Original Paper

Models of Codeswitching and a Lexical-Conceptual Approach to Bilingual Communication

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

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

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

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

Abstract

Unlike most previous studies of intrasentential codeswitching (CS for short) which focus on the descriptions of switched items and their surface configurations at the sentence level, this study investigates such a bilingual communicative strategy at a rather abstract level by adopting the Matrix Language Frame (MLF) Model (Myers-Scotton, 1993; Myers-Scotton & Jake, 1995) and the Bilingual Lemma Activation (BLA) Model (Wei, 2002, 2006, 2015). The MLF Model claims that the two participating languages play unequal roles in CS configurations, and the BLA Model claims that the bilingual mental lexicon contains language-specific lemmas in contact during CS and argues that it is the activation of language-specific lemmas which drives CS. Some naturally occurring CS instances selected for the study reveal that bilinguals switch content morphemes from one language into another language's sentential frame most probably because of cross-linguistic lexical-conceptual gaps. As specified in the MLF Model, bilinguals do not switch system (grammatical) morphemes because it is the Matrix Language (the host language) which provides the sentential frame and all system morphemes, and the Embedded Language (the guest language) inserts certain content morphemes as desired into the lexical slots in the ML frame. By exploring the nature and activity of the bilingual mental lexicon, this study regards CS as a lexical-conceptual solution to bilingual communication.

Keywords

bilingual, mental lexicon, content, system, morpheme, lemma, lexical-conceptual structure, predicate-argument structure

1. Introduction

As commonly observed in bilingual speech production, bilinguals may switch certain lexical items from one language into sentences in another language. This is recognized as the alternative use of two languages within sentence boundaries, which is called intrasentential codeswitching (hereafter, CS for short). For example, the speaker switches the English word "horses" into the Spanish sentence: veo los **horses** (I see the horses) (Apple & Muysken, 1987, p. 125). CS is commonly known as bilinguals' natural speech behavior as an outcome of language contact, which has its own linguistic manifestations and constraints (Wei, 2020, 2024). Numerous studies have been done to formulate the structural principles governing CS, but most of them remain at a rather superficial level of descriptions of the switched items and their grammatical configurations of CS. By adopting the Matrix Language Frame (MLF) Model (Myers-Scotton, 1993; Myers-Scotton & Jake, 1995) and the Bilingual Lemma Activation (BLA) Model (Wei, 2002, 2006, 2015), this study claims that CS is an outcome of the unequal activation of the two linguistic systems in contact during the bilingual speech production process and that CS is a lexical-conceptual solution to bilingual communication. To support such claims, this study cites many naturally occurring CS instances involving various language pairs. The descriptive and analytical evidence indicates that the languages involved in CS production play unequal roles in building the sentential frame into which certain items as conceptually desired by bilinguals are switched from another language.

2. Commonly Recognized Models of Codeswitching

Some commonly recognized models of CS are briefly reviewed as the theoretical background of this study. The structural principles and constraints as proposed by these models are comparatively reviewed to see which theory is more observationally, descriptively and explanatorily adequate than the others when applied to naturally occurring CS instances.

Surface-based Models

As the term "surface" indicates, surfaced-based models, by adopting a linear perspective, focus on the surface structure of sentences for clues about structural constraints on CS. A number of studies carried out under this heading have tended in the same direction, all claiming from various points of view that CS is a unified form of speech in which the grammars of the two languages involved in CS must be well preserved (e.g., Lipski, 1977; Pfaff, 1979; Poplack, 1980, 1981; Gumperz, 1982; Woolford, 1983; DiSciullo, Muysken & Singh, 1986; Poplack & Sankoff, 1988; Muysken, 1988, 1991). In other words, such studies argue that the two languages participating in CS must be identical to their use in monolingual speech and remain intact. Such studies aim to describe how monolingual grammars cooperate to form a single grammatical structure for CS by focusing on the surface configurations or structural linearity of utterances containing switched items. The outstanding claims of such studies and some key assumptions underlying surface-based models are reviewed below.

Lipski (1977) claims that a switch can only take place if the structural well-formedness of the languages involved in CS is well kept, and switches cannot possibly take place between certain indissolubly linked elements (e.g., switches are not possible between an article and a noun or between a pronoun subject and the verb). Similarly, Pfaff (1979) claims that in CS structural constraints permit only surface structures which are grammatical in both languages. That is, surface structures common to the participating languages are favored for switches. For example, according to her observation, in Spanish/English CS, the periphrastic verbal constructions raise no syntactic conflict; adjective switches are unrestricted when they take the form of predicate adjectives, but they are limited within the noun phrase; postnominal attributive adjectives do not in general occur in noun phrases with switched adjectives or nouns; prenominal adjective switching is restricted to a small class of typical limiting adjectives, which precede in both Spanish and English. Such observations lead to the conclusion that CS can only be generated by the two constituent grammatical systems whose rules must be capable of ensuring that the switch from one language to the other operates smoothly.

Similar to Lipsky's and Pfaff's claim that it is the like constituents matched in word order across the languages involved in CS which determine their switchability, Poplack (1980, 1981) claims that the two languages involved in CS must remain intact at all levels. Poplack (1980) and Poplack and Sankoff (1988) propose two constraints on CS: *the Free Morpheme Constraint* and *the Equivalence Constraint*, which are claimed to be general enough to account for all instances of CS and at the same time are restrict enough not to generate instances of non-occurring codeswitches. These two constraints are proposed as follows.

The Free Morpheme Constraint: "Codes may be switched after any constituent provided that the constituent is not a bound morpheme" (Poplack, 1980, p. 585). This means that switching cannot take place between two bound morphemes or between a bound morpheme and a free morpheme. As commonly defined, bound morphemes are the ones which are incapable of standing alone as words; they must be connected (bound) to other morphemes to form inflections (e.g., in English, nouns are inflected for plural marking, and verbs are inflected for tense/aspect/voice marking) and derivatives (e.g., in English, prefixes and suffixes are attached to other morphemes to derive one word from another, and they allow words to change parts of speech); free morphemes are the ones which can stand alone as words (i.e., free-standing items).

[1] ***eat**-iendo

    -ing

  "Eating"

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

Such a constraint prevents the switching in [1] from occurring because the Spanish bond morpheme *-iendo* (progressive *-ing*) is affixed to the English free morpheme *eat* (a verb root). However, an examination of the CS literature reveals numerous counterexamples in analytic and non-agglutinative languages. The following are just a few of them.

[2] a Ne mI **help**-e

    3 PL COP me -PRES PROG

    "They are helping me."

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

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

    I'm   have-ing pancake-s

    "I'm having pancakes."

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

[4] **vask**-ing

    wash

    "washing"

    (English/Danish; Petersen, 1988, p. 483)

These examples show that neither the bound vs. free morpheme distinction nor typological distinctions are relevant to permissible 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: "Code-switches will tend 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). However, the CS literature contains many counterexamples to such a surface-based constraint.

[5] j'ai vu un ancient **tilmid djali**.

