Original Paper

The Bilingual Lemma Activation Model as a Comparative Approach to Codeswitching

Longxing Wei[1*] & Xuexin Liu[2]

[1] Department of Linguistics, College of Humanities and Sciences, Montclair State University, New Jersey, USA

[2] Department of World Languages and Literature, Spelman College, Georgia, USA

[*] Longxing Wei, Department of Linguistics, College of Humanities and Sciences, Montclair State University, New Jersey, USA

Abstract

This paper explores intrasentential Codeswitching (CS) as a commonly observed bilingual speech behavior. Different from surface-based models, it investigates CS at an abstract level by relating the nature and activity of the bilingual mental lexicon in bilingual speech production to the structural principles governing CS. The Matrix Language Frame (MLF) Model is adopted for describing some fundamental structural principles governing CS, and the Bilingual Lemma Activation Model (BLA) is proposed for explaining the linguistic motivations for CS. Based on the analysis of some naturally occurring CS instances involving various language pairs, this study supports the claim that one of the bilingual's languages is activated as the Matrix Language (ML) and the other as the Embedded Language (EL), and content and system morphemes are unequally activated. It is the ML which provides the sentential frame for CS and the EL only provides content morphemes switched into this frame. It further argues that bilingual mental lexicon contains not only lexemes but also more abstract elements called "lemmas", and lemmas in the bilingual mental lexicon are language-specific and such lemmas are in contact in CS. This study provides evidence that only conceptually activated EL lemmas can be switched into the ML sentential frame.

Keywords

Codeswitching (CS), bilingual, mental lexicon, lemma, activation, conceptual, congruence

1. Introduction

As a commonly observed speech behavior, bilinguals may carry out a conversation in two languages, freely switching between two linguistic systems at will. That is, bilinguals may switch to another code at a certain point in their utterance production (i.e., codeswitching). CS may occur across sentence boundaries, that is, one sentence(s) is completely delivered in one language, and another sentence(s) is completely delivered in another language. This is called intersentential CS, involving a switch at a clause or sentence boundary, where each clause or sentence is well formed according to one of the languages involved within the stretch of speech in a discourse. This type of switching requires great proficiency in both languages. CS may also occur within a clause or sentence boundary, involving different types of constituents inserted into the syntactic slots provided by one of the languages known to the speaker. This type of switching is called intrasentential CS (Myers-Scotton, 1993).

This paper is a study of CS from some psychological and linguistic perspectives. This study adopts the Matrix Language Frame (MLF) model (Myers-Scotton, 1993 [1997]), which proposes some particular structural principles governing CS, and uses the Bilingual Lemma Activation (BLA) model (Wei, 2006b, 2009a, 2015), which explores the linguistic motivations for CS. The naturally occurring CS instances for the analysis and discussion involving different language pairs are selected from various published studies. The study reaches the conclusion that the languages involved in CS are not equally activated in structuring CS utterances; only content morphemes, rather than system morphemes, can be switched; lemmas contained in the bilingual mental lexicon are language-specific, and CS is driven by lexical-conceptual gap between the participating languages. To capture the explanatory adequacy of the BLA model compared with other models, a brief review of other models becomes necessary.

2. Models of Codeswitching

CS is a commonly observed linguistic phenomenon unique to bilingual speech production. Bilinguals may employ two (or more) linguistic systems within sentence boundaries, that is, they may switch certain items or constituents from one language into the syntactic slots provided by another language. For example, Tami/English: *Ellaam confused-aa irundadu* (Annamalai, 1989, p. 50) is an instance of CS where the English morpheme "confused" is switched into the Tami sentential frame Researchers like Poplack (1980), Joshi (1985), DiSciullo, Muysken and Singh (1986), Appel and Muysken (1987), Gardner-Chloros (1987), Azuma (1991), Myers-Scotton (1993 [1997]), Grosjean (1997), Jake and Myers-Scotton (1997, 2005), MacSwan (2000, 2009), Wei (2001b, 2002) have studied the phenomenon of CS from various perspectives. Most studies have focused on the analysis of specific grammatical structures of CS, that is, where in a sentence the speaker may switch from one linguistic variety to another and what linguistic items can be switched.

2.1 Surface-based Models

Poplack (1980, 1981), Poplack and Sankoff (1988) and Poplack and Meechan (1995) argue that the two languages involved in CS must remain intact at all levels. Poplack proposes two structural constraints on CS: *the Free Morpheme Constraint* and *the Equivalence Constraint. The Free Morpheme Constraint* specifies that "codes may be switched after any constituent provided that the constituent is not a bound morpheme" (Poplack, 1980, p. 585). According to this constraint, switches like the following are predicted not to occur. [1] is not well formed because the Spanish bound morpheme-*iendo* is connected to the English verb root *eat*.

[1] ***eat**-iendo

    -ing

   "Eating"

   (Spanish/English; Poplack, 1980, p. 586)

An examination of some naturally occurring CS instances reveal that this constraint is too general to account for numerous counter-examples in analytic and non-agglutinative languages.

[2] I'm **lav**-ing **pandekege**-s.

   I'm have-ing pancake-s

   "I'm having pancakes."

   (English/Danish; Petersen, 1988, p. 481)

[3] a Ne mI **help**-e.

   3 PL COP me -PRES PROG

   (Adaŋme/English; Nartey, 1982, p. 185)

[4] veo los **horses**.

   see the horses

   "I see the horses."

   (Spanish/English; Appel & Muysken, 1987, p. 125)

[5] Ellaam **confused**-aa irundadu.

   everything confused-ADV COP PAST

   "Everything was confused."

   (Tamil/English; Annamalai, 1989, p. 50)

In the English/Danish example, *-ing* (present progressive) and *-s* (plural) are English bound morphemes. In the Adaŋme/English example, *-e* (present progressive) is an Adaŋme bound morpheme. In the Spanish/English example, *los* is a Spanish bound morpheme. In the Tamil/English example, *-aa* is a Tamil bound morpheme. As these examples show, neither the bound vs. free morpheme distinction nor typological distinctions are directly relevant for permissible CS or intraword switching sites. "The key to acceptability is not whether a switch of languages may follow a bound morpheme (as the free-morpheme constraint proposes) but rather the source of the bound morpheme in question"

(Myers-Scotton, 1993, p. 33).

The Equivalence Constraint specifies that "code-switching will tent to occur at points in discourse where juxtaposition of L1 and L2 elements does not violate a syntactic rule of either language" (Poplack, 1980, p. 586).

[6] I   told him that so that he would bring it fast. (English)

    (Yo)le dije eso   pa′que   (él)   la   trjera   ligero. (Spanish)

In [6] the vertical lines indicate the places where switches between the two languages are possible, and the crossed lines indicate the places where switches are impossible. This constraint is restrictive enough not to allow switching to occur at points where the surface structures of the languages involved do not map onto each other.

Even some of the proponents of the surface-based models recognize apparent exceptions to this constraint. For example, Pfaff (1979) recognizes that switching may occur when the two languages do not share the common surface structures.

[7] **Me** hizo estudiar.

    "He made me study."

    (Spanish/English; Pfaff, 1979, p. 300)

[8] Tengo un **magazine** Nuevo.

    "I have a new magazine."

    (Spanish/English; Pfaff, 1979, p. 307)

English and Spanish do not share the same surface word order; however, the switches in these sentences are totally permitted. In [7] the object precedes the verb in Spanish but follows the verb in English. In [8] the adjective follows the noun in Spanish but precedes the noun in English.

The essence of such a surface-based model is that the surface category membership and the matched word order determine the switchability between the languages in question. The researchers analyzing CS data from various language pairs have come to very different conclusions regarding the "intactness" of the two participating languages and raised questions about the permissibility of a variety of switching points in CS (Bentahila & Davies, 1983; Gardner-Chloros,1987; Myers-Scotton, 1993). Wei (2015) offers some detailed critical discussions of all surface-based models.

2.2 Government and Binding Models

Unlike surface-based models, government and binding models stresses dependency rather than linearity. Klavans (1983), Woodford (1983), DiSciullo *et al*. (1986), Appel and Muysken (1987) and Pandit (1990) formulate this non-linear approach to structural constraints on CS. Appel and Muysken claim that "There cannot be a switch between two elements if they are lexically dependent on each other" (1987, p. 124). DiSciullo *et al*. (1986) claim that "No specific constraint needs to be stated to account for code-switching restriction. These fall out of general considerations of lexical integrity, constrained

by the government condition, which hold for all uses of natural languages, not just for code-switching" (1986, p. 4). What they emphasize is that the lexical governor and a governed maximal projection must have the same Lq (q is a language index). Thus, the government-binding models predict the acceptability of CS.

[9] veo los **horses**.

   see the horses

   "I see the horses."

[10] *veo **the horses**.

   (Spanish/English; Appel & Muysken, 1987, p. 125)

The switch in [9] would be acceptable because the Spanish *los* would make the whole NP Spanish, but the switch in [10] would be unacceptable because the whole NP, even though governed by the Spanish verb *veo*, is in English. As DiSciullo *et al*. (1986) speculated, switching is only possible at sites where nodes carry two indices (i.e., neutralization sites). The NP *los horses* carries two indices, but the NP *the horses* does not.

Though government and binding models look beyond surface linear ordering by identifying structural constraints on CS in terms of the phrase structure and government relations, they consider CS as basically a syntactic phenomenon following the same structural constraints evident in monolingual surface structure. Government and binding models fail to account for many naturally occurring CS instances.

[11] **Some Englishmen traditional Indian women**-ko passand karaten hain.

   some Englishmen traditional Indian women-ACCUS like do are

   "Some Englishmen like traditional Indian women."

   (Hindi/English; Pandit, 1990, p. 44)

[12] Kwetu sisi mtu hawezi kuleta **jokes** kama hizo …

   "At your place, a person can't bring jokes like these ones …"

   (Swahili/English; Myers-Scotton, 1993, p. 44)

[13] **Of all the places John has hidden** kuch **books bathroom** men.

   of all the places John has hidden some books bathroom in

   "Of all the places, John has hidden some books in the bathroom."

   (Hindi/English; Pandit, 1990, p. 53)

As shown in [11], CS can occur in the VP, where the English NP *traditional Indian women* plus the Hindi suffix for the accusative case marking is governed by the Hindi V *passand* (like) Also, in [12] the English NP *jokes* is governed by the Swahili V *kuleta* (bring) As shown in [13], CS can also occur in the PP, where the English NP *bathroom* is governed by the Hindi P *men* (in).

