UCL Working Papers in Linguistics 10 (1998)

Functional heads, Cantonese phrase structure and Cantonese-English code-switching[*]

BRIAN, HOK-SHING, CHAN

Abstract

This paper is concerned with the idea that the morpho-syntax of intra-sentential code-switching is strictly governed by the grammar of the *matrix language/ML*, whereas the *embedded language/EL* only contributes certain lexical elements within confines set by ML. A particular model embodying this idea - *the Lemma Congruence model* (Myers-Scotton and Jake 1995) - is discussed. The Lemma Congruence model states that an EL word can appear in code-switching provided its *lemma* (i.e. morpho-syntactic properties) is congruent with its ML equivalents. I argue that the Lemma Congruence model fails to explain recurrent patterns in Cantonese-English code-switching, where various English (i.e. EL) forms may appear despite their "incongruence" with their Cantonese (i.e. ML) counterparts. As an alternative, I suggest that code-switching may occur provided that the c-selection requirements of functional heads are satisfied. I argue that the latter approach - which I call the *Functional Head Selection* approach - is not only superior to the Lemma Congruence model, it also corroborates the idea that intra-sentential code-switching and monolingual sentences are governed by the same principles in Universal Grammar, a line of thought which has been explored most recently by MacSwan (1997) and Mahootian (1993).

1 Introduction

The issue of universal grammatical constraints on intra-sentential code-switching has attracted a great deal of debates in the past two decades. Various proposals have been put forward based on different language-pairs spoken in different communities (e.g. Di Sciullo, Muysken and Singh 1986, Belazi, Rubin and Toribio 1994, etc.). However, these

[*]This paper is part of work in progress (Chan, to appear). I would like to thank Professor Neil Smith and Dr. Ad Neeleman for many insightful comments they made on previous drafts. All remaining errors are mine.

proposals are soon found to be inadequate with the availability of more data from different language-pairs. There seems to be no single proposal which is widely agreed upon as the definitive, universal syntactic model of code-switching. New models are still proposed constantly (e.g. MacSwan 1997).

This paper focuses on one particular school of thought regarding the syntax of intra-sentential code-switching: One language is dominant and determines the overall morpho-syntactic properties of code-switched utterances. The other language only provides certain lexical categories, and the role of its grammar is severely limited. The most popular model to date which inherits such a tradition is *the Matrix Language Frame model* as developed by Myers-Scotton (1993, 1995). In such a model, the dominant language (i.e. *the matrix language/ML*) determines word order and supplies all the functional categories in the code-switched sentences. The other language (i.e. *the embedded language/EL*) only contributes lexical elements and fits them into appropriate slots in the structure with ML characteristics. In a more up-to-date version of the Matrix Language Frame model (Myers-Scotton and Jake 1995, Jake and Myers-Scotton 1997), it is proposed that an EL word (or, more abstractly, the *lemma* of this EL word)[1] has to be *congruent* with its equivalent (lemma) in ML so that it can be inserted into the code-switched sentence.

In this paper, I argue that such a *Lemma Congruence* model is overly restrictive. As an alternative, I propose that code-switching may take place if the switched element observes the *c-selection* requirement of its functional head. I will argue that such an approach, which I will call the *Functional Head Selection* approach, is superior to the Lemma Congruence model: Empirically, I will show that the Lemma Congruence model fails to account for code-switching data from Cantonese-English. Theoretically, the Lemma Congruence model presupposes an additional process which is unnecessary, thus making the whole model uneconomical.

This paper is organized as follows. In section 2, I will outline the Lemma Congruence model and the Functional Head Selection approach. In section 3, I will briefly outline the morpho-syntactic characteristics of Cantonese and give a general picture of Cantonese-English code-switching. In section 4 to section 8, I will apply these two approaches to analyze code-switching data from Cantonese-English. Section 9 is the conclusion.

---

[1] "Lemma" is a term borrowed from Levelt's (1989) model of speech production. It refers to the cluster of syntactic-morphological properties associated with a word. Muysken and de Rooij (1995) comment that the term "lemma" is not strictly defined in the Matrix Language Frame model/Lemma Congruence model. See Muysken and de Rooij (1995) for details.

2 Approaches to the syntax of code-switching

2.1 The Lemma Congruence model

The *Lemma Congruence model* has its roots in the *Matrix Language Frame model* (Myers-Scotton 1993, 1995). Crucially, the Matrix Language model posits a distinction between *the matrix language* and *the embedded language*. The matrix language plays a dominant role in supplying more morphemes to a particular code-switching discourse and in determining the overall morphological-syntactic properties of code-switched sentences. Whenever there is code-switching within a phrasal/maximal projection (i.e. a "mixed" constituent), the word order of that constituent comes from the matrix language. This idea is formalized as the Morpheme Order principle.

(1) **The Morpheme Order principle** (Myers-Scotton 1993: 83)
"In ML + EL constituents consisting of singly-occurring EL lexemes and any number of ML morphemes, surface morpheme order (reflecting syntactic relations) will be that of ML."

In addition, the matrix language also supplies the *system morphemes*. This idea is formalized as the System Morpheme principle.

(2) **The System Morpheme principle** (Myers-Scotton, 1993: 83)
"In ML + EL constituents, all system morphemes which have grammatical relations external to their head constituent (i.e. which participate in the sentence's thematic role grid) will come from the ML."

What are *system morphemes*? According to Myers-Scotton (1993), all morphemes are divided into two classes: *content morphemes* and *system morphemes*. System morphemes are defined as having one of the following features:

(3) **Defining features of system morphemes** (Myers-Scotton 1993: 99-101)
[+Quantificational] – System morphemes are quantificational, such as quantifiers, determiners and possessive adjectives.
[-Thematic Role Assigner] – System morphemes do not assign thematic roles.
[-Thematic Role Receiver] – System morphemes do not receive thematic roles.

On the other hand, content morphemes are non-quantificational; they assign or receive thematic roles. More concretely, content morphemes cover major word classes such as nouns, verbs, pronouns, most adjectives and prepositions. In Myers-Scotton 1995, some conjunctions (e.g. *because*) are considered content morphemes as well because they assign thematic roles to clauses at the discourse level (e.g. "*because*" assigns the role CAUSE to the clause it introduces).

In a nutshell, most of the system morphemes are what are assumed to be functional categories-determiners, quantifiers, modal verbs, inflectional affixes and complementizers. Some counter-examples have been attested to the above two principles - In one case, what are classified as EL content morphemes, for example, *pronouns*, do not always appear in mixed constituents.[2] To deal with this particular problem, *the Matrix Language Blocking hypothesis* has been proposed.

(4) **The ML Blocking hypothesis** (Myers-Scotton 1993: 120)
"In ML + EL constituents, a blocking filter blocks any EL content morpheme which is not congruent with the ML with respect to three levels of abstraction regarding subcategorization."[3]

For an EL morpheme (say, a pronoun) to appear in a mixed constituent, it is not sufficient that the EL pronoun is a content morpheme. According to the ML Blocking hypothesis, the ML counterpart of this EL morpheme (say, the ML pronoun) has to be "congruent". In other words, the ML pronoun has to be a content morpheme as well. In other words, the ML Blocking hypothesis blocks EL single pronouns (even when these EL pronouns are content morphemes) if the pronouns in ML are system morphemes.[4] It is argued that the ML Blocking hypothesis successfully accounts for the absence of EL pronouns in Spanish-English (Myers-Scotton 1993: 128).

Myers-Scotton and Jake (1995) elaborate the idea of "congruence", and state that the EL morpheme must be congruent with its ML counterpart in terms of *predication-argument structure* (i.e. thematic properties), *conceptual-lexical structure* (i.e. meaning)

---

[2]Actually, it has been argued that pronoun is a functional category in D (Longobardi 1994).

[3]The three levels of abstraction are *lexical-conceptual structure*, *predicate-argument structure* and *morphological realization patterns*. See below.

[4]Myers-Scotton (1993: 126) considers two types of pronouns as [-Thematic Role Receiver] and hence system morphemes: dummy pronouns (such as *it* and *there* in English) and clitics.

and *morphological realization patterns* (word forms).

When the EL morpheme is indeed incongruent with its ML counterpart, it may still appear in mixed constituents in either of these formats - *bare forms* or *EL islands* - which Jake and Myers-Scotton (1997) call *compromise strategies*. In bare forms, the EL morpheme appears in mixed constituents without being inflected at all. In EL islands, the EL morpheme triggers off EL morpho-syntactic procedures and ends up with an EL maximal projection. I will call this latest version of the Matrix Language Frame model *the Lemma Congruence model*.

