Original Paper

Codeswitching as Projection of Bilingual Lemmas in Contact

Longxing Wei[1][*]

[1] Department of Linguistics, Montclair State University, New Jersey, USA

[*] Longxing Wei, Department of Linguistics, Montclair State University, New Jersey, USA

Abstract

Unlike most previous studies of Codeswitching (CS) focused on describing surface configurations of switched items (i.e., where CS is structurally possible) or the switched items (i.e., what items from another language can be switched), this paper explores formulation processes of bilingual speech and the nature of the bilingual mental lexicon and its activity in CS. More specifically, it applies the Bilingual Lemma Activation Model (Wei, 2002, 2006b) to the data drawn from various naturally occurring CS instances. It claims that the mental lexicon does not simply contain lexemes and their meanings, but also lemmas, which are abstract entries in the mental lexicon that support the surface realization of actual lexemes. Lemmas are abstract in that they contain phonological, morphological, semantic, syntactic and pragmatic information about lexemes. It further claims that lemmas in the bilingual mental lexicon are language-specific and are in contact during a discourse involving CS at three levels of abstract lexical structure: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. The CS instances described and analyzed in this paper provide evidence that the bilingual speaker's two linguistic systems are unequally activated in CS, and CS is an outcome of bilingual lemmas in contact.

Keywords

codeswitching, activation, lemma, bilingual mental lexicon, content morpheme, system morpheme, congruence, island

1. Introduction

Most previous studies of intrasentential codeswitching (for short, CS) (e.g., Lipski, 1977; Pfaff, 1979; Poplack, 1980; Sridhar & Sridhar, 1980; Gumperz, 1982; Woodford, 1983; Azuma, 1993; MacSwan, 2000) focused on describing surface configurations of switched items, that is, where CS is structurally possible, or switched items themselves, that is, what items from another language can be switched. Confronting and expanding on Levelt's model of monolingual speech production (1989) and

Myers-Scotton's Matrix Language Frame (MLF) Model of bilingual speech involving CS (1993 [1997]), from some psycholinguistic perspectives, Wei (2002, 2006b) proposes the Bilingual Lemma Activation (BLA) Model to describe and explain CS at an abstract level in terms of the nature and activity of the bilingual mental lexicon during bilingual speech involving CS.

The BLA Model relies heavily on Levelt's notion of the mental lexicon and the abstract elements called "lemmas" underlying lexemes and applies Levelt's monolingual speech production model to the bilingual processes with a focus on the notion of bilingual lemma activation during CS. The BLA Model also employs Myers-Scotton's MLF Model to formalize certain "universal" structural constraints on CS in terms of the unequal activation of the languages involved and to explain why not all morphemes can be switched in terms of the unequal activation of morpheme types in the bilingual mental lexicon. The BLA Model is a psycholinguistic approach to the study of bilingual speech behavior and views CS as an outcome of bilingual lemmas in contact.

2. The Bilingual Mental Lexicon and Abstract Levels of Speech Production

"The mental lexicon" is generally defined as the store of information about particular words in one's language. From some psycholinguistic perspectives, "the mental lexicon" is defined under the assumption that something abstract is contained in any lexical item. Levelt defines that something abstract as a "lemma", which is the "nonphonological part of an item's lexical information" and claims that "it is the lemmas of the mental lexicon that conceptual information is linked to grammatical function" (1989, p. 162). According to Kempen and Huijbers (1983), Kempen and Hoenkamp (1987), Levelt (1989), Levelt, Roelofs and Meyer (1999), Myers-Scotton and Jake (1995), and Wei (2001a, 2001b, 2002), when speakers construct an utterance, they build a sentential frame without much regard for the phonological aspects of words 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"). 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 word's grammatical configuration or sentential frame). Thus, the mental lexicon does not simply contain lexemes and their meanings but more abstract elements called "lemmas". In other words, lemmas are 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 declarative knowledge about the word's meaning, and information about its syntax and morphology which is necessary for constructing the word's syntactic environment. For example, the lemmas for *she* require the word to be used of a female 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); the lemmas for *know* require a subject that expresses the theta role of EXPERIENCER, and an object that expresses the

41

theta role of PERCEPT (i.e., what is known), and these elements appear in a particular order. As evidenced in Richards (1976), Færch and Kasper (1984), Nation (1990), Ringbom (1987), Pienemann (1999) and others, lemmas also contain information about the word's spelling, pronunciation, its register, the kind of discourse it typically enters into, and it is pragmatic function. Thus, the mental lexicon can be defined as the speaker's internal representation of language specific knowledge about the surface forms, that is, lemmas.

Sharing the view of Kaplan and Bresnan (1982), Jackendoff (1983), Kempen and Hoenkamp (1987), Levelt posits that the lexical information stored with the entry for a particular verb has to communicate with the other constituents in the sentence as driven by the verb itself and explicates the lemma for *give*: conceptual specification: CAUSE (X, (GOposs (Y, (FROM/TO (X, Z))))), conceptual arguments: (X, Y, Z), syntactic category: V, grammatical functions: (SUBJ, DO, IO) (1989, p. 191). As explicated, *give* requires three conceptual arguments, each of which is assigned a specific thematic role, and they are mapped onto the grammatical functions:

X (Agent), Y (Theme), Z (Goal)

| | |

SUBJ   DO   IO (Pienemann, 1999, p. 63)

Thus, the "mental lexicon" differs from the "lexicon" in that it is not just an individual speaker's collection of words but deals with how those words are stored, activated, processed, and retrieved by each speaker. The activation of lemmas in the mental lexicon plays a central role in speech production. That is, lemma activation of [articular lexical items in the mental lexicon mediate between conceptualization and speech formulation as a necessary level of speech production.

One of the most crucial assumptions underlying this study is that lemmas are language-specific for lexicalization patterns of a particular language. For the notion of language-specific lemmas, one of the frequently cited examples is provided by Talmy (1985, p. 69): (English) *The bottle floated into the case* vs. (Spanish) *La botella entró a la Cueva flotando* (The bottle moved-into the cave floating). While in English, motion with manner or condition of movement can be conflated into a single lemma (i.e., FLOAT), in Spanish, a different pattern is required to express the notion of FLOATING periphrastically using the gerund.