It is apparent that government and binding theory alone is not adequate enough to account for CS because a purely syntactic approach is still too close to the surface-based analysis.

2.3 Subcategorization Models

Subcategorization models propose more abstract level equivalence to account for switchability in terms of the lexically based subcategorization restrictions. Bentahili and Davies claim that "all items must be used in such a way as to satisfy the [language-particular] subcategorization restrictions imposed on them," and "switching is freely permitted at all boundaries above that of the word, subject only to the condition that it entails no violation of the subcategorization restriction on particular lexical items of either language" (1983, p. 329). Muysken (1990, 1991) also claims that it is lexical subcategorizations that constrain switched elements. That is, in order for switching to occur, the demands made by individual lexical items with respect to their syntactic environment as expressed in subcategorization frames must be satisfied. He proposes that switching cannot occur between X and Y if X L-marks Y (L-marking refers to the special relation between a lexical item and the complement which it governs and theta-marks (cf. Chomsky, 1986, p. 15), and there is no equivalence between X and Y in both languages. That is, lexical elements impose certain requirements on their environments, and switched elements are constrained by lexical subcategorizations.

Azuma claims that "the subcategorization of the main verb is always preserved" and "the main verb provides a planning frame … content word insertion must be done within the specifications of the planning frame" (1991, p. 7). Thus, he proposes *the Frame-content Hypothesis*, which identifies two stages in CS: the planning frame-building where closed class items are accessed and retrieved, and the content word insertion state where content words are inserted in the planning frame. According to Azuma, the closed class elements include all grammatical (i.e., functional) items and inflectional morphemes and they are essential members of the planning frame; content words are not members of the planning frame but are inserted in a later stage. He proposes that CS only occurs in the stage of content word insertion, where the content word from the other language involved in CS is inserted into the available slot in the planning frame. This means that all closed class items must be from the "base" (or "host") language, and only content words from the "guest" language can be switched into the sentential frame projected by the "base" language. Below are two of the examples cited by Azuma in support of this hypothesis.

[14] Hata si-ku-**comment** …

    even 1S.NEG-NEG.PAST-comment …

    "I didn't even comment …"

    (Swahili/English; Scotton, 1983)

As shown in [14], the English verb *comment*, a content word, is inserted into the planning frame provided by Swahili, and all grammatical items are from Swahili.

[15] Elle desire **tzwez had I'am**.

    she wants she gets married this year

    "She wants to get married this year."

(French/Arabic: Bentahila & Davies, 1983)

According to Azuma, the switch in [15] is predicted not to occur because the French verb *desire* subcategorizes for an infinitive complement. That is, the finite complement switched from Arabic fails to fit into the French planning frame.

Azuma's subcategorization model identifies the unequal roles of the participating languages in CS, one providing the sentential frame, and the other inserting content words into this frame. Though subcategorization models look beyond surface configurations (i.e., surface linear orderings) of sentential elements contained in CS, they still consider CS as a syntactic and lexical phenomenon similar to those of monolingual surface structures.

2.4 The Minimalist Model

The major theoretical assumption underlying the Minimalist Model is that CS is simply a linguistic phenomenon where two monolingual systems are in contact, and each monolingual system must be preserved in order for the mixed constituents to be grammatical. This model considers grammaticality judgments as primary data in support of its theoretical assumption (MacSwan, 2000, 2009). MacSwan explains the ungrammaticality of examples like [16] and [17] in terms of a "PF Disjunction Theorem" which disallows CS at the level of PF (Phonetic Form): phonological system cannot be mixed" (2000, p. 45). The PF Disjunction Theorem "predict[s] code switches involving head movement should be ruled out since the movement results in the formation of complex Xs" (2000, p. 46); the aspectual verb and the lexical verb are reanalyzed as one unit. MacSwan restates the PF Disjunction Theorem as "The PF Interface Condition," again disallowing CS in head movement contexts (2009, p. 331).

[16] *The students had visto la pelicula italiana.

"The students had seen the Italian movie."

[17] *Los estudiantes habían seen the Italian movie.

"The students had seen the Italian movie."

(MacSwan, 2000, p. 42)

However, the CS literature and findings do not fully support the PF Interface Condition. Pfaff (1979) reports two examples of CS between *haber* (have) and an English past participle. Below is one of them.

[18] Yo creo que apenas se habia **washed out**.

"I think it has just …"

(Pfaff, 1979, p. 300)

In addition to the PF Disjunction Theorem or the PF Interface Condition, MacSwan that the *phi*-feature checking be a necessary condition of CS to occur. One particular set of features discussed by MacSwan (2000) in his explanation of patterns in Spanish/English CS concerns the realization of *phi*-features grammatical nominal features such as person, number, and gender on elements under the determiner node. Spanish and English differ in the overt realization of *phi*-features in NPs and also in the larger role that *phi*-features play in their respective agreement system. That is, the possible values of

83

phi-features are different in English and Spanish. While in Spanish noun *phi*-features are more specified, most determiners in English are not overtly distinguished on the basis of gender or number, except the demonstratives. In MacSwan's analysis, English determiners fail to occur within Spanish NPs because English determiners lack grammatical gender. Thus, only Spanish determiners can occur with Spanish NPs. As claimed, such a *phi*-feature approach can account for CS patterns with any "control structure" or any "mediating mechanisms" (MacSwan, 2000, p. 45). However, this approach is not sufficient enough to explain why English determiners do occasionally occur with Spanish NPs (see the examples in Jake, Myers-Scotton and Gross (2002, 2005)).

It seems that the forces patterning CS evidently cannot be reduced to *phi*-feature checking, lexical insertion, or surface configuration. The *phi*-feature checking approach resolving *phi*-feature mismatches in favor of the language in which the *phi*-features are strong or more complex does not necessarily make correct predictions for other language pairs involved in CS. For detailed discussion regarding the Minimalist Model in CS, see Myers-Scotton and Jake (2001), Jake *et al*. (2002, 2005), Myers-Scotton (2002) and Wei (2015).

2.5 The Matrix Language Frame Model

As an over-arching principle of the MLF Model, the Uniform Structure Principle specifies (Myers-Scotton, 2002, pp. 8-9):

> A given constituent type in any language has a uniform abstract structure and the requirements for well-formedness for this constituent type must be observed whenever the constituent appears. In bilingual speech, the structures of the Matrix Language are always preferred, but some Embedded Language structures are allowed if certain conditions are met.

This principle makes two interrelated asymmetries as structuring CS utterances (Myers-Scotton, 1991, 1993 [1997], 1994, 2002): The Matrix Language (ML) vs Embedded Language (EL) and the content vs. system morphemes. The ML vs. EL asymmetry specifies that unequal roles played by the languages participating in CS. It is the ML which sets the grammatical frame for CS. This frame determines the morpheme order and system morphemes. The system vs. content morpheme asymmetry specifies differential accessing of content vs. system morphemes in CS utterances. This asymmetry is crucial in accounting for three aspects of switchability: (i) All system morphemes having syntactic function external to their lexical heads must come from the ML. (ii) Morpheme insertion into the ML frame is possible only for those EL content morphemes which are congruent with the ML counterparts (Myers-Scotton, 1993, p. 120) for the issue of "congruence"). (iii) An intention to access an EL morpheme which is either a system morpheme or a content morpheme incongruent with its ML counterpart triggers an EL island.

Based on these two asymmetries, the MLF Model predicts two common types of constituents containing switched items:

> ML + EL constituents consist of any number of ML morphemes and (generally) singly occurring EL content morphemes.

[19] Ha-u ku-on-a a-ki-ni-**buy**-i-a **beer** siku hiyo?

NEG-2s NEG.PST-see-FV 3s-PROG-1s.OBJ-buy-APPL-FV beer day CL9. that

"Didn't you see him buying beer for me that day?"

(Swahili/English; Myers-Scotton, 2002, p. 98)

[20] wo zuijin hen **busy**, you san-fen **paper** bixu zai yue-di qian **finish**.

I recently very busy have three-CLASSIF paper must PREP/TIM month-end

"I'm very busy recently. I must finish three papers before the end of this month."

(Chinese/English; Wei, 1992)

[21] dore gurai koko ni **stay** suru no?

how long about here LOC stay do PARTIC/QUE

"About how long will you stay here?"

(Japanese/English; Wei, 2006b, p. 164)

In [19], *buy* and *beer* are content morphemes from English, the EL, but Swahili, the ML, provides the grammatical frame realizing how thematic content of the clause is mapped onto a clause well-formed in Swahili. In this example, *buy* takes two objects, but the grammatical frame of Swahili determines how those objects are realized; the beneficiary is realized as an object prefix on the verb *-ni-* and further mapped on the grammatical frame through the applied verbal suffix *-i-* on the EL verb from English. In [20], *hen busy* (very busy) is a mixed constituent, with the adjective *busy* from English, the EL; *san-fen paper* (three papers) is a mixed constituent, with the number and the noun classifier *san-fen* from Chinese, the ML, and the noun *paper* from the EL. It should also be noticed that the EL verb *finish* is switched into the ML grammatical frame where the object *san-fen paper* occurs before the verb. In [21], *stay* is a content morpheme from English, the EL, used in conjunction with *suru* (to do) from Japanese, the ML ("loan word + suru" is a typical Japanese verbal structure).

> EL islands consist of only EL morphemes, including EL system morphemes, and are well formed according to the EL grammar to show internal structural dependency relations.

[22] **The first one** que era elque llevaba para Maracaibo.

the first one COMP COP.S.IMP DEF.M.S. COMP go.3s.IMP PREP Maracaibo

"The first one, that was the one which was going to Maracaibo."

(Spanish/English; Blazquez-Domingo, 2000, cited in Jake *et al*., 2002, p. 81)

[23] Eb dann simmer go **le pentole** bring.

exactly then be.1.PL [we] go the.F.P. pan.P take-INF

"Exactly, and then we took the pans there."

(Swiss German/Italian; Preziosa-Di Quinzio, 1992, Appendix XXX)

[24] It's **totemo muzukashi** to find a convenient and **yasui** apartment here.

85

it's very difficult to find a convenient and cheap apartment here

"It's very difficult to find a convenient and cheap apartment here."

