Language Contact and Linguistic Aspects of Bilingualism

Language Contact and Linguistic Aspects of Bilingualism

By

Longxing Wei

Cambridge
Scholars
Publishing

Language Contact and Linguistic Aspects of Bilingualism

By Longxing Wei

This book first published 2024

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2024 by Longxing Wei

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN: 978-1-0364-0912-8
ISBN (Ebook): 978-1-0364-0913-5

Dedicated to my grandsons,

Dylan and Gavin

CONTENTS

PREFACE

This book proposes a model of contact linguistics and a relatively new approach to the study of bilingualism. It examines the universal principles governing grammatical structures of languages in contact, differentiates the lexical and grammatical features of morphemes in relation to lexical borrowing, mixed languages, bilingual processing and representation, codeswitching (CS), second language (L2) morpheme acquisition, and interlanguage (IL). Different from most existing models of contact linguistics, the approach presented in this book describes and explains some outstanding linguistic aspects of bilingualism with a focus on the mechanisms of the bilingual mind during bilingual processing and production at three levels of abstract lexical structure: lexical-conceptual structure, predicate-argument structure, and morphological realization patterns. It provides evidence of unequal activation of the languages in contact, one language being more activated than the other in grammatical frame building, and evidence of bilingual "lemmas" (i.e., language-specific abstract entries in the mental lexicon about lexemes) in contact and their activation and mutual influence during CS, L2 learning and IL development. To do so, in addition to various outstanding models of contact linguistics and bilingualism most relevant to the current topics, four specific models are adopted and tested for the studies presented in this book: (1) The Matrix Language Frame (MLF) Model (Myers-Scotton, 1993b [1997], 2002) proposes a shared, distributed, asymmetrical model for the bilingual mental lexicon, which claims that the bilingual's two languages are never equally activated, one being activated at both grammatical and lexical levels and playing an absolute role in providing the sentential frame of utterances involving CS, and the other being activated at a lexical level for some psycholinguistic and conceptual reasons and supplying only certain lexical items. (2) The 4-M Model (Myers-Scotton, 2002; Myers-Scotton and Jake, 2000a, 2000b, 2001, 2016) offers a different perspective on the critical differences among morpheme types by refining the content vs. system morpheme distinction and adding some precision to the treatment of morpheme types in CS under the MLF Model. This model emphasizes hypotheses on how morphemes differ in their election in language processing at an abstract level and aims to show how the surface roles of different types of morphemes can be linked to a model of language

production. (3) The Abstract Level (AL) Model (Myers-Scotton, 2002) addresses the nature of the critical distinction between whether the lemmas underlying morphemes are conceptually activated to convey speaker intentions (i.e., morphemes realizing speakers' intended semantic/pragmatic feature bundles) or structurally assigned to build grammatical structure (i.e., morphemes realizing grammatical functions). This model offers some important implications for bilingual phenomena, including the nature and activity of the bilingual mental lexicon and the bilingual processing and production. (4) The Bilingual Lemma Activation (BLA) Model (Wei, 2006a, 2006b, 2002, 2020) claims that lemmas in the bilingual mental lexicon are language-specific and are in contact during a discourse involving CS, L2 learning, and IL production. This model explains the role of lemma activation of particular lexical items in the bilingual mental lexicon and emphasizes that language-specific lemma activation must mediate between conceptualization and speech formulation as an indispensable level of bilingual processing and production.

From the perspective of psycholinguistics, this book presents some hypotheses about the abstract nature of bilingual mental organization of the two (or more) languages and its activity in bilingual production and endorses a lexical approach to explaining the basis of bilingual language structure. It tests the organizing principles governing all linguistic systems of languages in contact. The approach and studies introduced in this book move away from typological or deterministic models towards an explanatory model. Different from most previous studies, this book views language transfer as lemma transfer by endorsing the notion that the internal structure of lemmas is abstract and complex in that some aspects of abstract lexical structure are relevant to, or activated by, different psycholinguistic properties. What becomes most important in this book is that bilingualism, specifically including L2 morpheme acquisition, CS behavior and IL development, is regarded as a linguistic outcome of language contact at a rather abstract level.

This book demonstrates and explains why bilingualism should be studied in the domain of contact linguistics and why L2 acquisition and bilingual processing and production are governed by the same underlying organizing principles governing all linguistic systems of languages in contact. It is expected to contribute to the current research in contact linguistics in general and linguistic aspects of bilingualism resulting from language contact in particular.

—L. W.

ACKNOWLEDGMENTS

My greatest intellectual debt is to Carol Myers-Scotton and Janice L. Jake for their inspirational input to my research in contact linguistics, bilingualism, and second language acquisition. I have benefited tremendously from Myers-Scotton's Matrix Language Frame Model of codeswitching and language contact, Myers-Scotton and Jake's 4-M Model of speech production and Abstract Level Model of language processing, and Jake's Composite Matrix Language Model of second language acquisition. Their theoretical frameworks and linguistic analytical methods have significantly guided me in exploring the nature and activity of the bilingual mental lexicon in speech production and second language acquisition and in describing and explaining particular linguistic aspects of bilingualism resulting from abstract language contact. I am also grateful to them for their significant input to my initiation of the Bilingual Lemma Activation Model.

