

Original Paper

Bilingual Complex Abstract Lexical Structure and Its Relevance to Interlanguage Studies

Longxing Wei^{1*}

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

* Longxing Wei, E-mail: weil@mail.montclair.edu

Received: April 11, 2023

Accepted: May 13, 2023

Online Published: May 17, 2023

doi:10.22158/eltls.v5n2p162

URL: <http://dx.doi.org/10.22158/eltls.v5n2p162>

Abstract

This study adopts two assumptions about abstract lexical structure. One is that lexical structure is modular: lexical information is organized into subsystems pointing to different levels of linguistic structure, and parts of lexical structure can be split and recombined. The other concerns the sources of morphemes actually occurring in surface strings. As commonly recognized, abstract lexical structure contains three levels: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. This study argues that abstract lexical structure becomes “complex” because the bilingual mental lexicon contains lemmas (i.e., abstract entries about lexemes) from different sources, such as learners’ first language, their target language, and their interlanguage. By adopting a multi-layered speech production model (Levelt, 1989; Myers-Scotton & Jake, 2001; Wei, 2002, 2015, 2020), it further argues that different types of morphemes are accessed at different levels of language production, resulting in different degrees of learning difficulty (an implicational hierarchy of second language morpheme acquisition) and different types of learner errors in interlanguage production. The typical instances of learner errors for the study are collected from interlanguage performance by adult second language learners with various first language backgrounds. This study aims to explore the nature of the bilingual mental lexicon and mechanisms of interlanguage development.

Keywords

complex, abstract, lexical structure, morpheme, lexical-conceptual, predicate-argument, morphological realization, interlanguage, lemma, mental lexicon

1. Introduction

This study assumes two premises about abstract lexical knowledge. One is that lexical structure is modular in that lexical information is organized into several subsystems pointing to different levels of

linguistic structure, and parts of lexical structure can be split and recombined. The other premise concerns the sources of morphemes actually occurring in surface strings (cf. Talmy, 1985; Levelt, 1989, 1995; Jake, 1998; Myers-Scotton, 1993 [1997], 2002; Wei, 1996a, 1996b, 2000a, 2000b, 2015, 2020). Under these assumptions, this study considers two essential questions about the developing nature of interlanguage (IL): If second language (L2) learners build up and revise the developing or interim linguistic system (i.e., IL) by gradually increasing the complexity of the target language (TL) system, what is the origin of IL itself? How can the TL items be lexically projected and morph syntactically realized if L2 learners' knowledge of the TL at various linguistic levels is incomplete or insufficient? To answer such questions, this study, by adopting a multi-layered speech production model (Levelt, 1989; Myers-Scotton & Jake, 2001; Wei, 2002, 2015, 2020), investigates how different types of morphemes are accessed at different levels of language production in general, what may cause different degrees of L2 learning difficulty, and what are the sources of learner errors in IL production.

2. Complex Abstract Lexical Structure and Its Subsystems

It is assumed that lexical structure is composed of complex abstract subsystems (Jake, 1998; Wei, 1996a, 2000a, Myers-Scotton, 2002; among others). This assumption becomes necessary and useful in explaining psycholinguistic data. Levelt (1989) and Bock and Levelt (1994) have proposed that lexical structure contains three abstract levels: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. These three abstract levels are complex in that they are procedurally automated in speech production (Wei, 2015, 2020). The lexical-conceptual structure of a lexical item contains all the semantic and pragmatic feature bundles associated with a particular lexeme and pointers to other lexemes with which it frequently occurs (Jake, 1998). Semantic and pragmatic feature bundles of lexical-conceptual structure are mapped onto predicate-argument structure. For example, predicate-argument structure reflects the properties of verbs in terms of subcategorization and constrains the grammatical encoding of the expressed arguments in terms of their morphosyntactic procedures. To limit the domain of predicate-argument structure, morphological realization patterns spell out surface devices for word order, including directionality of thematic role assignment, case assignment, agreement, tense/aspect marking, voice marking, etc. (Levelt, 1989; de Bot & Schreuder, 1993; Myers-Scotton & Jake, 1995; Myers-Scotton & Jake 2001; Wei, 1996a, 1996b, 2002). The following examples of IL production (labeled IL) with equivalents in the speaker's L1, Chinese (labeled L1C) may illustrate such abstractness and complexity of lexical structure.

[1] IL: This boy hear radio.

L1C: ze-ge nanhai ting shouyinji.

DEM-CL boy hear/listen to radio

[2] (The interviewer said: I'm missing my friends in China.)

IL: Yes. I missing too.

L1C: shide. wo ye xiang.

yes I also miss

[3] IL: What you study?

L1C: ni xue shenme?

you study what

[4] IL: Could you tell me you how long have been here?

L1C: ni neng gaoshu wo nii zai zheli duou jiu le?

you could tell me you COP here how long PERF

(Chinese learners' English IL; Wei, 1995)