In Levelt's model of speech production (1989), semantic and syntactic information constitute the lemma of the lexical information while morphological and phonological information constitute the form of the lexical item. Levelt's model is 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. Myers-Scotton and Jake (1995) claim that lemmas are activated by language-specific semantic/pragmatic feature bundles that come from the CONCEPTUALIZER. Following the above lines of thinking, this study proposes that it is the preverbal message or the speaker's communicative intention that activates language-specific lemmas in the bilingual mental lexicon. In other words, it is

the semantic/pragmatic feature bundles selected by the CONCEPTUALIZERE at the conceptual level that trigger the appropriate lemmas into activity before the FORMULATOR has access to the relevant lexical items in the mental lexicon. Adapted from Levelt (1989), Myers-Scotton and Jake (2000a, 2000b) and Wei (2002), below is the simplified model of lemma activation in the bilingual mental lexicon.

(>: "before")

Conceptual Level: activation of speaker's preverbal message/intention → semantic/pragmatic feature bundles selected by the CONCEPTUALIZER > Lemma Level: activation of language-specific lemmas > Functional Level: activation of the FORMULATOR → projection of language-specific procedures > Position Level: projection of surface forms → morphological/phonological realization patterns.

One of the key questions being asked about the nature of the bilingual mental lexicon is whether the bilingual's two lexicons are linked to a shared conceptual store or separated in two conceptual stores (Keatley, 1992). Some recent bilingual processing and representation models assume that in the bilingual mental lexicon while phonological and morphosyntactic forms can be rather easily differentiated, lexical meanings or concepts are largely shared (Kroll and Stewart, 1994; Kroll and Sunderman, 2003; Costa, 2005). Accordingly, most studies of conceptual representation have focused on the links between word forms and meanings and factors that affect the bilingual conceptual processing, but not the nature of conceptual representation itself (Kroll & de Groot, 1997; de Groot, 2002; Kroll & Tokowicz, 2005). Other studies have focused on cross-linguistic differences in lexical/conceptual representation (Pavlenko, 2009; Jarvis, 2009; Altarriba & Basnight-Brown, 2009; Jiang, 1999, 2000).

Influenced by the above linguistic and psycholinguistic models of the bilingual lexical/conceptual representation and based on the BLA Model, this paper claims that lemmas in the bilingual mental lexicon are language-specific. This is because while the monolingual "mental lexicon represents a complex self-organizing system," the "bilingual mental lexicon, as opposed to the monolingual one, integrates the units of two linguistic systems and, therefore, ensures the processes of speech perception and production in two languages" (Leshchenko, Dotsenko, & Ostapenko, 2015, p. 1040). It further claims that language-specific lemmas in contact during a discourse involving CS. As assumed, CS is described and explained in terms of bilingual lemma activation at any level of abstract lexical structure. Based on some naturally occurring and commonly observed CS instances, this paper presents three arguments: (1) Lemmas are unequally activated in either monolingual or bilingual speech production. It is the unequal activation of lemmas from the bilingual mental lexicon which motivates and constraints the speaker's choice of morphemes in CS (Wei, 2002). (2) It is the structural procedures dictated by the Matrix Language (ML) (similar to the notion of "base" or "host" language) which set the sentential frame for sentences containing switched items the Embedded Language (EL) (similar to the notion of "guest" language) (Myers-Scotton, 1993). (3) Bilinguals can activate whichever language known to

43

them as the EL, but the activated lemmas from the EL must be sufficiently congruent with the ML counterparts at each of the three levels of abstract lexical structure: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. Thus, the fundamental assumption is that CS itself is a linguistic system and is governed and constrained by a set of morphosyntactic principles and rules governing any other linguistic system, and CS is an outcome of bilingual lemmas in contact.

3. Codeswitching as Bilingual Lemma Activation

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 (2006b) proposes the BLA Model to describe and explain CS in terms of the nature and activity of the bilingual mental lexicon. This model also draws on some contemporary proposals and assumptions about the nature and activity of bilingual language modes in speech production. For example, Færch and Kasper (1986) and Grosjean and Soares (1986) assume that the bilingual's language systems can be kept separate because they may be activated at different levels, depending on whether they are used at the moment. Green (1986) and Paradis (1981, 1997) propose that the language system of a bilingual is organized in two subsets, one for each language, that can be selectively activated, activated simultaneously to various degrees, or deactivated independently for one another. Thus, as assumed, when the bilingual is in a monolingual mode, the other language must be deactivated or inhibited. Green (1986) postulates that bilingual speakers who wish to select a particular language for the current communication must ensure that its activation exceeds that of the competing language(s) in their procession. According to Soares and Grosjean (1984), Green (1986) and Grosjean (1997), the deactivation of bilingual speakers' other language(s) is rarely total. This is clearly evidenced in the interference effects in bilinguals' production of the speaker-specific deviations from the target language being spoken due to the influence of the deactivated language. Paradis (1997) assumes that when bilingual speakers select one language rather than the other, the activation threshold of the nonselected language must be raised; however, the language not being selected for use at the moment is never totally deactivated. Paradis (1989) and Berg and Schade (1992) even claim that deactivation of one language may be so difficult that abundance involuntary mixing or hybridization may occur. According to Paradis (1997), deactivation or inhibition of one of a bilingual's languages is a matter of degree. However, "this deactivation has led to much theorizing and much controversy around the notion of a language switch or a monitor system," but "what is certain, however, is that bilinguals rarely deactivate the other language totally. This is seen in various types of production interference–the involuntary influence of one language on the other…" (Grosjean & Spares, 1986, p. 146). According to Grosjean (1997), if the bilingual is in the monolingual mode the guest language becomes deactivated as least as possible and as a consequence, the deactivated guest language does not act upon the base language

often; if the bilingual is in the bilingual mode, the guest language becomes activated to a relatively high level but less so than the base language, and thus the activated guest language intrudes upon the base language very often. Bilinguals find themselves at various points along the language mode continuum, which corresponds to different levels of activation of the two languages, and the base language is always more strongly activated than the guest language (Grosjean, 1985, 1989, 1994, 1997). Following these lines of thinking, Wei (2000a, 2000b, 2002, 2006b) assumes that when bilinguals are in the bilingual mode to produce CS, both of their languages are activated, but their ML is more strongly activated than their EL. According to Wei (2020), it is the language mode chosen and the relative degree of activation of the ML and the EL that cause the amount of CS and the morphosyntactic principles governing CS.