My intellectual debt is also to numerous outstanding linguists whose theoretical models and significant research findings which I have drawn on heavily. Without their important linguistic insights, this book would remain at a rather superficial and unbalanced level.

I am very grateful to the sponsored and anonymous reviewers of the book manuscript for their insightful comments, valuable input and helpful suggestions that have helped me avoid my personal pitfalls to a great extent. I thank them for offering me a rewarding experience in writing this book. My gratitude extends to Cambridge Scholar Publishing, the proofreading team, Adam Rummens, the Commissioning Editor, Amanda Millar, the Typesetting Manager, and Sophie Edminson, the Cover Designer, for their great encouragement, specific guidelines, and professional support.

I am also deeply grateful to my wife, Xuexin Liu, for her always being there for me. Without her persistent spiritual and phycological support and endless patience with me, the completion of this book would have been impossible.

With my gratitude to those having been there for me since the initiation of this book project, I will remain fully responsible for the present work.

—L. W.

SYMBOLS AND ABBREVIATIONS

/	or
+	carrying the feature
±	carrying or not carrying the feature
−	not carrying the feature
>	to be acquired before …
*	ungrammatical
4-M	four types of morpheme
ABS	absolutive
ACC, ACCUS	accusative
ADJ	adjective
ADVP	adverb phrase
AFFIRM	affirmative
AFFIX	affix
AL	abstract level
AP	adjective phrase
APPL	applied
ART	article
ASP	aspect
ASSOC	associative
AUX	auxiliary
BLA	bilingual lemma activation
CA	contrastive analysis
CAUS	causative
CL	noun class
CLASSIF	noun classifier
CM	concept mediation
COMP	complementizer
CON	connective
CONJ	conjunction
CONSEC	consecutive
COP	copula
CS	intrasentential codeswitching
CSM	concept selection model
DAT	dative
DEF	definite

DEM	demonstrative
EA	error analysis
EL	embedded language
EMPH	emphatic
ESL	English as a second language
FUT	future
FV	final vowel
IL	interlanguage
IMP	imperative, imperfect
INDEF	indefinite
INDIC	indictive
INF, INFIN	infinitive
INFL	inflection
INT	interjection
INTERROG	interrogative
IP	inflectional phrase
L1	first language
L2	second language
LFG	lexical functional grammar
LOC	location
MHM	Modified Hierarchical Model
ML	matrix language
MLF	matrix language frame
MP	minimalist program
N	noun
NEG	negative
NL	native language
NOM	nominative
NP	noun phrase
OB	object
PART	participle
PART, PARTIC	particle
PST, PAST	past
PERF	perfect
PL	plural
POSS	possessive
PP	preposition phrase
PREP	preposition
PRES	present
PROG	progressive
PRT	past participle

QUE	question
RECIP	recipient
RH	revised hierarchical
SAM	Shared Asymmetrical Model
SbPM	selection by proficiency model
SFP	sentence final particle
SG	singular
SLA	second language acquisition
SUBJ	subject
SUBJUNCT	subjunctive
SOV	subject object verb
SVO	subject verb object
TA	tense-agreement marker
TI	transitive inanimate verb
TIM	time
TL	target language
TNS	tense
TOP, TOPIC	topic
TRANS	transitive
UC	unified competition
UG	universal grammar
V	verb
VP	verb phrase
WA	word association

1

TOWARD A NEW APPROACH TO THE STUDY OF BILINGUALISM

1.0 Introduction

Language contact is a commonly observed social and linguistic phenomenon. It occurs when speakers of different languages or dialectal varieties of the same language interact with each other, leading to transfers of linguistic features. Thus, language contact is recognized as a major source of language change (Weinreich, 1953; Haugen, 1953; Gramley, 2019; Winford, 2003; Myers-Scotton, 2002). Major language contact phenomena include language convergence, lexical borrowing and relexification.[1] Major outcomes of language contact include attrition, pidgins, creoles, mixed languages, intrasentential codeswitching (CS for short)[2] and interlanguage (IL). Language change induced by language contact is commonly observed in pronunciations, vocabulary, and grammatical structures. As one of the pioneers of language contact studies, Weinreich (1953) is the first to notice that second language (L2) learners do not make a distinction between their first language (L1) and L2 linguistic forms, especially during early stages of L2 learning. Another outcome of language contact is that prolonged

[1] Language contact phenomena include lexical borrowing, grammatical convergence, first language attrition, mixed languages, pidgins, creoles, codeswitching, and so on. Contact linguistics is a critical investigation of what happens to the grammars of languages when speakers or language learners are involved in a language contact situation. This book discusses various language contact phenomena at a rather abstract level.

