun, the proportion is fairly low for both. Yet when the object is the first-person pronoun, predicates with benefactive verbs have a far lower tendency to be object-medial than those with capturing verbs. Both Fisher's exact test and bootstrap resampling show this difference to be highly significant between *bong1* and *gau3* of the benefactive group and and *wan2* and *zuk1* of the capturing group (see (24) and (25) of the appendix for p-values obtained).

Full nominals (kinship terms/proper nouns, definite NPs, indefinite NPs) are rarely fronted. The only example found that does not involve the verb *gu2* is the following:

(11)	*baan6*	*gin3*	*go3*	*gwai2-po4*	*m4*	*dou2*
	pretend	see	CL	ghost-old.lady	NEG	DOU

[9] An anonymous reviewer pointed out that *dou2* cannot be considered a verb based on its tone, as it never occurs with tone 2 in unambiguously verbal contexts.

'(I) pretended not to see the white woman.'[10]

3.2. The resultative

The second set of constructions examined was the resultative, where the V_2 has the object of V_1 as its subject and indicates the result of V_1. We examined four V_2s:

(12) Resultative predicates studied:

V_1	V_2
saat3 ('kill'), *ngo6* ('hungry'), *daa2* ('hit'), *haak3* ('scare'), *siu3* ('laugh')	*sei2* ('die')
fing6 ('fling'), *daa2* ('hit'), *mit1* ('pinch'), *sai2* ('wash'), *caat3* ('rub')	*lat1* ('come off')
ji1 ('heal'), *gaau3* ('teach'), *zou6* ('do'), *se2* ('write'), *hok6* ('learn')	*hou2* ('good')
fong3 ('release'), *daa2* ('hit'), *gaai2* ('untie'), *geu6* ('wrench'), *teoi1* ('push')	*hoi1* ('open')

Perhaps because of semantic reasons, very few constructions with pronominal objects were found, apart from those with the X ... *sei2* and *ji1* ... *hou2* predicates. Because of the high semantic similarity among the *sei2* predicates and apparent homogeneity in the data obtained, we treated all of the predicates as one homogeneous group: we combined all the predicates with *sei2*, regardless of V_2, and estimated the proportions of object-medial tokens for each object type. Relevant statistics are presented in (26) of the appendix.

When the V_2 is *sei2* ('die') and the object is pronominal, the objects are overwhelmingly in medial position, regardless of person. The case of *ji1* + Obj + *m4 hou2* is more equivocal, and this time, the third-person pronoun is the most likely to appear in medial position.

Regarding the rest of the constructions, there seems to be evidence that the object-medial *fing6 keoi5 m4 lat1* (eight tokens) and *se2 keoi5 m4 hou2* (nine tokens) are significantly preferred to the object-final *fing6 m4 lat1 keoi4* and *se2 m4 hou2 keoi5* (both unattested; Clopper-Pearson: (0.525543, 1) for *fing6 keoi5 m4 lat1* and (0.5644842, 1) for *se2 keoi5 m4 hou2*). Moreover, if we

[10]Taken from http://www.discuss.com.hk/viewthread.php?tid=26290759, accessed 4/19/2017

combine the data from the three persons, the construction *fong3 m4 hoi1* + Obj (11 tokens) is almost significantly more frequent than its object-medial counterpart (1 token) (Clopper-Pearson: (0.4996252, 0.9994696)).

For full NP objects (kinship terms, definite and indefinite NPs), the object-medial order is never attested, but we did find a good number of tokens with the object-final order. For such objects, then, the object-medial order seems categorically bad.

3.3. Highly lexicalised V_1-V_2 pair

Finally, we chose four predicates which are relatively lexicalised, in that they exhibit syntacto-semantic idiosyncrasies:

(13) Lexicalised predicates examined:

V_1	V_2	Meaning
ceoi1 ('blow')	*zoeng3* ('inflate')	to be helpless about (someone else)
ding2 ('endure')	*seon6* ('along the flow of')	to stand (someone's temper, etc.)
seon3 ('trust')	*gwo3* ('pass')	to trust (someone)
tai2 ('look')	*hei2* ('lift')	to look up to (someone)

The first two are highly idiomatic, as is clear from the disparity between the meanings of the individual verbs and the entire predicate. *Hei2* usually means 'finish' as a grammaticalised V_2, but here, the entire predicate *tai2 ... hei2* means 'to look up to (someone)'. The *seon3 ... gwo3* predicate is the least lexicalised of the four, but even it exhibits idiosyncrasies: *seon3* is probably the only V_1 that can take an non-topicalised object when *gwo3* is present. Consider *ni1 gin6 si6 zou6 m4 gwo3* (DEM CL thing do NEG GWO, 'this thing is not worth doing'), which is grammatical, and **zou6 m4 gwo3 keoi5* (do NEG GWO 3sg, 'it is not worth doing it'), which is not.[11]

Unlike the case of *dou2* or resultatives, there seems to be no categorical rule against the object-medial order when the object is a kinship term, proper noun or definite NP, and thus, we include confidence intervals for the first five points on our simplified Silverstein hierarchy in (24). From these

[11] Constructions like *daa2 m4 gwo3 keoi5* (hit NEG surpass 3sg) are attested on Google, but the *gwo3* does not carry the same meaning.

statistics, it is clear that *ceoi1 ... zoeng3* must always be object-medial regardless of object.[12] *Ding2 ... seon6* has a strong but non-categorical preference for the object-medial order for pronominal objects, and prefers the object-final order for full NP objects. *Seon3 ... gwo3* and *tai2 ... hei2* have no clear preference for pronominal objects, and prefer the object-final order for full NP objects.

