Lemma Congruence Checking for Verbs in Chinese/English Codeswitching

Longxing Wei
Montclair State University

1. Introduction

This study investigates codeswitching (hereafter CS for short) structural configurations. It is assumed that it is the unequal activation of lemmas in the bilingual mental lexicon that drives and constrains utterances containing switched constituents. It is further assumed that lemma congruence checking between the languages involved in CS is a semantic/pragmatic and morphosyntactic principle governing CS. The CS data for the study were collected from naturally occurring conversations carried out by Chinese/English bilinguals in various settings. The CS instances were tape-recorded, transcribed and analyzed by the investigator who has linguistic expertise in both languages. In this paper, only some typical CS examples containing switched verbs or verb phrases are selected for the focal discussion of relevant issues regarding lemma activation in the bilingual mental lexicon and lemma congruence checking between the two languages for permissible CS realizations.

Unlike most previous studies which focused on the description of surface structural constraints and configurations, this study, from some psycholinguistic perspectives, explores the nature of the bilingual mental lexicon in relation to the speech production process involving CS. Also, unlike most previous studies which focused on the observed CS product to formulate rules and principles governing CS, this study describes and explains CS in terms of bilingual production procedures with a special reference to bilingual mental activity at various levels of speech production process involving CS. Based on some empirical evidence as observed in the CS data for the study, this paper has reached several conclusions about the nature of the bilingual mental lexicon and underlying principles governing the bilingual CS production process, rather than the CS product itself.

2. Lemmas of the bilingual mental lexicon

A first step in speech production is to retrieve appropriate words from the speaker's mental lexicon. The mental lexicon stores declarative knowledge about words and idioms in a language. However, for each item, the mental lexicon contains its lemma information, or lemma for short, that is, declarative knowledge about the word's meaning, and information about its syntax and morphology. Such information is necessary for constructing the word's syntactic environment. For example, the lemma for *he* requires the word to be used of a male and that any following present-tense main verb must carry the suffix *–s* attached to it for subject-verb agreement, the lemma for *love* requires a subject that expresses the thematic role of EXPERIENCER and an object that expresses the thematic role of THEME, and the lemma for *give* requires a subject that expresses the thematic role of AGENT, a direct object that expresses the thematic role of THEME and an indirect object that expresses the thematic role of BENEFACTIVE or/and GOAL. The elements required by a lemma must appear in a particular order as specified. In addition, the lemma for a particular word contains information about the word's phonological composition, such as segments and its syllable and accent structure. Furthermore, the lemma may contain information about the word's appropriate register, a particular kind of discourse it typically enters into, and about its pragmatic, stylistic, and affective functions. Thus, the lemma for a particular word contains all the necessary pieces of information about the required linguistic environment in which the word is used. "It is in the lemmas of the mental lexicon that conceptual information is linked to grammatical function" (Levelt, 1989: 162). In other words,

© 2005 Longxing Wei. ISB4: Proceedings of the 4th International Symposium on Bilingualism, ed. James Cohen, Kara T. McAlister, Kellie Rolstad, and Jeff MacSwan, 2338-2348. Somerville, MA: Cascadilla Press.

lemmas are abstract entries in the mental lexicon and underlie surface configurations of speech production. The question is what happens to lemmas of the bilingual mental lexicon.

Although it is generally assumed that there is only one single mental lexicon for bilinguals, there is some disagreement about the exact nature of lemma representations in the bilingual mental lexicon (Kempen & Huijbers, 1983; Levelt, 1989, 1995; Roelofs, 1992; Bock & Levelt, 1994; Myers-Scotton & Jake, 1995; Myers-Scotton, 1997, 2002; Wei, 2001). This study follows the general assumption that lemmas are language-specific for lexicalization patterns of a particular language (cf. Talmy, 1985; Jackendoff, 1997; Wei, 2002). Talmy (1985) provides numerous examples of language-specific lemmas. Here is an often cited one: (English) *The bottle floated into the cave.* (Spanish) *La botella entró flotando en la cueva.* The underlying structure for these sentences is proposed as [DO [MOVE BOTTLE]:FIN [BOTTLE LOC IN CAVE]]:FLOAT BOTTLE (Bierwisch & Schreuder, 1992). It is apparent that the lexicalization pattern differs across the two languages. While in English MOVE and the manner of motion FLOAT can be conflated into a single lemma, in Spanish they cannot. In other words, in English a semantic form in the mental lexicon is available for both motion and the manner of motion to be expressed, but in Spanish a different lexicalization pattern comes into play. One of the most important assumptions underlying this study is that language-specific lemmas form interconnection between the lexical features and conceptual features, which map to and from syntax (cf. 'The distributed lexical/conceptual feature model' proposed by Kroll & De Groot, 1997). Another important assumption underlying this study is that each lemma of the bilingual mental lexicon is tagged for a specific language and supports the realization of an actual lexeme. Thus, according to this assumption, language-specific lemmas of the bilingual mental lexicon activate language-specific sets of morphosyntactic procedures and rules in the speech production Formulator. Figure 1 illustrates what pieces of information are contained in language-specific lemmas in the bilingual mental lexicon, and how lemma selection depends on such pieces of information. How language-specific lemmas are activated or deactivated in the bilingual mental lexicon in relation to the levels of speech production is discussed in the following section.