To summarize, the Lemma Congruence model predicts that for an EL word to appear in a mixed constituent, it (or its lemma) must be sufficiently "congruent" with its ML equivalent in terms of semantic, syntactic and morphological properties (namely, lexical-conceptual structure, predication-argument structure, and morphological realization patterns). Under this approach, there has to be an additional process which compares the EL morpheme with its ML counterpart for congruence before it is inserted into an ML-framed structure.

2.2 The Functional Head Selection approach

There is another approach which explicates the role of functional categories in code-switching. Based on their data from Tunisian Arabic-French and Spanish-English, Belazi, Rubin and Toribio (1994) observe that code-switching does not take place between functional heads and their complements. These authors seek to explain this observation by resorting to the close grammatical relationship between functional heads and their complements, as pointed out by previous work done on monolingual syntax (in particular, Abney 1987). It is proposed that a functional head not only "f-selects" specific complements with particular syntactic properties (such as phrasal category, etc.), it also "f-selects" a complement which is in the same language. To put it formally, Belazi et al. (1994) posit that a functional head carries a language feature which specifies its language (e.g. SPANISH or ENGLISH). This language feature has to be matched by the same language feature carried by the complement. This whole idea is captured by *the Functional Head constraint*.

(5) **The Functional Head constraint** (Belazi et al. 1994: 228)
"The language feature of the complement f-selected by a functional head, like all other relevant features, must match the corresponding feature of the functional

head."

Five functional heads are identified, with each one taking a different kind of complement.

(6) **Functional Heads and their complements** (Belazi et al. 1994: 228-230)
Complementizer (i.e. C) and IP
Inflection (i.e. I) and VP
Determiner (i.e. D) and NP
Quantifier (i.e. Q) and NP
Negation (i.e. NEG) and VP

Despite some support from spontaneous speech and judgment data, the Functional Head constraint has been challenged by data in other language pairs. Mahootian and Santorini (1996), in particular, point out a number of counter-examples to the Functional Head constraint.

(7) *Switching between C and IP*
I seen everything [$_{C}$ 'cause] [$_{IP}$ *no cogi na*]
I saw everything because not I took nothing
"I saw everything because I didn't take anything."
(English-*Spanish*, Sankoff and Poplack 1981: 6, as quoted in Mahootian and Santorini 1996: 465)

(8) Switching between I and VP
No parce que [$_{I}$ *hanno*] [$_{VP}$ donné des cours]
No because have given of the lectures
"No, because they gave lectures."
(Italian-*French*, Di Sciullo, Muysken and Singh 1986:15, as quoted in Mahootian and Santorini 1996: 466)

(9) *Switching between D and NP*
E wo [$_{NP}$ *green dress*] [$_{D}$ko]
He/she PAST TONE wear green dress ART
"She wore a green dress."
(Adaŋme-*English*, Nartey 1982: 187, as quoted in Mahootian and Santorini 1996:

471)

(10) *Switching between Q and NP*
I'll take [$_{Q}$ some] [$_{NP}$ *naemaek*]
I'll take some salt
"I'll take some salt."
(English-*Farsi*, Mahootian and Santorini 1996: 466)

Mahootian and Santorini (1996) do not give any instances of code-switching between NEG and the following VP. Nonetheless, such examples can be found in Cantonese-English.

(11) *Switching between NEG and VP*
nei5 zau6 wui5 gok3 dak1 keoi5 dei6 [$_{NEG}$ mou5] [$_{VP}$ *make progress*]
you LNK will feel ASP he/she PL not make progress
"You'll then feel that they haven't made progress."
(Cantonese-*English*, Chan 1992)

In addition to these data, there are some very well-documented corpuses in which code-switching consistently takes place between functional heads and their complements. For instance, in Halmari's (1997) corpus of Finnish-English, the majority of code-switching instances cited take place between functional heads and their complements (Halmari 1997: 93).[5]

In brief, the validity of the Functional Head constraint as a universal is seriously in doubt in view of the counter-examples. Notice that these counter-examples come from a lot of language-pairs, far more than the ones which Belazi et al. (1994) observed in formulating the Functional Head constraint. Furthermore, there exists data in which code-switching consistently takes place between functional heads and their complements (e.g. Finnish-English - Halmari 1997). Worse still, Halmari (1997: 94) reported that the Spanish-English bilinguals she consulted disagreed with the judgments in Belazi et al. (1994) and accepted code-switched sentences in which the Spanish complementizer "*que*" takes an English clause as complement, violating the Functional Head constraint.

[5]Halmari (1997: 54) included a statistical account of her data, but she did not give statistics on how much data are instances in which code-switching takes place within phrases - the instances which violate the Functional Head constraint.

The Functional Head constraint is at best descriptively adequate in the data Belazi et al. (1994) investigated.

I conclude that the Functional Head constraint is not empirically adequate and does not hold as a universal constraint on code-switching. Yet, this constraint brings out an interesting question regarding cases in which code-switching does occur between functional categories and their complements:

(12) *Do the grammatical relations between functional categories and their complements hold across elements from different languages*?

A quick review shows that the *c-selection* requirements of functional categories do indeed hold across their complements, even though these complements come from another language. In the above counter-examples to the Functional Head constraint, the complements are of the right syntactic or phrasal categories their functional head selects. Stronger evidence can be found in Morocaan Arabic-French code-switching (Bentahila and Davis 1983). For instance, the Arabic complementizer */baš/* selects finite clauses as complements. This requirement also holds in the following example where */baš/* takes a French finite clause.

(13) je peux le dire *had* le truc *hada*[$_{C}$ *baš*] [$_{IP}$ je commence à apprendre]
"I can say it this thing here in order that I start to learn."
(French-*Moroccan Arabic*, Bentahila and Davis 1983: 323, (101))

Hypothetical examples where */baš/* takes non-finite clauses in French are rejected and corrected so that the non-finite clause is replaced by a finite one (Bentahila and Davis 1983: 323). To sum up so far, intra-sentential code-switching can take place between a functional head and its complement. The only constraint is that a switched element has to observe the c-selection requirement of the functional head. Let's state this observation in the form of a constraint.

(14) **The Functional Head Selection constraint**
"When code-switching takes place between a functional head and its complements, the c-selection requirements that functional heads impose on their complements still hold cross-linguistically."

In fact, the Functional Head Selection constraint is not only superior to the Functional

Head constraint (Belazi et al. 1994) on empirical grounds.[6] Theoretically, the Functional Head Selection constraint also gets rid of the ad hoc stipulation of "language feature" in the Functional Head constraint (a point raised in MacSwan 1997: 63). On the other hand, it can be said that the Functional Head Selection constraint is not a genuine constraint on code-switching per se, but it falls out from a central principle of grammar - Functional heads c-selects the right complements, no matter in a monolingual or a code-switching context. In Chan (to appear), I will argue that there are no specific constraints on code-switching; the same grammar generates both monolingual and code-switching sentences alike, along the lines of Mahootian (1993) and MacSwan (1997).

In any case, under the present approach, a functional head "f-selects" its lexical complement (i.e. a morpheme/chunk of morphemes) from another language, by entering into a configuration with this complement. The two parts are joined directly by some general syntactic rule (say, *Merge* in the Minimalist framework) - the only requirement being that the c-selection requirements of a functional head are fully satisfied. There is no filter of the sort found in the Lemma Congruence model, where EL morphemes are compared with their ML counterparts. Nor are there further mechanisms which lead to the compromise strategies when an EL lemma is "not congruent".

3 A brief survey of Cantonese-English code-switching

In what follows, I will test the applicability of the Lemma Congruence model and the Functional Head Selection constraint with special reference to code-switching data from Cantonese-English code-switching. Cantonese is grammatically very different from English. Highlighted below are some of these differences with reference to the major categories (N, V, A, P).

---

[6]Notice that the Functional Head Selection constraint does not exclude the data in Belazi et al. (1994), in which there is no code-switching between functional heads and complements; it simply says nothing about it. Also, notice that the Functional Head Selection constraint is different from *the Subcategorization constraint* (Bentahila and Davis 1983), which rules that code-switching may take place provided that the language-specific subcategorization/selection requirements of all categories (lexical and functional alike) are satisfied. In Chan (to appear), I will show that the latter may not be correct.