(English/Japanese; Wei, 2006b, p. 167)

In [22], *the first one* is an EL island where both the system morphemes *the* and the content morphemes *first one* are from English. In [23], *le pentole* is an EL island where both the system morpheme *le* (the) and the content morpheme *pentole* (pans) are from Italian. In [24], *totemo muzukashi* (very difficult) is an EL island where both the system morpheme *totemo* (very) and the content morpheme *muzukashi* (difficult) are from Japanese.

Myers-Scotton (1993 [1997]) offers a detailed treatment of how the Matrix Language Hypothesis is supported by the two basic principles: The Morpheme Order Principle and the System Morpheme Principle. The Morpheme Order Principle predicts that the morpheme order in the mixed constituents must be that of the ML. The System Morpheme Principle specifies that all the syntactically relevant system morphemes must come from the ML in mixed constituents. EL system morphemes can only occur in EL islands, which must be embedded in the ML grammatical frame.

3. The Bilingual Lemma Activation Model

Adopting Levelt's (1989) speech production model, Myers-Scotton and Jake's (1995) bilingual language competence and production model and Wei's (2002) bilingual speech production model, Wei (2006a, 2006b, 2015) proposes the Bilingual Lemma Activation (BLA) Model to explain bilingual speech production involving CS. This model consists of four sequentially connected levels of speech production: the conceptual level → the lemma level → the functional level → the position level. Each level plays a particular role in the bilingual speech production process. At the conceptual level the CONCEPTUALIZER generates messages by attending to the speaker's communicative intention about the discourse mode, either the monolingual mode or the bilingual mode (cf. Grosjean, 1997) and preverbal message (cf. Myers-Scotton & Jake, 1995, 2000a) to be desired. If the speaker chooses the monolingual mode, no CS will occur; if the speaker chooses the bilingual mode, then he/she must decide whether intersentential or intrasentential CS should be performed. If the speaker decides to perform intrasentential CS, the CONCEPTUALIZER then generates the preverbal message about his/her choice of the language as the ML to be used and the semantic/pragmatic feature bundles to be desired for his/her communicative intention. The output of the CONCEPTUALIZER is the speaker's preverbal message which gives input to THE BILINGUAL MENTAL LEXICON at the lemma level. It is at this level that language-specific lemmas are in contact for activation.

However, the activation of language-specific lemmas alone is not sufficient enough for CS to occur. Lemma congruence checking between the languages involved in CS must come into play. Myers-Scotton and Jake (1995, p. 985) define lemma congruence as "a match between the ML and the EL at the lemma level with respect to linguistically relevant features" and regard lemma congruence

checking as an organizing principle for CS. The naturally occurring CS instances studied by Wei (2001b, 2002, 2006b) provide sufficient evidence that lemma congruence between the languages involved must be checked at the level of lexical-conceptual structure, at the level of predicate-argument structure, and at the level of morphological realization patterns. If the lemmas of the EL are congruent or sufficiently congruent with the counterparts of the ML at each of these levels, they can be activated for the speaker to proceed with the bilingual mode for CS; otherwise, the speaker must take some compromise strategies for possible CS realization or go back to the monolingual mode. Only when there is a match between the EL and ML lemmas or compromise strategies are taken, directions will be sent to the FORMULATOR at the functional level for morphosyntactic encoding by observing the structural principles governing CS. The successfully encoded morphosyntactic material will then be sent to the ARTICULATOR at the positional level for morphophonological encoding. The successfully encoded morphophonological material will then produce surface forms of word order and phonetic string, that is, speech output for speech comprehension.

Levelt defines a lemma as the "nonphonological part of an item's lexical information," including semantic, syntactic, and some aspects of morphological information, and claims that "it is the lemmas of the mental lexicon that conceptual information is linked to grammatical function" (1989, p. 162). In other words, lemmas are abstract entries in the mental lexicon and underlie surface configurations of speech production. Each lemma in the mental lexicon contains its own lemma specification, comprising declarative knowledge about the word's meaning as well as information about its syntax and morphology. For example, the lemma for the verb "like" requires a subject that expresses the thematic role of EXPERIENCER and an object that expresses the thematic role of THEME; the lemma for "she" specifies that the word must refer to a female and that any following present-tense main verb must have the inflectional morpheme "-s" for subject-verb agreement.

It seems obvious that the activation of lemmas in the mental lexicon plays a central role in speech production. The BLA Model (Wei, 2006b, 2009a, 2015) confronts and expands on Levelt's model of monolingual speech production by explaining and emphasizing the role of lemma activation. It claims that lemma activation of particular lexical items in the mental lexicon must mediate between conceptualization and speech formulation as an indispensable level of speech production. The role of lemma activation in speech production can be schematized as in Figure 1.

Levelt's model of speech production was designed for describing the major components and processes of monolingual speech production, and it must be adapted to account for bilingual speech behavior such as CS. The BLA Model proposes that the bilingual mental lexicon differs from the monolingual mental lexicon in that the former contains lexemes and their lemmas from both languages. Thus, it is the speaker's preverbal message/intention that activates language-specific lemmas in the bilingual mental lexicon. In other words, it is the semantic/pragmatic feature bundles selected by the CONCEPTUALIZER that trigger the appropriate lemmas into activity before the FORMULATOR has

87

access to the relevant lexical item in the mental lexicon. As the figure shows, lemmas in the mental lexicon, whether monolingual or bilingual, mediate between the CONCEPTUALIZER and the FORMULATOR (cf. Levelt, 1989; Myers-Scotton & Jake, 2000; Wei, 2002).
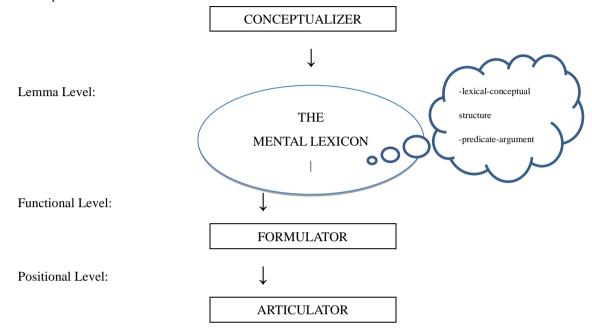
Conceptual Level:

CONCEPTUALIZER

↓

Lemma Level:

THE
MENTAL LEXICON
|

-lexical-conceptual structure
-predicate-argument

Functional Level:

↓

FORMULATOR

Positional Level:

↓

ARTICULATOR

Figure 1 Lemma Activation in Speech Production

(Adapted from Levelt, 1989; Myers-Scotton & Jake, 2000; Wei, 2009b)

3.1 Lemmas in the Mental Lexicon

The mental lexicon is generally defined as the store of information about particular words in one's language. As Richards (1976), Faerch and Kasper (1984), Talmy (1985), Ringbom (1987), Nation (1990), Wei (2001a, 2001b, 2002) and others have explained knowing a word means the ability to retrieve the word from the mental lexicon about its spelling and pronunciation, its meaning(s), its grammatical class and syntactic environment, its collocations and syntagmatic associations, its lexical and conceptual associations, and its registers. In speech production, speakers map what they intend to say onto words retrieved from lexical items currently stored in the mental lexicon. In other words, speakers conceptually retrieve the appropriate words from the mental lexicon to correctly express their intended meanings. A lexical item is retrieved from the mental lexicon not only contains its lexical content (i.e., its meaning) but also phonological, morphological, and syntactic information. Thus, a lexical item is a rather complex entity. When speakers construct an utterance, they build a sentential frame without much regard for the phonological aspects of words (see evidence in Levelt, 1989; Levelt, Roelofs, & Meyer, 1999) by using the syntactic information and aspects of the morphological information contained in the lexical items as retrieved from the mental lexicon. Such lexical

information is called *lemma information* (for short, *the lemma*) (cf. Kempen & Huijbers, 1983; Kempen & Hoenkamp, 1987). When we say that speakers have retrieved the lexical items from the mental lexicon, we mean they have acquired access to the lemmas that are relevant for the construction of the word's syntactic environment (i.e., the sentential frame or the grammatical configuration). Thus, the mental lexicon does not simply contain lexemes and their meanings but more abstract elements called "lemmas". Lemmas are defined as abstract entries in the mental lexicon that support the surface realization of actual lexemes. They are abstract in the sense that for each lexical item, the mental lexicon contains lemma information, that is, declarative knowledge about the word's meaning, and information about its syntax and morphology which is necessary for constructing the word's syntactic environment. Take, for instance, the lemmas of *know* require a subject that expresses the theta role of EXPERIENCER, and object that expresses the theta role of PERCEPT (i.e., what is known), and these elements appear in a particular order; the lemmas of *he* require the word to be used of a male and that the inflectional morpheme *-s* for the third person singular must be attached to the following present-tense main verb (i.e., inflectional morphology for tense marking). Lemmas also contain information about the word's phonological structure, syllabic composition, and accent structure. In addition, lemmas may contain information about the word's register, the kind of discourse it typically enters into, and its pragmatic function. Thus, the mental lexicon is defined as the speaker's internal representation of language specific knowledge about the surface forms, that is, lemmas.

Regarding lemma activation in speech production involving CS, the BLA Model draws on the theoretical assumptions presented by Myers-Scotton and Jake (1995, 2000a, 2000b) that all lemmas include three distinctive but related levels of abstract lexical structure. The first is the level of lexical-conceptual structure. At this level, lexical access takes place on the basis of the information contained in the speaker's preverbal message, that is, the speaker's intention before speech production. It is the speaker's preverbal message in the CONCEPTUALIZER which activates language-specific semantic/pragmatic feature bundles at the interface between the CONCEPTUALIER and the mental lexicon. These activated semantic/pragmatic features are then mapped onto lemmas in the mental lexicon as lexical-conceptual structure. Green (1986, 1989) also holds that a lemma is activated if it matches part of the conceptual structure created by the CONCEPTUALIZER. The second is the level of predicate-argument structure. At this level, the thematic structure of a particular verb is mapped onto grammatical relations (i.e., thematic role assignment). The third is the level of morphological realization patterns. At this level, surface grammatical relations, such as word order, agreement, inflectional morphology for tense/aspect marking, etc. are realized.