    I PAST see an old student of mine

    "I saw a former student of mine."

    (French/Moroccan Arabic; Bentahila and Davies, 1983, p. 319)

In Moroccan Arabic, adjectives follow the nouns they modify. Though this is also true of most French adjectives, there are other French adjectives which must proceed their nouns they modify. As predicted

17

by *the Equivalence Constraint*, such a switch in [5] would be impossible since the only word order shared by French and Moroccan Arabic is that adjectives follow the nouns they modify, but in [5] the French adjective *ancient* precedes the Moroccan Arabic NP (noun phrase) tilmid *djali*.

[6] Unaweza kumpata amevaa nguo nyingine **bright** …

 you can INF find AFFIRM wear clothes other

 "You can find wearing other bright clothes …"

 (Swahili/English; Myers-Scotton, 1993, p. 28)

[7] Anaonekana kama ni mtu **innocent**.

 he PRES look like as if COP person

 "He looks like (he) is an innocent person."

 (Swahili/English; Myers-Scotton, 1993, p. 29)

[6] and [7] show that though Swahili and English have different word orders with the NP: noun-adjective in Swahili, but adjective-noun in English, English adjectives are not prohibited from occurring in Swahili/English CS.

Numerous examples as cited in the CS literature show that inequivalent points do not prohibit CS from occurring. Thus, the "intactness" rule is rather questionable. The structural constraints or rules for switching as proposed in the surface-based models are only quantitative tendencies.

Government and Binding Models

Unlike surface-based models, government and binding models stress dependency rather than linearity (Klavans, 1983; Woolford, 1983; Pandit, 1990; DiSciullo, Muysken, & Singh, 1986). Researchers adopting such models motivate their CS constraints with the "government conditions" posted for monolingual data within the Government and Binding Theory of Chomsky (1981). They all claim that "There cannot be a switch between two elements if they are lexically dependent on each other" (Appel & Muysken, 1987, p. 124). DiSciullo, Muysken, and Singh (1986) propose a restriction in terms of "government", which specifies that whenever a constituent X governs a constituent Y, both constituents must come from the same language; however, governed elements can be switched when there is a neutralizing element such as a determiner. That is, if X governs Y, X and Y must share the same language index: … Xq … Yq… (q is a language index). This constraint holds at S-structure (i.e., the surface configuration of a sentence). This predicts that [8a] will be acceptable, but [8b] will not (Appel & Muysken, 1987, 125).

[8a] veo los **horses**.

   see the

   "I see the horses."

[8b] *veo the horses.

The switch in [8a] will be acceptable because the Spanish *los* makes the whole noun phrase (NP) Spanish, but the whole NP in [8b], even though governed by the Spanish verb *veo*, is in English, thus creating an unacceptable switch. DiSciullo *et al*. (1986) claim that switching is possible at naturalization sites (i.e., at nodes carrying two indices). In [8a] *los horses* carries two indices, but in [8b] *the horses* does not.

Though government and binding models look beyond surface linear ordering, they still consider CS as basically a syntactic phenomenon following the same structural constraints evident in monolingual surface structure. However, like surface-based models, there have been many counterexamples to the government-binding restrictions (e.g., Romaine, 1989; Pandit, 1990; Myers-Scotton, 1993; Wei, 2020). For example, Romaine (1989) finds that switching between V and its NP constituent occurs in her Panjabi/English CS data.

[9] **Parents** te **depend** hona.

   "It depends on the parents."

   (Panjabi/English; Romaine, 1989, p. 135)

This shows the compound-verb construction containing a Panjabi "operator verb" *honda*. She provides other similar such as "time waste kərna" and "exams pass kərna") (Romaine, 1989, p. 140). Therefore, government and binding theory alone does not adequately explain structural constraints on CS. Its major problem lies in the fact that a purely syntactic approach is still too close to the surface-based analysis.

Subcategorization Models

Because of the inadequacy of surface level equivalences as structural constraints on CS, other researchers propose more abstract equivalence to account for switchability (Bentahila & Davies, 1983; Muysken, 1990, 1991; Azuma, 1991). Their subcategorization models stress lexical fame restrictions as specified in syntactic theories. "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 restrictions on particular lexical items of either language" (Bentahila & Davies, 1983, p. 329). Similarly, according to Muysken (1991), "categorial equivalence" is required where specific categories are lexically specified in the subcategorization frame of a lexical item. He claims that lexical elements impose certain

19

requirements on their environments, and switched elements are constrained by lexical subcategorizations, but he does not offer any supporting evidence or further discussion.

"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" (Azuma, 1991, p. 7). Azuma (1991) proposes *the Frame-content Hypothesis*, which specifies two stages in CS: the frame-building state where closed class items are accessed and retrieved, ad the content-word insertion stage where content words are inserted in the planning frame. That is, closed class items are essential members of the planning frame, and content words are not members of the planning frame but are inserted in a later stage. Below is one of the examples he cites to support this hypothesis.

[10] Hata si-ku-**comment** …

    even 1S.NEG-NEG.PAST-comment …

    "I didn't even comment …"

    (Swahili/English; Scotton, 1983)

This example shows that *si* (pronoun) and *ku* (tense marker) are closed class items in Swahili; the switched *comment* is a verb stem, an open class item in English. Azuma (1991) further proposes that the two stages are serial and do not interact. The host language participates in the frame-building state to satisfy that language's subcategorization, and then the guest language participates in the content-word insertion stage to insert content words from that language. He emphasizes that the relation between the two stages is strictly serial and non-interactive. That is, the present frame should not be influenced by the subsequent stage of content-word insertion. This predicts that the guest language's subcategorization restrictions, if different from those of the host language, would not be met. Below are the unattested examples he cites to make the point.

[11] Je dois **nSeLi**.

    I must I pray

    "I must pray."

[12] Elle desire **tzwez had I?am**.

    she wants she gets married this year

    "She wants to get married this year."

    (French/Arabic; Bebtahila & Davies, 1983)

Such instances of CS are predicted not to occur because Arabic verbs subcategorize for finite complements, but French verbs subcategorize for infinitive complements. According to Azuma, the

subcategorization of the frame must always be preserved even after switching in the content-word insertion stage. Such a subcategorization model differs from other models in three specific ways. First, it relies on a traditional open vs. closed class diction. Second, it posits a two-stage process in which the subcategorization requirements of the frame-building stage take precedence. Third, it distinguishes the roles of the two languages involved in CS: one directing the frame building, and the other inserting content words.

Like government and binding models, subcategorization models, though they depart from surface-based models by looking beyond surface linear ordering, still consider CS as basically a syntactic and lexical phenomenon similar to those of monolingual surface structures.