Indefinite object NPs are still unattested at medial position; they are always found in final position. Even the *ceoi1 ... zoeng3* predicate, which categorically prefers the object-medial order for other NP types, was no exception:

(14)

Seng4	*go3*	*zing3fu2*	*dou1*	*ceoi1*	*m4*	*zoeng3*	*jat1*	*go3*	*Gwok3*	*Siu3*
whole	CL	government	FOC	blow	NEG	inflate	one	CL	Kwok	Siu

Git6?
Kit?

'The whole government could not do anything about one Alex Kwok?'

4. Discussion

4.1. Synchronic explanations

A synchronic explanation of the object preposing phenomenon may be able to account for syntactic variation through *functional motivations* for each of the two construction types. Each construction type should serve some purpose, such as facilitating the production or comprehension of the construction.

An obvious motivation in favour of the object-medial order is phonological weight. Heavier constituents are widely known to prefer the sentence-final position Hawkins (1994). The negated V_2 takes up two syllables, so it is preferable to append the NEG + V_2 pair to the end of the verb phrase. Yet this hypothesis predicts that if an expletive like *lan2* ('fucking') or *gwai2* ('bloody', lit. *ghost*) intrudes between the negative *m4* and the second verb, the object-medial order will be even more preferred - a prediction that seems false. A Google search for *X keoi5 m4 lan2 dou2* (X 3sg NEG fucking DOU) and *X keoi5 m4 gwai2 dou2* (X 3sg NEG bloody DOU) returns very few results, all of them with the verb *gu2* which, as we have seen above, categorically prefers the

[12] In fact, several instances of *ceoi1 m4 zoeng3* + definite NP were found; however, these constructions did not carry the idiomatic meaning we are targeting here, and were not counted.

object-medial order. By contrast, *X m4 lan2 dou2 keoi5* (X NEG fucking DOU 3sg) and *X m4 gwai2 dou2 keoi5* (X NEG bloody DOU 3sg) yield a fair number of results with a variety of V_1 types, such as *kau1* ('pursue (a girl)'), *teng1* ('hear'), *jing2* ('recognise'), *wan2* ('look for'), *bei2* ('give'), *zo2* ('obstruct'), *bong1* ('help') and *daa2* ('hit'). Consider the following examples:

(15) a.

Ngo5	*tung4*	*aa3mui2*	*heoi3*	*dou3*	*keoi5*	*hok6haau6*	*go2zan2,*	*[wan2*
1sg	CONJ	little.sister	go	DOU	3sg	school	TEMP,	find

m4	*gwai2*	*dou2*	*keoi5].*
NEG	bloody	DOU	3sg

'When my little sister and I went to her school, we could not bloody find her.'

b.

Daan6	*nei1*	*jat1*	*fong1min6*	*nei5*	*jau6*	*[bong1*	*m4*	*lan2*	*dou2*
but	DEM	one	aspect	you	also	help	NEG	fucking	DOU

keoi5]	*wo3.*
3sg	MIR.

'But you can't fucking help her in this regard.'

Since the influence of phonological weight seems small, a better phonological motivation is foot binarity Prince (1980); McCarthy & Prince (1995). Assuming that all elements are monosyllabic, [V_1 Obj *m* V_2] can be parsed as [(V_1 Obj) (*m* V_2)], where parentheses indicate foot boundaries. [V_1 *m* V_2 Obj], by contrast, would be parsed as [(V_1) (*m* V_2) (Obj)] to produce two degenerate feet, which is dispreferred for phonological reasons. If it were instead parsed as [(V_1 *m*) (V_2 Obj)], then there would be a mismatch between prosodic and syntactic parsing, an undesirable consequence for comprehension.

If this motivation were correct, then we would expect the effect to disappear when V_1 is disyllabic and the rest of the constituents remain monosyllabic, since both orders create a degenerate foot ([(V_1) (*m* V_2) (Obj)] vs. [(V_1) (Obj) (*m* V_2)]). Indeed, an informal Google search with the disyllabic verbs *faat3jin6* ('discover'), *lei5gaai2* ('understand') and *dung6jiu4* ('shake (of someone's goals, willpower, etc.)') as V_1, *dou2* as V_2 and *keoi5* (3sg) as Obj reveals that these verbs overwhelmingly prefer the object-final order: only one example of *faat3jin6 keoi5 m4 dou2* and one of *dung6jiu4 keoi5 m4 dou2* were found, and *lei5gaai2 keoi5 m4 dou2* was unattested.

Another possible motivation for object fronting in Cantonese is to highlight a topical object. We have seen that objects higher on the referential hierarchy are more likely to be preposed. With the exception of benefactive verbs, it can be seen from *dou2* predicates that the third-person singular

pronoun is less likely to be fronted than the the first or second persons, full NPs are less likely to be fronted than pronouns, and indefinite noun phrases are never fronted. This suggests a topicality effect. Consider the following pair of sentences with *seon3 ... gwo3*. (16a) has *keoi5* as the object and *keoi5* is acting as a resumptive pronoun for a left-dislocated proper noun, whereas (16b) has a definite NP as an object which has not appeared in previous discourse (it was likely definite by virtue of its presence in the shared 'aeroplane emergency' schema):

(16) a.

