A Lexical-conceptual Approach to Code Switching

Longxing Wei
Montclair State University
1 Normal Avenue, Montclair
New Jersey 07043, USA

Abstract

This study claims that the bilingual mental lexicon contains 'lemmas', defined as abstract entries in the mental lexicon about particular lexemes, and such lemmas are in contact in ICS. Lemmas contain information about lexical-conceptual structure, predicate-argument structure and morphological realization patterns of particular lexemes. It further claims that bilingual lemmas are language-specific, and it is the activation of language-specific lemmas which drives ICS. The ICS instances for the study reveal that bilinguals switch content morphemes from one language into another language's sentential frame because of cross-linguistic lexical-conceptual gaps. Bilinguals do not switch system morphemes because it is the Matrix Language (the host language) which provides the sentential frame, including all system morphemes, and the Embedded Language (the guest language) inserts certain content morphemes as desired into the lexical slots in the ML. The lexical-conceptual approach introduced in this study explores the nature and activity of the bilingual mental lexicon during ICS.

Keywords*: bilingual, mental lexicon, lemma, lexical-conceptual, content, system, morpheme, activation*

1. Introduction

In trasentential code switching (ICS) is the alternative use of two languages within a sentence boundary. For example, *veo los horses* (I see the horses) (Apple &Muysken, 1987, p. 125). The speaker switches 'horses' from English into the Spanish sentence. ICS is recognized as a bilingual phenomenon of language contact, but it has its own linguistic manifestations. Most previous studies noticed that the two (or more) languages participating in ICS remain identical to their use in monolingual speech and tried to formulate the structural principles governing ICS in which the languages involved remain intact (e.g., Lipski, 1977; Pfaff, 1979; Poplack, 1980; Poplack & Sankoff, 1988; Woolford, 1983; DiSciullo, Muysken, & Singh, 1986; Muysken, 1988, 1991). Such studies aimed to describe how monolingual grammars cooperate to form a single grammatical structure for ICS by focusing on the surface configurations or structural linearity of utterances containing switched items.

However, researchers differ in their opinions about general properties of ICS and have proposed various models and structural constraints to regularize utterances containing switched items. Unlike surface-based models, government and binding models stress dependency rather than linearity (Klavans, 1983; Woolford, 1983; Pandit, 1990; DiSciullo, Muysken, & Singh, 1986). They apply the Government and Binding Theory (Chomsky, 1981; Lasnik & Uriagereka, 1988; Haegeman, 1991) to their explicit formulation of the non-linear approach to structural constraints on ICS. 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). But, many instances of ICS reveal that the government and binding theory alone does not adequately explain ICS. Its major problem lies in the fact that a purely syntactic approach is still too close to the surface-based analysis. Because of the inadequacy of surface level equivalences as structural constraints on ICS, other researchers propose more abstract equivalence to account for switchability (Bentahila& Davies, 1983; Muysken, 1990, 1991; Azuma, 1991). Their subcategorization models stress lexical frame restrictions as specified in syntactic theories on ICS. Bentahila and Davies claim that "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" (1983, p. 329). Azuma suggests that "the subcategorization of the main verb is always preserved' and "the main verb provides a planning frame … content word insertion must be done within the specifications of the planning frame" (1991, p. 7).Though subcategorization models depart from surface-based models by looking beyond surface linear ordering, they still consider ICS as basically a syntactic and lexical phenomenon similar to

those of monolingual surface structures. Similar to these models, some researchers propose minimalist models to account for ICS (Toribio & Rubin 1996; MacSwan, 1997, 1999a, 1999b; Boeschoten & Huybregts, 1999; Ritchie & Bhatia, 1999). Employing an important aspect the Minimalist Program (Chomsky, 1995) that all syntactic variation is associated with the lexicon, minimalist models claim that no additional rule-based constraints are necessary because ICS arises from the insertion of lexical items from another language into the phrase structure. 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. Thus, ICS is seen simply as the consequence of two lexicons in contact in a mixed utterance. One obvious limitation of the minimalist approach to ICS is that it first 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. Wei (2015) offers some detailed comments on the previous models of ICS briefly introduced here.

Departing from any previous models, the Matrix Language Frame (MLF) Model (Myers-Scotton, 1993 [1997], 2002) makes two crucial distinctions in ICS: the Matrix Language (ML) (similar to the traditional notion of 'host' language) vs. the Embedded Language (EL) (similar to the traditional notion of 'guest' language) and content vs. system morphemes. The MLF Model proposes that one of the languages involved in ICS provides the sentential frame (i.e., the grammatical structure), and this language is identified as the ML; the other language is identified as the EL, which only provides certain 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 the languages involved in ICS 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 ICS, the EL only provides content morphemes.