Based on the model of lemma activation in speech production, the BLA Model claims that lemmas contained in the bilingual mental lexicon are language-specific and such language-specific lemmas are in contact during a discourse involving CS at these levels of abstract lexical structure. Accordingly, CS is described and explained in terms of activation of language-specific lemmas at any of these levels of

abstract lexical structure. The BLA Model specifically deals with the issues of unequal activation of bilingual lemmas from the language pairs involved in CS. It also discusses several issues of lemma congruence checking between the languages involved in CS as an organizing principle governing such a bilingual speech activity. The study of CS presented here leads to several specific hypotheses about structural principles governing the bilingual speech production process.

3.2 Unequal Activation of Bilingual Lemmas

The BLA Model assumes that CS juxtapositions which may surface do not have much to do with surface linear of typological correspondences between the participating languages. Instead, it proposes that CS juxtapositions originate with directions contained in lemmas (cf. Myers-Scotton & Jake, 1995; Wei, 2001a, 2001b). As introduced earlier, lemmas are abstract entries in the speaker's mental lexicon which support the surface realization of actual lexemes. This is because lemmas contain phonological, morphological, semantic, syntactic, and pragmatic information about lexemes stored in the mental lexicon. Thus, lemmas in the mental lexicon are defined as the speaker's internal representation of knowledge about surface forms. The BLA Model further assumes that lemmas in the bilingual mental lexicon are language-specific, and such lemmas are in contact in bilingual speech, especially in CS.

The BLA Model proposes that lemmas in the bilingual mental lexicon are language-specific and lexicalization patterns across languages reflect the fact that there are different configurations of semantic and pragmatic features across related lemmas in different languages. Hypotheses about cross-linguistic differences in how information is organized at the level of lexical-conceptual structure and at the level of predicate-argument structure, whether semantic or pragmatic or semantic with morphological consequences, affect the code choice and structures which will appear in CS.

3.2.1 Unequal Lemma Activation of Morphemes

According to Poulisse and Bongaerts (1994), each individual conceptual chunk is specified for a particular language in the speaker's preverbal message. What needs to be emphasized is that it is at the level of lexical-conceptual structure that the speaker seeks appropriate linguistic material for his/her communicative intention. It is at this level that the speaker conceptualizes a message and activates the appropriate concepts accordingly. The activated concepts will then spread activation to the corresponding lemmas in the mental lexicon. Sufficiently activated lemmas will then spread activation to the associated lexeme (cf. Roelofs, 1992; Levelt, 1995; Wei, 2001a, 2001b). Wei (2001b) provides evidence that the presence and conflation of universally available semantic and pragmatic features may vary cross-linguistically. Thus, the switched items in CS are seen as evidence of the relative importance of cross-linguistic lexical-conceptual differences in lemmas in the bilingual mental lexicon.

Commonly observed CS instances suggest that it is individual content morphemes, rather than system morphemes, which encode the speaker's specific communicative intention. However, at the conceptual level, the speaker does not produce surface level morphemes but rather makes choices about the semantic and pragmatic information which he/she wishes to convey. Wei (2001b, 2002) suggests that

90

one of the major reasons for content and system morphemes to be accessed differently is that it is content morphemes, rather than system morphemes, which contain semantic/pragmatic feature bundles. The speaker may switch to certain EL content morphemes at a certain point during a discourse to convey his/her intended meanings as generated in his/her preverbal message. It is in this sense that content and system morphemes are not equally activated, and it is also in this sense that certain language-specific lemmas are conceptually projected in CS.

Some typical instances of CS involving various language pairs show that EL content morphemes can be switched because they are projected from the EL lemmas which are sufficiently congruent with those of the ML.

[25] Kerran sä olit pannu si-tä mun **lunchbox**-iin.

once you had put it+PRT my lunchbox-IL

"You had once put it in my lunchbox."

(Finnish/English; Halmari, 1997, p. 59)

[26] Molemmat niinku teki ton **language**-in koulussa.

both-PL as/like do-PAST3SG that-ACC language-ACC school-in

"Both liked the language at school."

(Australian Finnish/English; Kovács, 2001, p. 152)

[27] Se sai semmose-n **stroke**-Ø.

s/he get-IMP3SG like-ACC stroke

"She had like a stroke."

[28] Mi tyala ghar ghyayla **persuade** kela la.

I he-DAT house to buy persuade did "to"

"I persuaded him to buy a house."

(Marathi/English; Joshi, 1985, p. 197)

[29] nei5 zou6 saai3 di **assignment** mei6.

you do ASP CL assignment SFP

"Have you done all the assignments?"

(Cantonese/English; Chan, 1998, p. 193)

[30] naan pooyi paaDuvein **Hindi song**-ei.

I go-INF sing Hindi song-ACC

"I will go and sing a Hindi song."

(Tamil/English; Sankoff, Poplack, & Vanniarajan, 1990, p. 79)

[31] I command you to do the **nokum**.

I command you to do the recording

"I command you to do the recording."

(English/Korean; Choi, 1991, p. 889)

[32] evet, **terras**-ta oturuyorlar.

    yes cafe-LOC sit-PROG.3PL

    "Yes, they are sitting at the outdoor cafe."

    (Turkish/Dutch; Backus, 1996, p. 140)

[33] Zachem ty na **grass**-e valjajesih'sja.

    what-for you.SG on grass-PREP.SG roll-around

    "Why are you rolling around on the grass?"

    (Russian/English; Schmitt, 2006)

[34] shi-bu-shi qu nei-ge **new library**?

    yes-not-yes go that-CLASSIF library

    "Are we going to that new library?"

    (Chinese/English; Wei, 2001b, p. 158)

[35] ima wa **summer course** o tot-te iru n.

    now PARTIC/TOP summer course PARTIC/OBJ take-PROG AUC/be PARTIC

    "(I)'m taking summer courses now."

    (Japanese/English; Wei, 2009b, p. 322)

In [25] *lunchbox* is a content morpheme from English, the EL, but it is marked with the appropriate Finnish case, an ML system morpheme. In [26] *language* is a content morpheme from English, the EL, but it is marked by both the deictic element preceding it and the case, the Australian Finnish system morphemes. In [27] the case marking is missing on *stroke*, an EL content morpheme. In [28] the complementizer *la* (to) is a system morpheme from Marathi, the ML. In [29] the noun classifier *di1*, a system morpheme, is from Cantonese, and also *assignment* is not inflected for the plural marking, which is not required in Cantonese. In [30] *Hindi song* is case marked by *-ei*, an ML system morpheme. In [31] *nokum* is an EL content morpheme from Korean, but the article *the*, a system morpheme, is from English, the ML. In [32] *terras* receives locative case from Turkish, the ML. In [33] English *grass* is inflected with prepositional case from Russian, the ML. In [34] the noun phrase *new library* appears with the ML demonstrative determiner *nei-ge* (that-CLASSIF). In [35] the noun phrase *summer course* is from the EL, but *o* marking the accusative case, a system morpheme, is from the ML.

The above examples provide the evidence that in CS, bilingual lemmas are in contact and are not equally activated. EL content morphemes can be freely activated to be switched for the speaker's communicative intention at a certain point during a discourse, but EL system morphemes cannot. It should be noticed that the System Morpheme Principle Myers-Scotton (1993 [1997]) applies to all language pairs involved in CS. That is, all system morphemes must come from the ML.

3.2.2 Unequal Lemma Activation of Morphosyntactic Procedures

The MLF Model emphasizes that CS occurs within the constraints of a sentential frame which is set by structural procedures (i.e., morphosyntactic rules) dictated by the ML word order and syntactically

92

relevant relational or functional elements (i.e., system morphemes). If the speaker chooses to engage in CS at a certain point during a discourse, he/she automatically activates one of the participating languages as the ML to provide the sentential frame. The BLA Model proposes that it is the activated language-specific lemmas which send directions to the FORMULATOR at the functional level for morphosyntactic encoding (see the Figure). Whichever language is chosen as the ML at the conceptual level, its corresponding language-specific lemmas are activated at the lemma level to realize the speaker's preverbal message, resulting in language-specific morphosyntax.

According to de Bot and Schreuder (1993) and Wei (2015), bilingual speakers are able to separate the systems of the languages they know and to mix them in a bilingual mode. According to Grosjean (1989, 1997), the amount of language mixing depends on the language mode the bilingual speaker is currently in, whether monolingual, bilingual, or anywhere else on the language-mode continuum between these two modes. Similarly, Wei (2009a, 2015) assumes that it is particular levels of activation of the ML and the EL which determine the bilingual speaker's adoption of a particular position on the language-mode continuum. This is because bilingual speakers know that the ML and the EL play unequal roles in CS. That is, bilingual speakers can clearly separate the two language systems and switch items from one language into the other. Bilingual speech is not so-called "mixed" speech but is governed by a set of underlying structural principles, such as those proposed in the MLF Model (Myers-Scotton, 1993 [1997], 2002; Myers-Scotton & Jake, 1995; Wei, 2001b) and the BLA Model (Wei, 2006b). Such a language separation in bilingual speech is evidenced in the CS examples involving different language pairs below.

[36] Ø-saa hi-yo i-na-**depend** na Ø-**certificate** z-ako.

   c.9-time DEM-c.9 c.9-non-PAST-depend with c.10-certificate c.10-your z-a Ø-shule c.10-ASSOC

   c.10-school

   "At this time, it depends on your school certificates."

   (Swahili/English; Myers-Scotton, 2004, p. 108)

[37] baceã nũ tusĩ **force** nəi kər sakde.

   children ACC you force NEG d

   "You can't force children."

   (Panjabi/English; Romaine, 1995, p. 140)

[38] want ou Tex laat ons daai **group** join.

   because old Tex make 1PL DEM group join

   "Because old Tex made us join that group."

   (Tsotsitaal/English; Slabbert and Myers-Scotton, 1997, p. 332)

[39] I have to **ttakē** my hand.

   I have to wash my hand

   "I have to wash my hand."

(English/Korean; Choi, 1991, p. 889)

[40] mula khurcyā **paint** kartāt.

boys chairs paint do+TNS

"Boys paint chairs."

(Marathi/English; Joshi, 1985, p. 193)

[41] n buka wo **understand** – noo.

1-SG TAM that understand – AUX

"I'm not able to understand that."

(Mandinka/English; Haust and Dittmar, 1998, p. 87)

[42] na wo yi dian **come to pick you up**.

so I one o'clock come to pick you up

"So, I'll come to pick you up at one o'clock."