Minimalist Models

Minimalist models draw on the theoretical assumption of a syntactic theory: the Minimalist Program (MP) (Chomsky, 1995). According to the MP, all surface differences in word order or morphological realization patterns are caused by the rearrangement of elements in the syntactic structure resulting from movement rules which are necessarily triggered by lexically encoded morphological features. Thus, the MP claims that "All learning is lexical, and all parameters are micro-parameters associate with individual lexical items" (MacSwan, 2000, p. 44). Building on the MP, some researchers propose minimalist models to account for CS (Toribio & Rubin, 1996; MacSwan, 1997, 1999a, 1999b; Boeschoten & Huybregts, 1999, Ritchie & Bhatia, 1999). Minimalist models claim that no additional rule-based constraints are necessary because CS arises from the insertion of lexical items from another language into the phrase structure. "Nothing constrains code switching apart from the requirements of the mixed grammars" (MacSwan, 2005, p. 69). A well-formed clause containing switched items results as long as that all lexical items satisfy all features, and there is convergence at both Logical Form and Phonetic Form. That is, CS is seen simply as the consequence of two lexicons in contact in a mixed utterance. This is because the bilingual features in CS necessitate the same checking as monolingual features, and such feature checking does not occur with any special mechanisms. In other words, no statements, rules, or principles of gramma are required for the mixed system. MacSwan posits that the principles and requirements of each participating grammar in any instance of CS, together with the principles and requirements of Universal Grammar (UG), are sufficient for explaining every aspect of CS.

MacSwan points out one main advantage of a minimalist approach to CS over any other approach: "Because the syntactic component of computational system (C_{HL}) may be assumed to e invariant cross-linguistically, no "control structure" or "third grammar" is required to mediate between contradictory requirements" (2000, p. 51). Contrary to this point, the MLF Model (Myers-Scotton, 1993 [1997], 2002) proposes that the key to resolving feature mismatches or structural conflicts in CS is the identification of one of the participating languages as the ML. *The Uniform Structure Principle* specifies that

> A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for tis 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 (Myers-Scotton, 2002, pp. 8-9).

Since bilingual speech involving CS is a natural linguistic phenomenon, when this principle is applied to CS, the ML-EL opposition comes into play: One of the languages involved in CS uniformly sets the morphosyntactic frame. In monolingual data, identifying the ML does not involve an ML-EL opposition; in bilingual data, the asymmetry in frame-building between participating languages becomes indispensable. In CS, only one of these languages is the ML, providing the morphosyntactic frame structuring any constituent including any item switched from the EL.

[13] wo you liang-fen **paper** mingtian bixu jiaoshangqu, ke wo xinazai yi-fen hai mei **finish** ne.

I have two-CL paper tomorrow must turn in, but I now one-CL yet not finish ART/AFFIRM

"I must turn in two papers tomorrow, but now I haven't finished one yet."

Chinese/English; Wei, 2009, p. 276)

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

good COP/be but the tuition COP/be very expensive

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

(Japanese/English; Wei, 2009, p. 277)

[15] Tena huwa mtu m-moja **stingy** sana hu-bani Ø-pesa z-ake.

once more 3s-COP-FV CL.1-person CL.1-one very 3s-sit on CL.10-money CL.10-3s/POSS

"Once more he is one very stingy person [who] sits on his money."

(Swahili/English; Myers-Scotton, 1988)

In [13], the noun *paper* precedes the verb *jiashangqu* (in turn), in accordance with the Chinese (the ML) word order. The verb *finish* also follows its object *yifen* (one-CL *[paper]*). In [14], the adjective *expensive* precedes the copula *desu* (be), in accordance with the Japanese (the ML) word order. In [15], the noun *mtu* (person) has two modifiers, *m-moja* (one) from Swahili, and *stingy* from English; however, they both are in accordance with the Swahili noun-modifier word order. Also, the adverb *sana* follows *stingy* which it modifies as expected in Swahili. Such examples provide evidence that it is the ML which controls the morphosyntactic frame into which the EL items are switched.

The minimalist approach sees CS simply as the consequence of two lexicons in contact in a mixed utterance. One obvious limitation of such an approach is that it attempts to account for switching

between monolingual constituents but dismisses the theoretically challenging problem of accounting for singly occurring forms from one language switched into a constituent structured by another language. Such an approach is insufficient to account for both singly occurring forms and full constituents switched from another language.

The Matrix Language Frame Model

The Matrix Language Frame (MLF) Model (Myers-Scotton, 1993 [1997], 2002) explains constraints on CS at an abstract, pre-syntactic level in sentence production. It makes two crucial interrelated asymmetries as structuring CS utterances: the Matrix Language (ML) (like "host language") vs. the Embedded Language (like "guest language") and content vs. system morphemes. It proposes that one of the languages involved in CS provides the sentential frame (i.e., the grammatical structure), and this language is identified as the ML, and the other language is identified as the EL, which only provides certain content morphemes to be inserted into the slots of the ML frame. Bilinguals can activate any language they know as the ML or choose any language they know as the EL, but the distinctive roles of the participating languages in CS should be clear. The MLF Model further proposes that morphemes are activated at different levels. Content morphemes (i.e., lexical morphemes) are activated at the conceptual level, but system morphemes (i.e., grammatical morphemes) are activated at the grammatical level. In CS, the EL only provides content morphemes or EL islands. The most important premise of the MLF Model is its *Matrix Language Hypothesis* and the two asymmetries indicated as two principles governing CS utterances (Myers-Scotton, 1993, pp. 6-7):

The Matrix Language Hypothesis: The ML provides the morphosyntactic frame of ML + EL constituents.

The Morpheme Order Principle: In ML + EL constituents, the surface morpheme order must not violate that of the ML.

The System Morpheme Principle: In ML + EL constituents, all system morphemes having grammatical relations external to their head constituents must come from the ML.

The key assumption underlying the MLF Model is that the languages participating in CS play unequal roles. One language is more central than the other in sentential frame-building. The ML is more activated than the EL in CS and the occurrence of its morphemes, either system or content morphemes, is more frequent and freer than that of the EL. Below are a few examples illustrating such an assumption.

[16] **Mailbox** li you nide xin.

    mailbox PREP/in exist your letter

    "There's a letter for you in the mailbox."

    (Chinese/English; Wei, 2015, p. 55)

[17] anata wa **registration** o shimashi-ta ka?

you TOP registration OBJ do-PERF PARTIC/QUE

"Have you done your registration?"

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

In [16], the English Noun Phrase (NP) *mailbox* is switched into the Chinese sentential frame for the Prepositional Phrase (PP), where the switched NP occurs before the Chinese preposition *li* (in) rather than after it as in English. The whole sentence is grammatically framed in Chinese as the ML. In [17], the English NP *registration* is switched into the Japanese predicate Verb Phrase (VP), where the switched NP occurs before the Japanese verb, following the Japanese Subject-Object-Verb (SOV) order. Thus, *the Morpheme Order Principle* is observed. The whole sentence is grammatically framed in Japanese as the ML. In both [16] and [17], the switched items are content morphemes from the ELs. The ML vs. EL distinction and the content vs. system morpheme distinction are further illustrated in the following examples.

[18] Hau-ku-on-a a-ki-ni-**buy**-i-a **beer** siku hi-yo?