Adopting the MLF Model, the Bilingual Lemma Activation (BLA) Model (Wei, 2006, 2015) further proposes that 'lemmas' (i.e., abstract entries in the mental lexicon about particular lexemes) in the bilingual mental lexicon are language-specific, and such language-specific lemmas are in contact in ICS. According to the BLA Model, content morphemes are activated at the lexical-conceptual level, and it is cross-linguistic lemma variations that motivate ICS. This study claims that it is the nature of the abstract lexical-conceptual structure in the bilingual mental lexicon, rather than surface structural configurations, which motivates and constrains ICS.

2. Differential Roles of Participating Languages in ICS

Different from any other model, the MLF Model proposed by Myers-Scotton (1993 [1997], 2002) explains ICS at an abstract, pre-syntactic level in sentence production. This model makes the theoretical claims having to do with two asymmetries between the participating languages. The first asymmetry is between the participating languages regarding their differential grammatical roles in constructing ICS utterances. The second asymmetry is between the participating languages regarding their respective contribution of morphemes to ICS utterances. *TheMatrix Language Hypothesis* is the most important premise of the MLF Model, and the two asymmetries are indicated as two crucial principles governing ICS utterances (Myers-Scotton, 1993, pp. 6-7):

> *The Matrix Language Hypothesis*: The ML provides the morphosyntactic frame for 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 morpheme having grammatical relations external to their head constituents must come from the ML.

The ML vs. EL distinction is more than a heuristic device for labeling constituents but underlies the differential participation of the ML and the EL in shaping ICS utterances. It is the ML which plays the dominant role by projecting the sentential frame and providing all system morphemes (i.e., the grammatical structure of ICS, such as the morpheme order, functional items and inflectional morphemes). The key assumption underlying the MLF Model is that the languages participating in ICS have unequal roles. One language is more central than the other in sentential frame-building. The ML is more activated than the EL in ICS and the occurrence of its morphemes, either system or content morphemes, is more frequent and freer than that of the EL

Thus, the identification of one of the participating languages as the ML becomes crucial in describing and explaining ICS configurations. The BLA Model (Wei, 2006, 2015) assumes that the ML is the language which the speakers engaged in ICS commonly identify as the 'main language' being used. It claims that it is the speaker's preverbal message at the conceptual level (i.e., intention before speech) that determines which language to be used as the ML in ICS. The ML chosen at the conceptual level, together with the semantic/pragmatic feature bundles as desired (i.e., the speaker's intended meaning), activates language-specific lemmas at the lemma level. According to the BLA Model, it is the activated language-specific lemmas which send directions to the speech production formulator at the functional level, where speech is morphosyntactically encoded. The overt speech product is realized at the positional level, where speech is morphophonologically encoded by the speech articulator.

Thus, the MLF Model identifies two distinctive but interrelated hierarchies: the ML vs. EL hierarchy, which specifies the unequal roles played by the languages participating in ICS, and the system vs. content morpheme hierarchy, which specifies differential accessing of morphemes in ICS utterances. These two hierarchies predict two types of constituents containing switched items :

1.*ML + EL constituents* consisting of any number of ML morphemes and (generally) singly occurring EL content morphemes inserted into the ML sentential frame.

[1] **Mail box** li you nidexin.
    mail box PREP/in exist your letter
    'There's a letter for you in the mail box.'
    (Chinese/English; Wei 1992 [2015: 55])

[2] anatawa**registration** o shimashi-ta ka?
    you TOP registration OBJ do-PERF PARTIC/QUE
    'Have you done your registration?'
    (Japanese/English; Wei 2006, p. 170)

[3] **Idea**bubua de gale asiwoa?
          another
    'Do you have another idea?'
    (Ewe/English; Dzameshie 1989, p. 2)

In [1], the English NP*mail box* is switched into the Chinese frame for the PP, where the switched NP occurs before the Chinese PREP/li rather than after it as in English. In [2], the English NP *registration* is switched into the Japanese predicate VP, where the switched NP occurs before the Japanese verb, following the Japanese SOV order.In [3], the English NP *idea* is switched into the Ewe sentential frame, where the OBJ NP being questioned occurs in the sentence initial position and follows the Ewe modifier *bubua*(another) in accordance with the Ewe morpheme order. These examples indicate that it is the ML which provides the sentential framefor ICS.