(Chinese/English; Wei, 2001b, p. 162)

[43] ni keyi ba zhe-zhang dade **sleeping sofa** fang zai xiaode **bedroom** li.

you can PREP/OBJ this-CLASSIF big sleeping sofa put PREP/LOC small bedroom PART/in

"You can put this big sleeping sofa in the small bedroom."

(Chinese/English; Wei, 2009b, p. 325)

[44] It's **totemo muzukashi** to find a convenient and **yasui** apartment here.

it's very difficult to find a convenient and cheap apartment here

"It's very difficult to find a convenient and cheap apartment here."

(English/Japanese; Wei, 2002, p. 280)

In [36] the order of *certificate* and its modifiers follow the Swahili word order, not that of English. In [37] *force* follows the Panjabi OV order. In [38] *join* follows the Tsotsitaal OV order. In [39] *ttakē* (wash) follows the English VO order. In [40] *paint* follows the Marathi OV order. In [41] *understand* follows the Mandinka OV order. In [42] the EL verb phrase *come to pick you up* is switched into the Chinese word order, where the adverbial of time *yi dian* (one o'clock) immediately proceeds the verb phrase. In [43] the prepositional phrase *zai xiaode bedroom li* (in the small bedroom), where the preposition *zai* signifying LOCATION closed by a particle *li* (in) indicates a specific location (other relevant particles may be used for other specific locations). In this example, the EL content morpheme *bedroom* is switched into the ML predicate argument- structure. In [44] the infinitive clause with the formal subject pronoun *it* is a typical English construction, into which the Japanese noun phrase *totemo muzukashi* (very difficult) is switched.

These examples show that in addition to the system morphemes from the ML, all the EL items are switched into the sentential frame set by the ML. It is obvious that the speakers performing CS know which language is being activated as the ML, that is, which language provides the word order.

4. Lemma Congruence Checking for Code-Switching Configurations

The naturally occurring CS instances discussed above demonstrate the unequal roles of the languages involved in CS. Furthermore, in order for CS configurations to be possible, lemma congruence between the language pairs must be checked. Myers-Scotton and Jake define lemma congruence as "a match between the ML and the EL at the lemma level with respect to linguistically relevant features" (1995, p. 985). The MLF Model and the BLA Model assume that at the conceptual level, the activation of the EL item must satisfy the semantic/pragmatic features of the speaker's intentions. Then at the level of the bilingual mental lexicon, three levels of abstract lexical structure must be checked for congruence. At the level of lexical-conceptual structure, lexical items relevant to the speaker's communicative intentions are activated, which is requisite for any CS. At the level of predicate-argument structure, thematic structure is mapped onto grammatical relations. At the level of morphological realization patterns, word order, agreement morphology, case marking, tense/aspect marking, and phonological forms are realized. Lemma congruence checking must take place at each of these levels in the bilingual mental lexicon, which determines the way an EL item may be integrated into an ML frame, either as a singly-occurring morpheme or as a larger EL island. Relevant to the current study is lemma congruence checking at the first two levels.

The BLA Model views the nature of congruence relevant to CS as more complex in that several different levels or subsystems have to be checked and articulates the relation of lemma congruence checking to a model of bilingual speech production. This model assumes that CS juxtapositions which may surface do not have much to do with superficial linear or typological correspondences between the participating languages. It argues that since lemmas are abstract entries in the mental lexicon (Levelt, 1989), CS juxtapositions originate with directions in the speaker's mental lexicon (Myers-Scotton & Jake, 1995). It further argues that lemma congruence is measured by the same universally-present metrics in all instances of CS, rather than by specific language-pair subcategorization patterns.

4.1 Lemma Congruence Checking for Lexical-Conceptual Structure

Grosjean (1982) reports that some code-switches are motivated by the lack of a particular word in one of the languages or by the greater availability of a word in the other language. Thus, such switches are deliberate or intentional. As Grosjean notes, such switches are motivated to "fill a linguistic need" or to use the word most available in the other language, "the most available word phenomenon" (1982, p. 151). Similarly, the BLA Model claims that because the language pairs involved may differ in their lemmas at the level of lexical-conceptual structure, it is this type of difference which motivates CS. In terms of lemma congruence between the languages involved, if such a difference is only partial, there is still sufficient cross-linguistic congruence. The BLA Model further suggests that a partial difference at the level of lexical-conceptual structure is one of the major reasons why certain morphemes are switched from the EL as chosen.

[45] wo xiawu qu jian wode **advisor**. wo bu neng he ni yiqi qu **mall** le.

95

I afternoon go see my advisor I not can with you together go mall PARTIC/AFFIRM

"I'm going to see my advisor this afternoon. I can't go to the mall with you."

[46] zhu zai zheli hen fanbian, meitian you **school bus**.

live PREP/LOC here very convenient everyday have school bus

"It's very convenient to live here (since) there is a school bus everyday."

[47] wo you liang-fen **paper** mintian bixu jiaoshangqu, ke wo xianzai yi-fen hai mei **finish** ne.

I have two-CLASSIF paper tomorrow must turn in but I at the moment one-CLASSIF yet not

finish PARTIC/AFFIRM

"I have two papers [which] I must turn in tomorrow, but at the moment I haven't finished one yet."

(Chinese/English; Wei, 2001b, p. 159)

These examples show that there exist semantic differences in the semantic/pragmatic feature bundles of conceptually-related lexemes. Although Chinese possesses conceptually similar lexemes, in these cases, related lexemes in English are selected as more appropriate. In [45] an English *advisor* assumes more responsibilities than a Chinese advisor. In the academic setting, an English advisor is a professor or instructor who gives advice or counsel to students regarding their academic progress or improvement, course requirements and sequential arrangements, thesis or dissertation writing, and so on. In addition, most English advisors are those who recommend their students to the job market or professional agencies. Though Chinese has the equivalent word *daoshi* (advisor), it does not necessarily mean that a *daoshi* assumes the same responsibilities as an English advisor. In China only a graduate student may have a *daoshi*, whose only or main responsibility is to guide the student in writing his/her thesis or dissertation. In [46] a *school bus* in English means a bus mainly for transporting students to and from a school. In China, the equivalent word *xiaoche* (school bus) usually only transports a school's sports or performance team or equipment. Most public schools even do not have *xiaoche*. In [47] *paper* in English may mean any written piece of work, such as an article, a report, an essay or a composition, but the Chinese equivalent word *zhi* (paper) itself only means a piece of paper to wrap things up in or to write something on.

[48] moshi Nihon ga soo iu **community force** mitaina no ga naku nattara Nihon mo **America** mitai ni nacchau no ja

nai ka?

If Japan PARTIC/NOM so say community force like PARTIC/NOM PARTIC/NOM no become

PERF if Japan also America same PREP/COND become PARTIC/NOM COP/be not

PARTIC/INTERROG

"If Japan had no such thing as a community force, would Japan become America?"

(Japanese/English; Wei, 2002, p. 282)

[49] anata wa **registration** o shimashita ka?

you PARTIC/TOP registration PARTIC/OBJ do-PERF PARTIC/INTERROG

"Have you done your registration?"

(Japanese/English; Wei, 2002, p. 283)

[50] futatsu no **bedroom** ga ate, hitori, Maria to iu ko wa hitori de **one bedroom** o mot-te imasu yo.

two POSS bedroom PARTIC/NOM COP one person and call person PARTIC/TOP one person

PREP/by one bedroom PARTIC/OBJ have-PROG AUX PARTIC/AFFIRM

"We have two bedrooms. One person, called Maria, had one bedroom."

(Japanese/English; Wei, 2006b, p. 170)

In [48] the concept of a *community force* may not only be American, but the general expression "community force" in the American context may include "neighborhood crime watch", "drug free zone", and so on. The Japanese expression similar to "community force" is *chouka* (neighborhood association), but such an association is mainly for organizing local social and cultural activities, overseeing environmental sanitation, taking care of the old, mediating a dispute, and so on. The speaker switches to *community force* probably to mean something beyond Japanese *chouka*. In [49] the speaker switches to *registration* for the possible reason that in Japanese universities/colleges, though students must register for the courses to take they are not free to select the courses which they are interested in taking. The speaker may choose the English word to express his/her intended meaning more accurately. In [50] the speaker switches to *bedroom* for the possible reason that the concept of "bedroom" is relatively new to Japanese. A traditional Japanese room is often used not only for sleeping but also for eating, studying, entertaining guests, or for other daily family activities.

Poulisse and Bongaerts's lexical access model (1994) explains how and why bilinguals differentiate between the lexical items of two or more language systems. The implications of their model for explaining CS are that EL lemmas may receive more activation than the corresponding ML lemmas when the speaker's preverbal message contains the specification [+EL] for some reason or other. In CS, the speaker uses an EL lexical item either because the ML has a similar but nonequivalent lexical item or because the ML has not lexicalized a particular concept at all. For the lemma to receive the most activation, it must meet a sufficient set of the conceptual specifications. The EL lexical items in the above CS instances indicate that their lemmas receive the most activation and are selected as appropriate thereby allowing access to the corresponding EL lexemes.

4.2 Lemma Congruence Checking for Predicate-Argument Structure

One of the characteristics of Chinese/English CS is that Chinese bilinguals tend to use the EL verbs and verb phrases as well as the EL nouns and noun phrases. One of the reasons for this is the fact that Chinese and English share the same basic V-O order. Since Chinese does not possess verb morphology for tense, aspect, voice, or person marking or grammatical devices such as the infinitive marker *to* and the dummy subject pronoun *it*, the bilingual speaker can easily switch the EL verb/verb phrases or nouns/noun phrases into the syntactic slots prepared by the ML. Most Chinese/English CS examples discussed here show that there is sufficient congruence between the EL and the ML lemmas underlying

97

the realization of an EL content morpheme, either a single verb or a verb with its complement noun.

[51] ta gong dao, ta dei xue **drive**.

　　he just arrive he must learn drive

　　"He just arrived, and he must learn how to drive."

[52] ni dei xiang banfa **make money**.

　　you must think way make money

　　"You must think of ways to make money."

[53] wode che you **give me trouble** le.

　　my car again give me trouble PARTIC/PERF

　　"My car has given me trouble again."

　　(Chinese/English; Wei, 2009a, p. 284)

[54] UT-de fanzi zai summer dou yao **demolish** le.