[2] "Intrasentential codeswitching" occurs within the same sentence (i.e., within the sentence boundary) or sentence fragments. Bilinguals may switch to another language (i.e., another code) at a certain point in their utterance production. "Intersentential codeswitching" occurs across sentences (i.e., across sentence boundaries). Bilinguals often engage in intersentential and/or intrasentential codeswitching (i.e., "language mixing" or "code mixing", the terms as used by other researchers) when communicating with another person who also speaks both languages.

language contact unavoidably leads to bilingualism or multilingualism, which becomes widespread in today's global phenomenon of language contact due to international exchange in all professions and bilingual education. The traditional count of language contact is that there must be some face-to-face interaction; otherwise, a contact explanation for synchronic variation or diachronic change will not be possible. Other possibilities for language contact in the modern world have also been recognized, such as novel means of worldwide travel, mass communication, and contacts through written language (Thomason, 2013). Language contact phenomena can be manifested in various domains, such as second language acquisition (SLA), bilingual processing and representation, IL production, bilingual conversational exchange, etc.

The study of language contact is called "contact linguistics". One of the most important notions underlying contact linguistics is cross-linguistic influence which drives language change. Language contact can frequently trigger or influence the development of grammar in a number of ways. Thus, as generally assumed, it is the grammatical replication through language contact which induces grammatical change (Heine and Kuteva, 2005). According to Matras (2009), the study of language contact becomes valuable if it describes and explains "grammar" in terms of its inner functions and structure resulting from language contact. According to Law (2014), a language contact phenomenon can be recognized if people have access to at least two linguistic codes in interaction, which results in one code being more similar to another code. Siemund (2008) holds that a realistic recognition of any outcome of language contact is that whatever linguistic material is transferred in a language contact situation must experience some perceivable modification.

Language contact is not a homogeneous phenomenon. It is driven by various social parameters and linguistic factors. From a social parametric perspective, there are various social factors involved in a language contact situation, such as the number of speakers of their respective languages in contact, the social status of the linguistic groups, the relative prestige of the languages involved, the length for the contact between the two social groups and the intensity of the contact (Siemund, 2008; Winford, 2003). It is such social factors which determine types of language contact situations and contact induced linguistic change (Thomason and Kaufman, 1988). From a cross-linguistic perspective, depending on particular language contact situations, a multitude of formal and functional constraints may determine and restrict the outcome of language contact. Markedness and complexity which may constrain linguistic change and the relevant mechanisms are among the issues being taken up by researchers in contact linguistics.

However, it has been hypothesized that these constraints are governed by a set of universal linguistic principles (Jake, 1998; Myers-Scotton, 1993 [1997], 2002; Wei, 2018, 2015). As studied in contact linguistics, language contact induced linguistic transfer or change may occur at various levels of language, such as phonology, morphology, syntax, vocabulary, mental lexicon,[3] etc. These levels of linguistic transfer or change will result in a wide variety of outcomes of language contact, such as pidgins, creoles, language attrition, lexical borrowing, mixed languages, CS, IL, etc. (Myers-Scotton, 2002; Winford, 2003; Heine and Kuteva, 2005; Siemund, 2008; Matras, 2009, among others).

As commonly recognized, language contact situations have a number of further consequences for the languages involved. Bilingualism is another outcome of languages in contact. As observed worldwide, two major kinds of bilingualism result through language contact. One usually sorts itself out with one language winning over the other (e.g., English over the other languages it has been in direct and prolonged contact with) and the other involves the languages entering some sort of equilibrium for social or political reasons (e.g., French and Flemish in Belgium, and English and Cantonese in Hongkong). Such a kind of bilingualism has been identified as a stable bilingualism called diglossia where two languages (e.g., Spanish and Guarani in Paraguay) or two distinct varieties of the same language (e.g., Swiss and High German in Switzerland) are used side by side for their particular functions in separate spheres of life. Most previous studies have focused on social and political aspects of bilingualism in society in terms of the direction of contact as determined by factors of social prestige (i.e., the superstrate language vs. the substrate language).

Social factors are, of course, directly relevant to contact induced language change or bilingualism. It would be impossible to understand the linguistic outcomes of language contact without exploring the social factors involved and the complexities of the social constellations. As well as being informed by the general scope of studies and insights in contact linguistics as outlined above, this book focuses on certain outstanding "linguistic" aspects, rather than "social" aspects, of bilingualism through language

[3] "The mental lexicon" stores the speaker's declarative knowledge about the words and idioms in his/her language. For each item known to the speaker, in addition to declarative knowledge about te word's meaning, the mental lexicon contains its lemma information about its syntax and morphology which is necessary for constructing the word's syntactic environment (Levelt, 1989, 6). Lemmas in the bilingual mental lexicon are assumed to be language-specific. In other words, each lemma is tagged for a specific language and supports the realization of an actual lexeme.