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)

[19] 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 finish

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

(Chinese/English; Wei, 2018, p. 85)

[20] 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, 2006, p 164)

In [18], the switched items *buy* and *beer* are content morphemes from English, the EL, but it is Swahili, the ML, which provides the sentential frame. In this example, *buy* takes two objects, but the sentential frame of Swahili determines how those objects are realized; the beneficiary is realized as an object prefix on the verb *-ni-* and further mapped onto the sentential frame through the applied verbal suffix *-i-* on the EL verb from English. In [19], the Adjective Phrase (AP) *hen busy* (very busy) is a mixed constituent, where the adjective *busy*, a content morpheme, is from English, the EL, and the degree

24

adverb *hen* (very), a system morpheme, from Chinese, the ML. The NP *san-fen paper* (three papers) is a mixed constituent, where the system morpheme *san* (three) and *fen* (noun classifier) are from Chinese, the ML, and the system morpheme for plural marking does not appear. The whole utterance is grammatically framed by the ML, where the object NP *san-fen paper* occurs before the PP *zai yue-di qian* (before the end of the month) and the verb *finish*. In [20], *stay*, a content morpheme from English, the EL, is switched into the Japanese *suru* (do) verbal construction, where the content verb is introduced in conjunction with *suru*.

The above examples illustrate how the ML vs. EL hierarchy distinguishes the differential roles of the languages participating in CS. The ML plays a dominant role in grammatically framing the utterances containing switched items and provides any number of content morphemes and all system morphemes, and the EL switches only content morphemes into the phrasal and sentential slots as provided by the ML.

In addition to the content vs. system morpheme distinction, the MLF Model further proposes that EL islands can be switched but only with EL content morphemes. "EL islands are inherently constrained: because they must show internal structural-dependency relations, all islands must be composed of at least two lexemes/morphemes in a hierarchical relationship" (Myers-Scotton, 1993, p. 138). Thus, EL islands are constituents consisting only EL system morphemes and content morphemes and are well formed according to the EL grammar.

[21] ta dui xueshen **very strict**.

    she to student very strict

    "She's very strict with students."

    (Chinese/English; Wei, 2015, 54)

[22] It's **totemo muzukashi** to find a convenient and **yasuyi** 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, 2006, p. 167)

[23] **The first one** que era elquellevaba 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, Myers-Scotton, & Gross, 2002, p. 81)

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

   exactly then ge.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)

In [21], *very strict* is an EL island where both the content morpheme *strict* and the system morpheme *very* are switched from English. As predicted, the degree word *very*, as a system morpheme, alone cannot appear by itself. In [22], *totemo muzukashi* (very difficult) is an EL island where both the content morpheme *muzukashi* (difficult) and the system morpheme *totemo* (very) are switched from Japanese. The degree word t*otemo*, as a system morpheme, alone cannot be switched. In [23], *the first one* is a EL island where both the content morphemes *first one* and the system morpheme *the* are switched from English. The determiner *the*, as a system morpheme, alone cannot be switched. In [24], *le pentole* is an EL island where both the content morpheme *pentole* (pans) and the system morpheme *le* are switched from Italian. The determiner *le*, as a system morpheme, alone cannot be switched.

These examples show that EL system morphemes may be switched with EL content morphemes in EL islands (i.e., EL lexical units), but they cannot appear without EL lexical heads. These examples also show that if EL islands are switched, like EL content morphemes, they must be inserted into the sentential frame provided by the ML.

The Bilingual Lemma Activation Model

The Bilingual Lemma Activation (BLA) Model (Wei, 2002, 2006, 2015) employs the notion of the mental lexicon and makes a distinction between the monolingual mental lexicon and the bilingual mental lexicon. The mental lexicon not only contains vocabulary items or lexemes but also more abstract information about particular words (Richards, 1976; Færch & Kasper, 1984; Talmy, 1985; Ringbom, 1987; Nation, 1990). In speech production, speakers retrieve/active the appropriate words stored in the mental lexicon to correctly express their intended meanings. The retrieved/activated word from the mental lexicon contains not only its semantic content but also information about its phonological structure, morphological structure, syntactic environment, pragmatic function, register, etc. Many scholars (e.g., Kempen & Huijbers, 1983; Kempen & Hoenkamp, 1987; Levelt, 1989; Myers-Scotton & Jake, 1995; Levelt, Roelofs, & Meyer, 1999; Wei, 2006) define such abstract information about a particular word or lexeme as "lemma information" ("lemmas" for short). Such information is rather abstract in the sense that lemmas contain abstract entries about a particular word or lexeme stored in the mental lexicon. For example, the lemmas of *give* require a subject that carries the thematic role of AGENT (i.e., the person who performs the act of giving), an indirect object that carries the thematic role of RECIPIENT (i.e., the person who receives what is given), and a direct object that carries the thematic role of THEME (i.e., the thing that is given), and the permissible word orders in which these NP arguments may appear. The lemmas of *he* require this pronoun to be used of a

male subject and the main (i.e., lexical) verb must be inflected by -s for the subject-verb agreement if the verb is in the present tense. 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).

The BLA Model draws on the insights in some linguistic and psycholinguistic studies of the bilingual lexical and conceptual representation and expands Levelt's model of monolingual speech production by explaining and emphasizing the role of bilingual lemma activation. The crucial assumption underlying the BLA Model is that the bilingual mental lexicon differs from the monolingual mental lexicon. As assumed, the lemmas in the bilingual mental lexicon are language-specific. This is because the monolingual "mental lexicon represents a complex self-organizing system", but 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, 2002, p. 1040). The BLA Model claims that CS is a natural outcome of bilingual lemmas in contact during bilingual speech production.

Moving away from most of the other models of CS as introduced above and drawing on the MLF Model, this study adopts the BLA Model to introduce a lexical-conceptual approach to CS.

3. Lexical-Conceptual Approach to Codeswitching

The BLA Model recognizes the distinction between content and system morphemes as one of the asymmetries as identified in the MLF Model but aims to explain why EL content morphemes, rather than EL system morphemes, are activated in CS. As generally assumed, a universal set of semantic/pragmatic feature bundles is available for the lexical-conceptual structuring of lemmas, but, as commonly observed, there is cross-linguistic variation in the presence and conflation of these feature bundles. That is, languages may differ in lemmas of certain lexemes. In addition to the structural constraints on CS as proposed in the various models, this study proposes that one of the major reasons for bilinguals to switch to some EL content morphemes is that at the lexical-conceptual level, certain EL content morphemes may encode their communicative intentions or intended meanings more accurately or desirably than the similar content morphemes available in their ML. This is because, at the conceptual level, bilinguals make appropriate lexical choices about the semantic/pragmatic feature bundles that they intend to convey. As mentioned earlier, bilingual lemmas of lexical items or lexemes are language specific. For this reason, bilingual lemmas are activated for language-specific lexical-conceptual structure to be lexically realized in bilingual communication, especially in CS.