Wo4gei3	*zan1*	*hai6*	*seon3*	*keoi5*	*m4*	*gwo3!*
HGC	real	is	trust	3sg	NEG	pass

'HGC really can't be trusted!'[13]

b.

ngo5	*gok3dak1*	*go3*	*gei1si1*	*seon3*	*m4*	*gwo3*	*go3*	*tou4sang1*
1sg	feel	CL	pilot	trust	NEG	pass	CL	escape.for.life

daan6se6	*hai6tung2...*
eject	system...

'I feel the pilot did not trust the emergency ejection system...'[14]

A slightly different interpretation of such a phenomenon is may hold that the *animacy* of objects favours fronting. This may be particularly relevant for the 'capturing' verbs and *saat3 ... sei2* (kill ... die), as many of the example sentences imply that the referent of the object is actively trying to avoid being affected by the action denoted V_1. Consider this caption of a video showing a footballer running around in circles to avoid another footballer, who was chasing him:

(17)

Tuen4tuen4zyun3,	*guk1faa1*	*jyun4!*	*Nei5*	*zuk1*	*ngo5*	*m4*	*dou2,*
going.round.in.circles	chrysanthemum	garden	2sg	catch	me	no	DOU

nei5	*zuk1*	*ngo5*	*m4*	*dou2!*
2sg	catch	me	no	DOU

'Going round in circles, chrysanthemum garden! You can't catch me, you can't catch me!'[15]

In contrast to motivations for the object-medial order, the motivation for the object-final order is fairly obvious. The two components of the complex predicate are conceptually close, and so placing them together creates an isomorphism between formal proximity and conceptual proximity

[13]Taken from http://www.hkepc.com/forum/viewthread.php?fid=12&tid=1505304, accessed 5/1/2017.
[14]Taken from http://hkgalden.com/view/106390/page/, accessed 5/1/2017.
[15]Taken from https://www.facebook.com/footballnewsexpress/videos/956113834427925, acessed 4/19/2017.

(Givón, 2001). For instance, *dou2* expresses the completeness and success of the V_1. This also explains why *dou2*-predicates favours object fronting more than resultatives do.

However, the synchronic account fails to account for idiosyncrasies. Why, for instance, does object preposing seem to be far more frequent in certain lexicalised predicates like *ding2 ... seon6* and *ceoi1 ... zoeng3*? If anything, we would expect the two verbs in these predicates to be conceptually close, and thus more likely to be adjacent to each other. A diachronic explanation may be necessary to explain some of the data.

4.2. Diachronic explanations

One of the most detailed elucidations of the diachrony of complex predicates is Givón (2009). Givón gives two main diachronic pathways of development: the V_2 may develop from a VP complement to another verb, or from the weaker verb of a serial verb construction through grammaticalisation or colexicalisation. He presents a three-stage diachronic process for the latter pathway:

(18) a. Stage I - Paratactic source: Main clause + chained complement clause
b. Stage II - Complex clause: Serial-verb clause
c. Stage III - Complex word: Complex verbal word

Givón also writes that in verb-serialising languages, it is harder for constructions to fully develop into complex predicates because of object intrusion between the two verbs. For instance, in Saramaccan, a VO language, the object-medial order is used with the V_2 *kaba* ('finish') (Givón, 2009):

(19)

a	*bi-fefi*	*di-wosu*	*kaba.*
3sh	TNS-paint	DEF-house	finish

'He finished painting the house.' (Saramaccan, Givón (2009))

It appears that Cantonese is passing from Stage II to III, i.e. going through the process of converting serial verb constructions into fully-fledged complex predicates through grammaticalisation and lexicalisation. During this process, the object must be moved to the end of the construction to allow for adjacency between the two verbs. Cantonese is more conservative than Mandarin

but more innovative than Yiyang in this regard. Unlike Mandarin, where the two verbs of the resultative have been inseparable since the fifteenth century (Shi, 2002), the older object-medial order does remain in Cantonese. But unlike Yiyang, as discussed above, the object-medial order in Cantonese has become the more marked order already. In the future, to complete this diachronic process, Cantonese speakers may generalise the object-final order to the remaining constructions through *extension* (Harris & Campbell, 1995), i.e. applying the syntactic rule governing most complex predicate constructions to outliers like *saat3* X *m4 sei2* and *ding2* X *m4 seon6*, at which stage the complex predicates can be said to be V-V compounds.

This agrees with known facts about the development of Chinese. Historically, at around the time of the Sui and early Tang Dynasties, complex-predicate constructions were nearly exclusively object-medial (Shi & Li, 2004):[16]

(20) a. *Xunmi* *Aniang* *bu* *jian.*
find mum no see

'(Mulian) could not find his mother.' (*Damuqianlian mingjian jiu mu bian wen*, Shi & Li (2004))

b. *Quan* *sui* *zha* *she* *si* *yan.*
dog then bite snake die SFP.