contact. As commonly observed and evidenced, the languages in contact, especially in direct and prolonged contact, are susceptible to contact effects. One of the fundamental assumptions is that languages are not arbitrary collections of properties at various linguistic levels but obey certain laws of language architecture governed by some universal language principles (Gass and Ard, 1984; Thomason, 2001; Croft and Cruse, 2004). Contact induced language change should also be restricted by certain universal language principles (Myers-Scotton, 2002; Siemund, 2008). This book focuses on the linguistic principles which govern and determine contact induced language change and bilingual mechanisms. Five specific linguistic aspects of bilingualism are to be discussed in this book: bilingualism and cross-linguistic influence, linguistic characteristics of adult bilingualism, CS, bilingualism and SLA, and IL. In addition to various applicable models of contact linguistics in general and bilingualism, SLA, CS and IL in particular, this book adopts and tests four specific models: the Matrix Language Frame (MLF) Model (Myers-Scotton, 1993 [1997]), the Four Types of Morpheme (4-M) Model (Myers-Scotton and Jake, 2000a, 2000b), the Abstract Level (AL) Model (Myers-Scotton, 2002), and the Bilingual Lemma Activation (BLA) Model (Wei, 2006a, 2006b, 2002, 2020).[4]

1.1 An Overview of Major Topics in Bilingualism and Key Issues

Bilingualism is commonly recognized as a linguistic outcome of extensive language contact or, to be more specific, a product of extensive contact between people speaking different languages or dialects. Bilingualism is a universal linguistic and sociolinguistic phenomenon and can be observed at the national, community and individual levels. Bilingualism becomes one of the most important research topics in modern linguistics and raises fundamental issues of the human language faculty, language acquisition, SLA, bilingual speech production process and bilingual language use. Some researchers examine bilingualism from the sociolinguistic perspective by treating bilingualism as a social or sociopolitical issue. Other researchers examine bilingualism from the theoretical linguistic perspective by treating bilingualism as a linguistic outcome of languages in contact. This book focuses on the outstanding linguistic aspects of bilingualism through

[4] Most naturally occurring instances of adult second language acquisition and interlanguage development for the discussions in this book were collected over a period of two years as part of the research project under NSF grant SBR-9319780 to Carol Myers-Scotton and Janice L. Jake.

language contact. The specific topics in bilingualism and key issues to be discussed in this book include the following.

Intrasentential codeswitching. CS is the alternative use of two linguistic systems within a clause, that is, one language's items are switched into another language within sentence boundaries. Thus, bilingual speech is defined as any clause that contains elements from two or more languages (Myers-Scotton, 2005). CS is observed as a common linguistic strategy employed widely by bilingual (or multilingual) speakers in their speech production. Researchers are interested in such a bilingual verbal strategy and investigate whether there is a "universal" grammar that presumably determines and delimits the range of CS, that is, whether there are grammatical constraints on code-switched utterances involving different language pairs. Some key issues remaining poorly understood include what does and does not occur in CS and how CS is grammatically framed or governed. Theoretical accounts differ in making predictions about the items to be switched and the location and the directionality of switched items due to the contradictory empirical evidence. Below are some outstanding models of CS.

The surface-based model. In early analysis of CS, Pfaff (1979) and Poplack (1980, 1981) propose the surface-based model and postulate two constraints: *the Free Morpheme Constraint* and *the Equivalence Constraint*. The former constrains the switchability of morphemes: free morphemes, unlike bound morphemes,[5] can be switched (Poplack, 1980; Poplack and Sancoff, 1988). The latter restricts CS to locations in the clause where surface structures of the languages involved remain intact: switching is not possible where the surface structures (i.e., surface word orders) of the two languages differ (Lipski, 1977; Pfaff, 1979; Gumperz, 1982; Woolford, 1983). Thus, surface-based models stress linearity. Many naturally occurring CS instances do not support these two constraints. Below are just a few of them.

[5] "Bound morphemes" cannot stand alone as words, that is, they must be attached (i.e., bound) to free-standing morphemes (i.e., words) to form inflections (e.g., in English, nouns are inflected for plural marking, adjectives and adverbs are inflected for degree marking, and verbs are inflected for tense/aspect/voice marking) and derivatives (e.g., in English, prefixes and suffixes are attached to free-standing morphemes to derive one word from another, and they allow words to change parts of speech). "Free morphemes" can stand alone as words (i.e., they are free-standing items).

[1] A Ne mI **help** -e.
3PL AUX be me help -PRES PROG
"They are helping me".
(Adaŋme/English; Nartey, 1982, 185)

[2] I'm **lav**-ing **pandekege**-s.
I'm have-PROG pandekege-PL
"I'm having pancakes".
(English/Danish; Peterson, 1988, 481)

[1] and [2] are the counter-examples to *the Free Morpheme Constraint*. In [1], *help* is a free morpheme but *-e* for aspect marking is a bound morpheme. In [2], *lav* (have) is a free morpheme but *-ing* for aspect marking is a bound morpheme. Such examples show that neither the bound vs. free distinction nor typological distinctions are relevant for permissible intraword switching sites.

