THAI-ENGLISH CODESWITCHING: A HAWAII CASE STUDY

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

LINGUISTICS

AUGUST 2005

By
Sumitra Suraratdecha

Dissertation Committee:

Michael L. Forman, Chairperson
Albert J. Schütz
Kenneth L. Rehg
Richard W. Schmidt
Barbara W. Andaya
© Copyright 2005
by
Sumitra Suraratdecha
This dissertation is dedicated to my dear family and my late grandma.
ACKNOWLEDGEMENTS

I would first like to thank my family for their unconditional support, and encouragement to pursue my study in linguistics, and for their enduring many years of separation. A heart-felt thank and appreciation goes to my sister, Supattra Suraratdecha, my mom, Kritcharat Suraratdecha, and my dear Huadchai, for always being there for me. I am grateful to my aunts, Nongnuch and Lawan Wongtapradit, and my late grandma, for their love and support. Thank you all for waiting patiently, for listening to my complaints and frustrations, and for believing in me all these years that we have had to be apart. Grandma, I wish you were still with us.

My achievement of completing this dissertation was due largely to the contributions of many people, to all of whom I owe deep gratitude. My thanks and appreciation to Dr. Michael Forman for persevering with me as my advisor and dissertation chair throughout the time it took me to complete my Master degree and write this dissertation for the Ph.D. degree. Dr. Forman was always there to meet and discuss a variety of topics, to read and edit my chapters, and to give me suggestions and advice. His suggestions, questions, and criticism have been invaluable. The members of my dissertation committee, Dr. Albert Schütz, Dr. Kenneth Rehg, Dr. Richard Schmidt, and Dr. Barbara Andaya, have generously given their time and expertise to better this dissertation. I thank them for their contribution and their good-natured support. I am also grateful to Dr. Miriam Meyerhoff who introduced me to the field of sociolinguistics, and whose work has inspired me. Thank you for all the inspiring conversations and classes, and for all the e-mails and talks of encouragement. A special thank also goes to Dr. William O’Grady for his encouragement throughout my years of study at the University
of Hawai‘i, for his valuable comments and suggestions on the syntactic part of this dissertation, and for his especially inspiring lectures in syntax. I would also like to thank Dr. George Grace for filling in at the last minute for my dissertation defense, and Dr. Amy Schafer and Dr. Christina Higgins for their time and suggestions.

My deep appreciation to the many friends, and colleagues who have enriched my life in Hawai‘i through their friendship and support. I would like to thank all my friends who have made my life in Hawai‘i easier and who have provided material and spiritual support at critical and opportune times. Thank you, Preena Kangkun, Laura Sacia and her husband Klaus Dragull, Piyalerg Kongsil, Valuncha and Jeremy Paterson, Napat Settachai and Wu Jing, Npasri Timyam, Montakan Bamrungruan and her brother Aphirak Bamrungruan, Pongpat Jeerangpitakkul, Woravudh Lekprathum, Aaron Tsang, Sunyoung Lee, Hsiu-chuan Liao, and Kasumi Yoshihara for all the phone calls of spiritual support, food deliveries, and fun time inside and outside the department. Your constant heartfelt support will never be forgotten. I would like to acknowledge the love and kindness of the Hansa family, especially Weera-anong Hansa whose friendship, hospitality, and wisdom have supported, enlightened, and entertained me over the many years of our friendship. I would also like to express my gratitude and deep appreciation to the immeasurable acts of kindness of the Thai people whom I met at the temple, Wat Buddhajakramongkolvararam, Pearl City. Thank you dear friends, aunties, and monks, who have participated as well as helped with this dissertation in one way or another. It would not have been possible without your help.
I would also like to thank the current secretary of the department, Jennifer Kanda, and another departmental staff, Tamara Zielinski, for all of their help in the past two years. Thanks also go to the former secretary, Wendy Onishi.

Finally, I would like to thank the Institute of Language and Culture for Rural Development, Mahidol University, and the Ministry of University Affairs for granting me the scholarship.

To all the above individuals and to several colleagues and friends whose names I cannot continue listing and who have assisted me, I feel very much grateful. Without your support, encouragement, and guidance, this dissertation would never have been possible.

While I believe that all of those mentioned have contributed to an improved final manuscript, none is, of course, responsible for remaining weaknesses. Without doubt there will be errors, omissions, and over-simplifications, for which I assume absolute responsibility.
ABSTRACT

Over the last three decades, the study of codeswitching has attracted many data-oriented and theory-oriented sociolinguists and syntacticians. It provides an avenue to understanding the relationship between social processes and linguistic forms. In this vein, the present study examines Thai/English codeswitching practiced by a group of Thai people in Hawai‘i from two perspectives: sociolinguistic as well as syntactic. In this study, different theoretical models are tested against each aspect of codeswitching.

With regard to the sociolinguistic aspect of codeswitching, there has been a debate concerning two models: the markedness model (Myers-Scotton 1993b, 1998) and the conversational codeswitching approach (Auer 1995, Li 1994). These two frameworks are distinctive in their ways of analysis. The former takes societal norms as its starting point, whereas the latter takes the face-to-face interaction or the conversation as its starting point. Accordingly, one can be seen as a top-down approach, and the other a bottom-up approach. Over the last decade, the markedness model has been criticized for being categorical and static, the conversational codeswitching approach for its inattention to existing societal norms. By examining ways in which Thai/English codeswitching is used to signal social, discourse, and referential meanings, the present study is in support of the markedness model, the model that unifies macro-level and micro-level approaches to the study of codeswitching.

In terms of the syntactic aspect of codeswitching, the present study employs the Matrix Language Frame model (Myers-Scotton 1993a, 2002) to investigate the grammatical constraints on Thai/English codeswitching. Most of the findings in the present study are in corroboration with the model. Even though there are a few examples
that seem to run counter to the System Morpheme Principle and the Morpheme Order Principle (Myers-Scotton 2002), the number of counter examples is too small to make a significant claim.

By considering both sociolinguistic and syntactic aspects of Thai/English codeswitching, the present study finds that social aspects of Thai/English codeswitching characterize its linguistic forms. In order to understand the social significance of codeswitching as well as its distinctive pattern in specific communities, ethnographic/sociolinguistic information regarding the participants and the community has to be taken into account.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... v

ABSTRACT ........................................................................................................................................ viii

TABLE OF CONTENTS ...................................................................................................................... x

LIST OF TABLES .................................................................................................................................. xiv

LIST OF FIGURES ............................................................................................................................... xv

CHAPTER 1: INTRODUCTION ............................................................................................................ 1

1.1 INTRODUCTION TO THE STUDY ............................................................................................ 1

1.2 PREVIOUS RESEARCH ON THAI/ENGLISH CODESWITCHING .......................................... 2

1.3 ASPECTS OF CODESWITCHING ............................................................................................ 7

1.3.1 Sociolinguistic Aspects of Codeswitching ............................................................................. 8

1.3.2 Grammatical Aspects of Codeswitching .............................................................................. 12

1.3.2.1 The linear approach ........................................................................................................ 12

1.3.2.2 The government and binding approach ........................................................................... 14

1.3.2.3 The insertion approach ................................................................................................... 16

1.4 DISTINCTION BETWEEN CODESWITCHING AND BORROWING ...................................... 19

1.5 OBJECTIVES OF THE STUDY ............................................................................................... 24

1.6 RESEARCH QUESTIONS .......................................................................................................... 26

1.7 THE PLAN OF THE STUDY ...................................................................................................... 26

CHAPTER 2: THEORETICAL FRAMEWORK ...................................................................................... 29
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 THE MARKEDNESS MODEL</td>
<td>29</td>
</tr>
<tr>
<td>2.2 THE CONVERSATIONAL CODESWITCHING APPROACH</td>
<td>38</td>
</tr>
<tr>
<td>2.3 THE MATRIX LANGUAGE FRAME MODEL</td>
<td>48</td>
</tr>
<tr>
<td>2.3.1 Matrix Language and Embedded Language distinction</td>
<td>50</td>
</tr>
<tr>
<td>2.3.1.1 The roles of Matrix Language</td>
<td>51</td>
</tr>
<tr>
<td>2.3.1.2 The Embedded Language Island</td>
<td>52</td>
</tr>
<tr>
<td>2.3.2 Content-system morpheme distinction</td>
<td>54</td>
</tr>
<tr>
<td>2.3.3 CP as the Unit of Analysis</td>
<td>55</td>
</tr>
<tr>
<td>2.3.4 The 4-M Model</td>
<td>57</td>
</tr>
<tr>
<td>2.3.5 The Abstract Level Model</td>
<td>60</td>
</tr>
<tr>
<td>CHAPTER 3: RESEARCH DESIGN AND METHODOLOGY</td>
<td>65</td>
</tr>
<tr>
<td>3.1 DATA COLLECTION</td>
<td>65</td>
</tr>
<tr>
<td>3.1.1 Participant Observation</td>
<td>65</td>
</tr>
<tr>
<td>3.1.2 Interviews</td>
<td>67</td>
</tr>
<tr>
<td>3.2 BACKGROUND INFORMATION</td>
<td>69</td>
</tr>
<tr>
<td>3.2.1 The Participants: Thai People in Hawai‘i</td>
<td>70</td>
</tr>
<tr>
<td>3.2.2 English in Thai Society</td>
<td>72</td>
</tr>
<tr>
<td>3.2.3 Thailand and Its Language</td>
<td>74</td>
</tr>
<tr>
<td>3.3 CONVENTIONS AND ABBREVIATIONS USED FOR PRESENTATION OF DATA</td>
<td>79</td>
</tr>
<tr>
<td>CHAPTER 4: FUNCTIONS OF THAI/ENGLISH CODESWITCHING</td>
<td>81</td>
</tr>
</tbody>
</table>
6.1 A SUMMARY OF FINDINGS ................................................................. 143
  6.1.1 Language Functions of Thai/English Codeswitching ....................... 143
  6.1.2 Syntactic Aspects of Thai/English Codeswitching .......................... 145

6.2 EVALUATING THE SOCIOLINGUISTIC FRAMEWORKS .......................... 148
  6.2.1 Criticisms of the Markedness Model ........................................ 148
  6.2.2 Criticisms of the Conversational Codeswitching Approach ................ 153

6.3 CONCLUSION ...................................................................................... 162

6.4 LIMITATIONS OF THE STUDY AND SUGGESTIONS FOR FURTHER
  STUDIES ............................................................................................... 167

APPENDIX A: A List of questions used in the interviews ............................. 169

APPENDIX B: Information regarding Thai population from U.S. Census Bureau, Census
  2002 ........................................................................................................ 170