'The dog then bit the snake, killing it.' (*Taiping Guangji* 43, Shi & Li (2004))

It thus appears not unlikely that the object-medial order in modern Cantonese - and by extension in other Chinese varieties exhibiting this alternation - are vestiges of this older word order.

The fact that highly lexicalised pairs discussed in 3.3 prefer the object-medial order is thus easily explained by frequency: As is well known, more frequent forms are more resistant to change (Croft, 2002), and as lexicalised V_1-V_2 pairs are more likely to co-occur, they also tend to preserve the older word order.

The reason why the order is better preserved in negative clauses than affirmative ones is clear too: Negative clauses are more resistant to change because of their greater cognitive complexity (Givón, 1979). Looking at the predicates with very strong preference for the object-medial order, many seem to appear fairly rarely in the affirmative. For instance, *gu2 keoi5 m4 dou2 laa1* (guess

[16] Object-final constructions did exist before Sui-Tang; the object-medial order appears to have reached its apex during the Tang Dynasty because of the rigidification of argument structure Shi & Li (2004).

3sg no DOU SFP) returns seven pages of results on Google; *gu2 dou2 keoi5 laa1* returns merely eight results. Similarly, *ceoi1 dak1 keoi5 zoeng3* (blow DAK 3sg inflate), the affirmative version of the idiom discussed above, yields only five pages, whereas its negative counterpart returns 34. This is perhaps no coincidence: If a predicate is more frequent in the negative form, then it is less likely for a speaker to generalise the verb-final order from the more innovative positive form to the more conservative negative form. Future research can further explore this hypothesis by regressing the relative frequency of object-medial tokens on the relative frequency of the negative form.

5. Conclusion and directions for future research

In this paper, we have shown that the object position alternation, rather than being completely random and unpredictable (in 'free variation'), does follow rules. Some of the rules are categorical, but most are gradient. This supports the views of writers who view grammatical competence as stochastic rather than deterministic (Bresnan et al., 2001; Manning, 2003; Bresnan, 2007).

In developing hypotheses that can explain the data collected, we made extensive use of phonological, semantic and pragmatic factors. This is in line with the position that grammar, far from being an isolated system, is actually shaped by other aspects of language, with semanto-pragmatic concepts like iconicity and topicality underlying syntactic choice (Givón, 2001; Croft, 2002).

Moreover, we discovered construction-specific idiosyncrasies such as the *gu2* + Obj + *m4 dou2* construction, suggesting that Cantonese complex predicates are best treated as a family of constructions (Goldberg & Jackendoff, 2004) rather a single group. This supports, in Goldberg and Jackendoff's terms, a 'constructional' view of grammar: there is a cline of grammatical phenomena ranging from general to idiosyncratic, and there's no sharp divide between lexicon and syntax.

Methodologically, we have demonstrated the value of nonparametric statistical methods such as the bootstrap and exact methods in investigating linguistic phenomena with sparse data. They allowed us to make inferences about grammaticality preferences, without relying on theoretical asymptotic assumptions like the Z-test and Pearson's χ^2-test do.

It must be noted that the observations made and conclusions drawn in this paper are necessarily tentative. As the phenomenon is rather poorly understood at present, our analysis of the data had to be exploratory in nature: we conducted *post-hoc* tests and constructed *post-hoc* confidence

intervals. Future research can take the points discussed above as starting points and subject them to rigorous testing in a hypothetico-deductive framework, whether with corpora or experimental work.

Furthermore, the alternation phenomenon is highly complex, and we have only touched on the tip of the iceberg: we investigated just a small inventory of monosyllabic verbs and three construction types. Future research can expand the range of constructions covered. There remains a large number of grammaticalised V_2s we are yet to cover, such as *saai3* ('all'), *maai4* ('as well'), *faan1* (literally 'return'; see Chor (2013)) and *cit3* ('on time'). It is also interesting to compare the resultative constructions and the corresponding constructions with the causative *zing2* ('make') as V_1.

In addition, it is worth investigating analogous alternations in the affirmative, whether involving the particle *dak* (21) or not (22):

(21) a. *daa2* *dak1* *keoi5* *sei2*
hit DAK 3sg die

'to be able to beat him to death'

b. *daa2* *dak1* *sei2* *keoi5*
hit DAK die 3sg

(22) a. *fong3* *keoi5* *zau2*
release 3sg go

'to release him'

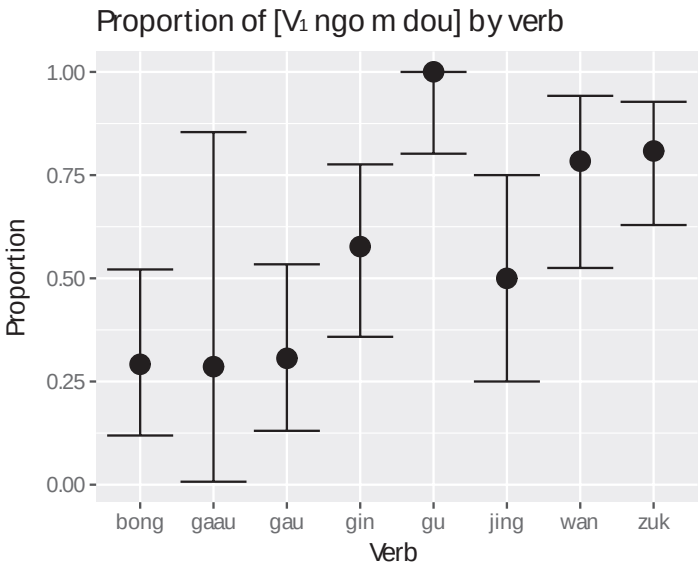
b. *fong3* *zau2* *keoi5*
release go 3sg

Ultimately, it is hoped that a complete model of Cantonese complex predicates can be built, taking into account all relevant factors and discarding all irrelevant ones.