[3] j'ai vu un ancient **tilmid ajali**.
I PAST see an old student of mine
"I saw a former student of mine".
(French/Moroccan Arabic; Bentahila and Davies, 1983, 319)

[4] Unaweza kumpata amevaa nguo nyingine bright …
you can INF find AFFIRM wear clothes other
"You can find her (she is) wearing other bright clothes …"
(Swahili/English; Myers-Scotton, 1993b, 28)

[3] and [4] are the counter-examples to *the Equivalence Constraint*. In [3], the French adjective *ancient* (old) precedes the Moroccan Arabic NP *tilmid djali* (student of mine), which violates the Moroccan Arabic word order where adjectives follow the nouns they modify. As shown in [4], though Swahili and English have a different order within the NP, noun-adjective order in Swahili, but adjective-noun order in English, CS involving English adjectives is not prohibited from occurring. It seems apparent that CS may occur at a considerable number of inequivalent points. The CS literature provides many naturally occurring code switches between the languages which do not share the common surface structures as long as there is a specific discourse reason for them to occur.

The government and binding model. Unlike the surface-based model, the government[6] and binding[7] model stresses dependency rather than linearity. Klavans (1983), Woolford (1983), DiSciullo, Muysken, and Singh (1986), Appel and Muysken (1987), and Pandit (1990) propose the explicit formulation of this non-linear approach to CS constraints. According to this approach, "There cannot be a switch between two elements if they are lexically dependent on each other" (Appel and Muysken, 1987, 124), and whenever a constituent X governs a constituent Y, both constituents must come from the same language; however, governed elements can be switched when there is a neutralizing element such as a determiner (DiSciullo, Muysken, and Singh, 1986).

Though the government and binding model looks beyond surface liner ordering, it still considers CS as basically a syntactic phenomenon following the same structural constraints evident in monolingual surface structure. Some naturally occurring CS instances do not support this model. Below are just a few of them.

[5] **Parents** te **depend** honda [E].
parents on depend be/become AUX
"It depends on the parents".

---

[6] Government theory deals in essence with the relationship between a head and its complements and defines relationships in other sub-theories. Without explaining the intricacies of government theory, we may simply note that government is a configurationally defined notion, which can be illustrated by a simplified and schematic representation of the structure for a passive sentence. There are two structures, one for the proposed D-structure and the other for the proposed S-structure, with lexical material distributed at each level:
D-structure 'e' was caught the burglar by the police ('e': empty slot)
S-structure the burglar was caught 't' by the police ('t': trace)
At D-structure, the sentence is identified as a finite tensed sentence, indicated by the labelling 'Infl(ection)' '+tns(tense)'. In this configuration, Infl(+tns) governs the subject NP, V governs the object NP, and the NP in the PP is governed by the preposition. At S-structure, the Case theory must be satisfied. The relation of government plays an important role in various sub-theories.

[7] "Binding" can be defined as: An element α binds an element β if and only if α c-commands β, and α and β corefer. For example, in *Mary saw her sister*, the NP *Mary* c-commands *her* because the first parent of the NP, S, contains *her*. *Mary* and *her* are also coreferential (i.e., referring to the same person). Therefore, *Mary* binds *her*. One the other hand, the sentence *The father of Mary likes herself* is ungrammatical because *Mary* does not c-command *herself*, that is, they have no binding relationship despite the fact that they corefer. "Binding" is used, along with particular binding principles, to explain the ungrammaticality.

[6] **Some Englishmen traditional Indian women**-ko passand karaten hain.
some English traditional Indian women-ACCUS like do are
"Some Englishmen like traditional Indian women".
(Hindi/English; Pandit, 1990, 44).

[7] Ulikuwa ukiongea **a lot of nonsense**.
"You were talking a lot of nonsense".
(Swahili/English; Myers-Scotton, 1993b, 44)

In [5], the switching between V and its NP constituent occurs, where the compound-verb construction contains a Panjabi operator verb *honda*. In [6], CS takes place in the VP: the English NP *traditional Indian women* plus the Hindi suffix for the accusative case are governed by the Hindi V *passand* (like). In [7], the English NP *a lot of nonsense* is governed by the Swahili V *ulikuwa ukiongea* (you were talking).

It seems apparent that the government and binding model is not adequate enough to explain CS. A purely syntactic approach is still too close to the surface-based approach.