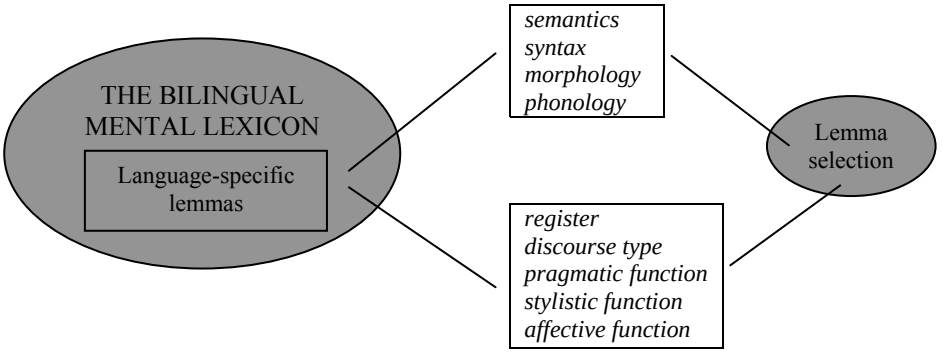

Figure 1: The bilingual mental lexicon

3. A bilingual lemma activation model

In Levelt's model of speech production (1989), semantic and syntactic information constitutes the lemma of the lexical item while morphological and phonological information constitutes the form of the lexical item. The conceptual information in the preverbal message activates the appropriate

lexical items during the formulation of a message. Levelt's model was designed for describing the major components and processes of speech production in general. It needs to be adapted to account for bilingual speech production involving CS. This is because it becomes necessary to explain why bilinguals may switch between the two or more language systems at a certain point during a discourse and how at the same time they keep the language systems apart.

One of the most important implications of language-specific lemmas of the bilingual mental lexicon and their activation in speech production for describing and explaining underlying configurational principles is that activation of language-specific lemmas mediates between the Conceptualizer at the conceptual level and the Formulator at the functional level. Such a mediation is schematized in Figure 2.

Slightly different from Levelt's model, a bilingual lemma activation model proposes four levels of speech production. At the conceptual level the Conceptualizer generates messages by attending to the bilingual's communicative intention in terms of the discourse mode and preverbal message to be desired. If the speaker chooses the monolingual mode, no CS will occur; if the speaker chooses the bilingual mode, then he/she has to decide whether CS will be performed. If the speaker decides to perform CS, the Conceptualizer then generates preverbal message containing the speaker's choice of the language as the Matrix Language (Myers-Scotton, 1997) (commonly known as the base language) to be used in CS and semantic/pragmatic feature bundles to be desired for the speaker's communicative intention. The semantic/pragmatic feature bundles selected by the Conceptualizer then trigger the appropriate lemmas into activity before the Formulator has access to the lexical items stored in the bilingual mental lexicon. Thus, language-specific lemma activation plays a crucial role in bilingual production involving CS and is treated as one of the levels of speech production in a bilingual lemma activation model. Another important feature of this model is the assumption that not all language-specific lemmas of the bilingual mental lexicon are activated equally, and such an unequal activation may result in CS. Thus, in order for CS to occur, lemma congruence checking between the languages involved must come into play. Lemma congruence is defined as "a match between the ML and the EL at the lemma level with respect to linguistically relevant features" (Myers-Scotton & Jake, 1995: 985) (ML: Matrix Language, EL: Embedded Language). Lemma congruence checking must take place at three levels of abstract lexical structure: at the level of lexical-conceptual structure, at the level of predicate-argument structure, and at the level of morphological realization patterns. Figure 2 shows that if the ML lemma and the EL lemma are congruent at each of these levels, the speaker can proceed with the bilingual mode for CS; if the ML lemma and the EL lemma are incongruent at any of these levels, the speaker has to give up the bilingual mode and go back to the monolingual mode instead. Only when there is a match between the ML and the EL at the lemma level, directions will be sent to the Formulator at the functional level for grammatical and phonological encoding by observing a set of structural principles governing CS; otherwise, compromise strategies must be taken in order for CS to occur. The successfully encoded material will then be sent to the Articulator at the position level for overt speech production, that is, surface forms including word order, agreement morphology, case marking and phonetic string for speech comprehension. Thus, lemma congruence checking is regarded as an organizing principle for CS (Wei, 2000b, 2002). The CS data to be studied in this paper provide empirical evidence for this principle. Relevant to the current study is lemma congruence checking at the first two levels of abstract lexical structure with special reference to verbs in Chinese/English CS.