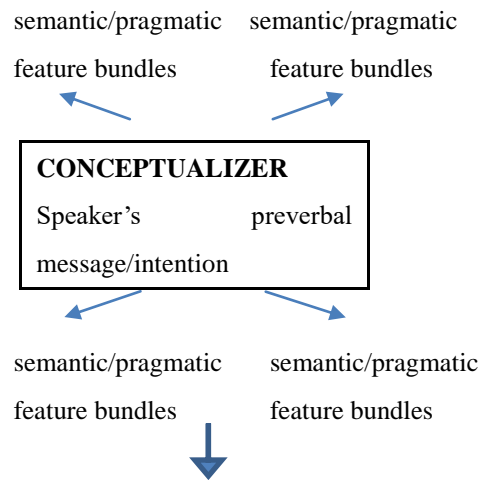
(1) Shows that the speaker assumes that the English lexical-conceptual structure matches that of the Chinese verb “*ting*” (hear/listen) which does not make a distinction between “*hear*” and “*listen to*”. In Chinese, “*ting*” optionally encodes intentionality on the part of the subject; further, it selects a nominal instead of prepositional complement. (2) shows that the speaker uses a Chinese predicate-argument structure in which the direct object of the verb “*xiang*” (miss) is optional; such null arguments do not occur in the TL. (3) and (4) show that the speaker mixes the Chinese (L1) word order with that of the TL. In (3), although the speaker places what according to the TL word order, he does not follow the TL rule of subject-verb inversion for question formation. The speaker also misses the auxiliary verb as required for question formation. In (4), although the speaker places the interrogative WH element “*how long*” according to the TL word order, he does not follow the TL WH-movement rule for question formation in the complement clause by using the L1 word order where the subject of the question precedes the WH element. Even when the WH element is appropriately fronted, its order with respect to other elements and the realization of those other elements do not match the TL. (3) and (4) show the learners' production of a mixed or composite word order.

Levelt's model of language production (1989) assumes two kinds of knowledge: declarative knowledge (similar to encyclopedic knowledge) and procedure knowledge (similar to automated knowledge). Figure 1 presents a multi-layered speech production model which depicts the relationship between these two kinds of knowledge. Speakers' declarative knowledge includes not only their knowledge about the world in general (encyclopedic knowledge) but also their knowledge about a given situation (situational knowledge) and/or their knowledge about a speech style appropriate for a circumstance (discourse knowledge) in particular. Speakers' declarative knowledge also includes the lexicon, which contains “lemmas” (i.e., abstract entries in the mental lexicon about particular lexemes). Speakers generate their intended messages by calling on declarative knowledge in the CONCEPTUALIZER to construct preverbal messages. The formation of a preverbal message enables the appropriate lemma to be activated. The FORMULATOR then automatically processes the message and send it to the ARTICULATOR to produce a surface morphological form as required for phonological encoding. The figure also depicts procedural knowledge as a succession of boxes indicating the actions automated in speech production. This automated knowledge is “procedural” in the sense that a series of conditions must be met in sequential stages as speakers for structures that provide input to successive levels of speech production.

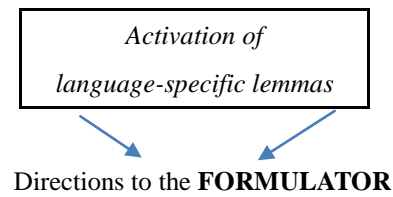
One of the distinctive features of this speech production model is the lemma level as an indispensable part of the speech production process. The mental lexicon does not simply contain lexemes, but more abstract elements called “lemmas” defined by abstract features affecting several structural levels (Levelt, 1989; Garrett, 1990; Myers-Scotton & Jake, 1995; Wei, 2002, 2020). Lemmas contain directions regarding all three subsystems of lexical structure and direct the FORMULATOR in the transformation of conceptual knowledge into linguistic knowledge during speech production. Surface forms based on lexical structure are the ultimate output of the FORMULATOR. This is because “lemmas are the driving force behind the speaker’s construction of the surface structure. It is the lemmas of the mental lexicon that conceptual information is linked to grammatical function” (1989, p. 162). For example, the lemma entry of a verb contains information about semantic and pragmatic selectional restrictions, argument structure, and directions for realizing projections for tense/aspect/voice/mood morphemes, case marking, and so on.

As schematized in Figure 1, in the multi-layered speech production model adopted and adapted for the present study, four levels interact to produce a surface form. At the conceptual level, which is prelinguistic, speaker intentions activate semantic and pragmatic feature bundles. These bundles lead to specific lemmas in the mental lexicon. In turn, at the lemma level, activated lemmas send directions to the FORMULATOR. At the functional level, the FORMULATOR constructs constituents out of the three subsystems of lexical structure encoded by lemmas. At the positional level, surface forms are realized. The following section discusses the relevance of this model to the classification of morphemes in relation to second language acquisition (SLA).

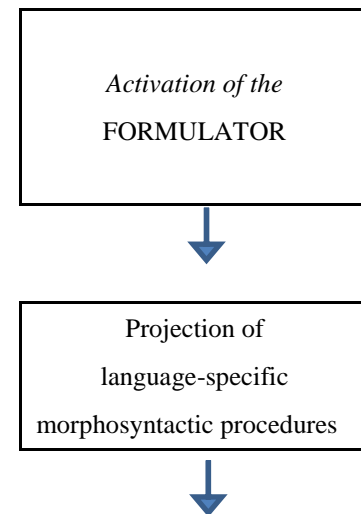
Conceptual Level



Lemma Level



Functional Level



Positional Level

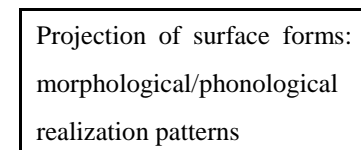


Figure 1. A Multi-layered Speech Production Model

(Adapted from Levelt, 1989; Myers-Scotton & Jake, 2001, 2014; Wei, 2002, 2015)

3. Morphemes Accessed at Different Levels of Speech Production

Morphemes can be classified based on how they are accessed in speech production. The 4-M Model (Myers-Scotton & Jake, 2000a) identifies 4 types of morphemes: content morphemes, early system morphemes, and two types of late system morphemes (also see Wei, 1996a, 1996b). Assumptions regarding different types of morphemes also refer to differences in the levels of abstract lexical structure. In the 4-M Model, the classification of morphemes is motivated by the Differential Access Hypothesis. “The different types of morphemes under the 4-M Model are accessed differently in the abstract level of the production process. Specifically, content morphemes and early system morphemes are accessed at the level of the Mental Lexicon, but late system morphemes do not become salient until the level of the Formulator” (Myers-Scotton, 2002: 17).