[4] Ha-u ku-on-a a-ki-ni-**buy**-i-a **beer**sikuhiyo?
    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)

[5] wo zuijin hen **busy**, you san-fen **paper** bixu zaiyue-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 this month.'
    (Chinese/English; Wei 2018, p. 85)

[6] doreguraikokoni**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 [4], 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 on the sentential frame through the applied verbal suffix *-i-* on the EL verb from English.

3

In [5], the AP *hen busy* (very busy) is a mixedconstituent, where the adjective *busy* is from English, the EL, and the degree word *very*, a system morpheme, from Chinese, the ML.The NP *san-fen paper* (three papers) is a mixed constituent, where the system morphemes *san* (three) and *fen* (noun classifier) are from Chinese, the ML, and the content morpheme *paper* is from English, the EL. It should be noticed that in this mixed NP the English '-s', a system morpheme, for plural marking does not appear. It is obvious that the whole utterance is grammatically framed by the ML, where the object NP *san-fen paper* occurs before the PP *zaiyue-di qian* (before the end of this month) and the verb *finish*.In [6],*stay*, a content morpheme from English, the EL, is switched into the Japanese *suru*(do) structural phrase, where the content verb is introduced in conjunction with *suru*.

The above examples of ICS illustrate how the ML vs. EL hierarchy distinguishes the differential roles of the languages participating in ICS. 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 prepared by the ML.

2. *EL islands* consisting of only EL morphemes and well formed according to the EL grammar to show internal structural dependency relations.

[7] ta dui xueshen**very strict**.
   she to student very strict
   'She's very strict to students.'
   (Chinse/English; Wei1992 [2015, p. 54])

[8] 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)

[9] It's **totemomuzukashi** to find a convenient and **yasui** apartment here.
   it's very difficult to find a convenient and cheap apartment here
   'It's very difficult to find a convenient and cheap apartment here'.
   (English/Japanese; Wei 2006, p. 167)

[10] **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)

In [7], the AP *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 [8], the NP *le pentole* is an EL island where both the system morpheme *le* and the content morpheme *pentole*(pans) are switched from Italian. The determiner *le*, as a system morpheme, cannot appear by itself. In [9], the AP *totemomuzukashi* (very strict) is an EL island where both the system morpheme *totemo* (very) and the content morpheme *muzukashi* (difficult) are switched from Japanese. The degree word *totemo*, as a system morpheme, alone cannot appear by itself. In [10], the NP *the first one* is an EL island where both the system morpheme *the* and the content morphemes *first one*are switched from English. The determiner *the*, as a system morpheme, alone cannot appear by itself.

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 structure provided by the ML. The speaker's possible motivations for producing EL islands are discussed in the following sections.

3. The Bilingual Mental Lexicon

The mental lexicon not only contains vocabulary items but also more abstract information about particular words (Richards, 1976; Faerch & Kasper, 1984; Talmy, 1985, Ringbom, 1987; Nation, 1990, Wei, 2001a, 2001b, 2002). In speech production, speakers retrieve the appropriate words stored in their mental lexicon to correctly express their intended meanings. Each word retrieved from the mental lexicon contains not only its semantic content (i.e., its meaning) but also information about its phonological structure, morphological structure, syntactic environment, pragmatic function, and so on.

Such abstract information about a particular word is defined as 'lemma information' (or 'lemmas' for short) (Kempen & Huijbers, 1983; Kempen & Hoenkamp, 1987; Levelt, 1989; Myers-Scotton & Jake, 1995; Levelt, Roelofs, & Meyer, 1999; Wei, 2006). In other words, lemmas are abstract entries about a particular lexeme stored in the mental lexicon. For example, in addition to every other piece of information, the lemmas of *give* require a subject that carries the thematic role of AGENT (i.e., the person who performs the act), an indirect object that carries the thematic role of THEME (i.e., the thing that is given), and a direct object that carries the thematic role of RECIPIENT (i.e., the person who receives what is given), and the permissible orders in which these items appear; the lemmas of *explode* require a subject that carries the thematic role of THEME (i.e., an entity that goes through the process) rather than an AGENT; the lemmas of *she* require this pronoun to be used of a female subject and the inflection morpheme *-s* must be attached to the main verb for the subject-verb agreement if the verb is in the present tense.