APPENDIX C: Thai consonants and vowels ............................................... 175

REFERENCES .......................................................................................... 175
LIST OF TABLES

Table 1: A List of Conventions and Abbreviations ......................................................... 80
Table 2: Ratio of Thai/English Codeswitching Frequency ................................................. 82
LIST OF FIGURES

Figure 1 Feature-based Classification of Morphemes in the 4-M Model............................59
Figure 2 Production Process Diagram: Lemma Activation..................................................61
CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION TO THE STUDY

In communities where there is a diversity of languages, people sometimes alternate between languages in a single conversation or even within a sentence boundary. This phenomenon is called codeswitching. The study of codeswitching is generally assumed to have begun in the early 50s. Haugen (1953), who coined the term 'codeswitching' in 1956, and Weinreich (1953) are among the first who addressed such phenomenon. There are variations in terminology referring to this phenomenon. The two terms employed most often are codeswitching and code-mixing, the former of which refers particularly to switches between sentences, and the latter exclusively to intrasentential switches. While some employ these two terms to discuss such alternations, the single term codeswitching is used here to cover both intersentential and intrasentential switches, for it has become the most widely used and understood since the writings of Haugen. Also, code-mixing can be understood as simply a common mode of codeswitching, serving the same functions, e.g., to mark identity, to mark group solidarity, to emphasize a message, to direct or redirect a message to particular addressees, and so on (see Wardhaugh 1986, Sridhar 1996). Other renowned scholars, such as Myers-Scotton (1993a, 1993b) and Milroy and Muysken (1995), also use the term codeswitching.

---

1 Before then, it was referred to as “interference”. Weinreich (1953:1) defined “interference” as “[t]hose instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language.” Weinreich also stated that a switch cannot occur (1) in a single sentence, and (2) in an unchanged speech situation. His statement has been invalidated by many codeswitching studies, e.g., Blom and Gumperz 1972, Poplack 1980, and Myers-Scotton 1993b. Codeswitching does occur intrasententially. According to Troike (2004), to this day, there are 900 publications on intrasentential codeswitching listed in the MLA bibliography. Works by Blom and Gumperz (1972), Poplack (1980), and Myers-Scotton (1993b) also show that codeswitching occurs when there is no change in speech situation.
as an umbrella term to cover both intrasentential and intersentential switches. Accordingly, the present study will employ the single term, codeswitching, to discuss the alternation between two languages or dialects of the same language within the same conversation.


1.2 PREVIOUS RESEARCH ON THAI/ENGLISH CODESWITCHING

In 1975, Surawan studied the speech of Thai students, 62 males and 34 females, residing in different states of the USA, including Texas, Oklahoma, Kansas, North Carolina, and California. The objectives of her study were twofold: (1) to identify and describe phonological and lexical interference in these Thai students' speech, and (2) to identify factors influencing Thai-English codeswitching in their speech. Regarding the
latter, she found that the speech of these students is conditioned by linguistic and extra-linguistic factors, both of which are interrelated. The extra-linguistic factors in her study include interlocutor, role-relation, topic, and style. According to her findings, each linguistic code is chosen at one time or another according to (1) the linguistic repertoires of the speakers present in conversations, (2) the concept of in-group or out-group (more switching occurs between close friends), (3) topics of conversation (more switching if the topics concern experiences and things in the United States), and (4) conversation genre (switching is found in quoting, emphasis, teasing or joking, and in avoidance of Thai taboo words). Surawan claimed that the linguistic factors found in her study are structural factors “which stem from the organization of linguistic forms into a definite system, different for every language and to a considerable degree independent of non-linguistic experience and behavior” (Surawan 1975:65). To explain, some items are more readily available than others in one language or the other, and that influences the switch between languages. Surawan suggested that her informants switch to English because they have absorbed English into their linguistic repertoire such that producing conversation in English has become their habit. Some of the English items included in this category are names of food, fruit, and beverages; proper names of persons and places; technical terms; routine expressions; exclamations; and loanwords. Surawan included all English-based words that appear in her data and do not have an entry as codeswitched words in Thai dictionaries, e.g. *hamburger, salad, vanilla, lipstick, violin, ski, lottery, America*, etc. This has been a customary practice among all other researchers on Thai-English codeswitching in Thailand. The present study takes an alternative standpoint,
differentiating between codeswitching and borrowing. §1.4 elaborates on the distinction between codeswitching and borrowing.

Pairat Warie (1977) explored formal characteristics of Thai-English code-mixing.\(^2\) She found that when a code-mixed unit is found in Thai discourse, it appears in the form of what she called unit insertion or unit hybridization.\(^3\) She also added that Thai-English code-mixing could be explained functionally with reference to what she identified as role identification, register identification, and elucidation. To explain, people with high socioeconomic background and university students associate code-mixing with modernization, high level of education, as well as sophistication. They code-mixed in order to be regarded by others as modern, highly educated, and sophisticated. Register identification refers to the Thai/English code-mixing of technical terms in professional fields, e.g., medical terms. The last function, elucidation, refers to the use of English words to explain concepts and ideas (see Warie 1977).

Siiha-umphai (1987) conducted a code-mixing study concerning the speech of Thai tour-guides in Thailand when on and off duty, taking into consideration gender, age, education, and working experience. His study showed that code-mixing is most common among college-educated 40-50-year-old male tour-guides while they are on duty. Medical doctors in Thailand, in the study by Rukthamying (1995), are found to mix English into Thai more when speaking with their peers (medical doctors) than when they are talking to others. Thitiwattana (1996) studied instructors’ attitudes toward code-mixing in different departments in Kasetsart University. She found that the reported code-mixing behavior

---

\(^2\) The source of her data is not specified.
\(^3\) Unit insertion is an insertion of an English word into a Thai sentence. Unit hybridization is a phrasal combination of a Thai morpheme and an English morpheme, e.g., [kʰəɾ tʰə:n ʔəɾən] ‘morning cartoon’ (Warie 1977: 28, 30).
does not always match the actual behavior. For example, those who have a positive attitude toward code-mixing do not mix languages when they converse with other Thais, as they are afraid that their interlocutors might not understand their mixed language or even reject it. Thitiwattana also found that English instructors at Kasetsart University tend to mix languages at the levels of word, phrase, and sentence, and code-mix more frequently than instructors from the political science and Thai departments. Wongpanitcharoen (1997) found that Thai FM radio broadcasters in Bangkok mix technical English terms, proper names, and other English words with Thai in their speech. In situations where the broadcaster is the sole speaker, the hosts of economic programs tend to mix languages the most, followed by educational programs and news programs. In situations in which two broadcasters interact with each other, the rate of the code-mixing is the highest when they talk about current news.

Furthermore, Maneepong (1997) studied Thai and Japanese people's attitudes toward language mixing. She found that Japanese people are more open to language mixing, while Thais are more resistant to language mixing. That is, Thai informants only allow cultural borrowings; they tend to use newly coined words, and those listed in the dictionaries rather than English loanwords. Thaatlek (1998) researched Thai-English code-mixing by taking into consideration socioeconomic factors, participants' attitudes toward code-mixing, and word types--whether the mixed words are technical terms or general English words. Her informants were university instructors, entrepreneurs, salespersons, and waitpersons in Bangkok, Thailand. Her study shows that the participants do not always mix their language, but usually do so with people who share

---

4 A similar result is also reported in Blom and Gumperz 1972.
5 In her study, cultural borrowings include loan place names and proper names (Maneepong 1997).
the same background and are not opposed to language mixing. Later, Boonkongsaen conducted a code-mixing study in Bangkok, by taking into account other social factors, such as different socioeconomic status and career level. In her study, questionnaires were used instead of the matched-guise technique to elicit information concerning code-mixing behavior. Her study shows that both career and language attitude play a role in code-mixing behavior (Boonkongsaen 1999). Finally, Suraratdecha (2003) examined Thai-English codeswitching by eight Thai students at the University of Hawai‘i at Mānoa by taking into account Bell’s audience-design factors (Bell 1984), speech-accommodation theory (Giles and Smith 1979; Giles and Coupland 1991), and other psycho-social factors. Similarly to previous codeswitching research, her study shows that codeswitching is an accommodative phenomenon. The codeswitching frequency of the participants depends on their evaluation of their own linguistic skills and their perceptions of others. By bringing together observations about who favors and who resists codeswitching, the topic of conversation, among other factors, emerges as the most important social variable across individuals. The result accordingly runs counter to the audience-design model proposed by Bell (1984). Bell claims that audience-design factors are more influential than non-audience-design factors in bilingual code choice.

By far, all the previous work focuses solely on social factors influencing Thai/English codeswitching; none address theoretical frameworks exclusively developed to account for the codeswitching phenomenon, e.g., the markedness model (Myers-Scotton 1993b, 1998, Myers-Scotton and Bolonyai 2001), the conversational codeswitching approach (Auer 1995, Wei 1994, 2002), and the Matrix Language Frame model (Myers-Scotton 1993a, 1995, 2002). It is a main objective of the present study to
investigate Thai/English codeswitching in terms of these theoretical frameworks, and to provide up-to-date information on the literature of Thai/English codeswitching.

1.3 ASPECTS OF CODESWITCHING

Linguists with varying backgrounds and interests have studied different aspects of codeswitching. In general, there are two main approaches to the study of codeswitching: the sociolinguistic or functional approach, and the grammatical approach. Those who take the sociolinguistic approach attempt to identify factors influencing codeswitching as well as meanings and functions of codeswitching, e.g., Blom and Gumperz (1972), Gumperz (1982), Gal (1988), Heller (1988), Myers-Scotton (1993b), Zentella (1997), Auer (1995), and Li (1994). Those who adopt the grammatical approach attempt to analyze codeswitching in terms of grammatical aspects—Given two different grammatical structures of the two languages involved in codeswitching, which constituents can be switched and under what syntactic environment? Such works include Pfaff 1979, Sankoff and Poplack 1981, Joshi 1985, Di Sciullo et al. 1986, Myers-Scotton 1993a, and Belazi et al. 1994.

This dissertation analyzes Thai-English codeswitching from two perspectives: (1) the sociolinguistic/functional perspective and (2) the syntactic perspective. The functions of Thai-English codeswitching will be investigated using the Markedness model (Myers-Scotton 1993b, 1998, Myers-Scotton and Bolonyai 2001) and the conversational codeswitching approach (Auer 1998, Li 1994, 2002). The grammatical constraints on Thai-English codeswitching will be examined with the assistance of the Matrix Language

---

6 This classification does not necessarily imply that sociolinguistics and syntax are two unrelated fields in linguistics.
Frame model (Myers-Scotton 1993a, 1995, 2002). Information regarding each framework will be given in Chapter 2.