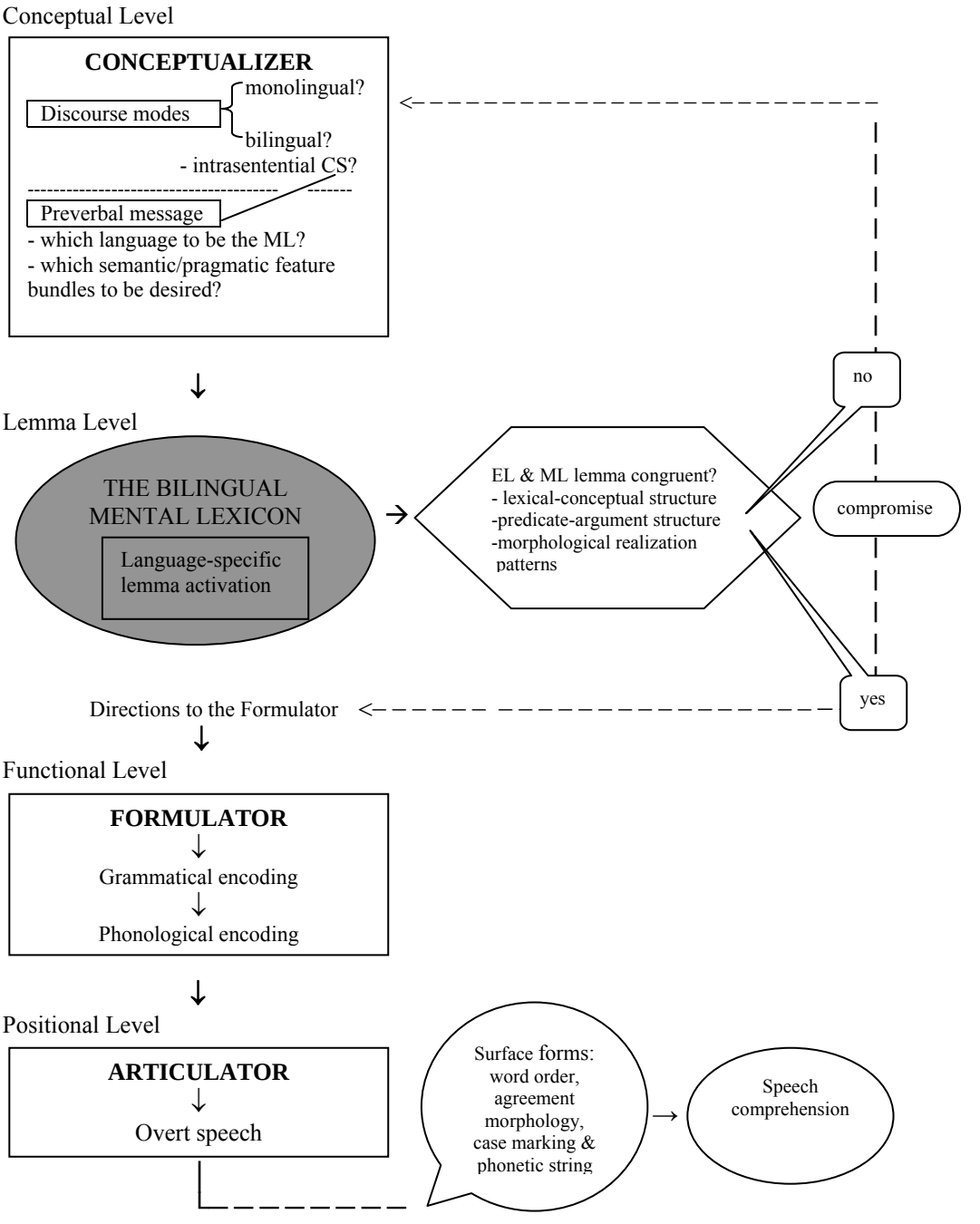

Figure 2: A bilingual lemma activation model
(adapted from Levelt, 1989; Myers-Scotton & Jake, 2000; Wei, 2002)

4. Unequal activation of lemmas and participating languages

Myers-Scotton (1993) proposed two fundamental distinctions in CS: the content vs. system morpheme distinction and the ML vs. EL distinction. The former determines what type of morphemes can be switched from another language, and the latter determines how switched morphemes can occur in a sentence. These two distinctions are assumed to constrain CS configurations. A brief explanation of such distinctions in relation to bilingual lemma activation becomes necessary before we deal with the issue of lemma congruence checking for verbs as observed in Chinese/English CS.