The BLA Model (Wei, 2006, 2015) claims that unlike the monolingual mental lexicon, the bilingual mental lexicon contains language-specific lemmas, and such language-specific lemmas are in contact during a discourse involving ICS at three distinctive but related levels of abstract lexical structure (Myers-Scotton & Jake, 1995, 2000a, 2000b). The first level is lexical-conceptual structure, where lexical access takes place on the basis of the information contained in the speaker's preverbal message (i.e., the speaker's communicative intention). It is the speaker's preverbal message which activates language-specific semantic/pragmatic feature bundles, which are then mapped onto lemmas for 'lemma activation' in the bilingual mental lexicon. According to Green (1986, 1998), a lemma is activated if it matches part of the conceptual structure created by the conceptualizer. The second level is predicate-argument structure, where the thematic structure of a particular verb is mapped onto grammatical relations (i.e., thematic role assignments). The third level is morphological realization patterns, where surface grammatical relations, such as word order, agreement, inflectional morphology for tense/aspect/voice/mood/case marking, etc., are realized. Accordingly, the BLA Model describes and explains ICS in terms of activation of language-specific lemmas at each of these levels of abstract lexical structure.

Drawing on the MLF Model (Myers-Scotton 1993 [1997]; Myers-Scotton & Jake, 1995), the BLA Model assumes that ICS juxtapositions do not have much to do with surface linear of typological correspondences between the participating languages but originate with directions contained in language-specific lemmas. The BLA Model further proposes that language-specific lemmas drive particular lexicalization patterns resulting in different configurations of semantic/pragmatic features across related lemmas in different languages. That is, it is cross-linguistic differences in abstract lexical structure which affect code choices and constrain ICS configurations.

4. Lexical-conceptual Projection of Lemmas in ICS

The BLA Model recognizes the distinction between content and system morphemes as one of the hierarchies under the MLF Model but aims to explain why content morphemes, rather than system morphemes, are activated from the bilingual mental lexicon. The BLA Model claims that it is the language-specific lemmas of certain EL content morphemes which are conceptually projected from the bilingual mental lexicon for the speaker's communicative intention. This is because a particular EL conceptual chunk is specified in the speaker's preverbal message (Poulisse & Bongaerts, 1994). In other words, the activated concepts in the speaker's preverbal message will then activate the corresponding lemmas, and the sufficiently activated lemmas will then activate the associated lexemes (Roelofs, 1992; Levelt, 1989, 1995; Myers-Scotton & Jake, 1995; Wei, 2001a). The BLA Model proposes that it is cross-linguistic variations in lexical-conceptual structure (i.e., semantic/pragmatic feature bundles) that motivate language-specific lemma activation in ICS. Thus, this study regards the switched items as evidence of the relative importance of cross-linguistic lexical-conceptual differences in the lemmas underlying particular lexemes in the bilingual mental lexicon.

As introduced earlier, one of the two hierarchies identified in the MLF Model is the distinction between content and system morphemes. As commonly observed in naturally occurring ICS instances, it is individual EL content morphemes, rather than EL system morphemes, which can be switched in ICS to encode the speaker's intended meaning. The BLA Model further proposes that it is at the conceptual level that the speaker makes choices about semantic/pragmatic information that satisfies his/her communicative intention (i.e., pre-verbal message).

The lexical-conceptual approach to ICS proposed in this study suggests that one of the major reasons for certain EL content morphemes to be switched is due to the fact that it is content morphemes, rather than system morphemes, which contain lexical content or, to be more specific, semantic/pragmatic feature bundles.

The speaker may prefer to use certain EL content morphemes because he/she is aware of the lemma specifications for these morphemes. It is in this sense that language-specific lemmas underlying certain EL content morphemes are lexical-conceptually projected from the bilingual mental lexicon. Below are some typical instances of lexical-conceptually projected lemmas underlying the content morphemes in ICS.

[11] naxie**visiting scholar** bu shi hen youqian ma, bi women **student** you qian duo le.
   those visiting scholar not/EMPH COP very rich AFFIRM PREP/than us student havemoney more AFFIRM
   'Aren't those visiting scholars very rich? They have a lot more money than us students.'

[12] wo xiawuqujianwode**advisor**.
   I afternoon go see my advisor
   'I'm going to see my advisor this afternoon.'

[13] zhuzaizheli hen fanbian, meitian you **school bus**.
   live PREP/LOC here very convenient everyday have school bus
   'It's very convenient to live here (since) here is a school bus everyday.'