The subcategorization model. To discount the importance of surface level equivalences, Muysken (1988, 1991) and Azuma (1991) propose the subcategorization model. According to Muysken (1988, 1991), differences in switching patterns are hard to explain in purely syntactic terms because switched elements are constrained by lexical subcategorizations. According to such a constraint, switching is possible if the demands made by individual lexical items with respect to their syntactic environment as expressed in their subcategorization frames are satisfied. Thus, "categorial equivalence" is required where specific categories are lexically specified in the subcategorization frame of a lexical item (Muysken, 1991). In other words, lexical elements impose certain requirements on their environments, and switched elements are constrained by lexical subcategorizations. Azuma (1991) proposes *the Frame-Content Hypothesis*, which identifies two stages in CS: the planning frame-building stage where closed class items are accessed and retrieved, and the content-word insertion stage where content words are inserted into the planning frame. According to this hypothesis, the closed class elements are essential members of the planning frame, and CS takes place in the stage of content-word insertion, where the content word from another language is inserted into the available syntactic slot in the planning frame. The key assumption underlying this model is that the languages involved in CS play unequal roles, one language participating in the frame-building stage to satisfy that language's subcategorization frames,

and the other language participating in the content-word insertion stage to provide content words from that language. What is emphasized in this model is that the two stages are strictly serial and non-interactive. Different from the other models, the subcategorization model relies on a traditional open-closed class distinction, posits a two-step process in which the subcategorization requirements of the frame-building take precedence and distinguishes the differential roles of the two languages involved in CS, only one language directing the frame building. However, like the government and binding model, this model, though it departs from the surface-based model by looking beyond surface linear ordering, still considers CS as basically a syntactic and lexical phenomenon as revealed in monolingual surface structures.

The matrix language frame model. Myers-Scotton (1993b [1997], 2002) proposes the Matrix Language Frame (MLF) Model to describe and explain "classic" CS, which is defined as the use of morphemes from dialects or varieties of two or more languages in the same conversation with the provision that speakers are proficient enough in these varieties to produce well-formed monolingual utterances in them. Substantially different from the earlier approaches to CS that consider CS phenomena as either surface, linear constructions, as typological correspondences, or as derived only from the same structural relations posited for monolingual syntax, the MLF Model explains CS at a rather abstract, pre-syntactic level in bilingual sentence production. This model is structured by several general principles and specific hypotheses. As proposed in this model, the constraints on CS are interrelated to each other and all types of CS constituents relate to the same constraints. In other words, all CS constituents are defined and constrained under the same categories (Myers-Scotton, 1990, 1991, 1993b; Myers-Scotton and Jake, 2000a, 2000b, 2001). The MLF Model is adopted as one of the most relevant and effective approaches to the study of linguistic aspects of bilingualism through language contact. 1.2 offers a more detailed overview of this model.

The minimalist model. MacSwan (1997, 1999, 2000, 2005) proposes the minimalist model of CS, which draws on the premise of a syntactic theory proposed by Chomsky (1995): the Minimalist Program (MP).[8] According

[8] The Minimalist Program (MP) published by Chomsky (1995) consists of four recent essays, including the previously unpublished "Categories and Transformations". These essays attempted to situate linguistic theory in the broader cognitive sciences. It is in these essays where the minimalist approach to linguistic theory is formulated and progressively developed. In the context of generative grammar, the MP draws on the minimalist approach of the principles and parameters model to explore the hypothetically optimal and computationally efficient design of the human language

to the MP, "All learning is lexical, and all parameters are micro-parameters associated with individual lexical items" (MacSwan, 2000, 44). "Nothing constrains code switching apart from the requirements of the mixed grammars" (MacSwan, 2005, 69).[9] Building on the MP, some researchers have approached CS from the perspective that CS occurs without any additional rule-based constraints, and clauses containing switched items arise from the insertion of lexical items from another language into the phrase structure (Toribio and Rubin, 1996; MacSwan, 1997, 1999; Ritchie and Bhatia, 1999). Researchers adopting the minimalist model share the claim that the principles and requirements of each participant grammar in any instance of CS, together with the principles and requirements of Universal Gramma (UG), are sufficient enough for explaining every aspect of CS.

The major theoretical assumption underlying the minimalist approach is simply a linguistic phenomenon where two monolingual systems are in contact, and each monolingual system must be preserved in order for the mixed constituents to be grammatical. If the minimalist approach regards well-formedness or ill-formedness of mixed constituents simply as an issue of *the Equivalence Constraint* (Poplack, 1980), it misinterprets the asymmetry between the Matrix Language (ML) and the Embedded

---

faculty. For "economy", the MP draws on the further development of peripheral aspects of transformational grammar involving "economy of derivation" and "economy of representation". Economy of derivation is a principle which governs movement. That is, movement only occurs to match interpretable features with uninterpretable features. For example, the plural inflection on regular English nouns (e.g., *cats*) is an interpretable feature. The word *cats* can only refer to two or several cats (i.e., more than one), rather than a single cat, and this morphological inflection contributes to meaning, making the noun interpretable. English grammar requires verbs in present tense to be inflected according to the number of their subject (e.g., *Cats eat fish* vs. *A cat eats fish*), but this information only becomes interpretable once a grammatical relationship is formed between the subject and the verb. Economy of representation is a principle which states that all grammatical structures must exist for a purpose (i.e., a sentence structure should be no larger or more complex than required to satisfy grammatical constraints). According to this principle, grammatical constraints are equivalent to constraints on the mapping between te conceptual/intentional and sensori-motor interfaces in the optimal system as explored in minimalism.