1.3.1 Sociolinguistic Aspects of Codeswitching

Researchers taking the sociolinguistic approach are interested in determining factors influencing codeswitching. Furthermore, they are interested in identifying the functions and interpreting the meaning of codeswitching. In their early work, Blom and Gumperz analyzed codeswitching as “situational” or “metaphorical.” Situational codeswitching occurs when there is an environmental or contextual change that does not involve a change in topic: speakers use one code (or language) in one situation and another in a different situation. On the other hand, metaphorical codeswitching occurs when situational or contextual features remain the same while external features, such as topics and interlocutors, change. This type of codeswitching calls for the distinction between marked and unmarked linguistic codes. The unmarked code is the norm or the expected code choice in certain speech situations. The marked code, on the other hand, is the unexpected or deviated code choice. In cases in which the marked code occurs, it is expected to deliver particular messages such as evoking specific feelings (Blom and Gumperz 1972). Sankoff (1972), however, questions the validity of the metaphorical codeswitching approach in practice. As mentioned above, metaphorical codeswitching requires the distinction between marked and unmarked codes. In her Huang case study, Sankoff found that it is not always possible to make such a distinction (Sankoff 1972). To address this shortcoming, Gumperz (1982) later developed the term “conversational codeswitching” to refer to situations when both languages are used without this distinction.
Conversational codeswitching refers to a norm in which two languages are used with no need for marked/unmarked code distinctions. Gumperz stated that the juxtapositions of two codes perform/signify conversational forms and functions, e.g., interjections, quotation, addressee specification, reiteration, message qualification, and personalization vs. objectivization (Gumperz 1982). His main source of data is naturally occurring data. He also considers language use a function of the dynamics of interaction:

Detailed observation of verbal strategies revealed that an individual’s choice of speech style has symbolic value and interpretive consequences that cannot be explained simply by correlating the incidence of linguistic variants with independently determined social and contextual categories. (Gumperz 1982:viii)

In response to Gumperz’s proposal concerning codeswitching functions, Poplack (1980) argued that functions could not always be assigned to the switching data, as Gumperz (1982) had claimed. She found it impossible to assign functions to her Spanish/English codeswitching data, not only because switching happens often, but also because it happens intrasententially. She concluded that codeswitching is a discrete mode of speaking. Intrasentential codeswitching does not necessarily entail conversational forms or functions as claimed by Gumperz. She further stated that this type of switching needs not be explained in terms of any social motivations, because the switching itself is a part of the repertoire of a speech community. According to Poplack, “situational motivations or consequences to specific intra-sentential switches... [have] little if any pertinence for the speakers themselves” (1980:614). The present study argues that many codeswitching studies have shown that situational motivations and consequences to switches play a key role in determining codeswitching, e.g. Myers-Scotton 1993b, Auer 1998, and Li 1994, 2002. It will also become clear in the present study that situational
motivations and consequences are significant to the speakers. This will be discussed further in Chapter 6.

Two prevailing frameworks that are of particular interest in the present study are the markedness model and the conversational codeswitching approach. These two models differ distinctly in their analyses of codeswitching data. There has been debate as to whether the former or the latter gives a better account of the study of codeswitching. It is a goal of the present study to contribute to clarifying these highly debated issues by testing each model against Thai/English data.

The markedness model (Myers-Scotton 1993b, 1997b, 1998, Myers-Scotton and Bolonyai 2001) is a model using socio-psychological motivations to account for speakers engaging in the act of codeswitching. Community norms are the starting-point for this model. The model assumes that, in a given community, each language or variety has its own value, attitude, and identity, usually contrasting one with another. This model also distinguishes between marked and unmarked code. In addition, the markedness model is speaker-oriented such that speakers are creative rational actors making code choices by weighing costs and rewards and associating codes with the rights and obligations set within a normative framework specific to their community. According to this model, sociolinguistic functions of codeswitching in conversation are both indexical and interpretative.

---

7 According to Heritage (1989), there are many studies that can be described as conversation analysis (CA). Within these studies, however, two strands are identifiable. The first focuses on the institution of interaction as an entity in its own right, “pure” CA. The second examines the management of social situations in interaction, “applied” CA. The conversational codeswitching approach falls in the latter type. It employs the terminologies used in the “pure” CA, e.g., turn transition, adjacency pairs, etc. to discuss conversational functions of codeswitching data by taking into account the sequential organization of the data. CA terminologies to the extent that a selection of these are used here will be briefly explained in §2.2.
In contrast, the conversational codeswitching approach centers upon the interactional effect of codeswitching in conversation. The starting point of the approach is the conversation. Interactional sociolinguists using the conversational codeswitching approach suggested that to fully understand codeswitching practices, primary and thorough attention has to be paid to its local production in the emerging context. They criticized other large-scale sociolinguistic models for failing to account for the local production and restricting their analysis to associating the social meaning of codeswitching phenomenon with reference to a priori given context; e.g., participants, topic, and setting (Auer 1984, 1998; Li 1994, 1998, 2002).

In addition, the conversational codeswitching approach considers social meanings of language use as a function of the situated contexts or the dynamics of interactions. This contrasts sharply with other large-scale studies considering linguistic variation, e.g., individual choices in particular, as derived from sociological attributes of the speaker and the situation. Accordingly, it is criticized for its inattention to the salience of sociological variables (Myers-Scotton 1995). It neglects the existence of societal norms. The approach views the speaker merely as a participant in an ongoing interaction, not as an identity-bearing individual.

This dissertation studies Thai/English codeswitching practiced by a group of Thai people on O'ahu. By bringing together the background information regarding the Thai participants, the status of English in Thai society, as well as the linguistic movement to purify the Thai language, the present study is in favor of the markedness model. The background information will be provided in Chapter 3. The application and the validity of each model will be discussed in Chapter 4 and Chapter 6, respectively.
1.3.2 Grammatical Aspects of Codeswitching

According to Appel and Muysken (1987) and Boumans (1998), the syntactic study of codeswitching can be divided into three approaches: the linear approach (Poplack 1980, 1988, 1990, Poplack and Sankoff 1981), the structural approach or the Government and Binding approach (Di Sciullo et al. 1986, Belazi et al. 1994), and the insertion approach (Joshi 1985, Myers-Scotton 1993a).

The linear approach examines the word categories (or constituents) between the languages in use and determines whether a switch can occur between the two languages. The Government and Binding approach predicts that a switch cannot occur under certain governor/governed relationships. These two approaches agree that insertion is a secondary mechanism: to the word order of the two relevant languages in the former, and to the hierarchical constituent structure or government relations in the latter. For other researchers, however, insertion is a fundamental mechanism. The insertion approach distinguishes between a matrix (or base) language and an embedded language. According to this approach, the matrix language is the main ordering force—that is, it is the language that provides a grammatical frame, in which embedded language constituents are inserted. Also, codeswitching regularities are attributed to psychological factors such as sentence planning and the organization of the mental lexicon (Boumans 1998).

1.3.2.1 The linear approach

According to Boumans (1998:12), “[t]he linear approach describes codeswitching basically by investigating the word categories or constituents between which there can or

---

8 The terms “matrix language” and “embedded language” may appear in lower case or upper case, depending on the original writings. The first letters of each word will be in upper case in the works of Myers-Scotton (1993a, 2002)—“Matrix Language” and “Embedded Language”—as she capitalises the first letters of these two terms in her writings. However, these will be in lower case elsewhere.
cannot be a switch. Word order is used to explain patterns found.” The linear approach proposes the existence of two main constraints: the free morpheme constraint and the equivalence constraint. According to the free morpheme constraint, a switch between a bound morpheme and a lexical item is inhibited “unless the latter has been phonologically integrated into the language of the bound morpheme” (Sankoff and Poplack 1981:5). The equivalence constraint suggests that constituents from two languages can be used in a conversation as long as the two languages have the same constituent order. Codeswitching is accordingly prohibited elsewhere. However, these two rules appeared to be applicable exclusively to the Spanish/English data, rather than applicable universally. Many counterexamples to the applicability of such constraints exist and have been found. The equivalence constraint in particular was disproved by codeswitching studies concerning two typologically different languages, such as Japanese and English (Nishimura 1997, Fotos 1995). To address the shortcomings of the equivalence constraint, Poplack and Sankoff introduced other strategies such as nonce borrowing, flagged switching, and constituent insertion, and concluded that codeswitching does not take place with any two languages with different word order, but borrowing does (see Poplack 1988). Even though the introduction of nonce borrowing or constituent insertion can explain many counterexamples and remedy the shortcomings of the equivalence constraint, it introduces another problem, namely the over-prediction of insertion, as these newly introduced strategies or constraints do not provide a restriction on the insertion of single function words. Despite the strategies developed by Poplack and Sankoff to cope with the shortcomings of their first two constraints, there are yet counterexamples to this roundabout approach. The absence of the distinction between
matrix and embedded languages results in a problem of ambiguity with their supplementary constituent insertion (Boumans 1998).

1.3.2.2 The Government and Binding approach

The most pronounced "government constraint," claimed by Di Sciullo et al. (1986:5), is formulated as follows:

- If X governs Y, ...Xq ...Yq ...

That is, if constituent X governs Y, both of the constituents must be from the same language. The fundamental idea behind Di Sciullo et al.'s model for codeswitching is that lexical items or syntactic constituents are endocentric. That is, head constituents project their categorical, semantic, and language index features in, and only within, the phrase. As such, a noun phrase derives from properties of the head noun, and a verb phrase derives from properties of the head verb. According to the government constraint, the switching between categories with the governor-governed relationship is disallowed.

Di Sciullo et al. (1986) define the concept of government as follows:

- X governs Y if the first node dominating X also dominates Y, where X is a major category N, V, A, P and no maximal boundary intervenes between X and Y. (Di Sciullo et al. 1986:6)

In this model, language index is assigned by the 'highest' lexical element in a maximal projection; called the Lq carrier. In principle, any function morpheme, such as a determiner, case marker, or plural marker, can be an Lq carrier.

- (a) If Lq carrier has index q, then Y max q.
  (b) In a maximal projection Y max, the Lq carrier is the lexical element which asymmetrically c-commands the other lexical elements or terminal phrase nodes dominated by Ymax.