The content vs. system morpheme distinction depends on two lexical features: [± thematic role assigner/receiver] and [± quantification] (Myers-Scotton, 1993). Those morphemes that assign or receive thematic roles are content morphemes. Nouns, descriptive adjectives, and most verbs are prototypical content morphemes because they can be characterized as categories that are potential thematic role assigners or receivers. In contrast to content morphemes, system morphemes do not assign or receive thematic roles. Prototypical system morphemes are quantifiers, specifiers, and inflectional affixes. The content vs. system morpheme distinction is now slightly recast in light of the sources of morphemes in terms of lemma entries. Assumptions regarding the sources of morphemes refer to differences in the levels of abstract lexical structure. According to Myers-Scotton and Jake (1995) and Wei (1996, 2000a, 2002), at the conceptual level, speaker intentions are mapped onto language-specific semantic/pragmatic feature bundles. This is the mapping of speaker intentions to lemma entries. It is assumed that entries in the mental lexicon support lexical entries. Lemmas contain all aspects of lexical information necessary for projecting a morphosyntactic frame. This is because they activate morphosyntactic procedures spelling out the lexical knowledge of the lexical entry. Such morphemes supported by those activated lemmas of the mental lexicon are content morphemes. In other words, those morphemes which are supported by lemma entries activated by the Conceptualizer are content morphemes.

The ML vs. EL distinction underlies the differential participation of the ML and the EL in shaping utterances containing CS. It is the ML which plays the dominant role. This is because in ML + EL constituents only the ML can project the sentential 'frame' (i.e., the sentence grammar, including word order, inflectional morphology, and other functional items). It is for this reason that the model proposed by Myers-Scotton (1993, 2002) is called the Matrix Language Frame (MLF) model. The assumption underlying this model is that the languages participating in CS have unequal roles, with one language more central than the other in sentential frame building. The ML is more activated than the EL in a discourse involving CS and the occurrence of its morphemes is much more frequent and freer than that of the EL. The ML is the language which the speakers engaged in CS will intuitively identify as the 'base' or 'main' language being used.

Figure 2 shows that if the speaker chooses the bilingual mode and CS, it is the speaker's preverbal message at the conceptual level that determines which language to be used as the ML. The ML chosen at the conceptual level, together with the semantic/pragmatic feature bundles selected as desired, activates language-specific lemmas of the bilingual mental lexicon at the lemma level in order to send directions to the Formulator for language encoding at the functional level.

Three main structural principles governing CS are essential in explaining structural constraints on CS and predicting possible ML + EL constituents. In the MLF model (Myers-Scotton, 1993, 2002), two principles are proposed under the Matrix Language Hypothesis: the system morpheme principle and the morpheme order principle. *The system morpheme principle* specifies that in mixed constituents, all syntactically relevant system morphemes must come from the ML, and the EL only supplies content morphemes to the sentential frame projected by the ML. *The morpheme order principle* specifies that in mixed constituents, surface morpheme order must follow that of the ML. Wei (2001, 2002) discussed compromise strategies in CS and proposed *the lemma congruence checking principle* as an organizing principle (cf. Myers-Scotton & Jake, 1995) for CS configurations. It specifies that lemmas activated from the EL must be congruent with the ML counterparts in order for CS to occur. Otherwise, compromise strategies must be taken for possible CS realization. These three essential principles are further tested in this study of the nature of the bilingual mental lexicon with a focus on the issue of lemma congruence checking.

5. Lemma congruence checking as an organizing principle

All lemmas include three levels of abstract lexical structure: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. This abstract lexical structure figures in explaining and predicting possible CS configurations. As mentioned earlier, the speaker's preverbal message (i.e., speaker intentions) in the Conceptualizer activates semantic/pragmatic feature bundles at the interface between the Conceptualizer and the bilingual mental lexicon, and these activated semantic/pragmatic feature bundles are then mapped onto language-specific lemmas in the bilingual mental lexicon as lexical-conceptual structure. At the level of predicate-argument structure thematic structure is mapped onto grammatical relations. At the level of morphological realization patterns surface grammatical relations, such as word order, agreement morphology, and case marking, including phonological forms, are realized. Thus, lemma congruence between languages involved in CS is regarded as an organizing principle governing CS configurations. Relevant to this study is lemma congruence checking at the first two levels of this structure with special reference to verbs in Chinese/English CS.