(15) **A brief comparison of the major categories in Cantonese and English**

a. *Nouns* - English nouns are inflected for number but Cantonese nouns are not.
b. *Verbs* - English verbs are inflected for tense, agreement and aspect, but Cantonese verbs are just inflected for aspect.
c. *Adjectives* - English adjectives may be derived forms with suffixes, and there is a distinction between a present participle (e.g. *interesting*) and a past participle (e.g. *interested*). Cantonese adjectives are not derived and thus do not have a contrast between present participles and past participles.
d. *Prepositions* - Many of the core prepositions in English (e.g. *in, out, before, after*) are encoded in *localizers* (Matthews and Yip 1994) which appear after the DP complement. These localizers are arguably nouns or postpositions.

Therefore, it seems that the lemmas of the Cantonese major categories are sufficiently dissimilar from the equivalents of English so that Cantonese-English seems a good testing ground for the Lemma Congruence model. Before going into the analysis, let me give a general picture of Cantonese-English and the data I am going to exploit.

Code-switching has been a common feature in the speech of Cantonese-English bilinguals in Hong Kong. These bilinguals mostly speak Cantonese as their first language. Gibbons (1987) elicits data from students in the University of Hong Kong. Leung (1988) investigates data from elder professionals. Chan (1992) examines data from university students and elder professionals most of whom have studied abroad, and data from radio phone-in programmes. Li (1996) focuses on written data of Cantonese-English by adult columnists. Pennington, Chan and Lau (1996) consists of data from the widest spectrum of speakers ever recorded for this variety of code-switching – the data come from seven disc-jockeys of various backgrounds, various guests and phone-in audiences. Of course, it is expected that variations may be found among different groups of bilingual speakers. In this paper, however, I would like to focus on the general syntactic patterns of Cantonese-English code-switching as spoken in Hong Kong. Therefore, I will consider all the corpora I have access to, instead of singling out a particular one corpus or one group of speakers.

In the following sections, I will discuss code-switching data which involve English verb/verb phrases (section 4), English noun/noun phrases (section 5), English adjectives (section 7) and English prepositions (section 8). Section 6 provides an account of the asymmetric behaviour between English noun/noun phrases and English verb/verb phrases in Cantonese-English. I will analyze the data and compare the two approaches - the Lemma Congruence model and the Functional Head Selection approach.

4 English verbs/verb phrases in Cantonese-English code-switching

One salient pattern which emerges from Cantonese-English data is that the English verbs are in most cases the infinitive or root forms, even though the linguistic context or discourse context would require inflected forms according to the grammar of English. One such context involves a single English verb modified by the Cantonese *perfective* aspect marker "*zo2*". This aspect marker signals an event which is "seen as a whole or as completed" (Matthews and Yip 1994: 204). Similar meaning is expressed by either the past tense or the present perfect tense in English. In code-switching, however, the English verbs remain in their infinitive forms.

(16) nei5 dei6 kei4 sat6 ***exclude*** **zo2** hou2 do1 jan4[7]
you PL in fact exclude ASP EMP many people
"You have in fact excluded many people."
(Cantonese-*English*, Chan 1992)

Another aspect marker which refers to past events is "*gwo3*", the *experiential* marker. This marker indicates events which have happened "at least once before" (Matthews and Yip 1994:206)[8]. Similar meaning is expressed by the present perfect tense in English. Nonetheless, the English verbs remain uninflected but are followed by the Cantonese aspect markers in code-switching.

(17) cyun4 bou6 din4 waa2 dou1 mou5 ***screen*** **gwo3**
all part telephone all no screen ASP
"All the telephone (calls) have not been screened."
(Cantonese-*English*, Pennington, Chan and Lau 1996)

The *continuous* aspect is expressed by *jyu6* in Cantonese, which "describes a continuous activity or state without change" (Matthews and Yip 1994: 206). Similar meaning is

[7]The transcriptions of Cantonese here follow the scheme of the Linguistic Society of Hong Kong (1993). The following are the symbols of the glosses: PL - plural, ASP - aspect, EMP - emphatic, PRT - particle, CL - classifier, NUM - numeral, DEM - demonstrative, COMP - complementizer, COV - coverb, P - preposition/postposition, LNK - linking particle. The numbers denote the distinctive tones of the syllable.

[8]See Matthews and Yip (1994: 206-207) for the difference between *zo2* and *gwo3*.

expressed by the continuous or progressive aspect - V + *ing* - in English. In code-switching, the English verb is uninflected but followed by the Cantonese continuous aspect marker, *jyu6*.

(18) nei5 ***hold*** **jyu6** laa1 maa3?
you hold ASP PRT PRT
"Are you holding (the line)?"
(Cantonese-*English*, Pennington, Chan and Lau 1996)

Apart from the perfective and experiential aspect markers, past events in Cantonese are also expressed by time adverbials. With Cantonese adverbials referring to past time, the English verbs remain in its infinitive form.

(19) keoi5 waa6 **ji5 ging1** ***make*** *up his mind*
s/he PL already make up his mind
"He says (he) already made up his mind."
(Cantonese-*English*, Leung 1988: 32, 6(a))

(20) jan1 wai6 ngo3 **saam1 nin4 cin4** ***retire***…
because I three years before retire
"Because I retired three years ago… "
(Cantonese-*English*, Chan 1992)

In some cases linguistic cues are not available. It is just the discourse context which indicates that the speakers are talking about past events. According to English grammar, the past tense is commonly used in such contexts. In the code-switching example below, the speaker was recounting past events, but the English verb remains in the infinitive form.

(21) nei5 jau5 mou5 ***make*** *an effort organize* keoi5 sin1?
you YES/NO make an effort organize it PRT
"Have you made an effort in organizing it?"
(Cantonese-*English*, Chan 1992)

Another context where English verbs are inflected is the description of present/ habitual events for third person singular subjects - According to the grammar of English the

present tense would be warranted. In Cantonese-English code-switching, the English verbs remain in their infinitive forms.

(22) **go3** ***goal guide*** ngo5 dei6 di1 *nurse* heoi3 zou6 je5
CL goal guide I PL CL nurse go do things
"The goal guides us nurses on how to do our job."
(Cantonese-*English*, Chan 1992)

(23) cyun4 kaau3 **keoi5** jat1 ci3 ***print*** sap6 fan6 bat3 fan6 cing2 syu1
all depend it one time print ten CL eight CL love letter
"It all depends on it (refer to a computer printer) which prints out ten or eight love letters at one time."
(A cartoon bubble - a guy was busy printing out love letters on computer)
(Cantonese-*English*, Li 1996: 196)

A very convenient explanation of this fact would be that English verbs are somehow adapted morphologically to Cantonese. They appear in *forms* and *positions* which are similar to Cantonese verbs in Cantonese (Chan 1992). Indeed, this is an explanation in the spirit of the Lemma Congruence model approach too: An EL item appears in a mixed constituent by establishing some sort of congruence with its ML counterpart in terms of morphological-syntactic properties. Since the English verb with inflections (supposedly the EL morphemes in our case[9]) is not congruent with the Cantonese verb without inflection for tense/agreement (supposedly the ML morphemes in our case) compromise strategies of some sort have to be taken - the English verb may appear as a *bare form*, with its distinct morphological-syntactic properties stripped away. In the above cases, the grammatical markers of English verbs - tense and agreement - are all removed so as to be congruent with their ML equivalents - i.e. the Cantonese verbs - and to appear in code-switching.

The Lemma Congruence model is not without problems here. Given the

[9]Myers-Scotton's (1993, 1995) definition of the *matrix language/ML* and the *embedded language/EL* has already been criticized by Bentahila (1995) and Muysken and de Rooij (1995). While I fully share the latter authors' reservations about defining ML in quantitative or discourse terms, I just *assume* the ML/EL dichotomy for the current discussion of the Lemma Congruence model. In most of the data I quote, it appears that Cantonese is always the ML (in terms of *the Morpheme Order principle* and *the System Morpheme principle*). This does not imply that I endorse the Matrix Language Frame model or the Lemma Congruence model.

"incongruence" between English verbs and Cantonese verbs (in morphological realization patterns), the Lemma Congruence model does allow English verbs to be inflected in *EL islands*, another compromise strategy. In our case, the EL island would be the whole English VP. In examples (19) and (21) above, the whole English VP appears, but the English main verbs remain in their infinitive forms. This is not predicted by the Lemma Congruence model.