Lexical structure is abstract in the sense that lemma entries in the mental lexicon are abstract. Thus, the notion of abstract lexical structure is similar to the matching principle that a Semantic Form (SF) triggers a particular lemma "if and only if there exists a complete match of all structures in the SF (i)

27

with all structures in the semantic representation of the lemma" (Bierwisch & Schreuder, 1992, p. 51). Lexical-conceptual structure is also abstract because the election of a particular lemma is based on whether it has all the primitives contained in the conceptual chunk to be lexicalized. According to de Bot and Schreuder (1993), because different languages may lexicalize in different ways, the language to be used in second language production must be specified before conceptual chunking takes place. Thus, before CS takes place, language-specific lemmas of lexical items must be triggered by the information concerning language choice in the speaker's preverbal message (i.e., the speaker's communicative intention). Thus, it is the cross-linguistic differences in semantic/pragmatic feature bundles encoded in abstract lexical-conceptual structure that trigger switches to certain EL lexical items during CS. According to Roelofs (1992), Levelt (1989, 1995), Myers-Scotton and Jake (1995), and Wei (2001a, 2001b), it is the activated concepts in the speaker's preverbal message that activate the corresponding lemmas, and it is the sufficiently activated lemmas that activate the associated lexemes.

Along the above lines of thinking, the BLA Model proposes that it is cross-linguistic differences in lexical-conceptual structure or semantic/pragmatic feature bundles that motivate language-specific lemma variation as conveyed in CS. This study regards the switched items as evidence of the relative importance of cross-linguistic lexical-conceptual differences in the lemmas of certain lexemes in the bilingual mental lexicon.

As commonly observed, CS becomes part of bilinguals' natural speech patterns. Bilinguals may switch to another language either intersententially or intrasententially during a discourse. If they switch to another language intersententially, only one of their languages is activated, and their other language is deactivated. If they switch to another language intrasententially, both of their languages are activated. Of course, if the latter happens, only one of their languages is activated as the ML, and the other is activated as the EL. It is the ML which provides the sentential frame and most morphemes, including both content and system morphemes. The EL only inserts certain content morphemes into the ML frame (Myers-Scotton, 1993; Myers-Scotton & Jake, 1995; Wei, 2000a, 2000b, 2002).

As reported by Grosjean (1982), Li (1996), Nishimura (1997) and Wei (2001b, 2002), some codeswitches may be caused by the lack of particular words in one of the languages or by the greater availability of words in the other language which speakers need to convey their intended meanings. That is why bilinguals may practice CS at a certain point during a discourse deliberately or intentionally to satisfy their lexical needs. The BLA Model attempts to describe and explain such a bilingual speech behavior at a rather abstract level by proposing that it is cross-linguistic differences in lemmas at the lexical-conceptual level which drive CS. It emphasizes that a partial lemma difference at the level of lexical-conceptual structure is one of the major motivations for certain EL content morphemes to be switched.

As introduced earlier, one of the two asymmetries as identified in the MLF Model is the distinction between content and system morphemes. Many naturally occurring CS instances indicate that

individual EL content morphemes, rather than EL system morphemes, can be switched to encode speakers' intended meanings. The BLA Model further proposes that it is at the conceptual level, rather than at the surface speech production level, that speakers make choices about the lemma specifications that satisfy their communicative interests.

Drawing on the MLF Model and the BLA Model, this study introduces a lexical-conceptual approach to CS. Beyond the commonly recognized structural constraints on CS as reviewed earlier, this approach suggests that one of the major reasons for certain EL content morphemes to be switched is that only content morphemes contain lexical content or, to be more specific, semantic/pragmatic feature bundles. Speakers may prefer to use certain EL content morphemes because they are aware of the cross-linguistic differences and the EL lemma specifications for these morphemes. Some typical naturally occurring examples of CS are selected for demonstrating how switched items are lexically and conceptually projected.

[25] wo you liang-fen **paper** mingtian bixu jiaoshangqu.

    I have two-CL paper tomorrow must turn in

    "I must turn in two papers tomorrow."

[26] wo xiawu qu jian wode **advisor**.

    I afternoon go see my advisor

    "I'm going to my advisor this afternoon."

[27] naxie **visiting scholar** bushi hen youqian ma, bi women **student** youqian duo le.

    those visiting scholar not/EMPH COP very rich AFFIRM PREP/than us student rich more AFFIRM

    "Aren't those visiting scholars very rich? They are much richer than us students."

[28] zhuzai zheli hen fanbian, meitian you **school bus**.

    live PREP/LOC here very convenient every day have school bus

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

    (Chinese/English; Wei, 2006, pp. 163-169)

In [25], the speaker switches to *paper*, an EL content morpheme, but *liang-fen* is a combined ML system morpheme consisting of *liang* (two) and the noun classifier *fen*. In English, a "paper" may mean any written piece of work, such as an essay, an article or a composition, but the Chinese equivalent word "zhi" (paper) itself only means a piece of paper for wrapping things up in or writing something on. It seems obvious that the speaker switches to *paper* for the English concept. In [26], the speaker

switches to *advisor*, an EL content morpheme, but the determiner *wode* (my) is a system morpheme from the ML. An "advisor" in the English academic setting means a professor or instructor who offers advice to students regarding their academic progress, improvement, course requirements and sequence, thesis or dissertation writing, research in progress, and so forth. In contrast, a Chinese "daoshi" (advisor) does not assume the same responsibilities as an English academic advisor. Chinese advisors' only or major responsibility is to supervise their students in wiring their theses or dissertations. The speaker switches to *advisor* for the intended meaning. In [27], *visiting scholars* and *student* are the EL content morphemes, but the demonstrative pronoun *naxie* (those) is an ML system morpheme, and there is no EL plural marking morpheme "-s" on *visiting scholar* or *student*. The speaker switches to *visiting scholar* probably because the English concept of "visiting scholar" is rather new to Chinese. One of the possible reasons for the speaker to switch to *student* in the same sentence is to realize the conceptual difference between "scholar" and "student". In [28], the speaker switches to the EL compound content morpheme *school bus*, but the EL system morpheme "a" is missing. A "school bus" in English mainly means a bus that transports students to and from a school. Some Chinese universities and schools in major cities may also have "xiaoche" (school bus), but a Chinese "xiaoche" means a bus that mainly transports a university's or school's sports or performance team, equipment or some faculty and staff members who need a ride from one campus to another. The speaker switches to *school bus* most probably because of the conceptual difference between "school bus" and "xiaoche".

[29] ima wa **summer course** o tot-teiru n.

now TOP summer course OBJ take-PROG AUX/be PART

"I'm taking a summer course now."

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

[30] moshi Nihon ga soo iu **community force** mitaina no ga naku Nihon mo **America** mitai ni

nacchau no ja nai ka?

If Japan PART/NOM so say community force like PART/NOM PART/NOM no become PERF if

Japan also America same PREP/COND become PART/NOM COP/be not PART/INTERROG

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

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

you TOP registration OBJ do-PERT QUE

"Have you done your registration?"