[14] wo you liang-fen **paper**mingtianbixujiaoshangqu.
   I have two-CL paper tomorrow must turn in
   'I must turn in two papers tomorrow.'
   (Chinese/English; Wei2006, pp.163-169)

In [11], *visiting scholar* and *student* are EL content morphemes, but the demonstrative pronoun 'naxie' (those) is an ML system morpheme, and there is no EL plural 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 also switches to *student* in the same sentence is to realize the conceptual difference between 'scholar' and 'student'. In [12], the speaker switches to advisor, an EL content morpheme, but the determiner 'wode' (my), a system morpheme, is 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 on. In addition, most English academic advisors also recommend their students to job markets or professional agencies. In contrast, a Chinese academic 'daoshi' (advisor) does not assume the same responsibilities as an English academic advisor. His/her major responsibility is to supervise his/her students in writing theses or dissertations. In [13], the speakers switches to *school bus*, an EL content morpheme, without the determiner 'a', an EL system morpheme. A 'school bus' in English mainly means a bus that transports students to and from a school. Some Chinese universities or schools in major cities 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 certain faculty. The speaker switches to *school bus* most probably because of the conceptual difference between 'school bus' and 'xiaoche'. In [14], the speaker switches to *paper*, an EL content morpheme, but 'liang-fen' is a combined ML system morpheme consisting of the numeral '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 lexeme 'zhi' (paper) itself only means a piece of paper for wrapping things up in or writing something on. It is obvious that the speaker switches to 'paper' for the English concept.

[15] imawa**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)

[16] anatawa**registration** o shimashi-ta ka?
   you TOP registration OBJ do-PERF QUE
   'Have you done your registration?'

[17] futatsu no **bedroom**ga ate, hitori, Maria to iu ko wahitori de **one bedroom** o mot-teimasuyo.
   two POSS bedroom NOM COP one person and call person TOP one person PERF/by onebedroom OBJ
   have-PROG AUX PART/ADDIRM

'We have two bedrooms. One person, called Maria, has one bedroom.'
(Japanese/English; Wei 2006, p. 170)

In [15], *summer course* is an English content morpheme phrase, but 'o' is an ML system morpheme marking the accusative case. The speaker switches to 'summer course' for the English academic concept that may not exist or is not popular in Japanese education settings. In [16], *registration* is an EL content morpheme, but 'o' is an ML system morpheme marking the accusative case. The speaker switches to 'registration' for the possible reason that in Japanese colleges/universities, although students must register for the courses to take, they are not free to select the courses which they are interested in and want to take. 'Registration' in the EL may carry a special meaning for the speaker to express his intended meaning accurately. In [17], *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' for the possible reason that in Japan the concept of 'bedroom' is relatively new. A traditional Japanese room is used not only for sleeping but also for eating, studying, meeting guests, or other daily family activities.

As observed in the above examples, bilingual speakers may switch to particular EL content morphemes at a certain point during a discourse because of cross-linguistic differences in lexical-conceptual structure, to be more accurate, cross-linguistic differences in the lemmas underlying particular lexemes. As observed by Li (1996), Nishimura (1997), and Wei (2001a, 2001b, 2002), language cues may have different values, and bilingual speakers may switch to particular EL content morphemes to realize their intended meanings. In other words, when the language cue specifies a particular language at a certain point during a discourse involving ICS, the lexical item from that language receives activation. This means that conceptual information and language cues must work together in activating language-specific lemmas in the bilingual mental lexicon.

The above ICS examples also provide evidence that all the system morphemes are from the ML, and all the content morphemes switched from the EL are inserted into the syntactic slots as framed by the ML.

Such instances of language-specific lemma activation are also observed in other language pairs involved in ICS. Below are a few representative ones selected among many.

[18] Kerransäolitpannusi-tämun**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)

[19] Mi tyalagharghyayla**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)

[20] nei5 zou6 saai3 di **assignment** mei6.
  you do ASP CL assignment SFP
  'Have you done all the assignments?'
  (Cantonese/English; Chan 1998, p. 193)

[21] 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)

[22] evet, **terras**-ta oturuyorlar.
  yes café-LOC sit-PROG.3PL
  'Yes, they are sitting at the outdoor café.'
  (Turkish/Dutch; Backus1996, p. 140)

[23] Ha-u ku-on-a a-ki-ni-**buy**-i-a **beer**sikuhiyo?
  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)