Sotillo's investigation of morphosyntactic features and sociolinguistic functions of CS in face-to-face and short message service conversations among Spanish-English and Tagalog-English bilinguals finds that the bilingual lemma activation "facilitates the interpretation of bilingual speech behavior and switches between languages as cognitively based strategies at an abstract level" (2016, p. 21). This finding supports Wei's proposal (2006b) that CS can be better explained at an abstract level in the sense that lemma activation of particular lexical items in the bilingual mental lexicon must mediate between the conceptual level, where specific semantic-pragmatic feature bundles are selected as desired for communication, and the functional level, where language-specific morphosyntactic procedures are projected before surface forms are realized at the positional level. The fundamental assumption underlying the BLA Model is that bilingual lemmas are never equally activated in CS, and the unequal activation of bilingual lemmas is the driving force for CS.

4. Codeswitching as Unequal Activation of Bilingual Lemmas

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

Regarding the bilingual mental lexicon, Green (1986, 1993) and de Bot and Schreuder (1993) propose that the lexical items belonging to different languages must be organized in subsets which can be either fully activated or deactivated. Myers-Scotton (1993) proposes that there is a universal set of semantic

and pragmatic feature bundles available for the lexical-conceptual structure of lemmas. Wei (2002, 2006b) proposes that lemmas in the bilingual mental lexicon are language-specific and lexicalization patterns across languages reflect that semantic and pragmatic feature bundles across related lemmas in different languages are presented in different configurations. Adopting these proposals, the BLA Model claims that it is cross-linguistic differences in how information is organized at the level of lexical-conceptual structure and at the level of predicate-argument structure that affect code choices and structures which are predicted to occur in CS. Such a claim implicates that it is the unequal activation of bilingual lemmas that drives CS and constrains its lexical and structural configurations.

4.1 Morphemes Unequally Activated in Codeswitching

One of the major assumptions underlying the BLA Model is that lemmas, in addition to other abstract entries about particular lexemes, contain semantic and pragmatic feature bundles which encode the lexical-conceptual structure representing the speaker's communicative intention as preverbal message generated by the CONCEPTUALIZER (Levelt, 1989; Poulisse & Bongaerts, 1994). This is because at the level of lexical-conceptual structure the speaker seeks appropriate linguistic material for his/her communicative intention, and it is at this level the speaker conceptualizes and activates the appropriate concepts. The activated concepts will then activate the corresponding lemmas in the mental lexicon. According to Roelofs (1992), Levelt (1995), and Wei (2002), it is sufficiently activated lemmas that activates the associated lexeme. According to Myers-Scotton (1993), there is a universal set of semantic and pragmatic features available for the lexical-conceptual structuring of lemmas. Wei (2001b) claims that the presence and conflation of universally available semantic and pragmatic features may vary cross-linguistically. The BLA Model proposes that the relative importance of cross-linguistic lexical-conceptual differences in lemmas in the bilingual mental lexicon and the consequences of CS should be considered as evidence of variation in semantic-pragmatic feature bundles.

As introduced above, lemmas in the bilingual mental lexicon are language-specific. In other words, bilingual lemmas are tagged for specific languages and support the realization of actual lexemes. The naturally occurring CS instances to be discussed indicate that it is content morphemes, rather than system morphemes, which encode the speaker's intended meaning. Wei (2001b, 2002) posits that the major reason for content and system morphemes to be accessed differently lies in the fact that only content morphemes contain semantic/pragmatic feature bundles, but system morphemes only play their abstract grammatical notions. At a certain point during a discourse, the speaker may switch to certain EL content morphemes to convey his/her intended or desired meanings as generated in his/her preverbal message. It is in this sense that certain language-specific lemmas are conceptually projected from the bilingual mental lexicon in a discourse involving CS.

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

[1] 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. 899)

[2] Kerran s äolit pannu si-t ämun **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)

[3] Zachem ty na **grass**-e valja jesih'sja.

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

"Why are you rolling around on the grass?"

(Russian/English; Schmitt, 2006)

[4] nei5 zou6 saai3 di1 **assignment** mei6.

you do ASP CL assignment SFP

"Have you done all the assignments?"

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

[5] 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; Kov ács, 2001, p. 152)

[6] 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, and Vanniarajan, 1990, p. 79)

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

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

"I persuaged him to buy a house."

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

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

yes café-LOC sit-PROG.3PL

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

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

[9] ni nei-pian **article** hai mei **finish** a?

You that-CL article yet not finish PART/AFFIRM-QUE

"You haven't finished that article yet?"

(Chinese/English; Wei, 2002, p. 696)

[10] ii desu keredomo **tuition** ga totemo **expensive** desu.

good COP/be but tuition PARTIC/NOM very expensive COP/be

"It's good, but the tuition is very expensive."

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

In [1] *nokum* is an EL content morpheme from Korean, but the article *the*, a system morpheme, is from English, the ML. In [2] *lunchbox* is a content morpheme from English, the EL, but it is marked with the appropriate Finnish case, an ML system morpheme. In [3] English *grass* is inflected with prepositional case from Russian, the ML. In [4] the noun classifier *di1*, a system morpheme, is from Cantonese, the ML, and also *assignment*, a content morpheme from English, the EL, is not inflected for the plural marking. In [5] *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 [6] *Hindi song* is from English, the EL, but it is case marked by *-ei*, a system morpheme from Tamil, the ML. In [7] *persuade* is from English, the EL, but the complementizer *la* is a system morpheme from Marathi, the ML. In [8] *terras* is from Dutch, the EL, but it receives locative case from Turkish, the ML. In [9] *article* and *finish* are content morphemes from English, the EL, but *nei* (equivalent to "that" in English) and *pian* (a Chinese noun classifier) going together as a determiner is from Chinese, the ML. Also, there is no perfect aspect marking on the English verb *finish*, since Chinese does not have verb morphology of any sort for this and other grammatical purposes. In [10] *tuition* and *expensive* are from English, the EL, but *totemo* (very) modifying expensive is a system morpheme from Japanese, 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. The evidence clearly indicates that it is the ML which provides all system morphemes.