Here, I would like to argue that the alternative analysis - the Functional Head Selection approach - can also explain the fact that English verbs, either single verbs or verbs in English VPs, are uninflected. Notice that neither the Cantonese aspect markers (i.e. examples (16) to (18)) nor the time adverbials (i.e. examples (19), (20)) trigger the English tense markers. Nor does the discourse context activate the tense marker (i.e. example (21)). The subjects - in their semantic/ syntactic properties as third person singular DPs (i.e. example (22) to (23)) - do not call for the English agreement markers. The occurrence of the English verbs in their infinitive forms appears to be controlled by some other grammatical factor. This factor, I would like to suggest, is the Cantonese *INFL*, a separate node in phrase structure which selects a bare VP (which dominates the English infinitive verb) in Cantonese-English code-switching.[10] The universal postulation of INFL as head of the sentence has been a standard assumption in (Chomskyan) generative grammar - the theory of Government and Binding (i.e. GB) and also Minimalism - though after Pollock (1989) views differ as to whether there is just one INFL or a split INFL in all languages, and whether all languages have the same INFL structure (Thráinsson 1997). I have no intention of getting into that debate, but it is necessary for me to spell out my assumptions about the INFL in Cantonese.

Following Cheng's (1991) analysis of Mandarin Chinese, I assume that Cantonese sentences are headed by INFL, which she actually names ASP (i.e. *aspect*). So, a Cantonese sentence is an *aspect phrase*. The aspect node contains features about the aspect of the whole sentence which are overtly realized by Cantonese aspect markers.[11]

---

[10]In Grimshaw's theory (1991), INFL selects VP by virtue of forming an *extended projection* with VP- INFL and VP share the categorial features of [+V, -N], with INFL being a functional head (bearing the functional feature of F_1) but VP being a lexical projection (bearing the functional feature of F_0).

[11]Problems arise as to how the aspect marker gets affixed to verbs. Cheng (1991) assumes an operation of aspect lowering, yet a lowering analysis would deviate from standard assumption that movements are leftwards (Kayne 1994). Further research is needed to sort out which analysis is preferable. The problem here is totally analogous to how tense/ agreement markers gets affixed to the verb in English, where English verbs are assumed to stay in VP lower than INFL (or AgrS and T - Pollock 1989). In the

Matthews and Yip (1994) observe that in some contexts the aspect marker is optional, say, in the presence of other linguistic cues on the aspect of the event.

(24) ngo5 dei6 cam4 maan2 heoi3 (zo2) taai2 jin1 faa1
I PL last night go ASP watch fireworks
"We went to watch the fireworks last night."
(Cantonese)

Back to Cantonese-English code-switching, I assume that in the above examples (i.e. examples (16) to (23)), the clauses are headed by the Cantonese aspect head. This aspect head *only* selects a bare VP (a VP which dominates an infinitive verb form), occasionally inserting a Cantonese aspect marker to attach to the English verb stem - that is the reason why the English verbs do not get inflected with the English tense/ agreement markers. The English tense/ agreement markers are not "stripped away" by a filter, as in the Lemma Congruence model; rather they are actually not called for - They are only licensed by an English INFL.

One may question why the examples I have cited all have the Cantonese INFL, or ASP as head, but not the English INFL. This may be due to the fact that Cantonese is the dominant language in discourses from which the above examples are taken. It does not mean that an English INFL is impossible in Cantonese-English code-switching. The following is one example, where the English auxiliary verb "-*'m*" is supposed to be in INFL:

(25) I'm speaking of *go3* cost, *m4 hai6* functionality
I'm speaking of CL cost not COP functionality
"I'm speaking of the cost, not functionality.
(English-*Cantonese*, Chan 1992, also cited in Chan 1998: 208, (38))

Unfortunately we do not have instances where Cantonese verbs are inflected with English tense/ agreement markers - such data may be found only in discourses where

---

minimalist framework (Chomsky 1995, ch.4), a word is supposed to carry its features when it enters derivations - This implies that the verb is already inflected with tense/ agreement markers when it merges with another syntactic object, say, an object DP.

English is the dominant language.[12] In the presence of such data (i.e. $V_{Cantonese}$ + *s/ed* $_{English}$), the Functional Head Selection approach would be much strengthened.[13]

5 English nouns/noun phrases in Cantonese-English code-switching

Let's turn to another prolific pattern in Cantonese-English code-switching. English nouns or noun phrases always occur in the position after Cantonese classifiers or quantifiers. This observation seems to prevail across different corpuses on Cantonese-English code-switching (Chan 1992, Leung 1988, Pennington, Chan and Lau 1996).
Let me first present some facts about Cantonese determiner phrases (DPs)[14] before we look at the code-switching data. The Cantonese determiner phrase consists of the basic structure of "numeral/NUM + classifier/CL + noun". For instance,

(26) jat1 bun2 syu1
NUM CL book
"one book"

Unlike English, Cantonese nouns do not inflect for number of plurality[15].

---

[12]I assume that *descriptively speaking* there may be a "dominant language" in code-switching which contributes more morphemes and displays more morpho-syntactic characteristics in code-switched sentences. This does not imply that I accept the MLF model (Myers-Scotton 1993) and its definition of *the matrix language*. Also see note 9.

[13]Ad Neeleman (personal communication) has suggested to me that a Cantonese verb may not combine with an English tense/agreement marker for some independent morphological reasons. I leave this possibility open here.

[14]Here, I just use the standard term *determiner phrase* to refer to *extended projections* of the noun in the sense of Grimshaw (1991). In Chan (to appear), I will assume that sometimes only a ClP (*classifier phrase)* is projected; sometimes a QP (*quantifier phrase*) is projected in which the quantifier head dominates a DP.

[15]Except pronouns. Like English or Mandarin Chinese, there are singular and plural forms for pronouns - i.e. "*ngo5*(I/me)" vs. "*ngo5 dei6*(we/us)", "*nei5*(you)" vs. "*nei5 dei6*(you(pl.))" and "*keoi5*(he/him/she/her)" vs. "*keoi5 dei6*(they/them)".

(27) loeng5 bun2 **syu1**
two CL book
"two books"

In Cantonese, there is a common classifier "*di1*", which, similar to "some" in English, takes both mass nouns (e.g. (28)) and count nouns (e.g. (29)).[16] When combined with count nouns, it refers to more than one entity or an object, even though the Cantonese noun is not inflected for number.

(28) ngo5 jam2 zo2 (jat1)[17] **di1 caa4**
I drink ASP (one) CL tea
"I drank **some tea**."

(29) zui3 gan6 dou3 zo2 (jat1) **di1 san1 syu1**
most recent arrive ASP (one) CL new book
"Most recently **some new books** have arrived."

Similar to "*di1*", the Cantonese quantifiers "*do1* (many/much)" and "*siu2* (few/little)" may combine with mass (i.e. (30)) or count nouns (i.e. (31)). When combined with count nouns, it is understood that the noun has plural reference (i.e. more than one of the referred object or entity).

(30) gam2 **do1** **faan6**, dim2 sik6?
So many/much rice how eat
"There is so **much rice**; how can I manage to eat it all?"

(31) zui3 gan6 ceot1 zo2 **hou2 do1 san1 syu1**
most near out ASP EMP many new book
"Most recently, **many new books** have come out (i.e. been published)."

---

[16]Although the distinction between mass and count nouns is not grammaticalized in Cantonese, there is evidence that child learners can interpret whether a noun is inherently count or mass (Teng 1997).

[17]As with other classifiers, the preceding numeral may be deleted when it means "*one*". Since "*di1*" only takes "*one*" as numeral, the numeral is optional.

Given the nouns do not inflect for number in Cantonese, for the sake of "congruence", one would expect that the English nouns would always appear as bare forms. Interestingly, this is not true. Though it is not always the case, there are quite a lot of examples in which English count nouns appear in their plural form after certain Cantonese numerals, classifiers or quantifiers.[18] This phenomenon was first observed by Leung (1988):

(32) daan6 hai6 **saam1** gaan1 ***colleges*** jau6 jau5 *short courses for language teachers*
But three CL colleges also have short courses for language teachers
"But the three colleges also have short courses for language teachers."
(Cantonese-*English*, Leung 1988: 35, (7))

In the above example (32), the English noun "*colleges*" is inflected for number in agreement with the Cantonese numeral "*saam1*(three)". The following are some other examples where English plural nouns appear after Cantonese numerals which are "more than ONE".