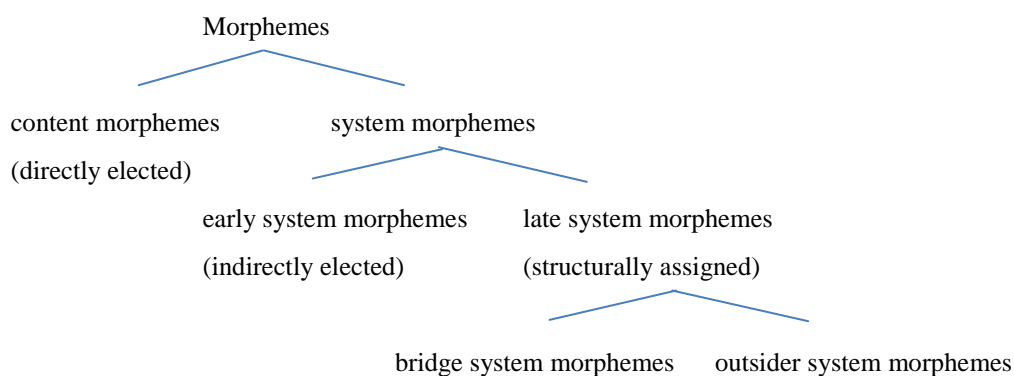


Figure 2. Types of Morphemes

(Adapted from Levelt, 1989; Myers-Scotton, 1993; Wei, 1996a, 2015).

3.1 Content vs. System Morphemes

Content and system morphemes are distinguished in terms of how they are elected at different levels of speech production. Content morphemes are “directly elected” by the semantic and pragmatic feature bundles of speaker intentions. “Directly elected” refers to activation at the lemma level, which sends directions to the FORMULATOR to project the predicate-argument structure of morphemes (Levelt, 1989; Wei, 1996a). More specifically, Myers-Scotton (1993 [1997]) defines content morphemes as either thematic role assigners or thematic role receivers within a complementizer phrase (CP), often a clause. Lexical verbs assign thematic roles, and lexical nouns receive thematic roles. Also, most prepositions assign thematic roles to lexical nouns. System morphemes are neither thematic role assigners nor thematic role receivers. Prototypical system morphemes include determiners, most auxiliary verbs, certain prepositions, and inflectional morphemes.

Content morphemes are activated at the conceptual level, where speaker intentions are mapped onto semantic and pragmatic feature bundles (Myers-Scotton & Jake, 1995; Wei, 1996a). Initial speaker intentions are not language-specific, but these semantic-pragmatic feature bundles are, corresponding to lemma entries in the mental lexicon. This is because lemmas contain all aspects of lexical information

necessary to project a morphosyntactic frame. In other words, to be more specific, content morphemes are “directly elected” morphemes in the sense that they are “directly elected” lemmas. It is “directly elected” lemmas that send directions about predicate-argument structure to the FORMULATOR, which activates morphosyntactic procedures to fulfill the lexical entry at the functional level. Based on these directions, and after phonological encoding, the ARTICULATOR produces surface level strings at the positional level.

Myers-Scotton and Jake (2000b) further distinguish between the directly elected morphemes and the additional semantically and pragmatically salient morphemes required to realize speaker intentions. Such additional morphemes are called “indirectly elected” early system morphemes. The content vs. system morpheme distinction captures the distinction between thematic and functional elements. For example, [5] The child looked at the picture.

In (5), “*child*”, “*look*”, and “*picture*” are content morphemes. “*Child*” receives the thematic role of EXPERIENCER, and “*picture*” receives the thematic role of STIMULUS. The remaining morphemes are system morphemes. The determiner “*the*” only indicates definiteness and presumably points to a specific known child and a particular picture; it does not assign or receive a thematic role. The morpheme “-ed” marks the past tense; it only plays an abstract grammatical function. The satellite preposition *at* fleshes out the meaning of “*look*”; it does not change the number or nature of the roles assigned to “*look*”, but it does allow the implied stimulus to surface. However, language may assign particular lexical concepts to content or system morphemes differently: not all members of a particular lexical category will pattern alike. This nuance justifies invoking the content vs. system morpheme distinction. Under this distinction, any pronouns that receive thematic roles directly are content morphemes. This is because such pronouns occur in the same argument positions as full NP (noun) arguments, and they exhibit the same properties as other content morphemes. By this criterion, English “*we/us*”, “*they/them*”, “*she/her*”, and “*he/him*” are content morphemes. The lexical pronoun “*it*” is also a content morpheme, but English dummy pronouns “*it*” and “*there*” are not. For example,

[6] She drinks milk.

[7] Who drinks milk?

[8] It is raining cats and dogs.

[9] *What/Who is raining cats and dogs?

In (6), “*she*” can be questioned as in (7), a possible sentence of English, since “*drink*” assigns a thematic role to a nominal in the subject position. In contrast, in (8), “*it*” cannot be questioned as in the unacceptable example (9), since “*rain*” does not assign a thematic role to its grammatical subject.