4.2 Morphosyntactic Procedures Unequally Activated in Codeswitching

In addition to the assumption that morphemes are unequally activated in CS, the other crucial

assumption underlying the BLA Model is that the language pairs involved in CS do not participate equally, one playing a more central role than the other in controlling morphosyntactic procedures. In other words, bilingual lemmas for morphosyntactic procedures are unequally activated in CS (Myers-Scotton, 1993). In a bilingual mode, although both language networks are activated, one is more activated than the other (Green, 1986; Grosjean, 1997). One of the main arguments of the MLF Model is that CS occurs within the constraints of a sentential frame which must be set by the morphosyntactic procedures dictated by the ML word order and syntactically relevant system morphemes (i.e., relational or functional elements) (Myers-Scotton, 1993). As assumed, at the conceptual level the discourse mode is chosen with one of the participating languages as the ML and then corresponding language-specific lemmas are activated at the lemma level to realize the speaker's preverbal message, resulting in CS morphosyntactically controlled by the ML.

Grosjean (1989, 1997) claims that 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. Wei (2015) further claims 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. de Bot and Schreuder (1993) and Wei (2009a) claim that bilingual speakers are able to separate the language systems they know and to mix them in a bilingual mode. This is because that bilingual speakers know that the ML and the EL play unequal roles in CS, and CS is not simply a so-called "mixed" speech but is governed by a set of structural principles, such as those proposed in the MLF Model (Myers-Scotton, 1993, 2002; Myers-Scotton & Jake, 1995; Wei, 2001b) and the BLA Model (2006b). Such a notion of language separation becomes crucial in explaining the structural principles governing CS and in exploring the nature and activity of the bilingual mental lexicon in a discourse involving CS.

Below are some CS instances involving various language pairs which provide strong evidence that it is the ML which provides all the morphosyntactic procedures for CS.

[11] 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)

[12] 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 & Myers-Scotton, 1997, p. 332)

[13] 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)

[14] I have to **ttakē** my hand.

I have to wash my hand

"I have to wash my hand."

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

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

children ACC you force NEG do

"You can't force children.'

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

[16] mula khurcyā **paint** kartāt.

boys chairs paint do+TNS

"Boys paint chairs."

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

[17] 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)

[18] ta meitian qu **office** huoshi qu **library**, hen shao zai jia gong zuo.

he everyday go office or go library very seldom PREP/at home work

"Everyday he either goes to the office or goes to the library, but very seldom works at home"

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

[19] **Supa** is close from here, but I have to walk **juugo fun gurai** to the bus stop.

supermarket is close from here but I have to walk fifteen minutes about to the bus stop

"The supermarket is close from here, but I have to walk about fifteen minutes to the bus stop."

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

50

In [11] *understand* follows the Mandinka OV order. In [12] *join* follows the Tsotsitaal OV order. In [13] the order of *certificate* and its modifiers follow the Swahili word order, not that of English. In [14] *ttakē* (wash) follows the English VO order. In [15] *force* follows the Panjabi OV order. In [16] *paint* follows the Matathi OV order. In [17] 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 [18] the verb phrases *qu office* (go office) and *qu library* (go library) show that unlike English maximal category projection of the preposition phrase where *to* as the head designates GOAL, in Chinese, GOAL is realized in the verb itself. In this example, the EL content morphemes *office* and *library* are switched into the ML predicate-argument structure. In [19] *walk juugo fun gurai to the bus stop* is the English verb initial construction, but in Japanese the verb final construction is always maintained.

The representative CS instances involving various language pairs discussed in the above sections reveal that bilingual speakers can activate any language known to them as the ML, and the ML is more activated than the EL. As assumed in the MLF Model and the BLA Model, it is the ML whose lemmas are fully activated for both content and system morphemes and morphosyntactic procedures, and the EL only supplies certain content morphemes to form ML+EL constituents. The BLA Model emphasizes that such an activation of deactivation occurs at the conceptual level of the speech production process. If at the conceptual level, bilingual speakers decide to choose the bilingual mode, they will then choose intersentential or intrasentential codeswitching, and if the latter is chosen, they must choose one of the languages as the ML. If the bilingual mode and intrasentential codeswitching (CS) are chosen, language-specific semantic/pragmatic feature bundles will activate language-specific lemmas in the bilingual mental lexicon at the lemma level. Such activated lemmas will send directions to the FORMULATER at the functional level for projection of language-specific morphosyntactic procedures to be realized at the position level, where the activated EL content morphemes are inserted into the grammatical frame provided by the ML.

5. Lemma Congruence Checking as an Organizing Principle Governing Codeswitching

Lemma congruence is defined as "a match between the ML and the EL at the lemma level with respect to linguistically relevant features" (Myers-Scotton and Jake, 1995, p. 985). In addition to the unequal roles of the ML and the EL in CS, for CS to be possible, lemma congruence between the languages involved much come into play. The MLF Model proposes the Blocking Hypothesis that a blocking filter blocks any EL content morpheme that is not sufficiently congruent with that of the ML (Myers-Scotton, 1993, p. 120). Myers-Scotton and Jake argue that all EL items must be checked twice, once at the CONCEPTUALIZER and once at the mental lexicon. As introduced earlier, all lemmas include three levels of abstract lexical structure, which figures in explaining and predicting possible CS configurations. 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

51

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 must be checked at each of these levels in the bilingual mental lexicon, which determines the way an EL item may be integrated into an ML frame. Relevant to the study presented in this paper is lemma congruence at the first two levels of abstract lexical structure.

The BLA Model views the nature of lemma congruence relevant to CS as more complex because several different levels or subsystems must be checked, and articulates the relation of lemma congruence checking to a model of bilingual speech production. The BLA 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 juxtaposition must originate with directions in the speaker's mental lexicon (Myers-Scotton & Jake, 1995). According to Wei (2001b, 2002, 2006b), lexicalization patterns across languages reflect the fact that there are different configurations of 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 pragmatic or semantic, or semantic with morphological consequences, affect the structures which will appear in CS. That is why the BLA Model regards lemma congruence between the languages as an organizing principle in CS production.