5.1 Lemma congruence checking at the level of lexical-conceptual structure

It is generally assumed that there is a universal set of semantic/pragmatic features available for the lexical structuring of lemmas, but there is also cross-linguistic variation in the presence and conflation of these features. Speakers select individual content morphemes to encode their communicative intentions. However, at the conceptual level speakers do not produce surface level morphemes but rather make appropriate choices about the semantic/pragmatic information that they intend to convey. The information chosen at this abstract level activates the lemmas in the bilingual mental lexicon which will support surface level morphemes. Most Chinese/English CS examples show sufficient congruence between the ML and the EL at the level of lexical-conceptual structure (i.e., congruence between the ML and the EL lemmas which support existing lemmas in both languages).

(1) ni nei-pian *article* hai mei *finish* a?
you that-CL yet no PART/AFFIRM-QUE
'You haven't yet finished that article?'

(2) wo *summer* bu *take course* le.
I not PART/AFFIRM
'I won't take any course in summer.'

(3) tingshuo ni *in May graduate*, shi ma?
hear you right PART/INTERROG
'(I) heard you'll graduate in May, won't you?'

As introduced earlier, *the morpheme order principle* specifies that it is the ML which governs surface morpheme order in mixed constituents, and *the system morpheme principle* specifies that in mixed constituents all system morphemes which have grammatical relations external to their head constituent (i.e., which participate in the sentence's thematic role grid) must come from the ML. Examples (1)-(3) show that English verbs and verbs with their complements or adjuncts are switched the ML sentential frame with all the system morphemes from the ML. Although Chinese and English share the same basic V-O order, Chinese is very flexible in the arrangement of sentential elements. In (1) the object (complement) 'nei-pian article' goes before the V rather than after it (i.e., VP: O V), in (2) the adverbial of time (adjunct) follows the subject rather than the V, and again in (3) the adverbial of time 'in May' (adjunct) is placed immediately before the V. Such word orders are not permitted in English, but they are just some of the typical Chinese grammatical patterns. One of the apparent reasons for English verbs to be switched into Chinese sentences lies in the fact that English verbs, whether a single verb (V), a verb with its complement (V"/VP) or a verb with its adjunct (V'), can fit into the Chinese flexible word order where verbs can occur.

To summarize briefly, to call ML + EL constituents 'mixed' obscures their highly systematic nature. The regularly consist of an ML morphosyntactic frame (i.e., system morphemes and morpheme order) into which content morphemes from the EL are inserted. This evidence leads us to claim that in bilingual speech involving CS, the constituent frame is constructed by higher order procedures called only by the ML. Either ML or EL lemmas may call the procedures which insert ML or EL content morphemes respectively into the frame, but only the EL lemmas which are congruent with the ML counterparts are possible.

5.2 Lemma congruence checking at the level of predicate-argument structure

The above examples show that it is the ML which sets the sentential frame into which the EL content morphemes are switched. Thus, only the ML controls the predicate-argument structure by supplying system morphemes, a subcategorization frame for the verb, and morpheme order. However, before morphosyntactic directions are sent to the Formulator, lemmas from both languages can be activated at a certain point during a discourse. Thus, the checking for congruence at the level of lexical-conceptual structure alone is not sufficient enough for CS to occur. The checking for congruence at the level of predicate-argument structure supported by an EL lemma and its ML counterpart must come into play.

The Chinese/English CS examples indicate that the speakers tend to use many EL verbs and verb phrases as well as EL nouns and noun phrases. One of the obvious reasons for this is that Chinese and English share the same basic V-O order. The speaker may simply switch the EL verbs and verb phrases or nouns and noun phrases into the syntactic slots prepared by the ML. Most examples for this study indicate that there is sufficient congruence between the EL and the ML lemmas underlying the realization of an EL content morpheme, either a single verb or a verb with its complement noun or a verb with its adjunct, in the ML morphosyntactic frame. Below are some typical examples.

(4) ta gang dao, ta dei xue *drive*.
he just arrive he must learn
'He just arrived, and he must learn how to drive.'

(5) ni dei xiang banfa *make money*.
you must think of way
'You must think of ways to make money.'

(6) wode che you *give me trouble* le.
my car again PART/PERF
'My car gave me trouble again.'

In (4) the infinitive V 'drive' is used here as the object of the main verb 'xue' (learn) in exactly the same word order as in English, except that Chinese does not possess the infinitive marker 'to'. In (5) the VP 'make money' follows the V-O order in both languages. In (6) the VP 'give me trouble' follows the V-O-O order (double object dative structure) in both languages.