(Japanese/English; Wei, 2002, pp. 282-283)

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

two POSS bedroom PART/NOM COP one person Maria call PART/TOP one person PREP/by one bedroom PART/OBJ have-PROGAUX PART/AFFIRM

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

(Japanese/English: Wei, 2006, p. 170)

In [29], *summer course* is an EL compound content morpheme, and *o* is an ML system morpheme marking the accusative case. The speaker switches to *summer course* for the English academic concept. Such a concept is rather new to Japanese students. Most probably, the speaker selects this EL compound content morpheme for the conceptual reason. In [30], the general semantic/pragmatic feature bundles of *community force* in the American social context may include "neighborhood crime watch", "drug free zone", "organized community activities" and so on. The Japanese expression like *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 forth. The speaker switches to *community force* probably to convey something more accurately beyond what "chouka" can lexical-conceptually realize. In [31], the speaker switches to *registration* for the probable reason that in Japanese colleges and universities, though students also register for the courses, they are not free to select the courses which they are truly interested in taking. The equivalent word in Japanese is xxx. The general concept of "registration" may be shared between English and Japanese, but the semantic/pragmatic feature bundles of "registration" and xxx are not the same. In [32], *bedroom* and *one bedroom* are the EL content morphemes. For *futatsu no bedroom* (two bedrooms), the EL system morpheme "-s" for plural marking does not appear, and *ga* is an ML system morpheme assigning the normative case to *futatsu no bedroom*. It should be noticed that the numeral *one*, an EL system morpheme, is activated together with the EL content morpheme *bedroom*. The speaker switches to *bedroom* because in Japan the concept of "bedroom" is relatively new or unpopular. A traditional Japanese room is used not only for sleeping but also for eating, studying, playing, meeting guests or other daily activities. In other words, in many traditional Japanese homes, a single room can be multifunctional, but in most American homes, a bedroom is a particular type of room designated for sleeping.

The above examples show that bilinguals may switch to certain EL content morphemes in CS. They do so not simply because of the content morphemes themselves but because of cross-linguistic differences in abstract lexical-conceptual structure. To be more specific, it is cross-linguistic differences in language-specific lemmas or semantic/pragmatic feature bundles of certain content morphemes which motivate CS. This is because language cues may have different values and are in competition in second language learning and production (Bates & MacWhinney, 1982; MacWhinney, 1987, 1989), and bilingual speakers may switch to certain EL content morphemes to convey their intended meanings (Li,

31

1996; Nishimura, 1997; Wei, 2001a, 2001b, 2002). That is, when the language cue specifies a particular language at a certain point during a discourse involving CS, the lexical item from that language receives activation. Thus, conceptual information and language cues must work together in activating language-specific lemmas in the bilingual mental lexicon.

The above examples also show that in CS all the system morphemes are from the ML, and all the content morphemes switched from the EL are inserted into the ML frame. This provides evidence in support of *the Matrix Language Hypothesis*, *the Morpheme Order Principe*, and *the System Morpheme Principle* (Myers-Scotton, 1993 [1997]).

Below are some representative examples selected from various language pairs to further demonstrate instances of language-specific lemma activation in CS.

[33] 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)

[34] Se sai semmose-n **stroke**-Ø.

she get-IMP3SG like-ACC stroke

"She had like a stroke."

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

[35] 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)

[36] nei5 zou6 saai3 di a**ssignment** mei6.

you do ASP CL assignment SFP

"Have you done all the assignments?"

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

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

yes caf éLOC sit-PROG.3PL

"Yes, they are sitting at the outdoor caf é"

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

[38] 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)

[39] Hau-ku-on-a a-ki-ni-**buy**-i-a beer siku hi-yo?

    2s/NEG-NEG/PAST-see-FV 3s-PROG-1s/OBJ-buy-APPL-FV beer day CL9-DEM

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

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

The examples in [33]-[39] show again all the switched items are the EL content morphemes selected by the speakers as lexical-conceptually appropriate to their intended meanings. Although it is difficult to tell what motives speakers for switching to certain EL content morphemes or only speakers themselves know why certain EL content morphemes are more desirable for their communicative intentions, what becomes clear is that bilingual lemmas are in contact and are not equally activated in CS. EL content morphemes can be lexical-conceptually activated for speakers' intended or preferred meanings at a certain point during a discourse, but EL system morphemes cannot.

The BLA Model assumes that it is the preverbal message at the conceptual level that motivates bilinguals to make appropriate choices of the semantic/pragmatic feature bundles that they intend to convey. It further assumes that if the bilingual mode is chosen at the conceptual level, but the lemmas activated from the EL do not sufficiently match the ML counterparts, but the speaker still prefers CS, some compromise strategies must be taken to overcome cross-linguistic differences in lexicalization patterns or predicate-argument structures.

4. Compromise Strategies for Producing EL Lexical-Conceptual Structure Islands

As commonly observed, different languages may lexicalize concepts in different ways. Thus, when EL lemmas are activated in CS, but they do not match their counterparts in the ML, speakers must take some compromise strategies for CS to occur. One of the compromise strategies is to produce "EL islands" (Myers-Scotton, 1993 [1997]; Jake & Myers-Scotton, 1997; Wei, 2001b). An EL island is defined as a constituent in which an EL content morpheme occurs with only other EL morphemes, including EL system morphemes. As assumed in the BLA Model, at the conceptual level bilinguals do not produce surface morphemes but make appropriate choices about the semantic/pragmatic information that they intend to convey (Wei, 2002). If certain EL lemmas are activated at the conceptual level, but they do not match the ML counterparts, speakers must produce EL islands in CS. Such a compromise strategy becomes necessary to overcome cross-linguistic differences in lexical-conceptual structure for possible CS realization. Below are some examples of EL islands

produced in CS.

[40] nali you wu-tai jiqi dan san-tai si **out of order**.

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

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

[41] ni neng-bu-neng **give me a ride**.

    you can-not-can give me a ride

    "Can you give me a ride?"

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

    then you tomorrow call me

    "Then, you call me tomorrow."

[43] 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, pp. 160-163)

In [40], *out of order* is an EL island. Instead of the Chinese equivalent expression "chu guzhang" (something going wrong), the speaker prefers the lexical-conceptual structure of *out of order*, which is activated as a single lexical-conceptual unit (i.e., a fixed idiomatic expression). In [41], *give me a ride* is an EL island, which is conceptually incongruent with the ML counterpart "song wo yixia". While in the EL, the means of transportation is conflated in the noun *ride* as the direct object of the verb *give*, in the ML, it is conflated in the verb "song" (send), but "song" itself may not contain the means of transportation at all. In other words, "song wo yixia" (send me once) does not necessarily means "give me a ride". The speaker prefers this EL island most probably to convey his intended meaning more clearly and specifically than he can be with the Chinese counterpart. In [42], *call me* is an EL island, where the lexical-conceptual features of "communicate with by telephone" are conflated in the verb *call*. Though the ML has the equivalent expression "da dainghua gei wo" (make telephone to me), the means of communication is expressed in the noun "telephone" itself. The speaker switches to the EL expression in favor of the lemmas underlying its lexical-conceptual structure. In [43], *come to pick you up* is an EL island. The speaker prefers the phrasal verb *pick up* for the probable reason that it contains the meaning of "to take on as a passenger". Though the verb "jie" (meet) in the ML may express the same meaning, but "jie" (e.g., to meet someone at a bus/train station or airport) 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, where the infinitive marker *to*, a system morpheme, is from the EL. The speaker prefers the EL phrasal verb *pick up* to express his intended meaning more accurately.