According to their model, the syntactic elements V-COMP, V-DET, V-P, V-Q, N-AP, N-PP, and P-NP must have the same index. Belazi et al. (1994) and Nishimura (1997) found this point of the claimed government constraint to be too restrictive.
Di Sciullo et al. can account for a construction in example (1) because il can function as an Lq marker in this case. However, it cannot account for a construction in which the governed constituent consists of a single content word, as in example (2).

Later, Belazi et al. (1994) distinguish between lexical heads like V and P and functional heads like Quantifier, Modal, and INFL. The model states that switching can occur between a lexical head and its complement, such as V and its object NP or P and its object NP, but not between a functional head and its complement, such as Q-N, and Neg-V. This is called the functional head constraint (Belazi et al. 1994). However, Nishimura (1986) found a counterexample to this claim: switching can occur between the subject NP and the tensed verb as well as between a negative and a verb in her Japanese/English data.

A more successful Government and Binding approach includes insertion. Researchers following this vein are Pandit (1990), Halmari (1993), Santorini and Mahootian (1995), and Muysken (1995). Whether or not their proposals predict correctly depends on which items are identified as phrasal heads. Pandit (1990) identifies the verbal inflection (INFL) and noun as a phrasal head. That is, she proposes that INFL-verb and N-DET must have the same index; that accounts for Swahili/English switches. Pandit, however, identifies the noun as a phrasal head for the relative order of noun and determiner and that formulation results in some incorrect predictions. According to Pandit (1990), N-DET must have the same index. This is not the case in Example (3) below.
He is a demonio. ‘He is a devil.’ (English/Spanish (Belazi et al. 1994:227)) According to Muysken (1995), the government constraint is simply too restrictive and thus cannot be maintained. Its predictions are borne out by all available data.

Boumans (1998) summarizes the merits and faults of Government and Binding approaches as follows:

...firstly, as long as functional categories are identified as head of phrases, and content word insertion as well as constituent insertion is allowed for, codeswitching data corroborate the government model. Secondly, government models make no predictions concerning the possibility of switching (inserting) constituents that are neither governors nor governed. Thirdly, in the case of inserted governing verbs, government models frequently make the wrong predictions. (Boumans 1998:24)

1.3.2.3 The insertion approach

According to Boumans (1998), three crucial observations have been identified in the analysis of codeswitching: (1) languages involved in codeswitching phenomena play unequal roles, (2) content words and inflection words differ in their status, and (3) structural congruence in categories between the languages is significant. These three observations are in fact indispensable for any insertion model, for it presupposes (1) a base or matrix language, (2) a frame for grammatical structure provided by the matrix language, and (3) the speaker’s perception of a congruency of the switched element between the matrix language and the embedded language.

The insertion embodies the asymmetrical roles of languages involved in codeswitching, the differentiation of content and function morphemes, and the aspect of categorical congruency. Therefore, the distinction between the matrix language and embedded language is required, aside from the distinction between the content and function morphemes. Boumans (1998) also calls this approach the “matrix language approach,” presumably because it distinguishes between the matrix and the embedded
language. According to Boumans, the matrix language is assumed to organize and express "the grammatical relations in the sentence by means of inflection, function words, and word order. For this reason, the distinction, roughly speaking between CONTENT WORDS on the one hand and INFLECTION and FUNCTION WORDS on the other is an indispensable ingredient of the matrix and embedded language approach" (Boumans 1998:34, original small caps). Researchers in this vein include Azuma (1993), Myers-Scotton (1993a), Joshi (1985), and Nishimura (1986).

Joshi (1985) first introduced the terms matrix language and embedded language to the study of codeswitching. He stressed the asymmetrical relation between the matrix and embedded language. He stated that inserted embedded language elements must be congruent to those in the matrix language. Unfortunately, his formulation is quite imprecise, and the notion of the matrix language is not specifically defined in his study. However, his ideas of matrix languages and congruence have inspired much of the later work, including the Matrix Language Frame model of Myers-Scotton (1993a).

Azuma's Frame Content Hypothesis, based on the insertional mechanism, explains syntactic constraints on codeswitching similarly to those of monolingual sentence processing and involves two stages: frame-building and content insertion. Frame-building includes accessing and retrieving closed-class items, while content insertion is the insertion of content words (Azuma 1993). This model, however, cannot account for phrase-level switches (Nishimura 1997).

The most influential model in this approach is the Matrix Language Frame model proposed by Myers-Scotton (1993a, 2002). The Matrix Language Frame model is a morphosyntactic model based on two oppositions: (1) the Matrix Language–Embedded
Language opposition and (2) the content-system morpheme opposition. According to this model, the Matrix Language is the language that provides the grammatical frame for the inserted Embedded Language constituents. The syntactic models reviewed thus far have one or more drawbacks; the Matrix Language Frame model is no exception. It has also been criticized by Muysken and De Rooij (1995). However, Myers-Scotton has made many important revisions and has clarified ambiguous notions in the model. The revised version of the revised Matrix Language Frame model (Myers-Scotton 2002) will be employed to explore the Thai/English codeswitching data in the present study. The criticisms of the model and its revision will be reviewed in § 2.3.

To sum up, of the three approaches, both the government and insertion approaches assume an unequal role for the two languages involved in codeswitching; the former approach associates one language with the governor and the other as being governed; the latter approach assumes one language to be a matrix or base language that provides the grammatical frame. The linear approach, however, assumes an equal role between a language pair involved in codeswitching. It further assumes that each sequence of language switches is explained by each language's own grammatical rules. Both the linear approach and the Government and Binding approach have been proven false in the previous literature. The insertion approach, particularly the Matrix Language Frame model, is thus the most promising tool to examine grammatical constraints of codeswitching.
1.4 DISTINCTION BETWEEN CODESWITCHING AND BORROWING

The study of codeswitching requires a discussion of the distinction between codeswitching and borrowing. In fact, there is a long-standing debate on such a distinction and different researchers' ideas vary on what counts as codeswitching forms or material. Generally, studies fall into one of two veins: those including both singly occurring embedded language elements as well as islands as codeswitching material, and those treating only fully embedded language constituents as such. The present study adopts the former practice: all instances of singly occurring embedded language elements as well as islands are considered codeswitching material. This is in compliance with Myers-Scotton's idea. Myers-Scotton (1997b, 2002) considers the distinction between codeswitching and borrowing in the synchronic study of codeswitching unnecessary, claiming that the Matrix Language Frame model gives a complete account to all embedded language elements appearing in the morphosyntactic frame of the matrix language or within the bilingual CP. Further, even though singly occurring embedded codeswitching elements and embedded language islands involve different activations, they are very similar in that "no Embedded Language material is accessed in a bilingual CP without some interaction between the Matrix Language and the Embedded Language; although Matrix Language features prevail, there is congruence checking between the two grammars at the abstract level of the mental lexicon" (Myers-Scotton 2002:154). She further states that Government and Binding practitioners, inter alia, Di Sciullo et al.

---

9 The Matrix Language is the base language that provides the morphosyntactic frame to codeswitched utterances. The Embedded Language is the "guest" language that is used in the Matrix Language frame.
10 An "island" is a stretch of Embedded Language constituents, as opposed to a singly occurring Embedded Language constituent.
11 CP, a complementizer phrase, is the maximal projection of a complementizer (COMP) in syntactic theories, e.g., Phrase Structure grammar, and Government and Binding theory (Crystal 1997).
Embedded Language elements as codeswitching forms simply because "their models are
devised to treat only phrasal constituents" (Myers-Scotton 2002:158).

Aside from which embedded language forms count as codeswitching elements,
there are also criteria to distinguish these elements from borrowed forms. Researchers
propose a wide variety of criteria, including phonological assimilation, morphosyntactic
and syntactic integration, frequency, dictionary entry, and congruency, some of which
have been disproved. In his work, Heath (1989) finds problems in distinguishing between
borrowing and codeswitching and comments that a binary distinction between the two is
quite idealistic. The use of pronunciation and frequency as parameters is problematic, for
they are not binary; rather they are gradient and difficult to determine. Boumans (1998)
concurs with this comment. Subsequently, Heath applies the term codeswitching to
elements showing little or no adaptation to the base language; and otherwise applies the
term borrowing.

Dictionary entry is the main criterion adopted by the Thai researchers.
Udomwong (1981) states that many of the English loanwords into Thai have gone
through phonological change and now follow Thai phonotactics and are listed in Thai
dictionaries. Boonkongsaen (1999) further comments that differentiating borrowing from
code-mixing is controversial. In her study of code-mixing and linguistic attitudes of
people from different socioeconomic backgrounds, she identifies all English words
occurring in a Thai context as code-mixed words. Similarly, to avoid the controversy,
Thitiwattana (1996) counts all the occurring English words as code-mixed words. The
Thai researchers' practice of including all English words as codeswitched words unless
they have an entry in a Thai dictionary has started since Grosjean’s (1982). Grosjean described borrowing as foreign language elements that appear in the receiver language and are recorded in the receiver language’s dictionary according to the receiver language’s phonological system. The phonological and semantic features of the borrowed items may be changed for the convenience of the receiver language. Therefore, the aforementioned Thai researchers classify all English words as code-mixed words to avoid the controversy regarding dictionary entry.

The present study, however, breaks from this practice. Instead, the occurrence of embedded language items will be considered in terms of congruency between the language pair, Thai and English, because quite a number of English words retain the phonotactics of the source language, are not listed in the dictionaries, and at least some are cultural loans, i.e., place names and proper nouns.

In another study, Boumans (1998) illustrates that phonological and morphological integrations do not work efficiently with two different language types: Moroccan Arabic/Dutch, and Turkish/Dutch. Both Moroccan Arabic and Turkish are immigrant languages in the Netherlands, however, these two language pairs show very different integration patterns, both phonologically and morphologically. Embedded Dutch constituents rarely reflect any morphological processes from Moroccan Arabic. However, Turkish morphological processes are very productive with embedded Dutch constituents. His own codeswitching analysis on Moroccan/Dutch conversation, therefore, does not employ this criterion to distinguish codeswitching from borrowing. However, it addresses some systematic differences between sets of “foreign lexemes” with regard to word
classes. In his Malay data, Nivens (2002) also shows that phonological assimilation and syntactic integration fail as indicators of codeswitching versus borrowing.