　　UT-POSS building PREP/in summer all will demolish PARTIC/AFFIRM

　　"All the buildings at UT (University Terrace) will be demolished in summer."

[55] **Complain** mei yong a.

　　complain not useful PARTIC/AFFIRM

　　"It's useless to complain."

　　(Chinese/English; Wei, 2009b, p. 332)

In [51] the infinitive verb *drive* is used as the direct object of the main verb *xue* (learn). This is allowed in Chinese, except that Chinese does not possess the infinitive marker *to*. In [52] the verb phrase *make money* fits the Chinese morphosyntactic frame, that is, the V-O order. In [53] the subcategorization frame for the verb *give*, that is, the V-O-O order, is congruent with that for the equivalent verb in Chinese. In [54] the verb *demolish* is congruent with the Chinese predicate-argument structure, except that Chinese does not possess the morphemes (as in English, inflectional morpheme for the verb and auxiliary verb *be*) for realizing the passive construction. In [55] the verb *complain* occupies the subject position in Chinese, which is congruent with the Chinese morphosyntactic pattern, except that Chinese does not possess the infinitive marker *to* to introduce the verb or the dummy pronoun *it* to balance the sentence as in English. These examples indicate that the switched items from the EL sufficiently match the ML lemma entries directing the morphosyntactic procedures to the FORMULATOR producing the fame into which they are switched.

5. Lemma Incongruence and Compromise Strategies

As commonly recognized, languages do not lexicalize concepts in the same way and may differ in grammatical patterns. Consequently, whenever an EL lemma is selected but it does not have a match with that of the ML, some compromise strategies must be taken for possible CS. One of such compromise strategies is the production of EL islands (Jack & Myers-Scotton, 1997; Wei, 2001b, 2002).

An EL island is a constituent in which an EL content morpheme occurs with only other EL morphemes, including EL system morphemes. This compromise strategy can be taken at the level of lexical-conceptual structure or at the level of predicate argument structure. If this compromise strategy is taken, the EL directs the FORMULATOR to activate only the EL morphosyntactic procedures. According to Levelt (1989), different procedures must be applied to the grammatical and phonological encoding of L1 and L2 production for typologically different languages. This also can be true if the language pairs involved in CS do not share the same morphosyntactic procedures. Because the speaker has two speech plans available for bilingual production, he/she may stop the encoding of one of them and continue with the other so as to solve the problem occurring in CS. Different from Levelt's model is that in CS the choice of one EL procedure versus another is determined by the larger ML frame.

5.1 Lemma Incongruence in Lexical-Conceptual Structure

The BLA Model assumes that at the conceptual level bilingual speakers do not produce surface morphemes but rather make appropriate choices about the semantic/pragmatic information that they intend to convey. If the bilingual mode is chosen at the conceptual level, but the lemmas activated from the EL do not sufficiently match the ML counterparts, some compromise strategies must be taken in order for CS to occur. One of the compromise strategies is the production of EL islands. Such a compromise strategy becomes necessary to overcome cross-linguistic differences in lexicalization patterns in bilingual speech involving CS.

[56] ni neng-bu-neng **give me a ride**?

    you can-not-can give me a ride

    "Can you give me a ride?"

[57] name ni mingtian **call me**.

    then you tomorrow call me

    "Then you call me tomorrow."

[58] na wo yidian **come to pick you up**.

    so I one o'clock come to pick you up

    "So, I'll come to pick you up at one o'clock."

    (Chinese/English; Wei, 2001b, p. 162)

In [56] *give me a ride* is incongruent with the ML counterpart *song wo yixia* (literally translated as "send me one time"). While in the EL the lexical-conceptual structure of the means of transportation is conflated in the noun *ride* as the direct object of the verb, in the ML it is conflated in the verb *song* because the verb itself may not contain the means of transportation at all. The speaker chooses the EL expression probably because he wants to be more specific than he can be with the Chinese counterpart. Thus, when the EL lemma is activated, the whole VP is accessed and produced as an EL island. In [57] the semantic features of "communicate with by telephone" are conflated in the verb *call*, but the Chinese equivalent to *call me* is *da dianhua gei wo* (literally translated as "make phone to me"). Since

the speaker chooses the EL lemma which activates the EL lexical-conceptual structure, the whole VP is accessed and produced as an EL island. In [58] *pick you up* occurs as an EL island because when the speaker chooses the EL lemma's lexical-conceptual structure, the whole VP with a pronominal object *you* before the particle satellite *up* is accessed. The speaker prefers *pick up* for the possible reason that this phrasal verb contains the meaning of "to take on as a passenger", but the Chinese equivalent *jie* usually does not. Chinese *jie* means "meet" (e.g., to go to the station to meet somebody), which does not necessarily involve providing personal transportation. It should also be noticed that *come* is accessed together with the infinitive phrase *to pick you up* as an EL island. The possible explanation is that the English infinitive marker *to*, a system morpheme, becomes obligatory if two successive verbs are activated and selected simultaneously. The speaker selects the EL phrasal verb *pick up* for its lexical-conceptual structure to realize this communicative intention more accurately.

It has also been observed that cross-linguistic differences in the conflation of semantic features of a predicate (cf. Talmy, 1985) may have morphosyntactic consequences in CS. The following examples illustrate such differences and consequences.

[59] ni jintian qu-bu-qu **library**?

　　you today go-not-go library

　　"Are you going to the library today?"

[60] jiao ni nu'er **come to** Xiao Ying de **birthday party**.

　　ask your daughter come to Xiao Ying POSS/'s birthday party

　　"Ask your daughter to come to Xiao Ying's birthday party."

　　(Chinese/English; Wei, 2001b, p. 160

In Chinese GOAL is not introduced by a preposition but conflated in verbs like *lai* (come) and *qu* (go). The speaker in [59] only selects the English noun *library* without using the preposition *to*. By contrast, once the speaker in [60] selects the English verb *come*, the preposition *to* indicating GOAL is simultaneously accessed. The consequence is the production of an EL island.

Another pair of examples illustrating how semantic features can be conflated differently with other semantic features of a predicate involves LOCATION.

[61] jiu zai qu feijichang de nei-ge **Chinese store** fujin.

　　just PREP/LOC go airport CP/ATTRIBUT that-CLASSIF Chinese store near

　　"(It's) just near that Chinese store (at the road) which goes to the airport."

[62] wo meitian dei **work in the lab** hao ji-ge xiaoshi.

　　I everyday have to work in the lab quite a few-CLASSIF hour

　　"Everyday I have to work in the lab for quite a few hours."

　　(Chinese/English; Wei, 2001b, p. 161)

In Chinese LOCATION is indicated by a preposition: *zai* (PREP) signifying LOCATION with a particle following the noun in question. In Chinese, such a particle can be equivalent to English *in*, *on*,

under, *near*, and so forth (e.g., *li* (in), *shang* (on), *xia* (under), *fujin* (near)). The speaker in [61] selects the Chinese lexical-conceptual structure in indicating LOCATION. By contrast, the speaker in [62] selects the English prepositional phrase *in the lab* based on the English lexical-conceptual structure where LOCATION is directly conflated in the preposition itself. The consequence is the production of an EL island.

Further examples showing cross-linguistic differences in semantic/pragmatic feature bundles are frequently found in bilingual speakers' choice of certain fixed/idiomatic expressions from the EL. If the same meaning or intention as contained in the EL lemma is not sufficiently congruent with that in the ML lemma, bilingual speakers engaging in CS tend to produce EL islands. Below are some examples of the EL fixed/idiomatic expressions realized in prepositional phrases.

[63] ni qu-bu-qu Kmart? tingshuo you xuduo dongxi **on sale**.

    you go-not-go Kmart hear have many things on sale

    "Are you going to Kmart? (I've) heard there are many things on sale."

[64] wo zuotian qu Kinko's fuyin ji-fen **paper**. Nali you wu-tai jiqi dan san-tai si **out of order**.

    I yesterday go Kinko's Xerox a few-CLASSIF paper there have five-CLASSIF machine   but

    three-CLASSIF COP/be out of order

    "Yesterday I went to Kinko's to zerox a few papers. There were five machines there, but three of

    them were out of order."

[65] wo bu neng baozheng arrive your home **on time** but I surely come.

    I not can guarantee arrive your home one time but I surely come

    "I can't guarantee (that I) will arrive at your home on time, but I'll surely come."

    (Chinese/English; Wei, 2001b, p. 163)

The Chinese equivalents to the EL fixed/idiomatic expressions in these examples are single lexical units without prepositions: *jian mai* literally means "cheap sale" (*on sale* in [63]), *chu guzhang* literally means "something going wrong" (*out of order* in [64]), and *zhunshi* literally means "punctually" (*on time* in [65]). Since such EL lemmas are activated in the speakers" bilingual mental lexicon, the output of the expressions follows the EL lexical-conceptual structure and morphosyntactic procedures.

Thus, one of the major reasons for EL islands to occur is that in the case of nonidiomatic expressions, the speaker's intentions at a certain point during a discourse cannot be realized in the ML because of the insufficient matching between the ML and the EL semantic/pragmatic feature bundles. The other major reason is that in the case of idiomatic expressions, a complete island is selected as a single unit forced by the complexity of the EL item as selected. In either case, the compromise strategy is to produce EL islands. That is, the EL semantic/pragmatic concept in questions is accessed as a single unit observing the EL morphosyntactic procedures.

5.2 Lemma Incongruence in Predicate-Argument Structure

As introduced earlier, in CS it is the ML which controls the morphosyntactic structure of the sentence

containing switched items. This is because the ML supplies system morphemes, subcategorization frames for verbs, and morpheme order. Although morphosyntactic procedures are realized by the FORMULATOR at the functional level, before morphosyntactic directions are sent to the FORMULATOR, lemmas from both languages can be activated at a certain point during a discourse. Thus, lemma congruence checking at the level of lexical-conceptual structure alone is not sufficient enough for CS to occur. Lemma congruence checking at the level of predicate-argument structure must also come into play. This is because in some cases, predicate-argument structures across the two languages may differ. If such incongruence occurs, but the semantic/pragmatic feature bundles desired by the speaker activate the EL lemma for his/her communicative intentions, a radical compromise strategy must be taken in order for the EL material to be accessed. In other words, even if the lexical-conceptual structures between the two languages are sufficiently congruent, the ML predicate-argument structure will reject the mapping if a particular EL predicate-argument structure does not match that of the ML.