3.2 Early vs. Late System Morphemes

The 4-M Model is designed to explain the distribution of different types of morphemes as a result of how they are accessed in speech production. Its premise is that lemmas supporting morphemes become salient at different levels in the process of language production. Some morphemes are selected from a limited number of choices via the projection of the structure required by other lexical items. Such morphemes are

“indirectly elected” along with “directly elected” elements (Bock & Levelt, 1994; Jake, 1994; Myers-Scotton, 1996; Wei, 1996a). They are indirectly elected as part of the realization of the predicate-argument structure and morphological realization patterns, but they do not represent lexical concepts independent of the directly elected elements with which they are accessed. As indicated in Figure 1, such morphemes are referred to as “early” system morphemes (Myers-Scotton & Jake, 1999a). For example, the directly elected morpheme *look* may indirectly elect one of several prepositions as its satellite (e.g., *look for*, *look at*, *look after*, *look down*, *look forward to*, *look into*, *look like*, *look out*, *look to*, etc.). Talmy uses the term “satellites” to refer to “certain immediate constituents of a verb root other than inflections auxiliaries, or nominal arguments. They relate to the verb root as periphery or (modifiers) to a head. A verb root together with its satellites forms a constituent in its own right, the “verb complex” ...” (1985, p. 102). Thus, satellites also include non-prepositions as in *look over*, *look on*, *look through*, etc. Unlike prepositions, such satellites are movable as in “He *looked over* the contract” or “He *looked* the contract *over*”, and such satellites are used with a head as a fixed lexical unit as in “I’m not playing; I’m just *looking on*”.

One piece of evidence for the distinction between directly elected and indirectly elected morphemes shows that L2 learners may miss certain indirectly elected morphemes, but their utterances still carry their intended meanings. For example, in (10), the speaker produces “*listening music*” without the preposition “*to*”, and in (11), the speaker produces “*looking pictures*” without the preposition “*at*”.

[10] You’re listening music?

[11] No. I’m looking pictures.

(Chinese learners’ English IL; Wei, 1995)

The process of “indirect election” described above reflects the relationship between speaker intentions and lemma entries. For example, in certain semantic/pragmatic contexts, the lemma supporting a noun directly elected by the semantic and pragmatic feature bundles activated by speaker intentions may also project a phrase structure calling for the definite article for “definiteness”, “specificity”, etc. In other words, the definite article is indirectly elected by projection, at the functional level, of the lemma supporting the noun. This means that indirectly elected lemmas only arise from the projections of directly elected lemmas (e.g., read *the* book, look at *the* picture, in *the* sun, listen to *the* radio, and go to *the* hospital). The lexical-conceptual structure of a particular noun’s lemma entry may require the article to realize certain semantic and pragmatic features, which in turn realize situational and discourse-specific, as well as encyclopedic, knowledge.

In addition to the distinction between content and system morphemes, the 4-M Model states that not all system morphemes have the same properties and that not all are early system morphemes, which are elected at the lemma level. Some system morphemes are only assigned at the functional level to satisfy grammatical requirements. Some other system morphemes do not reflect the semantic and pragmatic feature bundles conflated in the lexical-conceptual structure of lexemes but are only realized in the projection of morphological realization patterns at the positional level. Unlike early system morphemes,

both of these types are referred to as “late” system morphemes (Myers-Scotton & Jake, 1999a). They are called “late” system morphemes because their forms are not available until the FORMULATOR assembles morphemes into larger constituents based on directions from those morphemes elected at the lemma level (i.e., content and early system morphemes). Because late system morphemes are not salient at the lemma level, it is lemma entries which slot these morphemes. In other words, content and early system morphemes describe the morphological realization patterns in which late system morphemes occur and thus project structures that must be filled by other grammatical forms.

The 4-M Model further divides late systems into two types: “bridge” system morphemes and “outsider” system morphemes. Late system morphemes that assemble structure into larger constituents are referred to as “bridge” system morphemes. They are only required to satisfy requirements on well-formedness for a specific type of maximal projection. A maximal projection is the highest projection of a lexical item (e.g., VP is the maximal projection of its node V, NP is the maximal projection of its node N, etc.). For example, in *the friend of my uncle*, it is *of* which constructs a larger NP out of smaller a smaller nominal constituent. The second type of late system morphemes are called “outsider” system morphemes because they depend on information outside of their immediate maximal projection to determine their forms. For example, in English, subject-verb agreement requires the verb to look to its subject for its form. Thus, third-person singular “-s” is an outsider system morpheme.

In the context of IL studies, the distinction between bridge and outsider system morphemes is not as directly relevant as the more general distinction between content and system morphemes, the distinction between early and late system morphemes. The present study analyzes the differences between distribution of content morphemes and the two general types of system morphemes. The goal is to describe and explain differences in accuracy in adult SLA.

4. Complex Abstract Lexical Structure and Its Relevance to Interlanguage Studies

As introduced above, lexical structure is abstract and complex. It is abstract in the sense that it contains language-specific lemmas as abstract entries in the mental lexicon regarding particular lexemes, and such abstract entries affect three structural levels, such as lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. It is complex in the sense that these levels of abstract lexical structure are sequentially related in the speech production process. It is also complex in the sense that morphemes are classified into 4 types in terms of their lexical or grammatical functions in speech production.

The present study assumes that complex abstract lexical structure is universal and relevant to IL studies.

4.1 Content vs. System Morphemes in Interlanguage Production

Wei (1996a, 1996b, 2015) assumes that the content vs. system morpheme distinction underlies all linguistic systems, and IL is no exception. The division of morphemes into those which are directly elected and those which are not results in two major classes of morphemes: content vs. system morphemes. Directly elected morphemes are content morphemes because they are supported by lemma

entries activated by the CONCEPTUALIZER at the conceptual level. The lexical entries of system morphemes are “called” by the speech production FORMULATOR at the functional level in the mapping of lexical-conceptual structure onto predicate-argument structure or predicate-argument structure onto morphological realization patterns. Thus, both indirectly elected and structurally assigned morphemes are system morphemes, but their sources are different. That is, some system morphemes represent part of lexical-conceptual structure, but others do not.