In [18], *lunchbox* is a content morpheme from English, the EL, but it is marked with the Finnish case, an ML system morpheme. In [19], *persuade* is a content morpheme from English, the EL, but the complementizer 'la' (to), asystem morpheme, is from Marathi, the ML. In [20], *assignment* is a content morpheme from English, the EL, but the noun classifier 'di', a system morpheme, is from Cantonese. Also, *assignment* is not inflected for the plural marking in the EL. In [21], *nokum* (recording) is a content morpheme from Korean, the EL, but the article 'the', a system morpheme, is from English, the ML. In [22], *terras* (café) is a content morpheme from Dutch, the EL, but it receives the locative case from Turkish, the ML. In [23], *buy* and *beer* are the content morphemes from English, the EL, but Swahili, the ML, provides the sentential frame for the thematic content of the clause to be mapped onto a clause well-formed in Swahili. In this example, *buy* takes two objects, but the sentential frame of Swahili determines how those objects are realized; the BENIFICIARY is realized as an object prefix on the verb '-ni-' and is further mapped onto the sentential frame through the applied verbal suffix '-i-' on the EL verb from English.

Although it is difficult to explore speakers' exact motivations for switching to certain content morphemes from another language, what becomes clear is that bilingual lemmas are in contact and are not equally activated. EL content morphemes can be conceptually activated for speakers' intended or preferred communicative intention at a certain point during a discourse, but EL system morphemes cannot. It also becomes evident that *the Morpheme Order Principle* and *the System Morpheme Principle* (Myers-Scotton, 1993 [1997]) apply to all language pairs involved in ICS. That is, it is the ML which provides the morpheme order (i.e., the sentential frame), and all system morphemes must come from the ML.

5. Compromise of Lexical-conceptual Structure in ICS

The BLA Model assumes that languages may lexicalize concepts in different ways. Thus, when an EL lemmas are activated but they do not match those in the ML, the speaker must take some compromise strategies for ICS to occur. One of such compromise strategies is to produce EL 'islands' (Jake & Myers-Scotton, 1997; Wei, 2001b, 2002). An EL island is defined as a constituent in which an EL content morpheme occurs with only other EL morphemes, including EL system morphemes. The BLA Model proposes that at the conceptual level bilingual speakers do not produce surface morphemes but rather make appropriate choices about the semantic/pragmatic information that they intend to convey. If certain particular EL lemmas are activated at the conceptual level, but they do not match the ML counterparts, speakers must take the compromise strategy to produce EL islands. Such a compromise strategy becomes necessary to overcome cross-linguistic differences in lexical-conceptual structure for possible ICS.

[24] nineng-bu-neng**give me a ride**.
    you can-not-can give me a ride
    'Can you give me a ride?'

[25] name nimingtian**call me**.
    then you tomorrow call me
    'Then, you call me tomorrow.'

[26] 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.'

[27] jiaoninu'er**come to** Xiao Ying de **birthday party**.
    ask your daughter to come to Xiao Ying POSS/'s birthday party
    'Ask your daughter to come to Xiao Ying's birthday party.'

[28] wo meitiandei**work in the lab**hao ji-gexiaoshi.
    I everyday have to work in the lab quite a few hours
    'Everyday I have to work in the lab for quite a few hours.'

[29] nali you wu-tai 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, but three of them were out of order.'
    (Chinese/English; Wei 2001b, pp. 160-163)

In [24], *give me a ride* is an EL island, which is conceptually incongruent with the ML counterpart 'song woyixia'(send me one time). While the EL lexical-conceptual structure of 'give … a ride' contains the means of transportation conflated in the noun 'ride' as the direct object of the verb, the ML lexical-conceptual structure of 'song …' does not necessarily contain the means of transportation. The speaker switches to the EL expression probably because of its precise meaning. In other words, the lemmas underlying the whole EL verbal expression are activated as an EL island. In [25], *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 dianhuagei 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 [26], *pick you up* is an EL island. The speaker prefers 'pick up' for the possible reason that it contains the meaning of 'to take on as a passenger', but the ML equivalent expression 'jie' usually does not. The Chinese verb 'jie' means 'meet' (e.g., to meet someone at a bus/train station or airport), which does not necessarily involve providing personal transportation. It should also be noticed that *come* is accessed together with the infinitive phrase *to pick you up* as an EL island, where the infinitive marker 'to', a system morpheme, is from the EL. The speaker prefers the EL phrasal verb 'pick up' for its lexical-conceptual structure to express his intended meaning more accurately. In [27], *come to …* is an EL island. When the speaker selects the English verb 'come', the preposition 'to' indicating GOAL is simultaneously accessed. By contrast, in Chinese GOAL is not introduced by a preposition but conflated in verbs like 'lai' (come) and 'qu' (go) themselves. In [28], *in the lab* is an EL island, where LOCATION is directly conflated in the preposition 'in' itself. By contrast, in Chinese, LOCATION is indicated by a preposition: 'zai' signifying LOCATION with a particle following the noun in question. Chinese particles for the concept of location include 'li' (in), 'shang' (on), 'xia' (under), 'fujin' (near), and so for the. For example, 'zaishinianshe li' is equivalent to 'in the lab'. The speaker switches to the English preposition 'in' for the concept of location. The consequence is the production of an EL island. In [29], *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'. The consequence is the production of this idiomatic expression as an EL island.