[9] According to MacSwan, "The requirements are simply carried along with the lexical items of the respective systems. We may think of the system formally as one in which the grammar used for code switching consists of the union of the two lexicons plus the invariant computational system, with no mediating mechanisms needed" (MacSwan, 2000, 45).

Language (EL) as proposed in the MLF Model. The well-formedness or ill-formedness of the following CS examples needs to be explained in terms of this asymmetry.

[8] Uchi wa **whole chicken** o kau noyo.
we TOPIC whole chicken ACC buy tag
"We buy a whole chicken".

[9] *Watashi ga kata **the** hon wa takai.
I NOM bought the book TOPIC expensive
"The book I bought is expensive".
(Japanese/English, Azuma, 1993)

[10] *The students had visto la pelicula italiana.
"The students had seen the Italian movie".

[11] *Los estudiantes habían seen the Italian movie.
"The students had seen the Italian movie".
(MacSwan, 2000, 42)

In [8], *whole chicken* is switched into the morphosyntactic (i.e., grammatical) frame provided by Japanese as the ML, where the object NP introduced by the particle *o* (ACC) which is the ML system morpheme for case marking and occurs before the verb *kau* (buy), rather than after it. This bilingual CP is well-formed because it is the ML which controls the morphosyntactic frame and provides the system morpheme. In [9], *the* is the EL system morpheme, which cannot be switched alone, violating *the System Morpheme Principle* (see this principle introduced in 1.2), and occurs before *hon* (book), violating *the Morpheme Order Principle* (see this principle introduced in 1.2) because in Japanese, the ML, the article *the* (i.e., the particle in Japanese) must be placed right after the noun for case marking. [10] is predicted not to occur because the ML system morpheme *had* must be mapped onto the ML content morpheme *seen* to mark the past perfect aspect of the verb. Without the ML content morpheme, the ML system morpheme cannot appear by itself. [11] is predicted not to occur for the same reason. Without the ML verb, the ML system morpheme *habían* (had) alone cannot mark the past perfect aspect of the verb from another language. It is obvious that *the Equivalence Constraint* is not sufficient enough to explain such ill-formedness. The minimalism focuses on "minimal" use of theoretical assumptions and allows only those suppositions which correspond to "virtual conceptual necessity". Thus, "no code-switch-specific constraints"

are assumed under the minimalist approach (MacSwan, 2000, 51).[10] Consequently, the minimalist approach alone cannot explain what causes such ill-formedness as explained above.

The minimalist approach has its value, but to make such an approach account for the substantial body of CS data, it needs to be modified to incorporate a basic asymmetry between the participant languages in CS: only one language provides the morphosyntactic frame structuring clauses containing switched items. This language is recognized as the ML, and its morphosyntactic frame is referred to as the Matrix Language Frame (MLF). The limitation of the minimalist approach to CS is that it only attempts to account for switching from one monolingual constituent to another, dismissing constituent from one language switched into a constituent structured by another language. Without the construct of the ML, this approach alone is not adequate enough to account for both singly occurring forms and switched full constituents. For full discussions of the minimalist approach to CS, see Jake, Myers-Scotton and Gross (2002) and Wei (2015, 2020).

The bilingual lemma activation model. Wei (2006a, 2006b, 2002, 2020) proposes the Bilingual Lemmas Activation (BLA) Model[11] to explain bilingual speech production process involving CS. The BLA Model is derived from Levelt's (1989) speech production model, Myers-Scotton and Jake's (1995) bilingual language competence and production model and Wei's (2002) bilingual speech production model. This model consists of four sequentially connected levels of speech production: the conceptual level → the lemma level → the functional level → the positional level. Each level plays a particular role in the bilingual speech production process. The major premise of this model is that bilingual lemmas are language-specific which are in contact for activation in CS. The BLA Model is adopted as one

[10] In minimalism, parameters are restricted to the lexicon (Chomsky, 1991) and linguistic variation only occurs in the morphological properties of the lexicon (Borer, 1984). Two components are central to this model: a computational system of human language (C_{HL}), assumed to be universal, and a lexicon, which contains the idiosyncratic differences across languages. This suggests that I-language is fixed and invariant across languages. In this model, all surface differences in word order or morphological realization patterns are caused by the rearrangement of elements in the syntactic structure resulting from movement rules which are necessarily triggered by lexically encoded morphological features. Thus, according to the MP, "All learning is lexical, and all parameters are micro-parameters associated with individual lexical items" (MacSwan, 2000, 44).