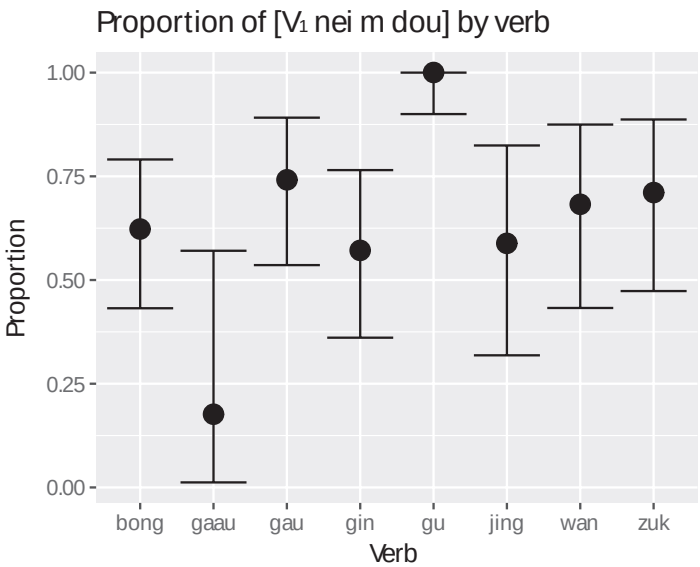
Appendix: Statistical results

Statistics for the dou construction

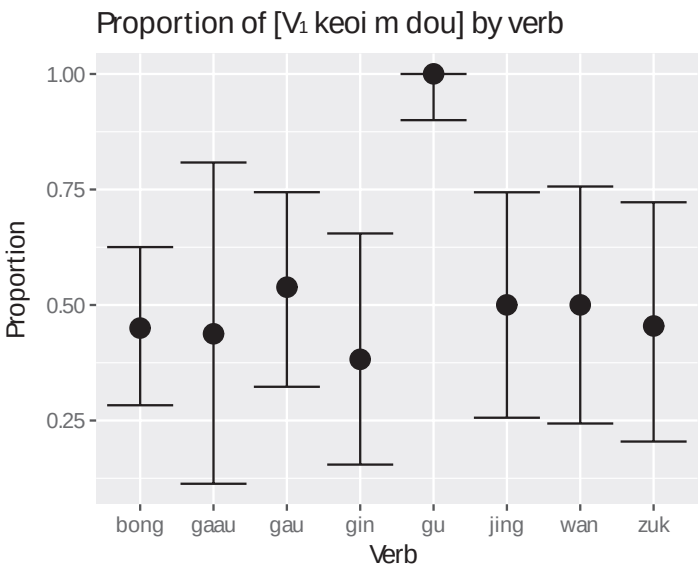
(23) Šidàk-corrected ($g = 24$) 90% Clopper-Pearson confidence intervals for the proportion of [V_1 Obj m dou] by verb:



x	$\hat{p}$	l	u	n
bong	0.29167	0.11917	0.52146	48
gaau	0.28571	0.00722	0.85423	7
gau	0.30612	0.13055	0.53394	49
gin	0.57692	0.35830	0.77602	52
gu	1.00000	0.80194	1.00000	31
jing	0.50000	0.24999	0.75001	36
wan	0.78378	0.52509	0.94213	37
zuk	0.80882	0.62904	0.92764	68



x	$\hat{p}$	l	u	n
bong	0.62319	0.43193	0.79067	69
gaau	0.17647	0.01213	0.57057	17
gau	0.74138	0.53595	0.89142	58
gin	0.57143	0.36099	0.76499	56
gu	1.00000	0.90008	1.00000	65
jing	0.58824	0.31852	0.82440	34
wan	0.68293	0.43262	0.87481	41
zuk	0.71111	0.47340	0.88705	45



x	$\hat{p}$	l	u	n
bong	0.45000	0.28301	0.62534	80
gaau	0.43750	0.11309	0.80843	16
gau	0.53846	0.32300	0.74416	52
gin	0.38235	0.15458	0.65487	34
gu	1.00000	0.90008	1.00000	65
jing	0.50000	0.25601	0.74399	38
wan	0.50000	0.24352	0.75648	34
zuk	0.45455	0.20447	0.72239	33

(24) Bonferroni-corrected ($g = 28$) p-values obtained by pairwise comparison of the tendency to use the object-medial order with *ngo* using Fisher's exact test:

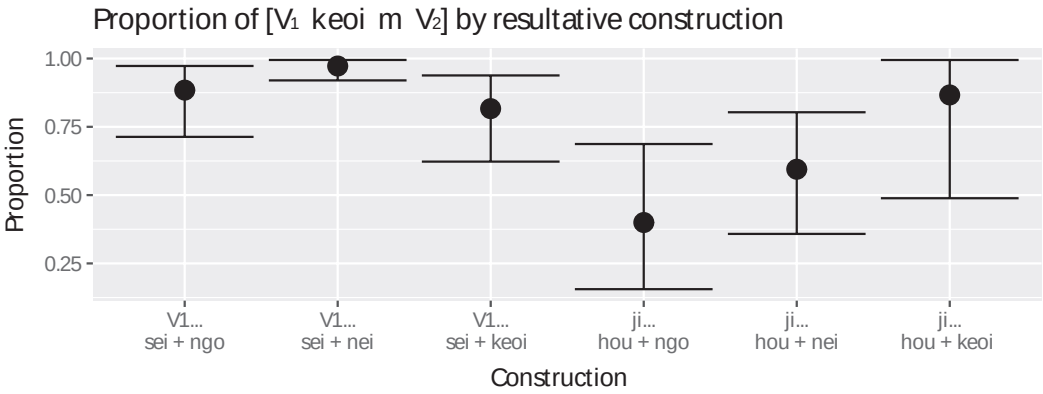
	bong	gaau	gau	gin	gu	jing	wan
gaau	1	-	-	-	-	-	-
gau	1	1	-	-	-	-	-
gin	0.13963	1	0.25274	-	-	-	-
gu	0***	0.00117***	0***	0.00013***	-	-	-
jing	1	1	1	1	0.00003***	-	-
wan	0.00024***	0.4774	0.00034***	1	0.17612	0.41978	-
zuk	0***	0.20828	0***	0.22745	0.22727	0.04405*	1

(25) p-values obtained by bootstrap (30000 resamples):

	bong	gaau	gau	gin	gu	jing	wan
gaau	0.92140	-	-	-	-	-	-
gau	0.88046	0.88080	-	-	-	-	-
gin	0.00366***	0.15600	0.00560***	-	-	-	-
gu	0***	0.01280**	0***	0***	-	-	-
jing	0.05806*	0.29900	0.07314	0.47950	0***	-	
wan	0***	0.02660**	0***	0.03820*	0***	0.01114**	-
zuk	0***	0.02086**	0***	0.00680***	0***	0.00166***	0.77760

Statistics for the resultative

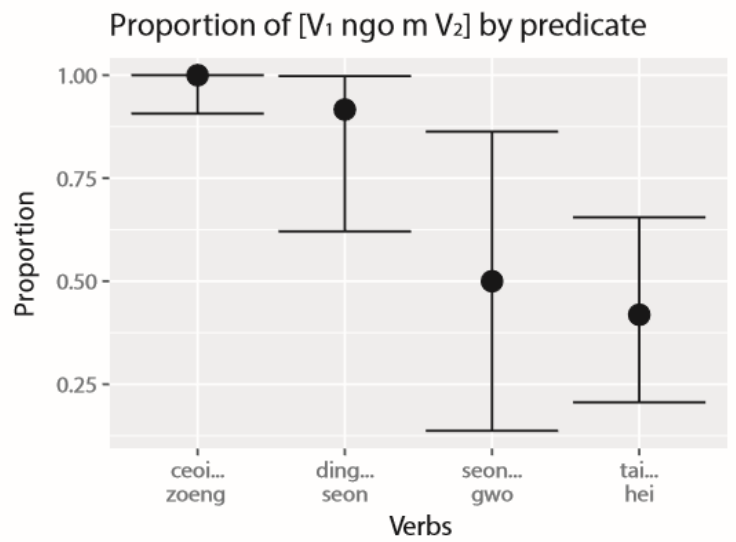
(26) Šidàk-corrected ($g = 9$) 90% Clopper-Pearson confidence intervals for the proportion of object-medial tokens by resultative construction:



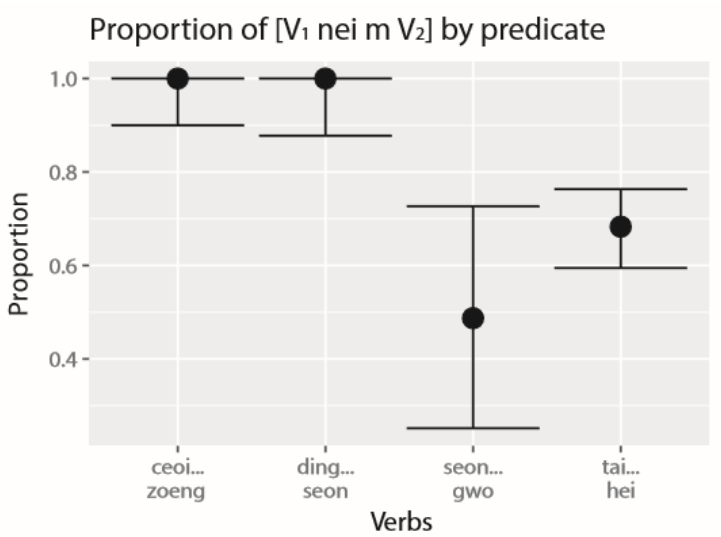
x	$\hat{p}$	l	u	n
V_1 ... *sei* + *ngo*	0.8846154	0.7131791	0.9727622	52
V_1 ... *sei* + *nei*	0.9725275	0.9199488	0.9947819	182
V_1 ... *sei* + *keoi*	0.8163265	0.6226495	0.9381407	49
ji ... *hou* + *ngo*	0.4000000	0.1556585	0.6870195	25
ji ... *hou* + *nei*	0.5945946	0.3579784	0.8033468	37
ji ... *hou* + *keoi*	0.8666667	0.4885575	0.9946766	15
fing ... *lat* + *keoi*	1.0000000	0.5255543	1.0000000	8
se ... *hou* + *keoi*	1.0000000	0.5644842	1.0000000	9
fong ... *hoi* + any Obj	0.0909091	0.0005304	0.5003748	12