It seems that one of the major reasons for EL islands to occur in ICS is that in the case of nonidiomatic expressions, the speaker's particular intention at a certain point during a discourse cannot be fully realized in the ML because of the insufficient matching between the ML and the EL semantic/pragmatic feature bundles. Thus, the speaker activates what he/she knows from the EL and switches it into the ML sentential frame. The other major reason is that in the case of idiomatic expressions, an EL island is activated as a single lexical-conceptual unit forced by the complexity of the EL item as selected. In either case, if the speaker prefers EL expressions, a compromise strategy must be taken in order for ICS to be possible. In other words, if particular EL lexical-conceptual features (i.e., EL semantic/pragmatic feature bundles) are preferred but are not sufficiently congruent with the EL counterparts, the speaker will activate them as single lexical-conceptual units and switch them into the ML sentential frame.

6. Compromise of Predicate-argument Structure in ICS

In addition to lexical-conceptual compromise of lemmas, different procedures must be applied to the encoding of predicate-argument structure in ICS involving typologically different language. In other words, if the language pairs involved in ICS do not share the same predicate-argument structure, because bilingual speakers have two speech plans available for them, they may stop the encoding of one of them and continue with the other so as to solve the problem in ICS. In ICS the choice of one EL predicate-argument structure versus another is determined by the larger ML sentential frame.

[30] wo keyi**wait for you**daoliangdian.
    I can wait for you till PREP/till two o'clock
    'I can wait for you till two o'clock.'
    (Chinese/English; Wei2001b, p. 166)

[31] tingshuonei-ge**professor** hen **crazy**. ta jingchang**fails students in exams**.
    hear that-CLASSIF professor very crazy she often fails students in exams
    '(I) heard that professor is very crazy. She often fails students in exams.'

[32] nibiyehoukeyi**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.'
      (Chiinese/English: Wei 2001b, p. 168)

[33] wo meitian de **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 [30], *wait for* is accessed as a single lexical unit, where the direct object, in this case, the THEME, is introduced by the preposition 'in'. In Chinese, the same meaning is expressed by a single verb 'deng' (wait). The activation ofan EL phrasal verb like this results in the maximal projection of an EL phrasal category, that is, an EL island. In [31], the verb phrase headed by *fail* is an EL island, where all the morphemes, including the system morphemes, are from the EL. In English, 'fail' can be used as a causative verb and thus takes the grammatical subject as the AGENT who makes the failure happen. By contrast, the Chinese equivalent verb 'shibai' (be defeated in …) is used only as a nonaccusative verb with the grammatical subject as the EXPERIENCER. The speaker switches to the predicate-argument structure initiated by the verb 'fail' due to the incongruence between the English 'fail' and the Chinese 'shibai'. In [32], *teach English to nonnative speakers* is an EL island, where 'nonnative speakers', the RECIPIENT, is introduced in the prepositional phrase headed by 'to', that is, the English indirect object dative construction. Chinese has the equivalent verb 'jiao' (teach), but it only permits the double object construction (e.g., jiaononnative speakers English). In this case, the speaker is aware of the incongruence between the EL predicate-argument structure and that of the ML but prefers the one initiated by the verb 'teach'. The consequence is the production of an EL island. In [33], *help her with her homework* is an EL island, where 'her homework' the THEME, is introduced in the prepositional phrase headed by 'with'. By contrast, in Chinese, the THEME must be introduced by a specific verb such as 'zuo' (do) in addition to the main verb 'bangzhu' (help) (e.g., wo meitian de bangzhu ta zuozuoye (I everyday have to help her do homework)). The speaker is aware of the incongruence between the predicate-argument structures driven by 'help' and 'bangzhu' but prefers the one driven by 'help'. The consequence is the production of an EL island.