The above examples of CS reveal that bilinguals may switch to lexical items from an EL, whether single EL lexical items or EL islands, most probably because of cross-linguistic differences in language-specific lemmas underlying the abstract lexical-conceptual structure of particular lexemes. The BLA Model assumes that bilinguals may switch to certain EL content morphemes and EL islands to realize their semantic/pragmatic intentions (Talmy, 1985; Li, 1996; Nishimura, 1997; Wei, 2001b, 2002). The BLA Model aims to describe and explain CS by assuming that lemmas are tagged with a language label (Green, 1986; de Bot & Schreuder, 1993; Poulisse & Bongaerts, 1994) and in CS, speakers switch to EL content morphemes and EL islands either because of lemma differences or because of the lack of the lexical items in the ML for particular semantic/pragmatic feature bundles as desired and preferred by speakers at a certain point during a discourse.

5. Compromise Strategies for Producing EL Predicate-Argument Structure Islands

In addition to the compromise strategies for producing EL lexical-conceptual structure islands, some compromise strategies for producing EL predicate-argument structure islands may come into play for possible CS configurations. As specified in the MLF Model, it is the ML which provides the sentential frame and controls the predicate-argument structure by supplying subcategorization frames for verbs and surface morpheme order. The BLA Model claims that when a certain EL verb is activated at the level of lexical-conceptual structure, but it generates a particular EL predicate-argument structure (i.e., the second level of abstract lexical structure), speakers may switch to the EL predicate-argument structure if they choose a particular EL verb as appropriate or desired. This is because "Language specific lemmas form the interconnection between the lexical-conceptual mappings to and from syntax" (Kroll & de Groot, 1997, p. 190). If the language pairs involved in CS do not share the same predicate-argument structure, speakers have two speech plans available to them. They may stop the encoding of one of them and continue with the other to solve the typological problem in CS. In CS, the choice of one EL predicate-argument structure versus another is determined by the larger ML sentential frame.

As observed in some instances of CS, a particular predicate-argument structure across the two languages differs, but the semantic/pragmatic feature bundles contained in the lexical-conceptual structure of the EL are activated by speakers for their communicative intentions. If this happens, the EL predicate-argument structure may also be activated as grammatically necessary for CS, resulting in EL predicate-argument structure islands, in which all the morphemes contained in the EL predicate-argument structure are from the EL, including the EL system morphemes. Below are some examples of EL predicate-argument structure islands.

[44] wo keyi **wait for you** dao liangdian.

　　I can wait for you till two o'clock

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

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

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

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

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

[46] ni biye how 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)

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

　　I every day have to help her with her homework

　　"I have to help her with her homework every day."

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

In [44], *wait for* is activated as a single lexical unit, where the direct object *you* is the THEME introduced by the preposition *for*. In Chinese, the same meaning is expressed by a single verb "deng" (wait) without the requirement of a preposition. The activation of an EL phrasal verb like this results in the maximal projection of an EL phrasal category, that is, an EL island. In [45], the verb phrase headed by *fail* is an EL island. In English, "fail" can be used as a causative verb and takes the grammatical subject as the AGENT (i.e., the causer) who makes the failure happen. Chinese has the equivalent verb "shibai" (be defeated in …), but it is used only as a nonaccusative verb with the grammatical subject as the EXPERIENCER. The speaker switches to the EL predicate-argument structure initiated by "fail" because of the incongruence between the English "fail" and the Chinese "shibai". In [46], *teach English to nonnative speakers* is an EL island, where the noun phrase, *nonnative speakers*, is the RECIPINET introduced in the prepositional phrase headed by *to*, which is the English indirect object dative construction. Chinese has the equivalent verb "jiao" (teach), but it is used only in the double object construction (e.g., jiao nonnative speakers English). The speaker is aware of the incongruence between the EL predicate-argument structure and that of the ML but switches to the EL one initiated by *teach*, resulting in an EL island. In [47], *help her with her homework* is an EL island, where *her homework* is the THEME introduced in the prepositional phrase headed by *with*. In Chinese, the THEME must be introduced directly by a specific verb such as "zuo" (do) in addition to a main verb

36

such as "bangzhu" (help) (e.g., we meitian de bangxhu ta zuo zuoye (I every day have to help her do homework.)). The speaker switches to the EL predicate-argument structure driven by "help", resulting in an EL island.

Such examples of EL predicate-argument structure islands reveal that when incongruence or insufficient congruence between the participating languages in CS at the level of predicate-argument structure occurs, but speakers prefer the predicate-argument structure to express their intended meanings, some compromise strategies for producing EL islands must be taken for CS configurations. It should be emphasized that every EL island is also lexical-conceptually driven, and every EL island must be switched into the ML frame.

6. Conclusion

Departing from the surface-based models of CS and drawing on the MLF Model and the BLA Model, this study proposes a lexical-conceptual approach to CS at an abstract level. One of the major theoretical assumptions underlying this approach is that though lexemes in the mental lexicon are universal, lemmas in the bilingual mental lexicon are language specific, and language-specific lemmas are in contact at any level of abstract lexical structure during a discourse involving CS. Under this assumption, this study claims that bilingual lemma activation is a crucial interface between bilinguals' communicative intention and CS. More importantly, it further claims that CS is lexical-conceptually driven and constrained by a set of asymmetries or hierarchies: ML vs. EL, and content vs. system morphemes. Many commonly observed naturally occurring examples of CS involving various language pairs provide evidence in support of this lexical-conceptually based approach to CS. This study reaches the following conclusions regarding the nature of the bilingual mental lexicon and its activity in CS.

1) During a discourse involving CS, bilinguals' both languages are "turned on", but only one of them must be activated as the ML, which is more strongly activated than the EL. Whichever language activated as the ML provides the sentential frame to structure the grammatical configuration of every CS utterance. The ML provides most content morphemes and all system morphemes. The EL plays an unequal role by supplying certain content morphemes to be switched into the slots in the ML frame. That is, the ML vs. EL hierarchy and the content vs. system morpheme hierarchy constrain CS configurations.

2) The bilingual mental lexicon contains language-specific lemmas in contact during CS. This is because though languages share universal lexemes, lemma information about certain lexemes may differ in different languages. The bilinguals engaged in CS may activate language-specific lemmas as desired for bilingual communication.