However, there also exist apparent differences between Chinese and English in some of their morphosyntactic features. In addition to the nonexistence of inflectional morphology for tense, aspect, voice, or person/number marking or grammatical devices like the infinitive marker 'to' and the dummy subject pronouns 'it' and 'there', Chinese shows two major differences from English in its morphosyntactic patterns. One is the category of the head of a maximal projection, and the other is the grammatical argument structure (Talmy, 1995; Wei, 2001, 2002). Such differences are discussed in the following section.

6. Lemma incongruence and compromise strategies

It has been commonly recognized that languages do not lexicalize concepts in the same way (i.e., lexical representations are language-specific) and languages may differ in grammatical patterns. Thus,

the existence of lack of sufficient congruence between languages requires some compromise strategies in order for CS to occur. One of the compromise strategies is the production of EL island (Myers-Scotton, 1993, 2002; Jake & Myers-Scotton, 1997; Wei, 2001). An EL island is a constituent containing an EL content morpheme with only other EL morphemes, including EL system morphemes. As shown in Figure 2, if lemma incongruence occurs between the language pairs involved in CS but the speaker does not want to 'compromise', he/she must give up CS and go back to the monolingual mode instead. This study assumes that incongruence between the language pairs involved in CS in regard to lexical-conceptual structure or predicate-argument structure are two of the major reasons for EL islands to be formed if the EL lemmas are selected and activated for speaker intentions.

6.1 Lemma incongruence in lexical-conceptual structure

The bilingual lemma activation model assumes that at the conceptual level speakers do not produce surface morphemes but rather make appropriate choices about the semantic/pragmatic information that they intend to convey. The semantic/pragmatic feature bundles chosen at the conceptual level activate the lemmas in the bilingual mental lexicon which will support surface morphemes, and the activated EL lemmas must be congruent with the ML lemmas in terms of their lexical-conceptual structure, predicate-argument structure, and morphological realization patterns in order for CS to occur. Most CS instances studied in this paper show sufficient congruence between the EL and the ML lemmas which support the existing lexemes in both languages. However, in some cases, the language pairs may differ in lexical-conceptual structure.

(7) na wo yi dian *come to pick you up*.
so I one o'clock
'So, I'll come to pick you up at one o'clock.'

(8) name ni mingtian *call me*.
then you tomorrow
'Then you call me tomorrow.'

(9) ni neng-bu-neng *give me a ride*?
you can-not-can
'Can you give me a ride?'

In (7) 'pick you up' occurs as an EL island because the speaker chooses the EL lemma's lexical-conceptual structure the whole VP with a pronominal object before the particle satellite 'up' is accessed. The speaker prefers 'pick up' for the possible reason that this phrasal verb contains the meaning of 'to take on as a passenger', but the Chinese equivalent 'jie' usually does not. Chinese 'jie' means 'meet' (e.g., to go to the station to meet somebody), which does not necessarily involve providing personal transportation. It should also be noticed that 'come' is accessed together with the infinitive phrase 'to pick you up' as an EL island. The possible explanation is that the English infinitive marker 'to', a system morpheme, becomes obligatory if two successive verbs are activated and selected simultaneously. The speaker selects the EL phrasal verb 'pick up' for its lexical-conceptual structure to realize his communicative intention more accurately. In (8) the semantic features of 'communicate with by telephone' are conflated in the verb 'call', but the Chinese equivalent to 'call me' is 'da dianhua gei wo' (literally translated as 'make phone to me'). Since the speaker chooses the EL lemma which activates the EL lexical-conceptual structure, the whole VP is accessed and produced as an EL island. In (9) 'give me a ride' is incongruent with the ML counterpart 'song wo yixia' (literally translated as 'send me one time'). While in the EL the lexical-conceptual structure of the means of transportation is conflated in the noun 'ride' as the direct object of the verb, in the ML it may be conflated in the verb 'song' because the verb itself may not contain the means of transportation at all. The speaker chooses the EL expression probably because he wants to be more specific than he can be with the Chinese counterpart. Thus, when the EL lemma is activated, the whole VP is accessed and produced as an EL island.

As commonly observed, bilinguals may switch to particular lexical items from another language at a certain point during a discourse because of cross-linguistic differences in lexical-conceptual structure (i.e., lemma incongruence between languages at the level of lexical-conceptual structure). This study assumes that although every language is capable of expressing its speakers' semantic/pragmatic meanings, it has its own lexicalization patterns (i.e., lexicalization patterns are language-specific). As observed by Li (1996), Nishimura (1997), and Wei (2001, 2002), in many cases exact cross-linguistic translation may be inaccurate or incomplete. This is because language cues may have different values, the speaker may switch to an EL lexical item to make his/her intended meaning explicit. When the language cue specifies a particular language at a certain point during a discourse involving CS, the lexical item from that language will receive more activation. This means that conceptual information and language cues must work together in activating language-specific lemmas in the bilingual mental lexicon for the speaker's particular communicative intention. If the activated lemmas are incongruent between the languages involved in CS, compromise strategies must be taken in order for CS to occur; otherwise, the speaker has to give up the bilingual mode for the ongoing discourse.