[66] tingshuo nei-ge professor hen **crazy**. ta jingchang **fails students in exams**.

hear that-CLASSIF professor very crazy she often fails students in exams

"(I) heard that professor is very crazy. She often fails students in exams."

[67] ni biye hou keyi **teach English to nonnative speakers**.

you graduate after can teach English to nonnative speakers

"After you graduate, you can teach English to nonnative speakers."

(Chinese/English; Wei, 2001b, p. 168)

[68] wo meitian dei **help her with her homework**.

I everyday have to help her with her homework

"Everyday I have to help her with her homework."

(Chinese/English; Wei, 2005, p. 2346)

[69] wo keyi **wait for you** dao liang dian.

I can wait for you PREP/till two o'clock

"I can wait for you till two o'clock."

(Chinese/English; Wei, 2001b, p. 166)

In [66] the verb phrase headed by *fail* is an EL island (i.e., with all the morphemes, including the system morphemes, from the EL). In English *fail* can be used as a causative verb and thus takes the grammatical subject as the AGENT who makes the failure happen, but the Chinese equivalent verb *shibai* means "be defeated in …" and is used only as a noncausative verb with the grammatical subject as the EXPERIENCER. The speaker prefers the EL concept, but there is incongruence between the EL and ML in predicate-argument structure. The consequence is the production of an EL island. It is possible for the speaker of a particular Chinese variety to say "tingshuo nei-ge professor hen crazy. ta jingchang <u>dang</u> students". In this case, the Chinese verb *dang* is equivalent to *fail* in terms of the

predicate-argument structure. Like *fail*, *dang* is used as a causative verb with the grammatical subject as the AGENT. If the speaker chooses *dang* rather than *fail*, of course, no EL island will be produced. In other words, in Chinese *shibai* and *dang* are two separate lexical entries. In [67] the RECIPIENT is introduced in the prepositional phrase headed by *to*, the English indirect object dative construction. By contrast, the equivalent Chinese verb phrase headed by *jiao* (teach) only permits the double object construction (e.g., *jiao ta English* (teach him English)). Again, since the speaker prefers the EL material, but the ML rejects the mapping which the EL prepositional phrase would project at the level of predicate-argument structure, the consequence is the production of an EL island. In [68] the THEME is introduced in the prepositional phrase headed by *with*. By contrast, in Chinese the THEME is always introduced by a specific verb such as *zhuo* (do). The speaker selects the EL verb *help* at the level of lexical-conceptual structure, but the EL and the ML are incongruent at the level of predicate-argument structure. Consequently, the whole VP in the EL is accessed and produced as an EL island. In [69] *wait for* is accessed as a single verbal unit. In English the direct object, in this case, the THEME, is introduced by the preposition in a phrasal verb like *wait for*, or introduced by the verb with a satellite as in *pick up* ([51]). Chinese has no such equivalent phrase structures. In Chinese these same meanings are expressed by singe verbs: *deng* for *wait for* and *jie* for *pick up*. It has been noticed that once such EL phrasal verbs are accessed, they may co-occur with further EL material, resulting in the maximal projection of an EL phrasal category, that is, an EL island.

6. Conclusion

The CS instances involving various language pairs discussed in this paper provide some empirical evidence for certain specifications about the nature and activity of the bilingual mental lexicon in relation to the bilingual speech production process. Different from the studies of CS at surface descriptive levels, the BLA Model describes and explains CS in terms of four abstract levels of bilingual speech production process (i.e., the conceptual level, the lemma level, the functional level, and the positional level), with a focus on bilingual lemma activation as a crucial interface between speaker intention at the conceptual level and language encoding at the functional level. This model claims that lemmas in the bilingual mental lexicon are language-specific, language-specific lemmas are in contact during a discourse involving CS, and such a contact occurs at three distinct but related levels of abstract lexical structure: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. Thus, CS is regarded as bilingual lemmas in contact at any level of abstract lexical structure. The major findings indicate that CS can be better accounted for at a deep or abstract level. The current study reaches several conclusions regarding the bilingual speech production process.

1). Bilingual speakers' languages are turned "on' during a discourse involving CS, but they are never equally activated at the same time. The ML is more strongly activated than the EL. It is the ML which

provides the sentential frame, controls morphosyntactic procedures, and provides all system morphemes as well as content morphemes at a much higher frequency. It is the speaker who chooses whichever language as the ML.

2). The bilingual mental lexicon contains lemmas from the languages known, these lemmas are tagged for their specific language, and language-specific lemmas are in contact in CS. In CS, the speaker may activate the language-specific lemmas as desired from his/her bilingual mental lexicon. However, lemmas are never equally activated in bilingual speech. The EL only supplies content morphemes as desired by the speaker to be switched into the ML sentential frame. This is because only conceptually activated EL lemmas in the bilingual mental lexicon can appear in ML + EL constituents to express the speaker's intended meaning.

3). Some switches are motivated by the lack of a particular word in the ML for the speaker's intentions. The naturally occurring CS instances for the current study show that bilinguals may use compensatory strategies to solve lexical problems caused by the lexical gap between the languages involved. When the speaker's intentions at the lexical-conceptual level call for an EL content morpheme, this selection activates the EL lemma supporting that morpheme. How the activated EL morpheme may appear in the ML morphosyntactic structure depends on the extent to which there is congruence between its lemma and an ML counterpart in the bilingual mental lexicon.

4). Bilingual speakers can activate lemmas from whichever language as the EL during a discourse involving CS, but the activated EL lemmas must be sufficiently congruent with the counterparts of the ML at the three levels of abstract lexical structure or some combination of these levels. If lemma incongruence of insufficient congruence occurs between the language pairs at any of these levels, but the speaker does not want to give up CS, radical compromise strategies, such as production of EL islands, must be taken in order for possible CS to occur.

The main purpose of this paper is to apply the BLA Model as well as the MLF Model to the explanations of the CS phenomenon. As assumed, different aspects of abstract lexical structure and bilingual lemmas in contact affect CS, and CS, like any natural language speech behaviors, is a rule governed bilingual behavior.

References

Annamalai, E. (1989). The Language Factor in Code Mixing. *International Journal of the Sociology of Language*, *75*, 47-54. https://doi.org/10.1515/ijsl.1989.75.47

Appel, R., & Muysken, P. (1987). *Language Contact and Bilingualism*. London: Edward Arnold.

Azuma, S. (1991). *Two Level Processing Hypothesis in Speech Production: Evidence from intrasentential codeswitching*. Paper presented at the 27th Chicago Linguistic Society.

Backus, A. (1996). *Two in One: Bilingual Speech of Turkish Immigrants in the Netherlands*. Tilburg, The Netherlands: Tilburg University Press.

Bentahila, A., & Davies, E. E. (1983). The Syntax of Arabic-French Code-Switching. *Lingua*, *59*, 301-330. https://doi.org/10.1016/0024-3841(83)90007-4

Blazquez-Domingo, R. (2000). *Spanish-English CS data. Unpublished corpus*. University of South Carolina, Columbia, SC.

Chan, B. H-S. (1998). Functional Heads, Cantonese Phrase Structure and Cantonese-English Code-Switching. *UCL Working Papers in Linguistics*, *10*, 253-284.

Choi, J. O. (1991). Korean-English Code Switching: Switch-alpha and linguistic constraints. *Linguistics*, *29*(5), 877-902. https://doi.org/10.1515/ling.1991.29.5.877

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

de Bot, K., & Schreuder, R. (1993). Word Production and the Bilingual Lexicon. In R. Schreuder, & B. Weltens (Eds.), *The Bilingual Lexicon* (pp. 191-214). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.6.10bot

DiSciullo, A., Muysken, P., & Singh, R. (1986). Government and Code-Switching. *Journal of Linguistics*, *22*, 1-24. https://doi.org/10.1017/S0022226700010537

Faerch, C., & Kasper, G. (1984). Two Ways of Defining Communication Strategies. *Language Learning*, *34*, 45-63. https://doi.org/10.1111/j.1467-1770.1984.tb00995.x

Gardner-Chloros, P. (1987). Code-Switching in Relation to Language Contact and Convergence. In G. Lüdi (Ed.), *Devenir Bilingue-parler Bilingue* (pp. 99-111). Tübingen: Niemeyer. https://doi.org/10.1515/9783111594149.99

Green, D. W. (1986). Control, Activation and Resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, *27*, 210-223. https://doi.org/10.1016/0093-934X(86)90016-7

Green, D. W. (1998). Mental Control of the Bilingual Lexico-Semantic System. *Bilingualism: Language and Cognition, 1*, 67-81. https://doi.org/10.1017/S1366728998000133

Grosjean, F. (1982). *Life with Two Languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.

Grosjean, F. (1989). Neurolinguists, Beware! The Bilingual is Not Two Monolinguals in One Person. *Developmental Psychology*, *25*, 577-588. https://doi.org/10.1016/0093-934X(89)90048-5

Grosjean, F. (1997). Processing Mixed Languages: Issues, Findings, and Models. In A. M. B. de Groot, & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic perspectives* (pp. 225-254).

Halmari, H. (1997). *Government and Codeswitching: Explaining American finnish*. Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.12

Haust, D., & Dittmar, N. (1998). Taxonomic or Functional Models in the Description of Codeswitching? Evidence from Mandinka and Wolof in African contact situations. In R. Jacobson (Ed.), *Codeswitching Worldwide* (pp. 79-90). Berlin and New York: Mouton de Gruyter. https://doi.org/10.1515/9783110812190.79

105

Jake, J. L., & Myers-Scotton, C. (1997). Codeswitching and Compromise Strategies: Implications for lexical structure. *International Journal of Bilingualism*, *1*, 25-39. https://doi.org/10.1177/136700699700100103

Jake, J. L., Myers-Scotton, C., & Gross, S. (2002). Making a Minimalist Approach to Codeswitching Work: Adding the matrix language. *Bilingualism: Language and Cognition*, *5*(1), 69-91. https://doi.org/10.1017/S1366728902000147

Jake, J. L., Myers-Scotton, C., & Gross, S. (2005). A Response to MacSwan (2005): Keeping the matrix language. *Bilingualism: Language and cognition*, *8*(3), 271-276. https://doi.org/10.1017/S1366728905002300

Joshi, A. K. (1985). Processing of Sentences with Intrasentential Code Switching. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural Language Parsing* (pp. 190-205). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511597855.006

Kempen, G., & Hoenkamp, E. (1987). An Incremental Procedural Grammar for Sentence Formulation. *Cognitive Science*, *11*, 201-258. https://doi.org/10.1207/s15516709cog1102_5

Kempen, G., & Huijbers, P. (1983). The Lexicalization Process in Sentence Production and Naming: Indirect election of words. *Cognition*, *14*, 185-209. https://doi.org/10.1016/0010-0277(83)90029-X

Klavans, J. L. (1983). The Syntax of Code-Switching: Spanish and English. In L. D. King, & C. A. Maley (Eds.), *Selected Papers from the 13th Linguistic Symposium on Romance Languages* (pp. 213-231). Amsterdam: John Benjamins.