5.1 Lemma Congruence in Lexical-Conceptual Structure

One obvious reason for bilingual speakers to engage in CS is that they select individual EL content morphemes to encode their communicative intentions. That is, at the conceptual level bilingual speakers do not produce surface level morphemes but rather make appropriate choices about the semantic/pragmatic feature bundles as desired. The semantic/pragmatic feature bundles chosen at this abstract level activate the lemmas in the bilingual mental lexicon to support surface level morphemes. However, the EL lemmas activated by the semantic/pragmatic feature bundles at the conceptual level must be sufficiently congruent with the counterparts of the ML in order for CS to occur. In other words, sufficient congruence between the EL and ML lemmas support the existing lexemes in both languages.

de Bot and Schreuder (1993) posit that because different languages may lexicalize in different ways, the language to be used in L2 production must be specified before chunking takes place. Relevant to CS is the implication that the language pairs involved may differ in semantic/pragmatic feature bundles encoded in lexical-conceptual structure. Grosjean (1982) reports that some codeswitches 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. Wei (2006b) claims that it is language-specific lemma differences at the level of lexical-conceptual structure which motivate CS. In terms of lemma congruence between the languages involved in CS, there is still sufficient cross-linguistic congruence if such lemma differences are only

partial. The BLA Model posits a partial lemma difference is one of the major reasons for a particular EL content morpheme to be activated and switched.

Below are some CS instances which illustrate the notion of "lemma congruence" checking in lexical-conceptual structure.

[20] 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."

[21] 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. There is a school bus every day."

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

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."

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

In [20] paper in English may mean any written piece of word, such as an article, a report, an essay or a composition, but the Chinese equivalent noun *zhi* (paper) itself only means a piece of paper to wrap things up in or to write something on. In [21] a school bus in English means a bus mainly for transporting students to and from a school, but the Chinese equivalent noun phrase *xiaoche* (school bus) usually means a bus for transporting a school's sports or performance team or equipment. In China, most schools even do not have *xiaoche*. In [22] an English *advisor* is expected to assume more responsibilities than a Chinese advisor. In the academic setting, an English advisor is an instructor or professor who offers advice or counsel to students regarding their academic weakness, improvement or progress, sequential course requirements, research projects, thesis or dissertation writing, and so on. Most English advisors are also those who recommend their students to the job market or professional agencies. Chinese does have the equivalent noun *daoshi* (advisor), but a *daoshi* does not necessarily assume the same responsibilities as an English advisor. In China, only a graduate student may have a *daoshi,* whose expected responsibility is to supervise a student in writing his/her thesis or dissertation. These CS instances reveal that there exist semantic differences in the semantic/pragmatic feature bundles of conceptually-related lexemes. Although Chinese possesses conceptually similar lexemes, in these CS instances, the related English lexemes are selected as more appropriate to convey the speakers' intended meanings.

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

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

"Have you done your registration/"

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

two POSS bedroom POP 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, has one bedroom."

[25] 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, 2006b, p. 170)

In [23] 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 convey his/her intended meaning more accurately. In [24] 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. In other words, the concept of "bedroom" in English is not the same as that in Japanese. In [25] 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*.

As commonly observed in naturally occurring CS instances, bilingual speakers may switch to particular lexical items of another language at a certain point during a discourse. This is most probably because of cross-linguistic differences in language-specific lemmas underlying particular lexemes. In terms of abstract lexical-conceptual structure, languages may possess similar lexical items, but language cues may have different values (Li, 1996; Nishimura, 1997; Wei, 2001b, 2002). When the language cue specifies a particular language at a certain point of bilingual speech production, its lexical item receives

54

activation and is thus easier to be selected. As suggested by Green (1986), lemmas are tagged with a language label. Poulisse and Bongaerts (1994) propose the lexical access model to explain 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. The speaker may switch to an EL lexical item because the ML has a similar but nonequivalent lexical item or the ML has not lexicalized a particular concept. 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.

The above instances show how cross-linguistic differences in semantic/pragmatic feature bundles may affect lexical selections in mixed constituents. Checking for lemma congruence across ML and EL elements becomes necessary, and the result of this checking has consequences for how the EL lexemes supported by the EL lemmas in question will appear in CS.

5.2 Lemma Congruence in Predicate-Argument Structure

As assumed in the MLF Model and the BLA Model, the lexical items stored in the mental lexicon contain syntactic, morphological, and phonological information about them in addition to the specification of their lexical content (i.e., semantic features). The activation of the appropriate lexical items frees the syntactic information about them, which will activate syntactic procedures. For grammatical encoding and phonological encoding (Levelt, 1989), the FORMULATOR must have access to the mental lexicon. Since lexical items from both languages involved in CS can be activated, the question becomes which of the two languages controls the syntactic procedures.

The MLF Model (Myers-Scotton, 1993, 2002) claims that whichever language activated as the ML provides the sentential frame into which the EL content morphemes are switched. Thus, it is the ML which controls the predicate-argument structure and supplies system morphemes, a subcategorization frame for the verb, and the morpheme order. It should be noted that there is an inseparable connection between the level of lexical-conceptual structure and the level of predicate-argument structure with regard to lemma selection from the mental lexicon. As stated by Kroll and de Groot, "language specific lemmas from the interconnection between the lexical-conceptual mappings to and from syntax" (1997, p. 190). The BLA Model (Wei, 2006b, 2009a, 2015, 2020) emphasizes that lemma congruence checking at the level of lexical-conceptual 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 order for the supported EL lexeme to be morphosyntactically integrated into the sentential frame set by the ML, its morphosyntactic pattern must be congruent with that of the ML lexeme whose lemma is sending the morphosyntactic directions to the FORMULATOR.

Wei (2009a, 2009b) finds that one of the frequently occurring instances of Chinese/English CS is that Chinese bilinguals tend to switch to English verbs/verb phrases as well as English nouns/ noun phrases.

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

you must think way make money

"You must think of ways to make money."

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

he just arrive he must learn drive

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

[28] wo de 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)

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

complain not useful PARTIC/AFFIRM

"It's useless to complain."

[30] UT-de fanzi 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."