6.2 Lemma incongruence in predicate-argument structure

As introduced earlier, it is the ML which controls the morphosyntactic structure of the sentence containing switched constituents. This is because the ML supplies system morphemes, subcategorization frames for verbs, and morpheme order. Although morphosyntactic procedures are realized by the Formulator at the functional level, before morphosyntactic directions are sent to the Formulator, lemmas from both languages can be activated at a certain point during a discourse. Thus, lemma congruence checking at the level of lexical-conceptual alone is not sufficient enough for CS to occur. Lemma congruence checking at the level of predicate-argument structure must come into play. This is because even if lemma congruence at the conceptual level provides a match between the EL lemma and its ML counterpart, the ML morphosyntactic structure may reject the mapping. Such incongruence has been observed in Chinese/English CS instances.

Although Chinese has the same basic V-O order as English, in some cases, predicate-argument structures across the two languages may differ. If such incongruence occurs, but the semantic/pragmatic feature bundles desired by the speaker activate the EL lemma for his/her communicative intention, a radical compromise strategy must be taken in order for the EL material to be accessed. That is, EL semantic/pragmatic feature bundles selected and activated at the conceptual level must be realized in EL islands at the functional level.

(10) tingshuo nei-ge *professor* hen *crazy*. ta jingchang *fails students in exams*.
hear that-CL very she often
'(I) heard that professor is very crazy. She often fails students in exams.'

(11) ni biye hou keyi *teach English to nonnative speakers*.
you graduate after can
'After you graduate you can teach English to nonnative speakers.'

(12) wo meitian dei *help her with her homework*.
I everyday have to
'Everyday I have to help her with her homework.'

In (10) the VP headed by 'fail' is an EL island (i.e., with all the system morphemes and the morpheme order from the EL). In English 'fail' can be used as a causative verb and thus takes the grammatical subject as the AGENT who makes the failure happen, but in Chinese 'shibai', an equivalent to 'fail', means 'be defeated in …', and is used only as a noncausative verb and thus takes the grammatical subject as the EXPERIENCER, rather than the AGENT. Since the speaker prefers the EL concept, but there is incongruence between the EL and the ML predicate-argument structure, the result is the production of an EL island. In (11) the RECEIPIENT is introduced by the PP headed by 'to': V-O-to-O (i.e., the indirect object dative structure). By contrast, the equivalent Chinese VP

headed by ‘jiao’ (teach) only permits the V-O-O structure (i.e., the double object structure). Since the speaker selects the EL material at the level of lexical-conceptual structure, but the ML does not accept the mapping which the EL PP would project at the level of predicate-argument structure, the results is the production of an EL island. In (12) the THEME is introduced by the preposition ‘with’. By contrast, in the ML the THEME is always introduced by a specific verb such as ‘zhuo’ (do). The speaker selects the EL verb ‘help’ at the level of lexical-conceptual structure, but the EL and the ML are in congruent at the level of predicate-argument structure. As a result, the whole VP in the EL is accessed and produced as an island.

The above examples show that lemma incongruence between the language pairs involved in CS, either at the level of lexical-conceptual structure or at the level of predicate-argument structure, requires a compromise strategy in order for CS to occur. Lemma incongruence at the level of lexical-conceptual structure is mainly caused by incomplete match between the EL and the ML lexemes, and lemma incongruence at the level of predicate-argument structure is caused by mismatch between the EL and the ML in their maximal projection of a category or grammatical argument structure. The radical compromise strategy for the EL material to be realized in CS is to produce EL islands switched into the ML sentential frame.

7. Conclusion

This study has demonstrated how naturally occurring CS instances provide some empirical evidence for certain specifications about the nature of the bilingual mental lexicon in relation to the bilingual mental activity in the speech production process involving CS. Unlike most previous studies of CS which remain at a surface descriptive level, this study explains code choice and code switch in terms of four abstract levels of bilingual speech production process (i.e., conceptual level, lemma level, functional level, and positional level), with a focus on bilingual lemma activation as a crucial interface between speaker intention (i.e., the speaker’s preverbal message activated at the conceptual level) and language encoding at the functional level. Although this study focused on verbs in Chinese/English CS, it has reached several conclusions regarding the bilingual speech production process.

First, the bilingual speech production process contains the same levels as those contained in the monolingual speech production process. However, since the bilingual has more than one language mode available to speech production, at the conceptual level the bilingual speaker makes several choices about the language mode, monolingual or bilingual, to be used and selects semantic/pragmatic feature bundles to convey his/her communicative intention. If the speaker chooses the bilingual mode, he/she still has to choose between intersentential an intrasentential CS (in this paper, CS is short for intrasentential CS because only this type of CS is relevant to the current study). It is assumed that speaker intention occurs before choice.