By far, the criterion that seems most promising to the study of codeswitching is congruency. The concept of congruency has been elaborately discussed by Nivens (2002) and Myers-Scotton and Jake (1995). Unlike Myers-Scotton and Jake (1995), Nivens (2002) discusses congruency not only at the abstract level, but also talks about congruency or what he calls "equivalence" in terms of economy and convenience, frequency, and context. One of the goals of his study is to draw attention to inconsistent linguistic behavior of humans, performance errors, and the effect of idiolect differences. In fact, he excludes some embedded language elements "induced" by discourse contexts, such as those in repetition and quotation, from codeswitching forms. The present study disagrees with this idea, as the author believes that a proficient bilingual speaker may always choose between Language A and Language B. As described by Nivens, discourse contexts do not necessarily induce a switch from Language A to Language B. Instead, the speaker may switch to another linguistic code in repetition or quotation for certain conversational effects. This point will be further discussed in Chapter 6.

Another recent claim views the phenomena of codeswitching and borrowing on a continuum. Heath (1989) looks at codeswitching as an avenue to borrowing. Codeswitching involves the use of a microstructure of foreign elements within a macrostructure of the native language. "Equally this type of interspersing of foreign material can be done sporadically or on a purely personal basis. "Borrowing", on the other hand, involves greater degree of formal adaptation to the base language and is more institutionalized within the community" (Heath 1989:viii). Furthermore, Myers-Scotton
(1992, 1997b) says that the motivation of these two distinct phenomena is the same: elements of language X are inserted into the grammatical frame provided by Language Y to fulfill speakers' expressive needs. However, they differ in terms of their speakers and their psycholinguistic status, and their structural and sociopolitical profile. To elaborate, (1) a monolingual speaker of Language X can only borrow from Language Y, but a bilingual speaker of Language X and Y can engage in codeswitching; (2) established borrowed forms of Language Y have an entry in the mental lexicon of Language X; and (3) borrowed forms tend to be incorporated into a less commanding language as opposed to a more socio-politically dominant language, but codeswitching of a more socio-politically dominant language tends to be sociolinguistically less commanding. Nevertheless, borrowing and codeswitching share an important similarity. That is, borrowed lexemes function like the embedded language in codeswitching utterances, i.e., they are morphosyntactically integrated into the matrix language. Additionally, core borrowings and codeswitching are part of a continuum. Unlike cultural borrowings, core borrowings do not fill lexical gaps - that is why a continuum exists. An explanation of the borrowing and codeswitching continuum is given below (Myers-Scotton 1997b:228):

... the life of a core borrowing into Language X begins as a form occurring either as singly occurring form in a mixed constituent or as part of what is referred to as an [Embedded Language] island above, a constituent entirely in the [Embedded Language]. When their frequency reaches an unknown threshold level, these [Embedded Language] lexemes move from being CS forms to becoming borrowed forms and therefore now part of the lexicon of the recipient language as well as the donor language.

In conclusion, although a wide variety of criteria may be used to differentiate between codeswitching and borrowing, the congruency between a language pair seems to be the most promising, therefore, it will be the criterion adopted in the present study. The present study also concurs with Haugen, with Myers-Scotton, and with Boumans in including singly occurring embedded language forms as codeswitching material, because
both codeswitching and borrowing involve congruency checking between the matrix and the embedded languages. Additionally, such a distinction does not play a key role in a synchronic study (Myers-Scotton 1997b, 2002) like the present one.

1.5 OBJECTIVES OF THE STUDY

The study of codeswitching has grown rapidly over the last twenty-five years. Fifty years ago, it was characterized as the speech of the imperfect bilingual (Weinreich 1953) as opposed to the fully competent speaker in either of the languages in his/her linguistic repertoire. With globalization and modernization taking place all over the world, there is increasing linguistic and social diversity in many places. To this day, there are nine hundred publications on the subject matter (Troike 2004). Given the pervasiveness of bilingualism and the projections for increasing diversity, continued research on the nature of bilingual speech is of crucial importance to understanding social and cultural diversity.

The present study focuses on codeswitching practiced by a group of Thai expatriates in Hawai'i. Most of the previous research on codeswitching concentrates around certain areas around the globe, e.g., Africa or European countries, and thus African and European languages such as Swahili, Spanish, French, etc. However, there are many more places where codeswitching occurs and in many other languages. Thai/English codeswitching is one of the codeswitching situations that have not received enough study. The present study will add the documentation of Thai/English codeswitching to the world literature of codeswitching. As will become clear later in the present study, the frequency of Thai/English codeswitching among the participants is
quite low. However, it seems likely that codeswitching will increase, not decrease, in Thai language usage in the future. If indeed that is the case, it will be valuable to have a study of what it was like at an early stage. All too often linguists look back and wish that some documentation might have been conducted at an earlier stage of some linguistic phenomenon, and can only regret that such studies were not done.

As mentioned earlier, most of the previous work on Thai/English codeswitching focused on identifying factors influencing Thai/English codeswitching. None as of yet have investigated Thai/English codeswitching in terms of its functions and meanings. It is a goal of the present study to add up-to-date information on such terms to the literature of Thai/English codeswitching.

The present study aims to explore the practice of Thai/English codeswitching in relation to the markedness model and the conversational codeswitching approach. It is also a goal of the present study to employ the Thai-English data and to contribute to clarifying and even settling some highly debated issues in this perspective of codeswitching study.

As far as the grammatical aspect of Thai/English codeswitching is concerned, a thorough syntactic analysis of Thai-English codeswitching has not yet been conducted. Previous researches merely listed the syntactic categories of switched items. This dissertation will thus be the first attempt to investigate grammatical constraints on Thai/English codeswitching. The Matrix Language Frame model will be employed to explore grammatical relations of the data.

---

12 Codeswitching frequency between other language pairs such as Chinese/English (Li 1994) is quite low as well. The low frequency in Thai/English codeswitching does not make the present study any less important than others. Rather, it is of valuable importance to find out what extra-linguistic factors characterize the Thai/English codeswitching pattern found in the speech of the Thai participants. This will be discussed further in Chapter 6.
1.6 RESEARCH QUESTIONS

The research questions below guided the present study. They are categorized by one of two perspectives, as follows:

1 Codeswitching from a sociolinguistic/functional perspective:
   1.1 What appear to be the functions of Thai-English codeswitching when data are analyzed according to the principles of the Markedness model?
   1.2 What appear to be the functions of Thai-English codeswitching when the data are analyzed according to the principles of the conversational codeswitching approach?
   1.3 Does the Markedness model give a better account of the Thai-English codeswitching data than the conversational codeswitching approach?

2 Codeswitching from a syntactic perspective:
   2.1 What are the grammatical constraints on Thai-English codeswitching?
   2.2 How well does Thai-English codeswitching data conform to the Matrix Language Frame model?

1.7 THE PLAN OF THE STUDY

This dissertation consists of six chapters. Chapter 1 has addressed the controversies in the study of codeswitching as well as the intended contributions of the dissertation. There has been considerable debate as to which of the different sociolinguistic models, particularly the markedness model and the conversational codeswitching approach, gives a more complete account of the study of codeswitching. Each of these two models is elaborately reviewed in, §2.1 and 2.2, respectively. Also, the
grammatical aspects of codeswitching are outlined in §1.3.2. There are three grammatical approaches in the study of codeswitching. The difference among these models lies in their predictions of codeswitching. The linear approach assumes an equal role between a language pair involved in codeswitching and assumes further that each sequence of language switches is explained by each language's own grammatical rules. The Government and Binding model predicts that there can only be a switch under the government of the same language. Finally, the insertion approach embodies the asymmetrical roles of languages and predicts that the matrix language is the language that provides the grammatical frame to codeswitching utterances. The section concludes that the Matrix Language Frame model, an insertion model, will be the focus of the present study, as it is by far the most promising one available.

Chapter 2 provides an elaborate review of the theoretical frameworks employed in the present study. The markedness model centers on the societal norm and the indexicality value of linguistic choices. The rationality of an individual is added to the model to explain deviated choices made by speakers. The conversational codeswitching approach takes the conversation as its starting point. With regard to the syntactic model, the Matrix Language Frame model predicts that codeswitching can occur in the boundary of bilingual CP (the maximal projection of a complementizer), within which a Matrix Language supplies the morphosyntactic frame to the bilingual constituents. It also predicts that certain constituents, e.g., non-congruent constituents, can never be switched unless they appear as Embedded Language islands—Embedded Language constituents comprised of more than one morpheme.
Chapter 3 offers an ethnographic/sociolinguistic overview of the Thai participants in the present study. It addresses the status of the English language in the Thai society, discusses the methodology of participant observation, as well as provides a list of conventions used in the study. An outline of linguistic characteristics of the Thai language is also provided in this chapter.

Chapter 4 and 5 present detailed analyses of the Thai/English codeswitching data. Chapter 4 concentrates on the two sociolinguistic models and the data. Chapter 5 focuses on the grammatical constraints of the data according to the Matrix Language Frame model. The present study concludes in Chapter 6 with a summary of findings from both sociolinguistic and grammatical aspects of the Thai/English codeswitching data and an evaluation of the theoretical frameworks.
CHAPTER 2
THEORETICAL FRAMEWORK

The present study examines both sociolinguistic and syntactic aspects of Thai/English codeswitching. As stated in Chapter 1, different models will be tested against different aspects of Thai/English codeswitching—the markedness model and the conversational codeswitching approach will be tested against the sociolinguistic aspect of the Thai/English data, and the Matrix Language Frame model will be tested against the syntactic aspect of the data. Chapter 2 reviews these theoretical frameworks: the markedness model, the conversational codeswitching approach, and the Matrix Language Frame model, respectively.

2.1 THE MARKEDNESS MODEL

The theory of markedness distinguishes between marked and unmarked linguistic varieties. Community norms are the starting-point for this model. The markedness model is speaker-oriented such that speakers are the creative rational actors making code choices by weighing costs and rewards and associating codes with the rights and obligations set within a normative framework specific to their community. The premise underlying the markedness model is that all linguistic codes have social and psychological associations for the speech communities in which they are used. Given these associations, linguistic choices can be viewed in terms of marked versus unmarked opposition. In Myers-Scotton’s own words,

...choices are labeled unmarked when they constitute predicted behavior, given the context, that is, they negotiate the unmarked rights and obligations set for that context. In some sense, they maintain the community’s status quo. Choices are marked when they can be seen as negotiations to invoke a
With exposure to language in use, particularly the use of marked and unmarked codes in actual community discourse, language users have the ability to (1) recognize the degree of markedness of particular linguistic codes and (2) to comprehend that marked and unmarked choices will result in different consequences and receptions; that is, they have the markedness evaluator. When speakers choose to use a marked linguistic variety, they want to convey certain additional messages of intentionality or implication. Speakers have a particular intention when choosing a certain linguistic code over the other, and they expect their interlocutors to recognize the additional message or the particular intention carried by the selected code choice (Myers-Scotton 1998).