(33) hai2 go3 fong1 leoi5 min6 jau5 **loeng5** go3 ***judges***
P CL room inside have two CL judges
"Inside the room were two judges."
(Cantonese-*English*, Chan 1992)

(34) ngo5 duk6 zo2 **gei2** go3 ***chapters***
I read ASP several CL chapters
"I've read several chapters."
(Cantonese-*English*, Chan 1992)

English plural nouns also appear after the Cantonese quantifiers "*hou2 do1*(many/ much)" or "*taai do1*(too many/too much)".

[18]I will discuss the cases in which plural English nouns do not appear after these Cantonese categories below.

(35) jau5 **hou2 do1** ***Jews***…
have EMP many/much Jews…
"There are many Jews…"
(Cantonese-*English*, Pennington, Chan and Lau 1996)

(36) cit3 gei6 **taai3 do1** ***details***
avoid too many/ much details
"Avoid too many details."
(Cantonese-*English*, Li 1996: 69 - from a newspaper column which gives tips for dressing up)

English plural nouns follow the Cantonese classifier "*di*1" as well.

(37) daai6 gaa1 ho5 ji6 *share* jat1 **di1** ***files*** (i.e. computer files).
We can share NUM CL files
"We can share some files."
(Cantonese-*English*, Chan 1992)

(38) …go3 **di1** ***coasters*** bui1 zin3 aa3
DEM CL coasters coasters PRT
" … those coasters, coasters."
(Cantonese-*English*, Pennington, Chan and Lau 1996, also cited in Pennington, Lau and Chan, to appear, (12))

The English nouns are not only sensitive to the classifiers immediately preceding them; they also agree with the Cantonese classifiers with intervening modifiers.

(39) giu3 **di1** *bilingual* ge3 ***tutors***…
ask CL bilingual LNK tutors…
"Ask the bilingual tutors…"
(Cantonese-*English*, Chan 1992, also cited in Chan 1998: 203, (34))

It is not the case that the plural English nouns only occur as lone nouns after Cantonese modifiers. They may appear after Cantonese quantifiers/ numerals/ classifiers and English noun phrases. Notice that in (41) the English noun phrase has an internal constituency of "head noun and postmodifier" - N + PP. This structure is distinctively

English in character, as Cantonese nouns are strictly final in DPs.

(40) nei5 ho2 ji5 gaau2 **jat1** go3 *farewell* ***party***
you can organize one CL farewell party
"You can organize a farewell party."
(Cantonese-*English*, Chan 1992)

(41) keoi5 hai6 gong2 gan2 **loeng2** go3 ***views*** *of world origin*
s/he COP talk ASP two CL views of world origin
"What he talks about is two views of world origin.
(Cantonese-*English*, Chan 1992, also cited in Chan 1998: 202, (29))

(42) syu1 zung1 lit6 gui2 liu2[19] **hen2 do1** *dueling* ***proverbs*** dik1 lai6 zi2
book inside list ASP EMP many dueling proverbs LNK example
"Inside the book (the author) listed many examples of dueling proverbs."
(Cantonese-*English*, Li 1996: 199 - from a newspaper column)

To summarize the above observations briefly, English nouns, in lone forms or as heads in English NPs, may be inflected in their plural forms after Cantonese numerals (more than ONE), classifiers (*di1*) and quantifiers (e.g. *hou2 do1*).

An interesting paradox arises between English nouns/noun phrases and English verbs/verb phrases in Cantonese-English code-switching. In the case of English verbs, the main verb in most cases must not be inflected. In the case of English noun phrases, however, the nouns are consistently plural-marked with certain quantifiers or classifiers, even though in Cantonese number is not grammaticalized on nouns.

In terms of the Lemma Congruence model, an EL plural marker in a mixed constituent is not the idiosyncratic property of Cantonese-English code-switching. The same phenomenon can also be attested in other language-pairs, for example, Lingala-French (Bokamba 1988: 37, Kamwangamalu 1989: 160, etc.[20]). Such examples can also be found in Myers-Scotton's own data:

[19]"*Liu2"* here is a counterpart of "*zo2*" in a formal or written register.

[20]See Myers-Scotton (1993: 112) and the references there for other language-pairs which show doubly marked plural nouns.

(43) …dzimwe dzenguva tinenge tichiita **ma-*game-s*** panze
"…sometimes we will be doing games outside"
(Shona-*English*, Myers-Scotton 1993: 132)

Myers-Scotton (1993) calls examples such as (57) *double morphology*. Double morphology is a problem for the Matrix Language Frame model because the plural marker is defined as a system morpheme. In the above example, the ML is Shona. The English plural marker, as an EL system morpheme, appears in a mixed constituent, violating the System Morpheme principle. Notice that the doubly marked stem is neither a good example of an EL island (i.e. it does not contain more than one word) nor a bare form (i.e. it is inflected), which Myers-Scotton (1993) and Jake and Myers-Scotton (1997) regard as possible "compromise strategies" as a result of non-congruence.

To circumvent the model, Myers-Scotton (1993: 132-135) suggests that the ML plural marker is accessed by error when the plural marker of the ML is activated by ML morphological-syntactic procedures. Obviously, this explanation is inadequate in explaining the English plural marker in Cantonese-English code-switching - there is simply no ML (i.e. Cantonese) plural marker being called up.

An alternative analysis is to keep to the primary observations and postulate that the quantifiers/ classifiers/ numerals license the plural English marker. One major query immediately arises - Why are the English nouns inflected for number, since Cantonese nouns are not. A tentative yet straightforward answer is that the English nouns do not need to inherit all the characteristics of Cantonese nouns in order to appear in code-switching. Contrary to the Lemma Congruence model, the English nouns (or the corresponding "lemmas") need not be compared with the Cantonese nouns. Let's try out the Functional Head Selection approach now.

Assuming the Functional Head Selection approach, one may straightforwardly say that quantifiers, numerals and classifiers are all functional heads which select the plural English nouns. More specifically, one may say that the Cantonese classifiers, numerals or quantifiers, heading extended projections of nouns in the sense of Grimshaw (1991), project a [+plural] feature. Given that the plural form projects a [+plural] feature and a singular form projects a [-plural] feature, the plural form is selected to project a matching [+plural] feature. The occurrence of the English nouns in plural forms can be seen as satisfying the selectional requirement of the Cantonese numerals/ classifiers/ quantifiers

or general principles of extended projections.[21]

(44) loeng5 go3 *judges*
two CL judges
NUM CL N
[+plural] [+plural]
"two judges"
(repeat from (33) above)

So far so good. Yet, a non-trivial problem arises as in some cases the English singular form is also selected.

(45) go2 **di1 *waterfall*** ne1
DEM CL waterfall PAR
"those waterfalls…"
(Cantonese-*English*, Pennington, Chan and Lau 1996)

But wait. Let's remember that Cantonese nouns are not inflected for number. This means that the "plural" Cantonese numerals/ classifiers/ quantifiers do not *obligatorily* select for a plural noun because they do *not* in a monolingual Cantonese context. In addition, it makes no sense to speak of Cantonese nouns as projecting a [-plural] feature simply because they may refer to more than one entity (e.g. *syu1(book))* in example (27) above). Taking these into consideration, we may say that the "plural" Cantonese categories select a noun which is either [+plural] (as in the plural English nouns) or *unspecified* for plurality (just like the Cantonese nouns). In code-switching, it may be expected that the "plural" Cantonese numerals/ classifiers/ quantifiers select a singular form of English noun which is not [-plural] but unspecified for plurality. This explains a number of other cases where the singular, count English nouns are used after the "plural" Cantonese categories (e.g. example (45)).[22]

---

[21]Grimshaw (1991) assumes that only functional heads and lexical heads may project syntactic/ semantic features (e.g. animacy, number). In case more than one head (functional or lexical) projects a similar feature, say [number] in DPs, these features will percolate up to the highest node of the extended projection, and they have to be matched in order to have an interpretation.