Second, the bilingual mental lexicon contains lemmas rather than lexemes from the languages known, and these lemmas are tagged for the specific languages (i.e., lemmas are language-specific). If the speaker chooses the bilingual mode and is engaged in intrasentential CS, he/she may activate the language-specific lemmas as preferred from his/her bilingual mental lexicon. Language-specific lemmas contain information about a particular word’s semantics, pragmatics, syntax, and morphology, and such information is necessary for using the word appropriately and for constructing its syntactic environment.

Third, although the bilingual’s language modes are turned ‘on’ all the time if the bilingual mode is chosen during a discourse, they are never equally activated at the same time. The ML is more activated than the EL in terms of morphosyntactic frame building and frequency of occurrence of types of morphemes. Whichever language is activated as the ML, it is the ML which controls morphosyntactic procedures and provides both content and system morphemes at a much higher frequency. The EL only supplies content morphemes desired by the speaker for semantic/pragmatic reasons to be switched into the ML sentential frame.

Fourth, the bilingual can activate lemmas from whichever language is the EL during a discourse involving CS, but these activated lemmas must be congruent with the ML counterparts at three levels of abstract lexical structure: the level of lexical-conceptual structure, the level of predicate-argument

structure, and the level of morphological realization patterns. If incongruence occurs between the language pairs involved in CS at any of these levels, compromise strategies must be taken in order for CS to occur. Otherwise, the speaker has to give up the bilingual mode and use the monolingual mode instead.

References

Bierwisch, Manfred, & Schreuder, Robert (1992). From concepts to lexical items. *Cognition*, 42, 23-60.

Bock, Kathryn, & Levelt, Willem J. M. (1994). Language production: Grammatical encoding. In Morton Ann Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). New York: Academic Press.

Jackendoff, Ray S. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.

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

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

Kroll, Judith F., & de Groot, Annette M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In Annette M. B. de Groot & Judith F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 169-199). Mahwah, NJ: Erlbaum.

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

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

Li, David C. S. (1996). *Issues in bilingualism and biculturalism: A Hong Kong case study*. New York: Peter Lang.

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

Myers-Scotton, Carol (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. New York: Oxford University Press.

Myers-Scotton, Carol, & Jake, Janice L. (1995). Matching lemmas in a bilingual language competence and production model: Evidence from intrasentential code switchinhg. *Linguistics*, 33, 981-1024.

Nishimura, Miwa (1997). *Japanese/English code-switching: Syntax and pragmatics*. New York: Peter Lang.

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

Talmy, Leonard (1985). Lexicalization patterns: Semantic structures in lexical form. In Timothy Shopen (Ed.), *Language typology and syntactic description III* (pp. 57-149). New York: Cambridge University Press.

Wei, Longxing (1996). Organizing principles behind codeswitching and interlanguage development in early adult second language acquisition. In Jennifer Arnold, Renée Blake, Brad Davidson, Scott Schwenter & Julie Solomon (Eds.), *Sociolinguistic variation: Data, theory, and analysis. Selected papers from NWAV 23* (pp. 417-431). Stanford, CA: Center for the Study of Language in Information.

Wei, Longxing (2000a). Unequal election of morphemes in adult second language acquisition. *Applied Linguistics*, 21(1), 106-140.

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

Wei, Longxing (2001). Lemma congruence checking between languages as an organizing principle in intrasentential codeswitching. *International Journal of Bilingualism*, 5, 153-173.

Wei, Longxing (2002). The bilingual mental lexicon and speech production process. *Brain and Language*, 81, 691-707.

ISB4: Proceedings of the 4th International Symposium on Bilingualism

edited by James Cohen, Kara T. McAlister, Kellie Rolstad, and Jeff MacSwan

Cascadilla Press Somerville, MA 2005

Copyright information

ISB4: Proceedings of the 4th International Symposium on Bilingualism
© 2005 Cascadilla Press, Somerville, MA. All rights reserved

ISBN 978-1-57473-210-8 CD-ROM
ISBN 978-1-57473-107-1 library binding (5-volume set)

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Press.

Ordering information

To order a copy of the proceedings, contact:

Cascadilla Press
P.O. Box 440355
Somerville, MA 02144, USA

phone: 1-617-776-2370
fax: 1-617-776-2271
sales@cascadilla.com
www.cascadilla.com

Web access and citation information

This paper is available from www.cascadilla.com/isb4.html and is identical to the version published by Cascadilla Press on CD-ROM and in library binding.