An important concept in the markedness model concerns the rights and obligations (RO) set, a theoretical construct for referring to what speakers can expect in any given conversational interaction. According to Myers-Scotton (1998), the RO set is another term for norm. Myers-Scotton (1993b) argued that code choices are indexical rather than symbolic, as “they have more than an arbitrary relationship with the RO sets with which they are identified” (1993b:86). For example, English in Kenya is associated with the unmarked RO set in both official encounters as well as means by which the high-level participants achieved their status. The mastery of English is directly linked to the educational system. In her own words:

The possibility of this indexicality derives from the fact that the different linguistic varieties in a community’s repertoire are linked with particular types of relationships, because they are regularly used in conversations involving such types. Through this type of accumulation, a code comes to index an RO set. (Myers-Scotton 1993b:85)

Additionally, the linguistic choices can be considered a type of negotiation in the sense that speakers make their code choices as goal-oriented actors. To explicate, each
linguistic choice offers different benefits; speakers would choose one giving them the most benefits relative to its costs. Marked and unmarked linguistic choices are then indexical of different RO sets between participants in a given interaction type (Myers-Scotton 1988, 1998). The following example from Myers-Scotton (1988) illustrates the linkage of a code with an RO set. In the example, a gatekeeper at the entrance of the IBM Nairobi head office converses with a visitor from the Luyia area of western Kenya in Swahili and Luyia. Speech in Luyia is italicized (original italics).

GUARD: Unataka kumwona nani?
‘Whom do you want to see?’

VISITOR: Ningependa kumwona Solomom I—.
‘I would like to see Solomon I—.’

‘Do you really know him? We have a Solomon A—. I think that’s the one [you mean].

VISITOR: Yule anayetoka Tiriki—yaani Mluyia.
‘That one who comes from Tiriki—that is, a Luyia person.’

GUARD: Solomon menyu wakhumanya valuhi?
‘Will Solomon know you?’

VISITOR: Yivi mulole umuvolere ndi Shem L— venyanga khukhulola.
‘You see him and tell him Shem L— wants to see you.’

GUARD: Yikhala yalia ulindi.
‘Sit here and wait.’

ANOTHER VISITOR (just appearing): Bwana K— yuko hapa?
‘Is Mr. K here?’

‘Yes, he’s here—he is doing something right now. He can’t leave until he finishes. Therefore you will wait here until he comes. You will wait about five or ten minutes.’

(Guard goes to look for Solomon A—.)

(Myers-Scotton, 1988:153-154)

13 Some examples of the costs are to reveal one’s ethnic background, or to discriminate or offend others by using an ethnic language that others do not understand.
The gatekeeper starts out his conversation in Swahili, the unmarked choice for two Kenyans (himself and the visitor) in this setting. He then switches to Luyia once the unmarked RO set between him and the first visitor changes; that is, once he knows that the visitor is also from Luyia.\textsuperscript{14} By switching to Luyia, the guard acknowledges the shared ethnic language and group membership—their new RO set. The guard then switches back to Swahili when a second visitor appears, indexing a more neutral RO set in this encounter.

According to Myers-Scotton (1988), the markedness model distinguishes four functions of code choice:

\textbf{Codeswitching as unmarked choice:}

a) Sequential unmarked choice: a switch from one unmarked choice to another, resulting in different RO balances when there is a redefinition of the exchange between participants. The switch of linguistic choice encodes the emerging unmarked relationship between participants and their recognition of this relationship.

b) Overall switching as the unmarked choice: a switch to another unmarked choice with no changes at all in the situation. This type of switch usually occurs between bilingual peers when the participants want more than one social identity to be salient in the current exchange.

\textsuperscript{14} Prof. Michael Forman questioned how the gatekeeper comes to know that the visitor is a Luyia person. This will be discussed in § 6.2.1.
Codeswitching as marked choice:

c) Codeswitching as a marked choice: a switch to the marked choice to negotiate different RO sets between participants. The switch can be either positive or negative depending on the situation.

d) “Permissible” marked choices: a switch to the marked choice that encodes marked RO sets but is considered unmarked in context. There are two types of this switch: (1) a switch for deferential purposes, and (2) a switch due to the lack of language proficiency in the unmarked choice.

There are also hypotheses and maxims underlying the markedness of a code choice:

The Unmarked-Choice Hypothesis (Myers-Scotton 1993b:89): A continuum of relative frequencies of occurrence exists so that one linguistic variety can be identified as the most unmarked index of a specific RO set in a specific interaction type, in comparison to other varieties also in use.

The unmarked-choice maxim, (Myers-Scotton 1993b:114): Make your code choice the unmarked index of the unmarked RO set in talk exchanges when you wish to establish or affirm that RO set.

The marked-choice maxim (Myers-Scotton 1993b:131): Make a marked choice which is not the unmarked index of the unmarked RO set in an interaction when you wish to establish a new RO set as unmarked for the current exchange.

The exploratory-choice maxim (Myers-Scotton 1993b:142): When an unmarked choice is not clear, use CS to make alternate exploratory choices as candidates for an unmarked choice and thereby as an index of an RO set which you favour.

The deference maxim (Myers-Scotton 1993b:147): Switch to a code which expresses deference to others when special respect is called for by circumstances.

The difference between the “overall codeswitching as the unmarked choice” type and the other types of codeswitching is that the overall pattern of the former provides the social message, while this message is provided at switching points in the other types of codeswitching. Individual switches in this type of codeswitching may have rhetorical functions; however, they do not index a new social message. Additionally, the overall
codeswitching is associated with familiarity in using those languages together, rather than other theoretical features, e.g., proficiency or indexicality of identity.

In her recent publications (Myers-Scotton 1998, Myers-Scotton and Bolonyai 2001), Myers-Scotton has argued that the rational choice model can explain why speakers select one linguistic variety over another. A quote from her follows:

Rationality functions both as a mechanism and as an explanation. As a mechanism, it directs actors to perform these three operations to find the best action; (i) actors consider their desires and values as well as prior beliefs; (ii) they confirm that these three elements are internally consistent; and (iii) finally, they make sure that their final desires, values, and beliefs take account of available evidence [including relevant norms and social meanings]. As an explanation, rationality tells us WHY choices are made. (Myers-Scotton and Bolonyai 2001:14, original small caps)

The rational choice model (Elster 1986) was developed by Jon Elster, a leading social scientist. The model draws a connection among different fields, including philosophy, history, social sciences, and cognitive psychology. The rational choice model stated that choices in interactions are best explained in terms of cognitive calculations depending on the actor's assessment as to which actions bring him/her the greatest utility. "That is, choices reflect a goal to enhance interpersonal relations and/or material or psychological rewards, and to minimize costs" (Myers-Scotton and Bolonyai 2001:6). This is in harmony with Myers-Scotton's idea of costs and rewards in her original markedness model. According to the rational choice view, linguistic choices are made by speakers as individuals. The available choices lie with the individual. This can explain why speakers from the same community make different linguistic choices. Much previous sociolinguistic research in the field of language variation has focused on describing and not explaining choices. Recent research in the variationist framework has found that

15 According to Melberg (on-line), "[Elster is a leading social scientist. He] has published on a wide range of subjects, including the theory of rational choice, Marxism, the methodology of the social sciences, justice, social order and political theory" (www.geocities.com/hmelberg/elster/elster.htm). This website also includes quotes from different reviews of Elster's work.
explanations for individual variation are not exclusively related to demographic variables. Choices are ultimately individually made; they reflect the speaker’s place in a social group. That is, the macro societal norms certainly affect choices. However, they do not determine actual choices. It is not sensible to think of the macro-societal norms as a mechanism for making choices. Besides, societal factors are not the only factors that play an important role in linguistic variation. Intra-speaker factors, such as motivations, are also determinant of code choices.

[The rational choice] theory is based on the assumption about human cognition that actors are oriented to seek optimality of an interpersonal nature in their actions, including their linguistic choices. The overall assumption is that the way speakers choose to speak reflects their cognitive calculations to present a specific persona that will give them the best “return” in their interactions with others, in whatever ways are important to them and are rationally grounded. (Myers-Scotton and Bolonyai 2001:23).

Moreover, Myers-Scotton also draws an analogy between her model and other premises developed in social science, law, and neurobiology regarding human intuition, the existence and use of social meanings in conversation, and decision-making for survival, respectively. These premises state that intuition is developed through experience (see Klein 1998); social meanings not only exist, they are dynamic and are used as a tool to accomplish a communicative goal; therefore, choosing certain words over others results in the signaling of one meaning rather than another (see Lessig 1995); and, lastly, human beings have an ability to call on previous experience to make a short-cut to their decision making (see Damasio 1998). To explicate, Myers-Scotton’s markedness evaluator is comparable to Klein’s and to Damasio’s ideas. All of these premises are grounded on cognitive calculations; they are all cognitively-based components that assist agents to make judgmental decisions by utilizing previous experience. Damasio (1998) developed the so called “somatic markers,” a device that helps actors make a decision by rapidly highlighting certain options and eliminating others based on their previous
experience. Analogously, according to the markedness model, speakers take account of the markedness evaluator, a sociolinguistic construct, in making linguistic choices. This particular sociolinguistic construct is also developed through speakers' experience in actual community discourse in which both marked and unmarked choices are exploited. As far as the existence of social meanings is concerned, the markedness model not only recognizes the existence of norms and social meanings, it also recognizes the dynamic quality as well as compelling and constraining nature of those norms.