[22]In Chan (to appear), I will analyze the structure of Cantonese DPs in greater detail and show how plurality is encoded by numerals, classifiers and quantifiers. Contrary to Ritter (1991), I will also argue

One problem I have ignored so far is that the "plural" Cantonese classifier "*di1*" and quantifier "*hou2 do1*" may combine with mass nouns from English. For instance,

(46) *computer* ho2 ji5 tau3 gwo3 *keyboard* tai4 gong1 jat1 **di1** ***feedback***
computer can through keyboard provide one CL feedback
"Computers can provide some feedback through the keyboard."
(English-Cantonese, Chan 1992)

(47) **hou2 do1** ***handicraft***
EMP many handicraft
"many handicrafts"
(Cantonese-*English*, Pennington, Chan and Lau 1996)

These classifiers/ quantifiers may not project *exactly* a [+plural] feature, which, strictly speaking, means "more than one countable object". Here, I just tentatively assume that they project a feature like [+quantity] (meaning a quantity of something), which has to be matched by a [+plural] feature from the noun (i.e. as in the case of plural English nouns), or *nothing* (as in the case of all Cantonese nouns and the uninflected English nouns). I assume the mass nouns in English, similar to the Cantonese nouns, are unspecified for plurality too.

To sum up briefly, the English plural nouns in Cantonese-English code-switching are licensed by certain "plural" Cantonese categories: numerals (more than ONE), classifiers (e.g. *di1*) or quantifiers (e.g. *taai2 do1)* - Numerals are specifiers of classifiers[23], whereas classifiers and quantifiers are functional heads. Apart from plural nouns, these plural Cantonese categories may also take complement nouns which are unspecified for plurality - which include all Cantonese nouns. In case the "plural" Cantonese categories take English count nouns in singular forms, these English forms are treated as being unspecified for plurality. Here, I would like to stress in these cases the English forms are not "adapted" or "compared" with their ML counterparts. They are selected by the Cantonese numerals/ classifiers/ quantifiers - categories in the "functional domain" of DP. Generally speaking, the selectional requirements between the Cantonese "functional

---

that numerals are specifiers of classifiers rather than a functional head in Cantonese. I skip the details here for limitations of space.

[23]See footnote 22 above.

categories" and English "lexical complements" are upheld (i.e. The "plural" Cantonese classifiers/ quantifiers select English nouns which are [+plural] or unspecified). The interesting thing here is that code-switching enlarges, not narrows, the selectional restrictions.

6 Towards a unified theory for English nouns and verbs in Cantonese-English code-switching

I have just suggested that the Functional Head Selection approach can explain the formal properties of English nouns and verbs in Cantonese-English code-switching. Notice that the English nouns and verbs are basically asymmetric - the English nouns may be inflected whereas the English verbs are always never inflected for tense/agreement. In order to maintain a uniform approach, I have implicitly made different assumptions about the nature of the inflectional markers of the nouns, namely, the plural marker, and those of the verb, namely (English) tense/ agreement and (Cantonese) aspect markers. What I have assumed is that the plural markers are more closely associated with the English nouns, and they surface when activated by the right context, no matter whether the context is English or Cantonese (e.g. as generic count nouns, after numerals, plural demonstratives, etc.). For English verbs, I have to assume that the tense and agreement markers are represented separately, namely, in the INFL node selecting the verb phrase. In the code-switching data we have examined, it is the Cantonese INFL, or ASP, which is selected in the first place, and therefore the English tense/ agreement markers do not appear with their English verb stems.

These assumptions are by no means arbitrary stipulations. Firstly, it has long been assumed that tense and agreement markers are associated with AUX / INFL, a node separate from the verb (in pre-GB theory, as in Chomsky 1965, in GB theory, as in Chomsky 1981, 1986a, 1986b, but not Minimalism, see footnote 10 above). The plural marker does not originate from another node, however. Descriptively, in English, the plural marker is attached to the head nouns of NPs mainly.[24] Nonetheless, the tense and

[24]Neil Smith suggests that the demonstratives "*these*" and "*those*" can also be analyzed as being inflected by the plural marker "*-s*". Yet, after these demonstratives the head nouns also have to be inflected by the plural marker "*-s*" as well. In this respect, the plural marker "*-s*" is still radically different from verbal inflection markers (e.g. tense/agreement markers) in that for the latter only one set of markers appear for a clause - if the modal takes up tense/agreement , the head verb has to be uninflected.

agreement marker may be realized by modal verbs or auxiliary verbs such as “have” or “be”, after which the verb appear in infinitive form.

Unlike English, Cantonese modal verbs do not inflect (for aspect). Yet, auxiliary elements like Cantonese *coverbs*/ *prepositions* may take up aspect, after which the following verb is uninflected. For instance,

(48) ngo3 **tung4** **zo2** keoi5 sik6 faan6
I with(COV/P) ASP him/her eat rice
“I had eaten rice with him/her (lit. I had a meal with him/her.).”

One may argue that “*tung4*(and/with)” is just a verb on a par with “*sik6*(eat)”. That is not correct. It is incomplete to say *“ngo3 tung4 keoi5*(lit. I and/with him/her)”, except when it is used as a fragment with the remaining part readily retrieved from the context (i.e. (49)).

(49) a. *ngo3 tung4 keoi5.
I with(COV/P) him/her
*“I with him/her”

b. Question:
nei5 tung4 bin1go3 sik6 faan6?
You with(COV/P) whom eat rice
“Whom did you eat rice with?”
(= Whom did you have a meal with?)

Answer:
ngo3 **tung4** aa3 maa1 (sik6 faan6)
I with(COV/P) mother (eat rice)
“I (had a meal) with Mom.” (= I had a meal with Mom.)

So, the coverb/preposition is different from a typical verb in Cantonese. When the former takes up aspect, the main verb cannot be inflected. In other words, the aspect marker in Cantonese can be assigned to either a coverb/preposition or a main verb, and hence it appears to originate from an independent *INFL* node.

Let’s look at the second piece of evidence, which comes from language acquisition. English children acquire the plural marker *-s* before they acquire the tense and agreement

markers of the verb (Radford 1990). Radford (1990) explains this phenomenon by postulating that the functional category INFL, which carries the tense/agreement markers, is developed later (i.e. after 24 months), though at that time the child seems to have mastered the verbs. Before that (i.e. after 20 months), the child can use the verbs, and the nouns with the plural suffix. The child only masters the tense/ agreement markers after his/her INFL is developed.

The third piece of evidence comes from speech errors. In one type of speech errors - the so-called *stranding* errors - lexical items change position, but inflectional markers do not move with their corresponding stems (Garrett 1980). For instance,

(50) You *order*ed up *end*ing some fish dish.
(Target: You *end*ed up *order*ing some fish dish.)
(Garrett, in press, quoted by Bock and Levelt 1994: 947)

If the plural marker is more closely associated with the noun than the tense/ agreement markers are with the verb, one would expect that the plural markers may move with their noun stems, while tense/agreement markers always strand. In other words, the plural marker may not stay in situ, but the tense/ agreement markers always remain in their positions even when their verb stems are dislocated in such stranding errors. This turns out to have empirical support, as first noted by Stemberger (1985) and later reported by Myers-Scotton (1993: 62). The following is an example:

(51) I presume you could get *light* in poor *pictures.*
(Target: I presume you could get *pictures* in poor *light*.)
(Stemberger 1985: 162, (10a))

The fourth piece of evidence comes from code-switching data in other language-pairs - instances of "double morphology" - in which a free morpheme is inflected with markers from both languages. In the Functional Head Selection approach I have been outlining, nominal functional heads (demonstratives/ quantifiers/ classifiers) select noun/noun phrase complements from another language. Since the plural marker is somehow inherently associated with the noun, it is activated after the right functional heads (i.e. demonstratives/ quantifiers/ classifiers which license plurality). Under this premise, "double morphology" which involves the plural marker is also expected - the plural markers from both languages are both licensed and activated after the right functional heads. On the other hand, I have been arguing that the INFL, hosting verbal inflectional

markers (such as tense and agreement in English), selects a bare VP which dominates an infinitive verb. If that were true, one would *not* expect double morphology which involves tense/agreement - There is only one INFL selecting one VP; only one set of verbal inflectional markers prevail. That appears to be true in the literature on double-morphology that I know (Bokamba 1989, Kamwangamalu 1989, Myers-Scotton 1993).