[11] Wei's BLA Model was initially proposed to relate CS to second language morpheme acquisition and IL development. This model applies the major principles of the MLF Model to IL studies at an abstract level (see Wei (2015)).

of the most relevant and effective approaches to the study of linguistic aspects of bilingualism through language contact. 1.5 offers a detailed overview of this model.

Second language morpheme acquisition. L2 grammatical morpheme acquisition in relation to bilingual development has been one of the major topics in bilingualism research. As commonly observed, L2 development appears to proceed in an orderly fashion. Early L2 researchers have discovered that there is a common or natural order of English grammatical morphemes to which L2 learners, despite their different ages and L1 backgrounds, adhere. Early studies of L2 grammatical morpheme acquisition have focused on discovering a distinct and consistent order in which L1 and L2 learners of English acquire a set of grammatical morphemes. Brown's (1973) longitudinal study of three children learning English as their L1 is recognized as the starting point. He identifies fourteen grammatical morphemes and their developmental sequence by the three children and finds there is a distinct acquisition order. Dulay and Burt (1973) investigate the grammatical morpheme acquisition order[12] of young learners of English as an L2 and find a distinct acquisition order. Dulay and Burt (1974) and Bailey, Madden and Krashen (1974) provide evidence in support of the existence of a universal natural order of grammatical morpheme acquisition regardless of learners' ages, L1 backgrounds, amount/duration of English as a second language instruction or exposure/access to English. Larsen-Freeman (1975) investigates the English grammatical morpheme acquisition by twenty-four adult students at an elementary level of English proficiency enrolled in an intensive English program at the English Language Institute of the University of Michigan. Those students represent four language backgrounds: Arabic, Japanese, Persian and Spanish, 6 subjects for each language. One of her findings is that despite some evidence of individual and language group variability, the coefficients of concordance produced among the language groups on the tasks within the study are significantly high. Thus, learners' L1 backgrounds does not seem to radically influence the acquisition order. Those studies led Krashen to propose the "Natural Order" of L2 grammatical morpheme acquisition (1977, 149). Among more recent researchers of L2 grammatical morpheme acquisition, Wei (1996b, 2000a, 2000b) investigates the English

[12] SLA researchers like Dulay and Burt (1973, 1974a) and Bailey, Madden, Krashen (1974) have claimed that there exists an acquisition order of English morphemes to which learners of English as an L2, despite their different ages and L1 backgrounds, adhere. Thus far, however, no single cause has been shown for this phenomenon, and an explanation for the occurrence of such a morpheme acquisition order has eluded researchers.

grammatical morpheme acquisition by 30 adult native speakers of Chinese and 30 adult native speakers of Japanese. The subjects are divided into three groups for each L1 background, each group consisting of 10 individuals based on their SLA stages: pre-basic, basic, and beyond-basic, like the learner stages identified in Klein and Perdue (1993, 30-4). Wei's findings support the natural order, but his explanations differ from those provided by other researchers. He relates the accuracy orders to the nature of different types of morphemes being acquired. He also relates the natural order to the mechanisms and processes of SLA. He claims that it is the unequal projection of different types of morphemes that determines the sequence of acquisition.

The central question of these studies is whether learners show a universal pattern in the acquisition order of grammatical morphemes. The view that learners follow a universal or natural order in the acquisition of L2 morphemes remains dominant to this day. What is not known is exactly what causes such a natural sequence or why certain morphemes are predicted to be acquired before others. The perceived order of L2 grammatical morpheme acquisition itself is not the reason for the order. However, in SLA studies, more and more researchers have established the L1 influence on L2 acquisition (e.g., Odlin 1989; Ellis, 2006; Jarvis and Pavlenko, 2007; Luk and Shirai, 2009; Ionin and Montrul, 2010; Murakami and Alexopoulou, 2015; Wei, 2015; Widianingsih and Gulö, 2016). Chapter 5 offers a detailed discussion of bilingualism and SLA with a focus on L2 morpheme acquisition.

Interlanguage. The influence or interference of learners' L1 in L2 learning and sources of learner errors in have been one of the intense debates in SLA research. "The greater the differences between two systems, i.e. the more numerous the mutually exclusive forms and patterns in each, the greater is the learning problem and the potential area of interference" (Weinreich, 1953, 1). Contrastive Analysis (CA) (Lado, 1957) contributes all learner errors to learners' L1 by contrasting L1 and L2 to see the differences between them at various linguistic levels. Researchers adopting CA claim that it is the differences between the two linguistic systems involved in L2 learning which cause learning difficulty and learner errors (Dušková, 1969; Wardhaugh, 1970; Larsen-Freeman and Long, 1991). One of the major criticisms of CA is that interlingual interference from L1 is not the only source of learner errors because cross-linguistic differences may not always cause learning difficulty and learner errors may not be related to learners' L1. That is, cross-linguistic differences cannot always predict learner errors.