In addition, the amended markedness model has three additional sequential filters, adapted from Elster’s two filters in his decision-making model (Elster 1989 cited in Myers-Scotton and Bolonyai 2001). The first filter is the social context and situational factors contributing to speakers’ linguistic repertoire. The second filter includes the markedness evaluator and speakers’ experience in terms of linguistic choices. The third filter is rationality, the mechanisms responsible for a selection of the actual linguistic choice. Rationality also offers an explanation as to why a certain choice is made. To sum up:

Choices in a rationally based model of linguistic variation pass through several filters. They begin with the external constraints on speakers: their linguistic repertoires, which in turn are constrained by large-scale societal factors and the discourse structure of their communities. They are also filtered to internal constraints, the innately available architectures (a markedness evaluator, somatic markers) that bias choices based on experience. Finally, choices pass through a third filter in which a social mechanism, rationality, is the centerpiece. To act rationally means that speakers take account of their own beliefs, values, and goals, and that they assess these in regard to internal consistency and available evidence. (Myers-Scotton and Bolonyai 2001:22)

The markedness model has been used as a framework to interpret codeswitching practiced by many researchers. In recent work, Herbert (2001) carried out research on the use of codeswitching in Johannesburg, a city in South Africa well known for its linguistic and ethnic complexity. In particular, Herbert found that, among the above types of codeswitching proposed by Myers-Scotton, the type called "overall codeswitching as the
unmarked choice” was the most common type of codeswitching found in the speech of his participants (see Herbert 2001). In another study, Myers-Scotton (1988) found that a conversation between two University of Nairobi students shows the use of codeswitching between Swahili and English as the unmarked choice. Myers-Scotton also found the use of codeswitching as a marked choice in two excerpts of conversations between a bus conductor and a passenger, the first in which the passenger spoke in Lwidakho (instead of Swahili) to negotiate for solidarity with the bus conductor, and the second in which the passenger switched to English when talking to the bus conductor to encode authority and educational status. Another use of codeswitching as a marked choice in a positive sense is found in her study in the conversation between a young Luyia woman who switched to Maragoli\textsuperscript{16} when talking to a Luyia gatekeeper to narrow social distance (see Myers-Scotton 1988). Goyvearts and Zembele (1992) also make use of the markedness model in their study of codeswitching in Bukavu, the capital of the Kivu region in Africa. They found that their participants use codeswitching both as a marked and unmarked choice; to indicate power or mark solidarity in the first case; and to signal their multiple identities between peers or to promote solidarity or maintain ethnic balance in the latter case (see Goyvearts and Zembele 1992).

Although the markedness model has become one of the most influential models of codeswitching study and has been employed by many analysts to examine codeswitching data, as shown above, it has also been criticized for its lack of interactional interpretation for codeswitching. Li (1998) commented that the “markedness” theory of codeswitching is the most influential model of codeswitching in terms of social and pragmatic aspects. It

\textsuperscript{16} According to Myers-Scotton (1998), Swahili is the unmarked code, and Maragoli is the shared ethnic language between the woman and the gatekeeper.
provides a framework to analyze the social and symbolic values of languages that are brought along into codeswitching practices in a postcolonial speech community or a speech community where diglossia or ethnic conflicts exist. Accordingly, the framework is a convenient tool for analysts to interpret and predict certain linguistic choices according to the presumptive social value attached to particular instances of codeswitching. Nevertheless, it explicitly rejects the idea that there are locally and creatively constructed social meanings of linguistic choices. Li further mentioned that the theory of markedness perceives codeswitching practices to be mono-directional such that situation, social meaning, and indexical value of codeswitching are brought into the conversation as given factors.

2.2 THE CONVERSATIONAL CODESWITCHING APPROACH

Gumperz (1982) focused on the communicative aspects of codeswitching. He argued that the association between group identity and communicative style is symbolic and does not predict the actual use of codeswitching, and that codeswitching signals contextual information equivalent to what in monolingual settings is conveyed through prosody or other syntactic or lexical processes. It has a communicative function in actual conversations; specifically, it serves as a contextualization cue and is commonly used in quotations, addressee specification, interjections, reiteration, and so forth (Gumperz 1982). In Gumperz’s own words (1982:57),

...language shift reflects basic changes in the structure of interpersonal relations rather than mere macro-alterations in the extralinguistic environment. We have suggested that the linguistic factors involved here are best studied at the discourse level in terms of cues which members use to signal the non-objective content of messages and evaluate the importance of what is said.
According to Gumperz (1982:131), a contextualization cue is "...any feature of linguistic form that contributes to the signaling of contextual presuppositions." Examples of contextualization cue features are code, dialect, style switching, prosodic phenomena, lexical choice, syntactic options, formulaic expression, and so forth. The term "contextualization" is referred to by Gumperz (1992:230) as "...speakers' and listeners' use of verbal and nonverbal signs to relate what is said at any one time and in any one place to knowledge acquired through past experience, in order to retrieve the presuppositions they must rely on to maintain conversational involvement and assess what is intended." The fact that the speakers understand the semantic information encoded in the codeswitched utterances and agree on what is accomplished in particular settings suggests the speakers not only share the codes but also share principles of interpretation within the group. The speakers have their own socially defined notions of code or grammatical system and this shared knowledge or background assumptions are learned through prolonged intensive interactions among the in-group members. This also explains why bilinguals do not always codeswitch with other out-group bilinguals until the backgrounds and attitudes of their interlocutors are known. To do otherwise may result in serious misunderstanding (Gumperz 1982).

Gumperz's idea of conversational codeswitching has developed into the "interactional model" or "conversation analysis of codeswitching." In this particular analytic framework, situations are not predetermined, but are interactively achieved by participants of the conversations. That is, the meaning of a given switch has to be interpreted with reference to language choices used in conversational interactions by the speakers; i.e. how social factors such as identity and attitude are presented, understood,
negotiated, and accepted or rejected in the process of interaction. Following Gumperz, Auer, a conversational analyst, concurred (1988, 1995, and 1998) that situation is not predetermined, but interactively achieved by participants of the conversations. This is also in accordance to Duranti and Goodwin (1992). That is, the meaning of codeswitching has to be interpreted with reference to language choices used in conversational interactions by the speakers. Accordingly, Auer proposed a conversation analytical approach to account for interactive meaning of codeswitching in term of indexicality. He suggested that a codeswitching analyst has to regard codeswitching as a conversational activity and must pay attention to the sequential development of interaction in order to truly understand the meaning of the phenomenon.

The conversational codeswitching approach focuses on the member’s procedures of local meaning of linguistic choices—the meaning that is constructed in the sequential organization of a face-to-face interaction. A comment on the approach from Li (1998:163) is instructive:

Those who adopt the CA approach to code-switching argue that we must not assume that, in any given conversation, speakers switch languages to ‘index’ speaker identity, attitudes, power relations, formality, etc.; rather, we must be able to demonstrate how such things as identity, attitude and relationship are presented, understood, accepted or rejected, and changed in the process of interaction.

Auer’s conversational codeswitching is extended to include background information in its analysis, unlike the original conversation analysis in which an analysis is based solely on the sequential order of the transcript. He said, “[i]n order to pinpoint the conversational meaning of such a case of [codeswitching], we need to know about the ‘episode-external’ preferences of speakers for one language or the other, or about the community norms for that particular kind of interaction.” (1995:121).
Auer (1984, 1995) stated that sequential organization of language choice provides a frame of reference for the interpretation of functions or meanings of conversational codeswitching. He subsequently proposed four sequential patterns of codeswitching, as follows:

(1) Discourse related codeswitching: this type of codeswitching contributes to the organization of discourse. It is usually found in the contextualizing of a shift in topic, addressee selection, etc. This pattern is schematized as followed. Type Ib shows that a switch to another language can occur within a single turn: Speaker 1 uses Language A (A1) and so does Speaker 2 (A2) until a switching point, marked by //. At the switching point Speaker 1 switches to Language B within a single turn (A1//B1). Speaker 2 subsequently accepts Language B (B2) as the language of interaction.

Type Ia: A1 A2 A1 A2//B1 B2 B1 B2

(2) Preference related codeswitching: this type of switching reflects the participants’ preference in term of code choices.


(3) Semi-discourse-related and semi-participant-related codeswitching: this type of switching usually may have a conversational function, such as to emphasize or to mark topic/comment.

Type IIIa: AB1 AB2 AB1 AB2
Type IIIb: AB1//A2 A1 A2
(4) Transfer: This type of switching is considered "transfer" by Auer (1984, 1995). It describes an insertion of another language in the middle of a speaker's turn without any effect on the language choice for the interaction. It can be considered discourse- or participant-related. It displays the participants' bilingual competence.

Type IV: \[ A[1][B]A1 \]

Pattern I is what is usually considered "conversational codeswitching" in the literature. Pattern II and III are not considered so. Rather, they are matters of language choice and language negotiation determined by societal or psycholinguistic factors, and thus usually analyzed in term of ethnographic description.

In summary, Auer proposed a distinction between discourse-related and participant-related switching: the former concerns the organization of the ongoing interaction, e.g., addressee selection, shift in topic, activity type, etc., while the latter concerns participants' preference for and competence in a language. However, the distinction is not always clear-cut. Type III switching is a mix between both discourse- and participant-related switching. Auer further suggested that participant-related codeswitching should be analyzed in terms of contextualization cues. However, its function as such goes beyond discourse structure, for it involves social attributes and relationship of the participants. Contextualization cues have to be analyzed according to their sequential organization, as their meanings are conveyed as part of the interactive process.

Another interactional sociolinguist, Li, asserted that codeswitching is a discourse strategy. It is one of many linguistic resources functioning as contextualization cues. According to Li, codeswitching is used to signal turn handovers. That is, participants
codeswitch at turn boundaries to indicate turn completion and turn transition. Codeswitching contextualizes turn transition by building up a contrast between code choices. A common type of codeswitching found, the tag-switching, is an example to support such a claim. He also found that dispreferred response is marked by codeswitching in his Chinese/English bilingual data, while it is marked by conversational features like pauses, apologies, prefaces like *well* in monolingual English conversations. Interestingly, this is found mainly in inter-generational conversation and less frequently among speakers of the same generation. Codeswitching is also used to contextualize self-initiated repair, by repeating in another language or by inserting words from another language. Li concluded that “there is no simple, one-to-one relationship between code-switching structure and community-level language preference. The latter does not constrain individual instances of code-switching, rather it offers a frame of reference for participants--and analysts too for that matter--to interpret the meaning of language choice in conversation” (Li 1994:178).

Interactional sociolinguists pay close attention to a fine-grained turn-by-turn analysis. Turns are constructed of turn-constructional units, including lexical, phrasal, clausal, and sentence constructions, signaling the end of a single turn. Examples of such signals include rising or falling intonation, choice of words at the end like *you know*, a gaze at the addressee, or other gestures.

One of the basic observations made by conversation analysts is that conversation interaction is characterized by an orderly sharing of speakership. In order to achieve smooth and frequent transitions from one speaker to another, conversational participants employ a ‘local management system’ – a set of rules with ordered options which operates on a turn by-turn-basis. (Li 1994:154)

Another commonly employed term in this analytic framework is “adjacency pair.” “Adjacency pair” includes question-answer, offer-acceptance, greeting-greeting, and so
on. Generally, “adjacency pair” is consisted of the first pair part and the second pair part. The occurrence of the second pair part is constrained by that of the first pair part. To explicate, if a question is asked, an answer is expected. The question is the first pair part that sets up a conditional relevant expectation which the second pair part, the answer, fulfills.