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

In [26] the speaker switches to *make money*, which fits the Chinese morphosyntactic frame for the V-O order. In [27] the infinitive verb *drive* is switched into the Chinese verb phrase as its direct object, which is allowed in the Chinese morphosyntactic frame. In [28] *give me trouble* is switched into the Chinese subcategorization frame for the V-O-O order. In [29] *complain* is switched into the subject position, which is congruent with the Chinese sentential frame, 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. In [30] demolish is congruent with the Chinese predicate-argument structure, except that Chinese does not possess the morphemes as in English for realizing the passive construction. Such CS instances indicate that the switched items from the EL are sufficiently congruent with the ML lemma entries which direct the morphosyntactic procedures to the FORMULATOR to produce the frame into which they are switched.

6. Lemma Incongruence as a Driving Force for Codeswitching

As commonly observed, there are cross-linguistic differences at any level of abstract lexical structure (i.e., at the level of lexical-conceptual structure, predicate-argument structure or morphological realization patterns). Consequently, when an EL lemma is selected, but it does not match that of the ML, some compromising strategies must be taken in order for CS to occur. One of the compromising strategies is for the speaker to produce EL islands (Jake & 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. EL islands may be produced at the level of lexical-conceptual structure or at the level of predicate-argument structure when lemma incongruence occurs but the speaker still wants to be engaged in CS.

Such a compromising strategy makes CS possible because in producing an EL island, it is the EL which directs the FORMULATOR to activate only the EL morphosyntactic procedures. According to Levelt (1989), in speech production, it is the FORMULATOR which gives language-specific directions for the grammatical and phonological encoding. For possible CS, if the language pairs cannot possibly be encoded by means of the same morphosyntactic procedures, the speaker may stop the encoding of one of them and continue with the other so as to solve the problem of lemma incongruence. Different from Levelt's model of speech production, the MLF Model specifies that in CS the choice of one EL procedure versus another is determined by the larger ML frame. The BLA Model claims that lemma incongruence between the Language pairs involved in CS in regard to lexical-conceptual structure and predicate-argument structure is one of the major reasons for EL islands to be produced if the EL lemmas are selected for the speaker's communicative intentions. The BLA Model regards lemma incongruence as a driving force for CS.

6.1 Lemma Incongruence in Lexical-Conceptual Structure

As commonly observed, languages do not lexicalize concepts in the same way due to the fact that lexical representations are language-specific. As assumed in the BLA Model, if the bilingual mode is chosen at the conceptual level, speakers do not produce surface morphemes but rather make appropriate choices about the semantic/pragmatic information that they intend to convey. If the EL lemmas for some particular semantic/pragmatic feature bundles are activated but do not match the ML counterparts, speakers must take the compromising strategy by producing EL islands for possible CS configurations.

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

then you tomorrow call me

"Then you call me tomorrow."

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

you can-not-can give me a ride

"Can you give me a ride?"

[33] 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 [31] the English 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* ("make phone to me"). Since the speaker chooses the EL lemma which activates the EL lexical-conceptual structure, the whole verb phrase "call me" is activated and produced as an EL island. In [32] the verb phrase *give me a ride* is incongruent with the ML counterpart *song wo yixia* ("send me one time"). The incongruence between these two verb phrases lies in their individual lexical-conceptual structure. While in the EL lexical-conceptual structure of the means of transportation is conflated in the noun *ride* as the direct object of the verb *give*, in the ML lexical-conceptual structure it is conflated in the verb *song*, which itself may not contain the means of transportation at all. The speaker chooses the EL expression most probably to make his intended meaning explicit and specific. Thus, when the EL lemma underlying "ride" is activated, the whole EL verb phrase is accessed and produced as an EL island. In [33] *pick you up* is accessed as an EL island, in which the pronominal object *you* is placed between the verb *pick* and the particle satellite *up*. The speaker chooses *pick up* most probably because the EL phrasal verb contains the meaning of "to take on as a passenger", but the ML counterpart *jie* usually does not. In Chinese, the verb *jie* means "meet" (e.g., to go to the bus/train station or airport to meet somebody"), which does not necessarily involve providing personal transportation. The speaker chooses the English phrasal verb *pick up* to realize his communication intention more accurately. Also, in [33] *come* is accessed with the infinitive maker *to*, an EL system morpheme, to introduce the phrasal verb *pick you up*. This is because in English the infinitive marker becomes obligatory if two successive verbs are activated and selected simultaneously. Thus, *come to pick you up* is switched as a larger EL island into the ML sentential frame.

Cross-linguistic differences in semantic/pragmatic feature bundles are also found in bilingual speakers' choice of certain fixed/idiomatic expressions from the EL. If the meaning or intention contained in the EL lemma is not sufficiently congruent with that in the ML lemma, and if bilingual speakers choose the one in the EL lemma, they will produce an EL island in CS.

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

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

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

[35] nali you wutai jiqi dan san-tai si **out of order**.

there have five-CLASSIF machine but three-CLASSIF COP/be out of order

"There were five machines there, but three of them were out of order.

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

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

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

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

In [34] *on sale* is the EL fixed/idiomatic expression. The Chinese equivalent to on sale is *jian mai* ("cheap sale"). In [35] *out of order* is the EL fixed/idiomatic expression. The Chinese equivalent to *out of order* is *chu guzhang* ("something going wrong"). In [36] *on time* is the EL fixed/idiomatic expression. The Chinese equivalent to *on time* is *zhunshi* ("punctually"). The EL fixed/idiomatic expressions in these examples are all introduced by particular prepositions, but the Chinese equivalents are not. Since Such EL lemmas are activated in the bilingual speakers' mental lexicon, the output of the expressions follows the EL lexical-conceptual structure and morphosyntactic procedures.

It seems that one of the major reasons for EL islands to be switched is that in the case of non-idiomatic expressions, the speaker's intended meaning at a certain point during a discourse cannot be realized in the ML because of the insufficient congruence between the ML semantic/pragmatic feature bundles and those of the EL. The other major reason is that in the case of fixed/idiomatic expressions, a complete EL island is activated as a single lexical unit for the speaker's preferred EL lexical-conceptual structure. In either case, the compromising strategy is to produce EL islands. In other words, a particular EL semantic/pragmatic concept is accessed as a single unit observing the specific EL morphosyntactic procedures.