Kovács, M. (2001). *Code-Switching and Language Shift in Australian Finnish in Comparison with Australian Hungarian*. Abo: Abo Akademi University Press.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1995). The Ability to Speak: From intention to spoken words. *European Review*, *3*, 13-23. https://doi.org/10.1017/S1062798700001290

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences*, *22*, 1-75. https://doi.org/10.1017/S0140525X99001776

MacSwan, J. (2000). The Architecture of the Bilingual Language Faculty: Evidence from intrasentential code switching. *Bilingualism: Language and Cognition*, *3*, 37-54. https://doi.org/10.1017/S1366728900000122

MacSwan, J. (2009). Generative Approaches to Codeswitching. In B. E. Bullock, & A. J. Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-switching* (pp. 309-335). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511576331.019

Muysken, P. (1990). Concepts, Methodology, and Data in Language Contact Research: Ten remarks from the perspective of grammatical theory. In *Papers for the Workshop on Concepts, Methodology and Data* (pp. 15-30). Strasbourg: European Science Foundation.

Muysken, P. (1991). Needed: A comparative approach. In *Papers for the Symposium on Code-Switching in Bilingual Studies: Theory, significance and perspectives* (pp. 253-272). Strasbourg: European Science Foundation.

Myers-Scotton, C. (1991). Whither Code-Switching? Prospects for Cross-Field Collaboration: Production-based models of code-switching. In *Papers for the Symposium on Code-Switching in Bilingual Studies: Theory, significance and perspectives* (pp. 207-232). Strasbourg: European Science Foundation.

Myers-Scotton, C. (1993 [1997]). *Duelling Languages: Grammatical structure in codeswitching*. Oxford: Clarendon Press.

Myers-Scotton, C. (1994). A Lexically-Based Production Model of Codeswitching. In L. Milroy, & P. Muysken (Eds.), *One Speaker, Two Languages: Cross-disciplinary Perspectives on Code-switching* (pp. 233-256). Cambridge, UK and New York: Cambridge University Press.

Myers-Scotton, C. (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299530.001.0001

Myers-Scotton, C. (2004). Precision Tuning of the Matrix Language Frame (MLF) Model of Codeswitching. *Sociolinguistica*, *18*, 106-117. https://doi.org/10.1515/9783484604650.106

Myers-Scotton, C., & Jake, J. L. (1995). Matching Lemmas in a Bilingual Language Competence and Production Model: Evidence from intrasentential code switching. *Linguistics*, *33*(5), 981-1024. https://doi.org/10.1515/ling.1995.33.5.981

Myers-Scotton, C., & Jake, J. L. (2000a). Four Types of Morpheme: Evidence from aphasia, codeswitching, and second language acquisition. *Linguistics*, *38*(6), 1053-1100. https://doi.org/10.1515/ling.2000.021

Myers-Scotton, C., & Jake, J. L. (2000b). Testing a Model of Morpheme Classification with Language Contact Data. *International Journal of Bilingualism*, *4*(1), 1-8. https://doi.org/10.1177/13670069000040010101

Myers-Scotton, C., & Jake, J. L. (2001). Explaining Aspects of Codeswitching and Their Implications. In J. Nicol (Ed.), *One Mind, Two Languages: Bilingual language processing* (pp. 84-116). Oxford: Blackwell.

Myers-Scotton, C., & Jake, J. L. (2017). Revising the 4-M Model: Codeswitching and morpheme election at the abstract level. *International Journal of Bilingualism*, *21*(3), 340-366. https://doi.org/10.1177/1367006915626588

Nartey, J. (1982). Code-Switching, Interference or Faddism? Language Use among Educated Ghanaians. *Anthropological Linguistics*, *24*(2), 183-192.

Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.

Pandit, I. (1990). Grammaticality in Code-Switching. In R. Jacobson (Ed.), *Codeswitching as a Worldwide Phenomenon* (pp. 33-69). New York: Peter Lang.

Petersen, J. (1988). Word-Internal Code-Switching Constraints in a Child's Grammar. *Linguistics*, *26*, 479-493. https://doi.org/10.1515/ling.1988.26.3.479

Pfaff, C. (1979). Constraints on Language Mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language*, *55*, 291-318. https://doi.org/10.2307/412586

Poplack, S. (1980). Sometimes I'll Start a Sentence in Spanish Termino en Español: Toward a typology of code-switching. *Linguistics*, *18*, 581-618. https://doi.org/10.1515/ling.1980.18.7-8.581

Poplack, S. (1981). Syntactic Structure and Social Function in Code-Switching. In R. P. Duran (Ed.), *Latino Language and Communicative Behavior* (pp. 169-184). Norwood, NJ: Ablex.

Poplack, S., & Meechan, M. (1995). Patterns of Language Mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In L. Milroy, & P. Muysken (Eds.), *One Speaker, Two Languages: Cross-disciplinary Perspectives on Code-switching* (pp. 199-232). Cambridge, UK and New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511620867.010

Poplack, S., & Sankoff, D. (1988). Code-Switching. In U. Ammon, N. Dittmar, & K. J. Mattheir (Eds.), *Sociolinguistics, an International Handbook of the Science of Language and Society* (pp. 1174-1180). Berlin: Mouton de Gruyter.

Poulisse, N., & Bongaerts, T. (1994). First Language Use in Second Language Acquisition. *Applied Linguistics 15*, 36-57. https://doi.org/10.1093/applin/15.1.36

Preziosa-Di Q. I. (1992). *Teoreticamentela Firma la Indietro: Frammistione di Italiano e Schwyzertüsch Nella Conversazione one di Figli di Emigrati*. Universitá di Zurigo, Facoltá de Lettere e di Filosofia: Vavoro di Licenza in Linguistica Italiana.

Richards, J. C. (1976). The Role of Vocabulary Teaching. *TESOL Quarterly*, *10*, 77-89. https://doi.org/10.2307/3585941

Ringbom, H. (1987). *The Role of the First Language in Foreign Language Learning*. Clevedon: Multilingual Matters.

Roelofs, A. (1992). A Spreading-Activation Theory of Lemma Retrieval in Speaking. *Cognition*, *42*, 107-142. https://doi.org/10.1016/0010-0277(92)90041-F

Romaine, S. (1995). *Bilingualism* (2nd ed.). Oxford: Blackwell.

Sankoff, D., Poplack, S., & Vanniarajan, S. (1990). The Case of the Nonce Loan in Tamil. *Language Variation and Change*, 2.1, 71-101. https://doi.org/10.1017/S0954394500000272

Schmitt, E. (2006). *Russian-English Corpus. Unpublished*.

Scotton, C. M. (1983). The Negotiation of Identities in Conversation: A theory of markedness and code choice. *International Journal of the Sociology of Language*, *44*, 115-136.

Slabbert, S., & Myers-Scotton, C. (1997). The Structure of Tsotsitaal and Iscamtho: Code switching and in-group identity in South African townships. *Linguistics*, *35*, 317-342. https://doi.org/10.1515/ling.1997.35.2.317

Talmy, L. (1985). Lexicalization Patterns: Semantic structure in lexical forms. In T. Shopen (Ed.),

Language Typology and Syntactic Description, Vol. 3, pp. 57-149. Cambridge: Cambridge University Press.

Wei, L. (1992). *Chinese/English Codeswitching Corpus*. Unpublished.

Wei, L. (2000). Types of Morphemes and Their Implications for Second Language Morpheme Acquisition. *International Journal of Bilingualism*, *4*(1), 29-43. https://doi.org/10.1177/13670069000040010301

Wei, L. (2001a). The Multilingual Mental Lexicon: Language separation/activation in trilinguals. In *Papers Selected from the Second International Conference on Third Language Acquisition and Trilingualism: Interactive CD-ROM* (pp. 13-5).

Wei, L. (2001b). Lemma Congruence Checking between Languages as an Organizing Principle in Intrasentential Codeswitching. *International Journal of Bilingualism*, *5*, 153-173. https://doi.org/10.1177/13670069010050020201

Wei, L. (2002). The Bilingual Mental Lexicon and Speech Production Process. *Brain and Language*, *81*, 691-707.

Wei, L. (2005). Verbs in Chinese/English codeswitching. In *The Proceedings of the 4th International Symposium on Bilingualism CD-ROM*. Cascadilla Press.

Wei, L. (2006a). The Multilingual Mental Lexicon and Lemma Transfer in Third Language Learning. *International Journal of Multilingualism*, *3*(2), 88-104. https://doi.org/10.1080/14790710608668390

Wei, L. (2006b). Intrasentential Codeswitching as Conceptual Projection of Lemmas in the Bilingual Mental Lexicon. *Journal of Cognitive Science*, *7*(2), 149-178.

Wei, L. (2009a). Code-Switching and the Bilingual Mental Lexicon. In B. E. Bullock, & A. J. Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-Switching* (pp. 270-288). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511576331.017

Wei, L. (2009b). Intrasentential Codeswitching: Bilingual lemmas in contact. *Concentric: Studies in Linguistics*, *35*(2), 307-344.

Wei, L. (2009c). The Composite Nature of Interlanguage as a Developing System. *Research in Language*, *7*, 5-30. https://doi.org/10.2478/v10015-009-0002-9

Wei, L. (2015). *Interlanguage: The abstract level in language acquisition*. New York: The Edwin Mellen Press.

Woodford, E. (1983). Bilingual Code-Switching and Syntactic Theory. *Linguistic Inquiry*, *14*, 520-536.