What I have attempted to do is give evidence that the tense/ agreement markers in English and the aspect markers in Cantonese are more "detached" from the verb than the plural markers are from the nouns. This motivates and justifies a standard assumption - Verbal inflectional elements (such as tense/ agreement markers in English or aspect markers in Cantonese) are represented in an INFL node, separate from the verb. Consequently, when an INFL selects a VP, the verb is not attached with any inflectional elements until after an additional operation, be it I-lowering or Verb-raising, Nominal inflectional elements, on the other hand, are linked more closely to the head noun. Accordingly, when a determiner/ classifier/ quantifier selects a plural noun, the noun may already be inflected with the plural marker. This assumption is essential if we want to maintain a uniform approach, namely, the Functional Head Selection approach in explaining the asymmetric behavior of English nouns and verbs in Cantonese-English code-switching.

7 English adjectives/adjective phrases in Cantonese-English code-switching

English adjectives in Cantonese-English code-switching provide additional evidence against the Lemma Congruence approach. These English adjectives contain a root and a suffix, whereas Cantonese adjectives do not carry any affixes. For Myers-Scotton and Jake (1995), the English adjectives are "incongruent" with their Cantonese ML counterparts in terms of morphological realization patterns. The English adjective can just be avoided, or else its suffix has to be stripped off in order to pass the filter and appear in code-switching. This is not the case, however.

(52) ngo5 gwo3 di1 sang1 wut6 hou2 ***relaxing***_
I live CL life very relaxing
"The life I've been leading is very relaxing."
(Cantonese-*English*, Chan 1992)

(53) jau2 jat1 sai3 gei2 mut6 jau5 fong3 gong1 hoei3 *have a drink*
have one century no have after work go have a drink
gam2 gok3 hou2 ***deprived***
feel very deprived
"For ages (I) haven't gone to have a drink after work; (I) feel very deprived."
(Cantonese-*English*, Li 1996: 190 - newspaper column)

Facing similar forms in Arabic-English, Myers-Scotton and Jake (1995: 998-999) suggest that participles should be exempted from blocking by ML because they are "category-internal" and "syntactically irrelevant" to external syntactic relations. True, the derivational suffix is category-internal in the sense that without it the word would not be an adjective anymore. Nevertheless, it is hard to see how the suffix is "syntactically irrelevant" here - Notice that the alternation of *-ing* and *-ed* forms is sensitive to the interpretation of the Cantonese subjects external to the English adjectives.[25]

The Functional Head Selection approach predicts straight away that these adjectives appear in their "full" English form: Only then would they be of the recognized category - adjectives - which are called for in those particular positions. Obviously, they appear not because they are morphologically and syntactically "adapted" to Cantonese - their derivational suffix endings are not shared by their ML counterparts, that is, the Cantonese adjectives.

A question arises as to what selects the adjectives then. Following Abney (1987), I assume that what have traditionally been called adjective phrases are actually *degree phrases/DegP*. With reference to examples (52) to (53), the English adjectives are preceded by the Cantonese "*hou2*", which carries the meaning of "*very*".

[25]In (52), the subject is the external argument of the predicate "*relax*", from which the adjective "*relaxing*" is derived (i.e. *Something* relaxes me → *Something* is relaxing). In (53), the subject is the internal argument of the predicate "*deprive*", from which the adjective/past participle "*deprived*" is derived (e.g. Somebody deprives *me* (of something) → *I* am deprived (of something).). In Cantonese, the adjectives are not derived from verbs; therefore, it is difficult to speak of equivalent thematic relations between subjects, adjectives/participles and their corresponding predicates/verbs. Actually, in (53), the equivalent word for "*deprive*" in Cantonese - "*mok1 soek3"* - is used as a full verb rather than an adjective. In this light, it is even doubtful whether the English adjectives/participles in (52) and (53) are "congruent" with their ML equivalents in terms of *predicate-argument* structure (assuming that Cantonese is the ML), further violating the Lemma Congruence model.

8 English prepositions/prepositional phrases in Cantonese-English code-switching

The appearance of English prepositions in Cantonese-English code-switching provides yet another counter-example to the Lemma Congruence approach. In the case of English plural nouns or derived adjectives, they should not appear under the Lemma Congruence model because they are "incongruent" with their ML counterparts - that is, Cantonese nouns and adjectives - in terms of morphological realization patterns. In this case, English prepositions should not appear in Cantonese-English code-switching because they are "incongruent" with their ML counterparts - Cantonese prepositions - in terms of "lexico-conceptual" structures. Let me elaborate on this point now.

English is prepositional, and the prepositional phrase(or PP) consists of a preposition and a following determiner phrase (DP). It is the preposition itself which carries the meaning of position or location. For instance,

(54) $[_{PP}\,[_{P}$ After$]\,[_{DP}$ this review$]]$, we will submit this review to the government for reference.

Cantonese has also been analyzed as prepositional. Nevertheless, the internal structure of PP is drastically different from that of English - In Cantonese, the meaning of location or position is carried by a noun (N), which is preceded by another determiner phrase (say DP). What justifies the categorial status of the N is that one can insert a linking particle *ge3* or *zi1* between DP and N. This particle is used to link adjectives or noun phrase modifiers with the head noun (Matthews and Yip 1994: 88, 158-159).[26] The preposition *hai2* does not carry the meaning of location.[27] The internal structure of PPs in Cantonese

[26]Gasde and Paul (1996) have recently analyzed *the linking particle and the N* (i.e. in (55)) as a postposition in Mandarin Chinese. Matthews and Yip (1994) call N (in (97)) a *localizer*, and do not commit themselves to any particular categorial status of these localizers. Here I assume that the localizer is a noun because the linking particle always precedes nouns in other contexts. Though Gasde and Paul (1996) have convincingly shown that the localizers are not typical nouns, I think one can still maintain that the localizers are *defective* nouns, or [+N] elements. The DP and N in (55) form a complex DP. More will be said about these issues in Chan (to appear). No matter which way one analyses these localizers - as nouns or postpositions - they carry information which is typically encoded in the prepositions of English. This renders English prepositions "incongruent" with Cantonese prepositions(i.e. *hai2*), the latter not encoding a specific locative meaning.

[27]Sometimes, the preposition *hai2* appears without a localizer, where it means roughly "in" or "at" in English. For instance, "*keoi5* ***hai2*** *Ying1gwok3* (lit. He/She is in England.)." However, *hai2* can be

is thus as follows:

(55) P DP (ge3/zi1) N
hai2 LNK **localizer**

The following is the Cantonese translation of the English PP - "after this review":

(56) [$_{PP}$ [$_{P}$ hai2] [$_{DP}$ ni1 go3 gim2 tou2] zi1 [$_{N}$ hou6]]]
P DEM CL review LNK after
"After this review…"

Below is an example in which an English preposition appears in Cantonese-English code-switching - an example which would have been erroneously ruled out by the Lemma Congruence model.

(57) *After* ni1 go2 *review* zi1 hou6, zau6 bei2 zing3 fu2 caam1 haau2
After DEM CL review LNK after, then give government refer
"After this review, (we) will submit (this review) to the government for reference."
(A police commander was asked about what the police would do after an accident.)
(*English*-Cantonese, Chan 1992, also cited in Chan 1998: 196, (18))

Considering the Functional Head Selection approach, one may ask which functional categories select the prepositions. That is not a relevant question for me here, as I follow Grimshaw (1991) in assuming that prepositions are functional categories. What is relevant here is that the English preposition, as in the code-switching example (57), selects a Cantonese DP as complement. Although the Cantonese DP (complement of P) is very different from the English DP (complement of P) in the sense that the former encodes the meaning of location, it seems that code-switching may take place if the c-selection restrictions between functional heads and their complements are satisfied. That

---

analyzed a verb in this context as it may form A-not-A questions on a par with other verbs - "*keoi5* ***hai2-m4-hai2***(A-not-A) *Ying1 gwok3?* (lit. "Is he/she in England?). In any case, there is no data in which an English preposition "in" or "at" appears in the position of *hai2* - **keoi4* **in/at** *Ying1 gwok2* (He/she is in England.).

is, the English preposition selects a DP.

9 Conclusions

In this paper, I have compared two approaches to the syntax of intra-sentential code-switching - the Lemma Congruence model and the Functional Head Selection approach - and argue that the latter is superior to the former both empirically and theoretically. The Functional Head Selection approach not only accounts for data from Cantonese-English code-switching which are erroneously ruled out by the Lemma Congruence model, it is also more economical as a theory in entailing no additional "code-switching" grammar. Assuming the Functional Head Selection approach is indeed correct, I would like to highlight a number of important implications (which I have mentioned at various places in this paper).