The distinction between content and system morphemes is relevant to IL studies, as are differences in the sources of system morphemes. Specifically, the level, functional or positional, at which a particular system morpheme is activated is also relevant to IL studies. In other words, the distinction between content and system morphemes and their different sources (i.e., directly elected, indirectly elected or structurally assigned morphemes) inform the principles determining the differing contributions of the three linguistic systems involved in IL development: the learners’ L1, the T, and the developing IL system itself. It is these different sources of morphemes which affect IL forms because their structural and psycholinguistic characteristics determine what each linguistic system involved in IL can contribute.

The present study finds that even though IL shows little actual bilingual speech, learners, especially at the early stage of learning, may frequently structure their L2 utterances in terms of more than one language. As observed, learners may produce utterances by making use of their L1 morphological realization patterns.

[12] I English not speak well.

[13] I come USC just one year.

(Chinese learners’ English IL; Wei, 1995)

[14] I really this picture like.

[15] I’m sorry I not understand.

(Japanese learners’ English IL; Wei, 1995)

In (12), the speaker uses the Chinese word order for topicalization and Chinese pre-verbal negation. In (13), the speaker does not use the TL system morpheme for past tense marking. In (14), The speaker uses the typical Japanese verb-final word order. In (15), the speaker does not use the TL auxiliary verb for negation, but only the pre-verbal negative particle occurs as in Japanese.

As introduced earlier, one crucial opposition structuring utterances is the distinction between content and system morphemes. Most nouns, verbs, adjectives and prepositions are prototypical content morphemes. They can be characterized as thematic role-assigners or thematic role-receivers, but potential thematic role-assigners that do not actually assign a thematic role (e.g., the copula *be*) are system morphemes. In contrast to content morphemes, system morphemes do not assign or receive a thematic role of any kind. Prototypical system morphemes are auxiliary verbs, quantifiers, specifiers, and inflectional affixes (Myers-Scotton, 1993, [1997], 2002).

The distinction between content and system morphemes and their potential sources also apply to IL. This is because the abstract lexical structure from learners’ L1 appears in initial or early IL production though

not the L1 form itself. Furthermore, this distinction supports the Implicational Hierarchy Principle (Wei, 1996b, 2000a, 2000b) that target-like use of TL content morphemes comes before target-like use of TL system morphemes. It is obvious that no actual morphemes from learners' L1 occur (i.e., all surface forms are from the TL), the abstract lexical structure, including its subsystems: lexical-conceptual structure, predicate-argument structure and morphological realization patterns, may come from learners' L1.

As evidenced in early research in SLA, the distinction between content and system morphemes clearly applies in the initial stage of L2 learning. One commonly recognized characterization of initial IL is "telegraphic speech" (Ellis, 1985). Beginning learners' use of TL words become predominant for immediate communicative purposes. In the general linguistic literature, these words are commonly labeled "open class words" (e.g., Joshi, 1985), which include "high information" words (Dale, 1976: 37), such as nouns, verbs, adjectives, adverbs, and some prepositions. This is because L2 learners initially have only a limited number of lexical items and restricted means available to indicate their interrelation within more complex units. In order to express their most immediate communicative needs, they attempt to be understood using their learner language, whose grammar is limited and different from that of the TL. The following examples reveal some characteristics of early IL production.

[16] Yesterday daddy buy me two story book.

[17] You go too? We have three State Fair ticket.

[18] Outside cold, inside warm.

[19] Tomorrow we no go school. Today Mrs. Murry ... she tell class.

[20] You doing homework?

[21] You not go library, I go.

(Chinese learners' English IL; Wei, 1995)

1. The elimination of semantically redundant morphology (i.e., system morphemes), as in (16) and (17).
2. The absence of certain system morphemes' such as inflections for tense articles, auxiliary verbs, possessives, and the copula *be*, and the substitution of adverbials or contexts for time reference, as in (18), (19) and (20).
3. The pre-verbal negation, as in (19) and (21).
4. The lack of movement rules (i.e., word order transformation) for question formation, as in (20).

These examples show that beginning L2 learners can communicate with certain words from the TL. Such words are content morphemes, which have constant semantic and pragmatic properties, but they do not show morphological inflections as required in the TL (e.g., verb inflections for tense or aspect marking noun inflection for number, etc.). As commonly observed, early IL forms are mainly content morphemes. Thus, we can speak with more specificity about early IL if we look at its nature and sources in terms of the distinction between content and system morphemes.

4.2 Developing Nature of Complex Abstract Lexical Structure in Interlanguage

The present study holds that IL is a "composite" developing linguistic system (Wei, 2009, 2021) in the sense that it contains not only incompletely acquired L2 elements but also elements from Learners' L1. In

building the IL composite, learners' L1 may contribute complex abstract lexical structure at the level lexical-conceptual structure, predicate-argument structure, and/or morphological realization patterns. The examples below illustrate how learners' L1 may contribute complex abstract lexical structure to the IL composite.

4.2.1 L1 Lexical-Conceptual Structure in the IL Composite

[22] Yesterday in library I *look* Japanese magazine.

(Japanese learners' English IL, 1996a)

[23] When I've cold I *eat* medicine, cold medicine.

[24] In Japan students *do* many tests and exams in class.