Statistics for lexicalised V_1-V_2 pairs

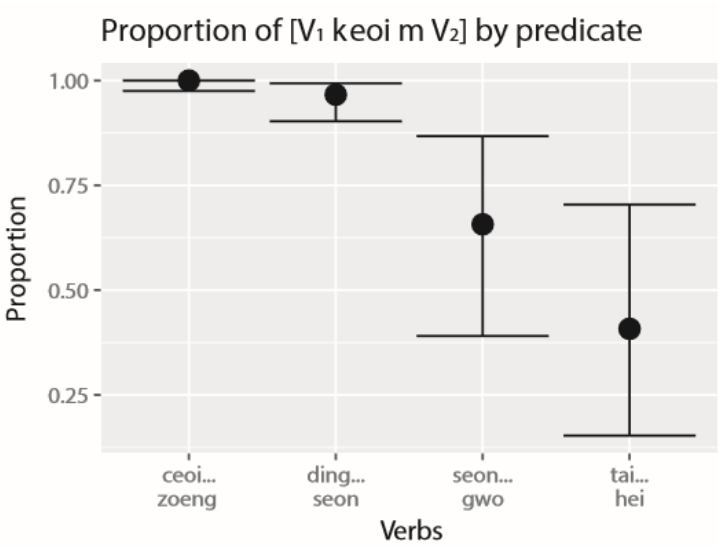
(27) Šidàk-corrected ($g = 20$) 90% Clopper-Pearson confidence intervals for the proportion of [V_1 Obj *m* V_2] by construction:



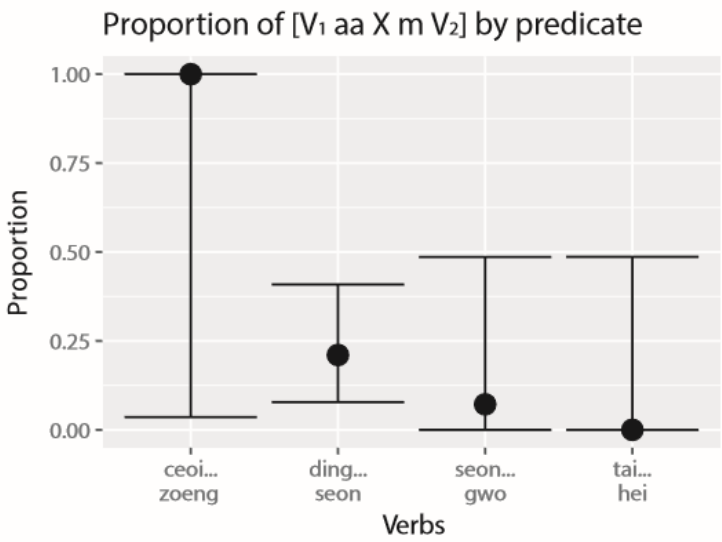
x	$\hat{p}$	l	u	n
ceoi... zoeng	1.00000	0.90670	1.00000	68
ding... seon	0.91667	0.62098	0.99781	24
seon... gwo	0.50000	0.13701	0.86299	14
tai... hei	0.41860	0.20622	0.65503	43



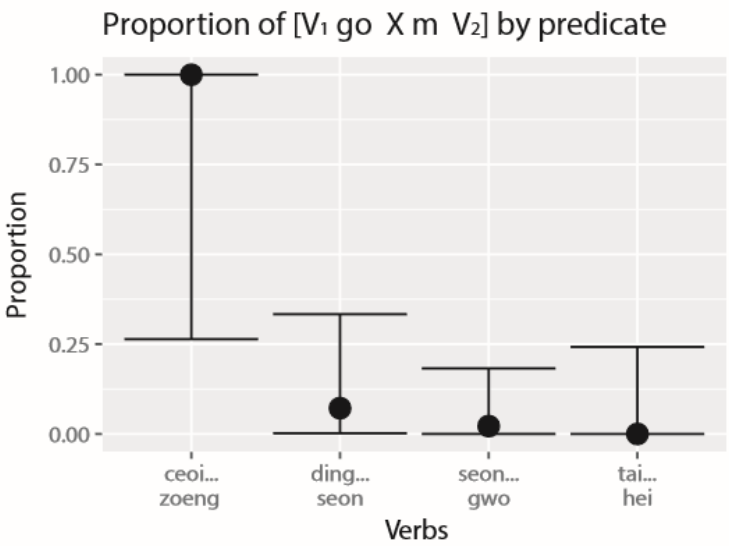
x	$\hat{p}$	l	u	n
ceoi... zoeng	1.00000	0.89968	1.00000	63
ding... seon	1.00000	0.87757	1.00000	51
seon... gwo	0.48718	0.25178	0.72673	39
tai... hei	0.68310	0.59459	0.76321	284