We need to rethink the necessity of a matrix language/ML vs. embedded language/EL dichotomy in the morpho-syntax of intra-sentential code-switching, even though in many code-switching discourses such a distinction is apparent (e.g. ML contributes more morphemes; the sentence structure is basically that of the ML, etc.). In those cases which I have analyzed as exceptions to the Lemma Congruence model - the English nouns, adjectives and prepositions - Cantonese appears dominant. Indeed, in most of these cases the whole sentence just contains one single English word. Supposedly Cantonese is ML and English is EL. Nonetheless, it is still possible for distinct features of the English grammar (i.e. the plural marker *-s*, the morphological alternation of *-ing/-ed* forms) to appear in a code-switched sentence. It follows that the Lemma Congruence model and the Matrix Language Frame model (which the former derives from) must be excessively restrictive.

The data and analysis presented in this paper has preferred certain assumptions in the theory of phrase structure which are still subject to controversy. Remember that there is an asymmetry between English verbs and nouns in their behaviour in Cantonese-English code-switching: the former must not be inflected while the latter may be inflected in appropriate contexts. To explain this asymmetry by the Functional Head Selection approach, one has to assign a different status to noun inflection and verb inflection, which seems not widely assumed in the literature. On the other hand, to explain the fact that the English verb is always uninflected in Cantonese-English code-switching, one has to assume that English tense/agreement markers (and the Cantonese aspect marker as well) are separately represented in syntactic structure in INFL. This is contrary to a

lexicalist account of English tense/agreement markers (e.g. Chomsky 1995, ch.4 and others) - that these markers have been affixed to a word when it enters syntactic derivations.

Last but not least, constraints on intra-sentential code-switching can be reduced to principles which govern monolingual utterances as well. Accordingly, monolinguals, bilinguals and polyglots essentially have the same "grammar" which generates monolingual and code-switched (involving two or more languages) utterances alike. In this paper, the relevant principle is the c-selection of functional categories. Other grammatical principles which have been discussed in this light include word order between heads and complements (Mahootian 1993) and feature-checking (MacSwan 1997).

References

Abney, S. (1987) *The English Noun Phrase and its Sentential Aspects*. Doctoral dissertation. MIT.

Belazi, H., Rubin, E. and Toribio, J. (1994) Code switching and X-bar theory: The Functional Head constraint. *Linguistic Inquiry* 25, 221-237.

Bentahila, A. (1995) Review of 'Duelling languages: Grammatical structure in codeswitching'. *Language* 71, 1, 135-140.

Bentahila, A. and Davis, E.D. (1983) The syntax of Arabic-French code-switching. *Lingua* 59, 301-330.

Bock, K. and Levelt, W. (1994) Language production: Grammatical encoding. In M. A. Gernsbacher (ed.), *Handbook of Psycholinguistics*, 945-984. London: Academic Press.

Bokamba, E. G. (1989) Are there syntactic constraints on code-mixing? *World Englishes* 8, 277-292.

Chan, B. (1992) *Code-mixing in Hong Kong Cantonese-English bilinguals: Constraints and processes.* Unpublished M.A. dissertation. Department of English, Chinese University of Hong Kong.

Chan, B. (1998) How does Cantonese-English code-mixing work? In M.C. Pennington(ed.), *Language use in Hong Kong at century's end.* Hong Kong: Hong Kong University Press.

Chan, B. (to appear) *Aspects of the syntax and pragmatics of Cantonese-English intra-sentential code-switching*. University College London.

Cheng, L. (1991) *On the typology of wh-questions.* Doctoral dissertation. MIT.

Chomsky, N. (1965) *Aspects of the theory of syntax.* Cambridge, Massachusetts: MIT Press.

Chomsky, N. (1981) *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, N. (1986a) *Knowledge of language.* New York: Praeger.

Chomsky, N. (1986b) *Barriers.* Cambridge, Massachusetts: MIT Press.

Chomsky, N. (1995) *The Minimalist Program.* Cambridge, Massachusetts: MIT Press.

Di Sciullo, A., Muysken, P. and Singh, R. (1986) Government and code-mixing. *Journal of Linguistics* 22, 1-24.

Garrett, M. F. (1980) Levels of processing in sentence production. In B. Butterworth (ed.), *Langauge Production Vol.1: Speech and Talk*, 177-221. Academic Press.

Gasde H-D. and Paul, W. (1996) Functional categories, topic prominence and complex sentences in Mandarin Chinese. *Linguistics* 34, 263-294.

Gibbons, J. (1987) *Code-mixing and Code choice: A Hong Kong Case Study*. Clevedon, Avon: Multilingual Matters.

Grimshaw, J. (1991) *Extended Projection*. Ms. Brandeis University.

Halmari, H. (1997) *Government and code-switching: explaining American-Finnish.* John Benjamins.

Jake, J. and Myers-Scotton, C. (1997) Codeswsitching and compromise strategies: Implications for lexical structure. *International Journal of Bilingualism* 1 (1), 25-40.

Kamwangamalu, N. M. (1989) The morphosyntactic aspects of French/English-Bantu code-mixing: Evidence for Universal Constraints. In *CLS 25, Parasession on Language in Context*, 157-170. Chicago, IL: Chicago Linguistic Society.

Kayne, R. (1994) *The Antisymmetry of Syntax*. Cambridge, Massachusetts: MIT Press.

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. Hong Kong: University of Hong Kong.

Levelt, W. (1989) *Speaking*. Cambridge, MA: MIT Press.

Li, D.C-S. (1996) *Issues in Bilingualism and Biculturalism: a Hong Kong case study*. New York: Peter Lang.

Longobardi, G. (1994) Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25(4), 609-665.

MacSwan, J. (1997) *A Minimalist approach to intra-sentential code-switching: Spanish-Nahuatl Bilingualism in Central Mexico.* Doctoral dissertation. University of California, Los Angeles.

Mahootian, S. (1993) *A null theory of code-switching.* Doctoral dissertation. Northwestern University.

Mahootian, S. and Santorini, B. (1996) Code switching and the complement/adjunct distinction. *Linguistic Inquiry* 27, 3, 464-479.

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

Muysken, P. and de Rooij, V. (1995) Review of 'Social Motivations for Code-switching: Evidence from Africa' and 'Duelling Languages: Grammatical Structure in Codeswitching'. *Linguistics* 33, 1043-1066.

Myers-Scotton, C. (1993) *Duelling languages: Grammatical structure in codeswitching.* Oxford: Oxford University Press.

Myers-Scotton, C. (1995) A lexically-based production model in code-switching. In Milroy, L. and Muysken, P. (eds.) *One speaker: Two languages: Cross-disciplinary perspectives*, 233-256. Cambridge: Cambridge University Press.

Myers-Scotton, C. and Jake, J. (1995) Matching lemmas in a bilingual language competence and production model: evidence from intrasentential code-switching. *Linguistics* 33, 981-1024.

Pennington, M. C., Chan, B. and Lau, L. (1996) Unpublished data of Cantonese-English code-switching. *Hong Kong bilingual radio: a changing profile*. (Strategic Grants Research from the City University of Hong Kong, with M. Pennington as principal investigator, B. Chan as assistant investigator and L. Lau as senior research assistant.)

Pennignton, M. C., Lau, L. and Chan, B. (to appear) Establishing discourse coherence bilingually: A

comparison of code-switching and code-mixing modes. *Conference Proceedings for XVIth International Congress of Linguists, Paris, July 20-25, 1997*.

Pollock, J.-Y. (1989) Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20, 365-424.

Radford, A. (1990) *Syntactic Theory and the Acquisition of English Syntax.* Basil Blackwell.

Ritter, E. (1991) Two functional categories in noun phrases: evidence from modern Hebrew.
Syntax and Semantics, 25: Perspectives on Phrase Structure, 37-62. Academic Press.

Stemberger, J.P. (1985) An interactive activation model of language production. In A. W. Ellis (ed.), *Progress in the Psychology of Language*, 143-186. Norwood, NJ: Erlbaum.

Teng, Anne, Y-Y. (1997) *Count/Mass distinction in Cantonese and its acquisition.* Ms. University College London.

Thráinsson, H. (1997) On the (non-)universality of functional categories. In W. Abraham, S. Epstein, H. Thráinsson, J-W. Zwart (eds.), *Minimal Ideas*, 253-281.