(Japanese learners' English IL, 2003)

[25] She now *do* meal.

[26] Open air condition.

(Chinese learners' English IL, Wei, 1995)

[27] yoru ni denwa o *ageru*.

“(I) will give you a call in the evening.”

[28] watashi wa tennis o *asobu*.

“I play tennis.”

[29] kare wa shiken o *toru*.

“He'll take the test.”

[30] watashi wa mai nichu juuni ji ni hirugohan ga *aru*.

“I have lunch at 12 o'clock every day.”

(English learners' Japanese IL, Wei, 2003)

In (22), the learner produces “*look*” instead of “*read*”. This does not necessarily indicate that the learner has not acquired the English verb “*read*”. The learner does so probably because the lemmas underlying the Japanese equivalent verb “*look*” includes “*read*”, “*see*”, “*look at*”, “*visit*” or “*observe*”. In (23), the learner produces “*eat medicine*” rather than “*take medicine*” probably because the lemma underlying the Japanese equivalent verb “*eat*” includes the concept of taking medicine. In (24), the learner produces “*do many tests and exams*” rather than “*take many tests and exams*” probably because the learner uses the Japanese equivalent verb “*do*”, whose lemmas already include the concept of “*take*”. In (25), the learner produces “*do meal*” probably because the lemmas underlying the Chinese equivalent verb “*do*” include the concepts such as “*cook*”, “*play*”, “*work*”, “*act*”, etc. In (26), the learner uses the verb “*open*” instead of “*turn on*”. It is possible that the learner has not acquired the English expression yet, but the lemmas underlying the Chinese equivalent verb “*open*” allows the learner to use it for the same concept. In (27), the learner produces “*ageru*” (give) rather than “*kakeru*” as lexical-conceptually required in Japanese because of the lemma underlying “*give*” as in “*give you a call*”. In (28), the learner produces “*asobu*” (play) as influenced by the English express rather than “*suru*” (do) as required in Japanese. In (29), the learner produces “*toru*” (take) rather than “*ukeru*” (receive) for the English equivalent express “*take the*

tesu". In (30), the learner produces "*aru*" (have) as used in English for "*have lunch*" rather than "*taberu*" (eat) for the same concept in Japanese.

The above examples of L1 lexical-conceptual structure in IL production reveal that in the early stage of L2 learning learners acquire certain simple TL lexical items first which match up possible L1 conflation categories of semantic notions. These examples also reveal that learners have not acquired or have not completely acquired certain particular TL lexical items with their language-specific lemmas underlying the TL lexical-conceptual structure of particular lexemes. Thus, the IL lexical-conceptual becomes composite which contains elements of learners' L1 abstract lexical-conceptual structure (Pinker, 1989a, 1989b; Jake, 1994; Wei, 1996a, 1996b, 2015, 2021).

4.2.2 L1 Predicate-Argument Structure in the IL Composite

In addition to the above examples of L1 abstract lexical structure in the IL composite, the IL data for the present study also shows particular instances of L1 influence at the level of predicate-argument structure in the IL composite.

[31] He is funny. His words in class laugh me.

[32] He busy. He not help my homework.

[33] I can wait you here.

[34] Why you ask many questions for me?

(Japanese learners' English IL; Wei, 1995)

[35] She cost me hundred dollar ... bad tooth.

[36] Today he help dinner.

[37] She yesterday look you.

[38] Please help me look my child.

[39] You're listening music?

(Chinese learners' English IL; Wei, 1995)

In (31), the semantic feature of the CAUSE is incorrectly extended to the TL verb "*laugh*". The resulting causative lexical-conceptual structure affects the predicate-argument structure and the morphological realization patterns. In the IL composite, the CAUSEE is the PATIENT "*me*", rather than the AGENT "*I*". As in the TL predicate-argument structure, "*his words*" should be the STIMULUS introduced in a PP as in "*I laughed at his words in class*". In (32), the verb "*help*" assigns the THEME to a direct object without the BENEFICIARY. In the TL, a double complement predicate-argument structure is required for a direct object BENEFICIARY and a THEME introduced in PP as in "*He doesn't help me with my homework*". In (33), the THEME "*you*" is introduced by the verb "*wait*" without the preposition "*for*", but the TL predicate-argument structure for the verb "*wait*" requires this sentence to be "*I can wait for you*". In 34), the GOAL or INTERLOCUTOR is introduced in a PP, structurally subordinate to the internal object THEME. Such a predicate-argument structure reflects the L1 counterpart of "*ask*" (*suru*) which projects the GOAL as a postpositional object with the postposition "*ni*", rather than the internal object with the particle "*o*". In the TL, the GOAL and the THEME occur in a double-object construction as in "*ask*"

someone a question". In (35), the verb *cost* occurs with the POSSESSOR "she" of the THEME "tooth" as the subject, rather than the THEME itself (the thing on which the money is spent), as in the TL "Her bad tooth cost me a hundred dollars". In (36), the verb "help" assigns the THEME to a direct object, but the BENEFICIARY as required in the TL does not appear, as in "Today he helped me with dinner". In (37), the THEME "you" is introduced directly by the verb "wait" without the preposition "for", since the Chinese counterpart of "look" (zhao) does not need a preposition to introduce the THEME, but the TL predicate-argument structure requires the THEME to be introduced by the preposition "for" as in "She looked for you yesterday". In (38), the TL preposition "after" does not appear to introduce the THEME "my child", since the Chinese counterpart of "look" (zhaoliao) does not need a preposition to introduce the THEME, but the TL predicate-argument structure does not allow the preposition to be left out to introduce the THEME, as in "Please help me look after my child". In (39), the THEME of "listen" is the internal or direct object, since the Chinese counterpart of "listen" (ting) does not need a preposition like to introduced the THEME, but the TL predicate-argument structure requires the preposition "to" to introduce the THEME, as in "Are you listening to the music?"