6.2 Lemma Incongruence in Predicate-Argument Structure

As specified in the MLF Model, it is the ML which controls the morphosyntactic structure of the sentence containing the switched items from the EL by supplying system morphemes, subcategorization frames for verbs, and morpheme order. As claimed in the BLA Model, morphosyntactic procedures are realized by the FORMUALTOR at the functional level, but before they are sent to the FORMULATOR, lemmas from both languages can be activated at a certain point during a discourse. Thus, in addition to lemma congruence checking at the level of lexical-conceptual structure, lemma congruence checking at the level of predicate-argument structure must also come into play for possible CS. The BLA Model posits that 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 is not sufficiently congruent with that of the ML. If this happens, but the speaker prefers the EL concept as realized in the EL predicate-argument structure, a compromising strategy much be taken for possible CS. That is, the speaker must produce an EL island.

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

I can wait for you till two o'clock

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

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

[38] ta jingchang **fails students in exams**.

she often fails students in exams

"She often fails students in exams."

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

you graduate CONJ/after can teach English to nonnative speakers

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

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

[40] 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)

In [37] *wait for* is accessed as a single phrasal verb., an EL island. While in English the object *you*, the THEME, is introduced by the preposition, in Chinese the same object is introduced by a single verb *deng* (wait). In [38] the verb phrase headed by *fail* is an EL island. While in English *fail* can be used as a causative verb and this takes the grammatical subject as the AGENT who makes the failure happen, in Chinese the equivalent verb *shibai* means "be defeated in…" and is used only as a noncausative verb with the grammatical subject as the EXPERIENCER. In [39] *nonnative speakers*, the RECIPIENT, is introduced in the English indirect object dative construction headed by preposition *to*. By contrast, the equivalent Chinese verb *jiao* (teach) only permits the double object construction (e.g., *jiao ta English* (teach him English)). In [40] *her homework*, the THEME, is introduced by the preposition *with*. By contrast, in Chinese the THEME is always introduced by a specific verb such as *zhuo* (do).

Such instances of CS reveal that if speakers select certain particular EL verbs at the level of lexical-conceptual structure, but the EL and the ML are incongruent at the level of predicate-argument structure, they may activate the whole verb phrases in the EL and produce them as EL islands.

60

7. Conclusion

This study explains CS at some abstract levels of the bilingual speech production process with a focus on bilingual lemma activation as a crucial interface between speaker intention and code activation. It 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, this study regards CS as projection of bilingual lemmas in contact. Based on the linguistic analysis of some naturally occurring CS instances involving various language pairs, this study reaches several conclusions reading the nature and activity of the bilingual mental lexicon as observed in the bilingual speech production process.

(1) While the bilingual speaker's languages are turned "on" during a discourse involving CS, they are not equally activated at the same time. Whichever language that the bilingual speaker chooses as the ML is more strongly activated than the EL. It is only the ML which provides the sentential frame, controls morphosyntactic procedures, and provides all system morphemes as well as most content morphemes.

(2) The bilingual mental lexicon contains language-specific lemmas, which are in contact in CS. If the speaker chooses the bilingual mode to engage in CS, he/she may activate certain language-specific lemmas as desired for his/her intended meaning. 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 can appear in CS to express the speaker's intended meaning.

(3) The intentional use of the EL system often manifests itself as CS. Some switches are motivated by the lack of particular words in the ML for the speaker's communicative intentions. The speaker may take compensatory strategies to solve lexical problems caused by the lexical gap between the languages involved. When the speaker's communicative intention contained in the preverbal message call for a particular EL content morpheme, this selection activates the EL lemma supporting that morpheme.

(4) For a possible CS realization, 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: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. If lemma incongruence or insufficient congruence occurs between the language pairs at any of these levels, but the speaker does not want to give up CS, a radical compromising strategy, such as the production of EL islands, must be taken in order for CS to be possible.

As proposed in the MLF Model and the BLA Model, the unequal roles played by the ML and the EL, different aspects of abstract lexical structure, and bilingual lemmas in contact affect CS. This study posits that CS, like any natural speech behavior, is a rule governed bilingual behavior and regards CS as projection of bilingual lemmas in contact.

References

Altarriba, J., & Basnight-Brown, D. M. (2009). An Overview of Semantic Processing in Bilinguals. In A. Pavlenko (Ed.), *The Bilingual Mental Lexicon: Interdisciplinary Approaches* (pp. 79-98). Bristol: Multilingual Matters. https://doi.org/10.21832/9781847691262-006

Azuma, S. (1993). The Frame-Content Speech Hypothesis in Speech Production: Evidence from Intrasentential Code Switching. *Linguistics*, *31*, 1071-1093. https://doi.org/10.1515/ling.1993.31.6.1071

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

Berg, T., & Schade, U. (1992). The Role of Inhibition in a Spreading-Activation Model of Language Production. *Journal of Psycholinguistic Research*, *21*, 405-462. https://doi.org/10.1007/BF01067523

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

Costa, A. (2005). Lexical Access in Bilingual Production. In J. F. Kroll, & A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 308-325). Oxford: Oxford University 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

de Groot, A. M. B. (2002). Lexical Representation and Lexical Processing in the L2 User. In V. J. Cook (Ed.), *Portraits of the L2 User* (pp. 32-63). Clevedon: Multilingual Matters. https://doi.org/10.21832/9781853595851-004

Færch, 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

Færch, C., & Kasper, G. (1986). Cognitive Dimensions of Language Transfer. In E. Kellerman, & M. Sharwood Smith (Eds.), *Crosslinguistic Influence in Second Language Acquisition* (pp. 49-65). Elmsford, NY: Pergamon Press.

Green, D. W. (1986). Control, Activation and Resource: A Framework and Model for the Control of Speech in Bilinguals. *Brain and Language*, *27*(2), 210-223. https://doi.org/10.1016/0093-934X(86)90016-7

Green, D. W. (1993). Toward a Model of L2 Comprehension and Production. In R. Schreuder, & B. Weltens (Eds.), *The Bilingual Lexicon* (pp. 249-277). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.6.12gre

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

Grosjean, F. (1985). The Bilingual as a Competent but Specific Speaker-Hearer. *Journal of Multilingual and Multicultural Development* 6: 467-77. https://doi.org/10.1080/01434632.1985.9994221

Grosjean, F. (1989). Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person. *Developmental Psychology*, *25*, 577-588.