3) One of the major reasons for bilinguals to switch to certain EL content morphemes is that the switched EL content morphemes are lexical-conceptually motivated for their intended meanings. The purpose of switching to EL content morphemes is to "fill a linguistic need" or to use the word most

available in the other language (Grosjean, 1982, p. 151). This is because bilinguals are aware of the cross-linguistic lemma differences (i.e., cross-linguistic lexical-conceptual differences or gaps) between the languages involved in CS. It is their communicative intention contained in the preverbal message which calls for certain EL content morphemes to solve cross-linguistic lexical-conceptual problem. Thus, CS can be regarded as a linguistic solution to bilingual communication.

4) Bilinguals can activate language-specific lemmas from whichever language as the EL during CS, but the activated lemmas must be sufficiently congruent with the counterparts of the ML at the lexical-conceptual structure and predicate-argument structure levels of abstract lexical structure or some combination of these levels. If lemma incongruence or insufficient congruence occurs between the language pairs, but speakers do not want to give up CS, some radical compromise strategies must be taken for possible CS to occur. One of the compromise strategies is to produce EL lexical-conceptual structure islands, and the other is to produce EL predicate-argument structure islands. All EL islands must be headed by EL content morphemes. Thus, all EL islands are lexical-conceptual driven.

References

Apple, 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.

Bates, E., & MacWhinney, B. (1982). Functional approaches to grammar. In E. Wanner, & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 157-193). New York: Cambridge University Press.

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

Bierwisch, M., & Schreuder, R. (1992). From concepts to lexical items. *Cognition*, *41*, 23-60.

Boeschoten, R., & Huybregts, H. (1999). *A minimalist view of bilingual grammar: Code-switching can do without switching codes*. Paper presented at International Symposium on Bilingualism, Newcastle, 4/99.

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.

Chomsky, N. (1981). Principles and parameters in syntactic theory. In N. Hornstein, & D. Lightfoot (Eds.), *Explanations in linguistics: The logical problem of language acquisition* (pp. 32-75).

London: Longman.

Chomsky, N. (1995). *The minimalist program*. 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.

DiSciullo, A., Muysken, P., & Singh, R. (1986). Government and code-switching. *Journal of Linguistics*, *22*, 1-24.

Færch, C., & Kasper, G. (1984). Two ways of defining communication strategies. *Language Learning*, *34*, 45-63.

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.

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

Gumperz, J. J. (1982). Conversational code-switching. In J. J. Gumperz (Ed.), *Discourse strategies* (pp. 55-99). Cambridge University Press.

Halmari, H. (1997). *Government and codeswitching: Exploring American Finnish*. Amsterdam: John Benjamins.

Jake, J. L., & Myers-Scotton, C. (1997). Codeswitching and compromise strategies: Implications for lexical structure. *International Journal of Bilingualism*, *1*, 25-39.

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.

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.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, *11*, 201-258.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, *14*, 185-209.

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.

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

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.

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.

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.

MacWhinney, B. (1987). Applying the competition model to bilingualism. *Applied Psycholinguistics*, *8*(4), 315-327.

MacWhinney, B. (1989). Competition and lexical categorization. In R. Corrigan, F. R. Eckman, & M. Noonan (Eds.), *Linguistic categorization* (195-242). New York: John Benjamins.

MacSwan, J. (1997). *A minimalist approach to intrasentential code switching: Spanish-Nahuatl bilingualism in Central Mexico*. Unpublished doctoral dissertation, UCLA.

MacSwan, J. (1999a). The argument status of NPs in Southeast Nahuatl: Comments on the polysynthesis parameter. *Southwest Journal of Linguistics*, *17*(2), 101-114.

MacSwan, J. (1999b). *A minimalist approach to intrasentential code switching*. New York: Garland.

MacSwan, J. (2000). The architecture of the bilingual language faculty: Evidence from intrasentential code switching. *Bilingualism, Language and Cognition*, *3*, 37-54.

MacSwan, J. (2005). Précis of a minimalist approach to intrasentential code switching. *Rivista di Linguistica*, *17*(1), 55-92

Muysken, P. (1988). *A unified theory of grammar contact*. Paper presented at Contact and Conflict Conference. Brussels.

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. (1988). *Swahili/English Nairobi corpus*.

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.

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.

Nartey, J. (1982). Code-switching, interference or faddism? Language use among educated Ghanaians. *Anthropological Linguistics*, *24*(2), 183-192.

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.

Pandit, I. (1990). Grammaticality in code-switching. In R. Jacoson (Ed.), *Codeswitching as a worldwide phenomenon* (pp. 33-69). New York: Peter Lang.

Peterson, J. (1988). Word-internal code-switching constraints in a child's grammar. *Linguistics*, *25*, 479-493.

Pfaff, C. (1979). Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language*, *55*, 291-318.

Poplack, S. (1980). "Sometimes, I'll start a sentence in Spanish termino en espa ñol": Toward a typology of code-switching. *Linguistics*, *18*, 581-618.

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., & 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.

Preziosa-Di Quinzio, I. (1992). *Teoreticamentelafirma la indietro: Frammistione di Italiano e Schwyzert ütschnella conversazione one di figli di emigrati*. Universit à di Zurigo, Facolt à de Lettere e di Filosofia: Vavoro di licenza in linguisticaItaliana.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, *10*, 77-89.

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Clevedon: Multilingual Matters.

Ritchie, W., & Bhatia, Y. (1999). Codeswitching, grammar, and sentence production: The problem of light verbs. In E. C. Klein, & G. Martoharjono (Eds.), *The development of second language grammar* (pp. 269-287). Amsterdam: Benjamins.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107-142.

Romaine, S. (1989). *Bilingualism*. Oxford: Blackwell.

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.

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.

Toribo, A., & Rubin, E. (1996). Code-switching in generative grammar. In A. Roca, & J. Jensen (Eds.), *Spanish in contact: Issues in bilingualism* (pp. 203-226). Somerville, MA: Cascadilla.

Wei, L. (2000a). Unequal election of morphemes in adult second language acquisition. *Applied Linguistics*, *2*, 106-140.

Wei, L. (2000b). Types of morphemes and their implications for second language morpheme acquisition. *International Journal of bilingualism*, *4*(1), 29-43.

Wei, L. (2001a). The multilingual mental lexicon: Language separation/activation in trilinguals. *Papers Selected from the Second International Conference on Third Language Acquisition and Trilingualism: Interactive CD-ROM*, 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.

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. *The Proceedings of the 4th International Symposium 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. (2009). 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.

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

Wei, L. (2018). The bilingual lemma activation model as a comparative approach to codeswitching. *Studies in Linguistics and Literature*, *2*(2), 77-109.

Wei, L. (2020). *The bilingual mental Lexicon*. UK: Cambridge Scholars Publishing.

Wei, L. (2024). *Language contact and linguistic aspects of bilingualism*. UK: Cambridge Scholars Publishing.

Woodford, E. (1983). Bilingual code-switching and syntactic theory. *Linguistic Inquiry*, *14*, 520-536.