The above examples illustrate L1 influence on the IL composite at the level of predicate-argument structure. Such early IL predicate-argument structures are affected by the L1 abstract lexical structures because of learners' incomplete access to TL lemmas underlying particular TL lexical items.

4.2.3 L1 Morphological Realization Patterns in the IL Composite

Morphological realization patterns must be realized at the positional level to satisfy the surface requirement for speech production. As commonly observed, in the early stage of learning, learners may turn back on their L1 morphological realization patterns in IL production.

[40] I from Japan arrive.

[41] In Japan student English junior high school start.

[42] I in Japan my city like.

(Japanese learners' English IL; Wei, 1995)

[43] Sorry, only little English I know.

[44] Tomorrow to New York we'll go with some friend.

(Japanese learners' English IL; Wei, 2003)

[45] I like ... accounting people job because I study English, just more study English.

[46] You how long have been here?

[47] I and my husband go to Washington D.C., New York, Atlanta last year.

[48] I go to Logan School 1994 September.

[49] It's small city, but China Xian is big city.

(Chinese learners' English IL; Wei, 1995)

[50] You go too? We have three ticket.

[51] Go swim? No. Parent no go, you no go swim.

(Chinese learners' English IL; Wei, 1996b)

[52] watashi wa nouu kakiowatta watashino repooto.

“I already finished my paper.”

[53] watashitachi wa shigoto ni iku mainichi.

“We go to work every day.”

(English learners' Japanese IL; Wei, 2003)

In (40), the verb occurs at the end of the sentence. The verb final structure is typical of the learner's Japanese morphological realization patterns. Such a word order influenced by learners' L1 are also observed in (41)-(44). In addition to the L1 influenced word order, inflectional morphemes do not appear for past tense marking as in (40) and (42), and for plural marking as in (44). In (45), the adverb “*more*” appears before “*study English*”, which is a rather normal word order in Chinese. In (46), the WH element “*how long*” is not moved to the pre-subject position, which is a normal Chinese word order in question formation. In (47), “*I and my husband*”, instead of “*my husband and I*”, appears since in Chinese the first person is ordered before other persons. In (48), the year precedes the month. In (49), the name of the country precedes that of the city. In (50) and (51), no auxiliary verb is used to form questions. It should also be noticed that some system morphemes are missing, such as the indefinite article in (49) “*(a) small city*”, “*(a) big city*”, and the inflection morpheme “-s” for plural marking in (50) “*three ticket(s)*”. Also, system morphemes like auxiliary verbs for negation do not appear in (51). Instead, the learner uses “*no*” to negate the verb directly, which is rather morphologically normal. In (52), the sentence elements are ordered in the typical English Subject-Verb-Object order, where the object follows the predicate verb. In (53), the sentence elements basically follow the Japanese Subject-Object-Verb order, but the adverbial of time “*mainichi*” (every day) is placed in the sentence final position, which is lowed in English, but not in Japanese.

The above examples provide the evidence that L1 may become part of the morphological realization patterns in the IL composite. It should also be noticed that the L1 influence at this level may reflect when the L1 provides aspects of abstract lexical structure at other levels. For example, in (31) above, thematic role assignment impacts predicate argument structure, and this has consequences for word order and other surface realizations of grammatical structure.

5. Conclusion

The nature of complex abstract lexical structure in relation to the roles of its subsystems at different but successive levels of speech production has ramifications for IL constructions. The present study reaches several implications about IL processes and constraints.

1. The content vs. system morpheme distinction operates in IL constructions: TL content morphemes are acquired before TL system morphemes. TL content morphemes are predicted to be acquired or to have a more TL-like distribution before TL system morphemes because in the early stage of SLA, learners mainly use TL content morphemes to express their intended meanings. Presumably, this reflects direct election of content morphemes within the larger context of language production as well as the relatively

greater transparency of abstract semantic and pragmatic features that support content morphemes.

2. The sources of morphemes determine the sequence of acquisition of the original TL complex abstract lexical structure and affect IL constructions. Directly elected morphemes (i.e., content morphemes) are acquired before indirectly elected morphemes (i.e., early system morphemes), which in turn supersede structurally assigned morphemes (i.e., late system morphemes) (Wei, 1996a, 1996b, 2000a, 2000b). Thus, structurally assigned morphemes are predicted to be the most difficult. This implies identification of various stages or sequences within developing ILs. As such, this approach to the analysis of IL makes falsifiable claims in terms of relative frequency of occurrence and accuracy of independently identifiable morpheme types.

3. Every natural language has its abstract lexical structure, and IL is no exception, but IL abstract lexical structure is composite. This is because before learners have full access to the TL abstract lexical structure, they may contribute their L1 abstract lexical structure to each of the subsystems: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. In other words, IL abstract lexical structure contains elements from learners' L1, incompletely acquired morphemes from the TL, and completely acquired morphemes from the TL.

4. TL lexical entries of any type (i.e., both content and system morphemes) preferentially supply lexical structures in the construction of the IL composite. This is because in the early stage of SLA, in acquiring TL content morphemes, learners are able to produce TL content morphemes with certain degrees of accuracy, but they may not have total access to the abstract lexical structure of these TL content morphemes. Consequently, when these content morphemes become part of the IL, they may have somewhat different lexical structures from their TL counterparts. This is possible because lexical structure can be split and rebuilt in the process of acquisition. As predicted, learners' L2 product is getting more and more target-like while moving along the IL continuum progressively and successfully.