Such instances reveal that when incongruence or insufficient congruence between the languages involved in ICS at the level of predicate-argument structure, but the speaker prefers the predicate-argument structure to express his/her intended meaning, a compromise strategy, such as the production of EL islands, must be taken for ICS to occur. It should be noticed that a particular EL island is also lexical-conceptually driven, and any EL island must be switched into the sentential frame provided by the ML.

7. Conclusion

Unlike most models which focus on surface grammatical constraints on ICS, this study adopts the MLF Model and the BLA Model and proposes a lexical-conceptual approach to ICS at a deep or abstract level. The major theoretical assumption underlying this lexical-conceptual approach is that the bilingual mental lexicon does not simply contain lexemes but also language-specific lemmas (i.e., abstract entries about lexemes) and language-specific lemmas are in contact during a discourse involving ICS. Thus, this study regards ICS as a linguistic phenomenon of bilingual lemmas in contact at any level of abstract lexical structure. It claims that bilingual lemma activation is a crucial interface between speaker intention and code switch. The representative ICS examples from various language pairs discussed in this study provide evidence that for the arguments and claims about the structural principles governing ICS in relation to the nature and activity of the bilingual mental lexicon. This study reaches several conclusions regarding the bilingual mental activity in ICS.

1. The bilingual speaker's languages are 'turned on' during a discourse involving ICS, but they are never equally activated at the same time. The ML is more strongly activated than the EL. Whichever language activated as the ML provides the sentential frame for the structural configuration of ICS. It is the ML which provides all system morphemes and most content morphemes. The EL only provides certain content morphemes to be switched into the larger ML frame.

2. The bilingual speaker's mental lexicon contains lemmas from the languages known, these lemmas are tagged for their specific languages, and language-specific lemmas are in contact during a discourse involving ICS. If the bilingual speaker is engaged in ICS, he/she may activate the EL lemmas as desired.

10

The major reason for the speaker to switch to certain EL content morphemes is that such EL content morphemes are lexical-conceptually activated for his/her intended meaning. 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: 151).

This is because the bilingual speaker is aware of the lexical-conceptual differences (i.e., lemma differences) between certain ML and EL content morphemes. The bilingual speaker may also solve lexical-conceptual problems caused by the lexical-conceptual gap between the languages involved. It is the bilingual speaker's intention contained in the preverbal message which calls for an EL content morpheme, and this selection activates the EL lemma supporting that content morpheme.

3.The bilingual speaker can activate lemmas from which language as the EL during a discourse involving ICS, but the activated EL lemmas must be sufficiently congruent with the counterparts of the ML at each level of abstract lexical structure. Otherwise, compromise strategies must be taken in order for ICS to occur. One of the compromise strategies is to produce EL islands. Whether a compromise strategy is taken at the level of lexical-conceptual structure or at the level of predicate-argument structure, all EL islands are conceptually 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.

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

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

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). *Lectures on government and binding*. Dordrecht: Foris.

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

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

Dzameshie, A. (1989). Codeswitching in EWE and English. Paper presented at African Linguistics Conference. Urbana, Ill.

Faerch, 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.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67-81.

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

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

Haegeman, L. (1991). *Introduction to government and binding theory*. Oxford: Blackwell.

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* (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.

Lasnik, H., &Uriagereka, J. (1988). *A course in GB syntax: Lectures on binding and empty categories*. Cambridge, MA: MIT Press.

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

Levelt, W. J. M. (1995). The ability to speak: From intention to spoken words. *European Review*, *3*, 13-23.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1-38.

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. (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 Puebla 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.

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. (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 competence and production model: Evidence from intrasentential code switching. *Linguistics*, *33*(5), 981-1024.

Myers-Scotton, C., & Jake, J. L. (2000a). Four types of morpheme: Evidence from aphasia, codeswitching, and second language acquisition. *Linguistics*, *38*(6), 1053-1100.

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, I. S. 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. Jacobson (Ed.), *Codeswitching as a worldwide phenomenon* (pp. 33-69). New York: Peter Lang.

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 TerminoenEspañol: Toward a typology of code-switching. *Linguistics*, 18, 581-618.

Poplack, S, &Sankoff, D. (1988). Code-switching. In U. Ammom& 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 Schwyzertüschnella conversazione one di figli di emigrati. Universitá di Zurigo, Facoltá 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. and Bhatia, T. (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. 269-87). Amsterdam: Benjamins.

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

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. (1992). Chinese/English codeswitching corpus. Unpublished.

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* (pp. 13-15). FryskeAkademy (2002).

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 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. (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.

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