x	$\hat{p}$	l	u	n
ceoi... zoeng	1.00000	0.97499	1.00000	263
ding... seon	0.96591	0.90272	0.99332	176
seon... gwo	0.65714	0.39056	0.86722	35
tai... hei	0.40741	0.15300	0.70413	27



x	$\hat{p}$	l	u	n
ceoi... zoeng	1.00000	0.03579	1.00000	2
ding... seon	0.21053	0.07790	0.40882	57
seon... gwo	0.07143	0.00009	0.48595	14
tai... hei	0.00000	0.00000	0.48626	10



x	$\hat{p}$	l	u	n
ceoi... zoeng	1.00000	0.26393	1.00000	5
ding... seon	0.07143	0.00187	0.33343	28
seon... gwo	0.02222	0.00003	0.18239	45
tai... hei	0.00000	0.00000	0.24234	24

References

Arppe, A. (2009). Linguistic choices vs. probabilities–how much and what can linguistic theory explain? In S. Featherston & S. Winkler (Eds.), *The fruits of empirical linguistics, Vol. 1: Processes*, Studies in Generative Grammar (pp. 1–24).

Bresnan, J. (2007). Is syntactic knowledge probabilistic? experiments with the english dative alternation. In S. F. W. Sternefeld (Ed.), *Roots: Linguistics in search of its evidential base*, Studies in Generative Grammar (pp. 75–96). Berlin; Boston: De Gruyter Mouton.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In *Cognitive foundations of interpretation* (pp. 69–94). Amsterdam: Royal Netherlands Academy of Science.

Bresnan, J., Dingare, S., & Manning, C. D. (2001). Soft constraints mirror hard constraints: Voice and person in English and Lummi. In Butt, M. & King, T. H. (Eds.), *Proceedings of the LFG01 Conference*, (pp. 13–32). Stanford: CSLI Publications.

Cheng, L. L.-S. & Sybesma, R. (2005). Classifiers in four varieties of Chinese. In *The Oxford Handbook of comparative syntax* (pp. 259–292).

Chor, W. (2013). From 'direction' to 'positive evaluation': On the grammaticalization, subjectification and intersubjectification of faan1 'return' in Cantonese. *Language and Linguistics*, *14*(1), 91.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems.* Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Croft, W. (2002). *Typology and universals.* Cambridge: Cambridge University Press.

Givón, T. (1979). *On understanding grammar.* New York: Academic Press.

Givón, T. (2001). *Syntax: an introduction*, volume 1. Amsterdam: John Benjamins Publishing.

Givón, T. (2009). *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-cognition, evolution.* Amsterdam: John Benjamins Publishing.

Goldberg, A. E. & Jackendoff, R. (2004). The English resultative as a family of constructions. *Language*, 532–568.

Gries, S. T. (2006). *Multifactorial analysis in corpus linguistics: A study of particle placement.* New York; London: Continuum.

Harris, A. C. & Campbell, L. (1995). *Historical syntax in cross-linguistic perspective*, volume 74. Cambrdige: Cambridge University Press.

Hawkins, J. A. (1994). *A performance theory of order and constituency.* Cambridge studies in linguistics. Cambridge: Cambridge University Press.

Li, R. (2013). Guanyu binyu qianzhi [On the prepositioned object]. *Hanyu Xuebao*, *2013*(4), 63–71.

Manning, C. D. (2003). Probabilistic syntax. In J. H. Rens Bod & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 289–341). Cambridge, MA: MIT Press.

Matthews, S. (2006). On serial verb constructions in Cantonese. In A. Y. Aikhenvald & R. M. V. Dixon (Eds.), *Serial Verb Constructions. A Cross-Linguistic Typology* (pp. 69–87). Oxford: Oxford University Press.

McCarthy, J. J. & Prince, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. W. Dickey, & S. Urbanczyk (Eds.), *Papers in Optimality Theory*, University of Massachusetts Occasional Papers. Amherst, MA: Graduate Linguistic Student Association.

Prince, A. S. (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry*, 511–562.

Shi, Y. (2002). *The establishment of modern Chinese grammar: The formation of the resultative construction and its effects*, volume 59. Amsterdam: John Benjamins Publishing.

Shi, Y. & Li, C. N. (2004). *Hanyu yufahua de licheng: xingtai jufa fazhan de dongyin he jizhi [The Historical Morphosyntax of Chinese].* Beijing: Beijing University Press.

Silverstein, M. (1987). Cognitive implications of a referential hierarchy. In *Social and functional approaches to language and thought* (pp. 125–164). Orlando, FL: Academic Press.

Talmy, L. (2007). Lexical typologies. In T. Shopen (Ed.), *Language Typology and Syntactic Description, Vol. 3: Grammatical Categories and the Lexicon* (pp. 66–168). Cambridge: Cambridge University Press.

Tang, S.-W., Kwok, F., Lee, T. H.-T., Lun, C., Luke, K. K., Tung, P., & Cheung, K. H. (2002). *Guide to LSHK Cantonese romanization of Chinese characters.* Hong Kong: Linguistic Society of Hong Kong.

Xu, H. (2001). *Yiyang Fangyan Yufa Yanjiu [A Study of the Yiyang dialect].* Changsha, Hunan: Hunan Educational Press.

Index of Authors