References

- Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). New York: Academic Press.
- Dale, P. S. (1976). *Language development: Structure and function*. New York: Holt, Rinehart and Winston.
- de Bot, K., & Schreuder, R. (1993). Word production and the bilingual lexicon. In R. Schreuder, & B. Weltens (Eds.), *The bilingual lexicon* (pp. 191-214). Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.6.10bot>
- Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Garret, M. F. (1990). Sentence processing. In D. N. Osherson, & H. Lasnik (Eds.), *An invitation to cognitive science* (pp. 133-175). Cambridge, MA: MIT Press.
- Jake, J. L. (1994). Intrasentential codeswitching and pronouns: On the categorical status of functional elements. *Linguistics*, 31, 271-298. <https://doi.org/10.1515/ling.1994.32.2.271>

- Jake, J. L. (1998). Constructing interlanguage: Building a composite matrix language. *Linguistics*, 36, 333-381. <https://doi.org/10.1515/ling.1998.36.2.333>
- Joshi, A. K. (1985). Processing of sentences with intrasentential code switching. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing* (pp. 190-205). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511597855.006>
- 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*, 13-23. <https://doi.org/10.1017/S1062798700001290>
- Myers-Scotton, C. (1993 [1997]). *Duelling languages: Grammatical structure in codeswitching* (1997 edition with a new Afterword). Oxford: Clarendon Press.
- Myers-Scotton, C. (1996). *Compromise structural strategies in codeswitching*. Paper presented at Symposium on Immigrant Languages in Europe. University of Tilburg, The Netherlands.
- Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. New York: Mouton de Gruyter. <https://doi.org/10.1093/acprof:oso/9780198299530.001.0001>
- Myers-Scotton, C., & Jake, J. L. (1995). Matching lemmas in a bilingual language competence and production model: Evidence from intrasentential code switching. *Linguistics*, 33(5), 981-1024. <https://doi.org/10.1515/ling.1995.33.5.981>
- Myers-Scotton, C., & Jake, J. L. (1999a). Chichewa/English codeswitching: The “do” verb construction. In R. Finlayson (Ed.), *African mosaic* (pp. 406-417). Pretoria: University of South Africa Press.
- Myers-Scotton, C., & Jake, J. L. (1999b). *Giving structure to creoles*. Paper presented at the Annual Conference on Pidgin and Creole Language, Los Angeles.
- Myers-Scotton, C., & Jake, J. L. (2000a). Four types of morpheme: Evidence from aphasia, codeswitching, and second language acquisition. *Linguistics*, 38(6), 1053-1100. <https://doi.org/10.1515/ling.2000.021>
- Myers-Scotton, C., & Jake, J. L. (2000b). Testing a model of morpheme classification with language contact data. *International Journal of Bilingualism*, 4(1), 1-8.
- Myers-Scotton, C., & Jake, J. L. (2001). Explaining aspects of codeswitching and their implications. In J. Nicol (ed.), *One mind, two languages; Bilingual language processing* (pp. 91-125). Oxford: Blackwell.
- Pinker, S. (1989a). *Learnability and cognition*. Cambridge, MA: MIT Press.
- Pinker, S. (1989b). Resolving a learnability paradox in the acquisition of the verb lexicon. In M. L. Rice, & R. L. Schiefelbusch (Eds.), *The teachability of language* (pp. 13-61). Baltimore, MD: Brushwood Graphics.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description* (Vol. 3, pp. 57-149). Cambridge: Cambridge University Press.
- Wei, L. (1995). *Chinese and Japanese learners' interlanguage data* (Unpublished corpus). University of

- South Carolina, Columbia, SC.
- Wei, L. (1996a). Organizing principle behind codeswitching and interlanguage development in early adult second language acquisition. In J. Arnold, R. Blake, B. Davidson, S. Schwenter, & J. Solomon (Eds.), *Sociolinguistic variation: Data, theory, and analysis* (pp. 417-4310). Stanford, CA: CSLI Publishers.
- Wei, L. (1996b). *Variation in the acquisition of morpheme types in the interlanguage of Chinese and Japanese learners of English as a second language* (Unpublished doctoral dissertation). University of South Carolina, Columbia, SC.
- Wei, L. (2000a). Unequal election of morphemes in adult second language acquisition. *Applied Linguistics*, 2, 106-140. <https://doi.org/10.1093/applin/21.1.106>
- Wei, L. (2000b). Types of morphemes and their implications for second language morpheme acquisition. *International Journal of Bilingualism*, 4(1), 29-43. <https://doi.org/10.1177/13670069000040010301>
- Wei, L. (2002). The bilingual mental lexicon and speech production process. *Brain and Language*, 81, 691-707. <https://doi.org/10.1006/brln.2001.2557>
- Wei, L. (2003). Activation of lemmas in the multilingual mental lexicon and transfer in third language learning. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *The multilingual lexicon* (pp. 57-70). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-0-306-48367-7_5
- Wei, L. (2009). Code-switching and the bilingual mental lexicon. In B. E. Bullock, & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code-switching* (pp. 270-288). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511576331.017>
- Wei, L. (2015). *Interlanguage: The abstract level in language acquisition*. New York: The Edwin Mellen Press.
- Wei, L. (2020). *The bilingual mental lexicon*. UK: Cambridge Scholars Publishing.
- Wei, L. (2021). Composite abstract lexical structure in interlanguage production. *Studies in Linguistics and Literature*, 5(1), 81-93. <https://doi.org/10.22158/sll.v5n1p81>