Grosjean, F. (1994). Individual bilingualism. In R. E. Asher, & J. M. Y. Simpson (Eds.), *The Encyclopedia of Language and Linguistics* (pp. 1656-1660). Oxford, UK: Pergamon.

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). Mahwah, NJ: Erlbaum.

Grosjean, F., & Soares, C. (1986). Processing Mixed Language: Some Preliminary Findings. In J. Vaid (Ed.), *Language Processing in Bilinguals: Psycholinguistic and Neuropsychological Perspectives* (pp. 145-179). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gumperz, J. J. (1982). Conversational Code-switching. In J. J. Gumperz (Ed.), *Discourse Strategies* (pp. 55-99). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511611834

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.

Jackendoff, R. (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.

Jarvis, S. (2009). Lexical Transfer. In A. Pavlenko (Ed.), *The Bilingual Mental Lexicon: Interdisciplinary Approaches* (pp. 99-124). Bristol: Multilingual Matters. https://doi.org/10.21832/9781847691262-007

Jiang, N. (1999). Testing Processing Explanations for the Asymmetry in Masked Cross-language Priming. *Bilingualism: Language and Cognition*, *2*, 59-75. https://doi.org/10.1017/S1366728999000152

Jiang, N. (2000). Lexical Representation and Development in a Second Language. *Applied Linguistics*, *21*(1), 47-77. https://doi.org/10.1093/applin/21.1.47

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

Kaplan, R. M., & Bresnan, J. W. (1982). Lexical-Functional Grammar: A Formal System for

Grammatical Representation. In J. W. Bresnan (Ed.), *The Mental Representation of Grammatical Relations* (pp, 173-281). Cambridge, MA and London: MIT Press.

Keatley, C. W. (1992). History of Bilingualism Research in Cognitive Psychology. In R. J. Harris (Ed.), *Cognitive Processing in Bilinguals* (pp. 15-49). Amsterdam: North-Holland. https://doi.org/10.1016/S0166-4115(08)61486-7

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

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

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical and Conceptual Memory in the Bilingual: Mapping Form to Meaning in Two Languages. In A. M. B. de Groot, & J. F. Kroll (Eds.), *Tutorials in Bilingualism*: *Psycholinguistic Perspectives* (pp. 169-200). Mahwah, NJ: Erlbaum.

Kroll, Judith F., and Erika Stewart. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections between Bilingual Memory Representations." *Journal of Memory and Language*, *33*, 149-174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J. F., & Sunderman, G. (2003). Cognitive Processing in Second Language Learners and Bilinguals: The Development of Lexical and Conceptual Representations. In C. J. Doughty, & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 104-129). Malden, MA: Blackwell. https://doi.org/10.1002/9780470756492.ch5

Kroll, J. F., & Tokowicz, N. (2005). Models of Bilingual Representation and Processing: Looking Back and to the Future. In J. F. Kroll, & A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 531-553). Oxford: Oxford University Press.

Leshchenko, Y., Dotsenko, T., & Ostapenko, T. (2015). Cross-linguistic Interactions in Bilingual Mental Lexicon and Professional Linguistic Competence Formation: An Experimental Research with Native Speakers of the Komi-Permyak and Russian Languages. *Procedia–Social and Behavioral Sciences*, *214*, 1039-1047. https://doi.org/10.1016/j.sbspro.2015.11.700

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

Li, D. C. S. (1996). *Issues in Bilingualism and Biculturalism: A Hong Kong Case Study*. New York: Peter Lang.

Lipski, J. (1977). Codeswitching and the Problem of Bilingual Competence. In M. Paradis (Ed.), *Aspects of Bilingualism* (pp. 250-265). Columbia, SC: Hornbeam Press.

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

Myers-Scotton, C. (1993 [1997]). *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon 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 Turning 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.

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

Nishimura, M. (1997). *Japanese/English Code-switching: Syntax and Pragmatics*. New York: Peter Lang.

Paradis, M. (1981). Neurolinguistic Organization of a Bilingual's Two Languages. In J. E. Copeland (Ed.), *The Seventh LACUS Forum*. Columbia, SC: Hornbeam Press.

Paradis, M. (1989). Bilingual and Polyglot Aphasia. In F. Boller, & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 2, pp. 117-140). Amsterdam: Elsevier.

Paradis, M. (1997). The Cognitive Neuropsychology of Bilingualism. In A. M. B. de Groot, & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 331-354). Mahwah. NJ: Erlbaum.

Pavlenko, A. (2009). Conceptual Representation in the Bilingual Lexicon and Second Language Vocabulary Learning. In A. Pavelenko (Ed.), *The Bilingual Mental Lexicon: Interdisciplinary Approaches* (pp. 125-160). Bristol: Multilingual Matters. https://doi.org/10.21832/9781847691262-008

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

Pienemann, M. (1999). *Language Processing and Second Language Development*. Amsterdam: John

Benjamins. https://doi.org/10.1075/sibil.15

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

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

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.* Oxford: Blackwell. https://doi.org/10.1016/B978-012589042-7/50019-0

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). Unpublished Russian-English corpus.

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

Soares, C., & Grosjean, F. (1984). Bilinguals in a Monolingual and a Bilingual Speech Mode: The Effect on Lexical Access. *Memory and Cognition*, *12*, 380-386. https://doi.org/10.3758/BF03198298

Sotillo, S. M. (2016). Structural Features and Sociolinguistic Functions in Bilinguals' Code Switched Face-to-Face and Short Message Service Conversations. *International Journal of Language and Linguistics*, *3*(4), 11-23.

Sridhar, S. N., & Sridhar, K. (1980). The Syntax and Psycholinguistics of Bilingual Code-switching. *Canadian Journal of Psychology*, *34*, 7-16. https://doi.org/10.1037/h0081105

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. (2000a). Unequal Election of Morphemes in Adult Second Language Acquisition. *Applied Linguistics*, *2*, 106-140. https://doi.org/10.1093/applin/21.1.106

Wei, L. (2000b). 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. 13-15). Fryske Akademy.

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. https://doi.org/10.1006/brln.2001.2557

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

Wei, L. (2006). 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. (2015). *Interlanguage: The Abstract Level in Language Acquisition*. New York: The Edwin Mellen Press.

Wei, L. (2020). *The Bilingual Mental Lexicon*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

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