Nevertheless, as Auer pointed out, the conversational codeswitching approach alone cannot give a complete account to codeswitching phenomena. He stated that “such a theory of bilingual conversation obviously has to be complemented by another theory which explains who switches in a given community, why and when... Needless to say, bilingual work on any concrete bilingual community has to refer both to micro and macro theories of [codeswitching] and to their interdependencies’ (Auer 1995:116). To remedy this shortcoming, Li (1994:180) accordingly integrated the concept of social network to the conversational codeswitching approach. In his own words:

[0]ne particular important point emerging from the analysis is that social networks affect and are affected by their members language behaviour. On the one hand, social networks of individual speakers constrain their opportunities to learn and use particular languages or language varieties; on the other, use of certain language or language variety can contribute to the development of network contacts of individual speakers.

Such (inter-speaker) social dimensions of linguistic variation can be analyzed according to the social network approach which assumes that speakers’ language use is influenced by the social contact that they have, and try to maintain, with others. Two important implications of the social network approach are (1) that members of a given network have their distinctive patterns and they expect their members to conform to these patterns in order to maintain their membership, and (2) that members are the active contributors in constituting the social relations and social structure through their interactive behavior. Li’s Chinese/English data have shown that the parent and grandparent generations mainly
use Chinese and only occasionally switch to English to contextualize turn-allocation and repair initiators in conversations with their younger generations. On the other hand, the children mainly use English in their daily interaction with their peers. The children usually use codeswitching to contextualize dispreferred responses with the older generations. Li (1994) claimed that this approach is more feasible with those who are considered economically marginal, and those in homogeneous as well as territorially well-defined neighborhoods. He further claimed the conversational codeswitching approach as a bridge between the macro and micro approaches as well as the interrelation between them.

The rest of this section reviews the studies of researchers whose work follows the conversational codeswitching approach.

In his study of the children of Italian migrant workers, Auer (1988, 1998) found that his participants use codeswitching to cue the change in addressee selection, in mode of interaction (e.g., between a formal interview and casual talk); in topic, between informative and evaluative talk, and for sequential contrast. Furthermore, Auer found that these children also use codeswitching as a contextualization cue to set off “non-first firsts” (e.g., a repetition of questions), reformulations or elaborations, prefaces from stories, “setting” and “events” in narratives, and to distinguish between different types of information in conversation (e.g. “new” and “given”). In another study, Li (1998) proposed that codeswitching functions as a contextualization cue for turn competition, pre-sequences and embedded sequences, and preference organization. In other words, “[codeswitching] can help the speaker to restart a conversation at the end of an interactive episode, or to change conversational direction; it also helps the participants to
keep track of the main ‘drift’ of the interaction by mapping out complex nested structural patterns in the conversation” (Li 1998:169). Alfonzetti (1998) reported a similar observation: codeswitching in Sicily contextualizes self-repair, pre-closings, and sequential subordination by signaling the beginning and the end. It is also a contextualization cue for story-telling in relation to the preface, evaluative talk, comments, and climax of a story, topic change, and setting off quotations in terms of the change of footing. Some of the findings in Moyer’s (1998) work echo the results in the previous two studies. Moyer found that codeswitching in Gibraltar functions as a contextualization cue for topic, alignment of speakers to either English or Spanish, and humor generated by codeswitching, all of which play a role in portraying the ambivalence and the multifaceted identity of Gibraltarians. According to Moyer, there are three basic bilingual strategies used in this particular speech community: a selection of main language for the interaction (either Spanish or English), negotiations of language or languages between turn, and preference for various intra-sentential codeswitching structures. The choice of languages can show formality of the conversation with reference to a particular setting and context as well as the speaker’s alignment with Spanish or British identities. Similarly, the findings in Sebba and Wootton (1998) show that speakers’ identities are reflected by their choice of linguistic codes, namely London Jamaican Creole and London English. Their identities are not a priori brought along into the conversation as given factors. Rather speakers’ identities are constructed and negotiated throughout the conversation and its content and context by the use of codeswitching as a contextualization cue for turn taking, quotations, and so forth. In a quite recent work, Ben-Rafael (2001) studied the French-Hebrew (Franbreu)
codeswitching by francophone immigrants in Israel. She found that the immigrants in her study use codeswitching as a contextualization cue to differentiate the selection of addressees, end sequences, broach new subjects, reformulate, and so forth (see Ben-Rafael 2001).

In sum, previous work has shown that codeswitching can function as a contextualization cue and play a role in portraying the multifaceted identity of the speakers. Therefore, according to Auer (1995:120), it can and should be analyzed on a conversational level. The following is his outline of the activities in which bilinguals tend to switch from one language to another (original italics):

(i) reported speech
(ii) change of participant constellation, particularly addressee selection – this includes the use of code-switching in order to include/exclude/marginalize co-participants or bystanders
(iii) parentheses or side-comments
(iv) repetitions, i.e. quasi-translations into the other language, for example for the purpose of putting emphasis on demands or requests, or for purposes of clarification, or for attracting attention, e.g. in the regulation of turn-taking (also called ‘translations’, ‘repetitions’, or ‘recycling’)
(v) change of activity type, also called ‘mode shift’ or ‘role shift’
(vi) topic shift
(vii) puns, language play, shift of ‘key’
(viii) topicalization, topic/comment structure

In his Chinese/English data, Li (1994) found codeswitching at three different levels: A, B, and C. Level A is the level where switch is found at turn boundaries. This level of switching involves participants’ different attitudes and language abilities. Level B involves a switch within a single turn at a sentence boundary. Both the level A and B may be a result of the language preference of different generations. Level C is the intrasentential codeswitching within a single turn. It is found most prominently in the bilingual communities with an established history of language contact. Li additionally pointed out that, unlike his sequential approach, other approaches, including Myers-Scotton’s, do not take into account the former two levels concerning inter-sentential
codeswitching; they mainly focus on the intrasentential one. In his work, he examines codeswitching in terms of the organization and the management of conversation.

The conversational codeswitching approach is a promising trend in the study of codeswitching. It provides a framework to answer questions such as: (a) “What are the ‘codes’ in code-switching?” and (b) “How does conversational codeswitching relate to its wider ethnographically reconstructed (social and cultural) context?” (Auer 1998:2). However, like the markedness model, the conversational codeswitching approach has also been criticized. Myers-Scotton (1998) contended that in addition to the referential messages emerging from the analysis of sequential organization using the conversational codeswitching approach, there are critical messages or intentional messages that cannot be found by looking at the alignment of surface structures. The critical messages in discourse are at the cognitive level, particularly in the mental calculation of what is being inferred. According to Myers-Scotton (1998:22), “… speakers “know” that they can employ utterances for more than referential messages and … that addressees know this also and therefore will be looking for organization at this level.”

Chapter 4 discusses the application of these two competing models to Thai/English data.

2.3 THE MATRIX LANGUAGE FRAME MODEL

The Matrix Language Frame model is developed by Myers-Scotton (1993a, 2002) to account for grammatical constraints on codeswitching. Myers-Scotton (2002:14) claims that syntactic models devised for monolingual data cannot explain bilingual data, particularly codeswitching structures. In her own words, (original italics):
Of course such models ought to be able to account for the phrase structures of the participating languages; but they account for these structures as they are used in monolingual CPs. It is quite another task to account for how many elements from two languages can be combined in the same (bilingual) CP.

According to Myers-Scotton (2002), there are four premises underlying the Matrix Language Frame model. First, the Matrix Language Principle states that there is always a morphosyntactic frame set by a Matrix Language. It also predicts that content morphemes—elements in an open class such as noun and verb—are more likely to be codeswitched than system morphemes—elements in a closed class such as articles (Myers-Scotton 1992). Second, the Uniform Structure Principle states that “[a] given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this constituent type must be observed whenever the constituents appear” (Myers-Scotton 2002:8). That is, in bilingual speech, the morphosyntactic frame set by the Matrix Language is always preferred; however, that of the Embedded Language can be allowed under certain conditions. Third, the Asymmetry Principle for Bilingual Frames states that there is an asymmetry of participating languages in terms of the source of Matrix Language. Fourth, the Morpheme-Sorting Principle states that not all morphemes are equal. There is an asymmetry of different types of morphemes in participating varieties; only content morphemes can occur freely.

Theoretically, the Matrix Language Frame model is based on two oppositions: (1) the Matrix Language–Embedded Language opposition and (2) the content-system morpheme opposition. The members of these oppositions do not contribute equally to codeswitching structures. The language making a larger contribution is the Matrix

---

17 The Matrix Language versus Embedded Language opposition is relevant only when at least two linguistic codes are present within a CP showing codeswitching (Myers-Scotton 2002). This implies that the Matrix Language is not the language that supplies the morphosyntactic frame to all CPs. The grammatical frame of a monolingual CP comes from the language that occurs in the monolingual CP.
Language, and the other is the Embedded Language. Regarding the second opposition, content morphemes and system morphemes perform different functions: the former convey semantic and pragmatic messages, and the latter provide relations between the content morphemes.

As mentioned earlier in § 1.3.2.3, the Matrix Language Frame model has received some criticisms by other researchers, e.g., Muysken and De Rooij (1995). Myers-Scotton has made many important revisions as well as clarified ambiguous notions in the model. These major revisions include the definitions of the Matrix and Embedded Language, the distinction between the content-system morpheme, and the change of the unit of analysis.

### 2.3.1 Matrix Language and Embedded Language distinction

The Matrix Language Frame model requires the identification of a Matrix and an Embedded Language. In the early version of the Matrix Language Frame model, the Matrix Language was defined as “the language which sets the morphosyntactic frame for code-switching utterances.” The Embedded Language was the language that “appears in code-switching discourse in a frame provided by the Matrix language” (Myers-Scotton 1992:19). These definitions have been subsequently discarded due to their ambiguity in application. The Matrix Language is presently defined as the language that provides a grammatical frame in mixed constituents. The Embedded Language is then defined as the other language(s) used in codeswitching (Myers-Scotton 2002). The Matrix Language Frame model provides two principles in identifying the Matrix Language and the Embedded Language, as follows (Myers-Scotton 2002:59):

**The Morpheme Order Principle:** in Matrix Language + Embedded Language constituents consisting of singly occurring Embedded Language lexemes and any number of Matrix Language morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the Matrix Language.
The System Morpheme Principle: in Matrix Language + Embedded Language constituents, all system morphemes which have grammatical relations external to their head constituent (i.e. which participate in the sentence’s thematic role grid) will come from the Matrix Language.

A language satisfying the terms of both principles is the Matrix Language. According to Myers-Scotton (1997a), the Matrix Language can change within an utterance or even within a sentence, but cannot do so frequently. Essentially, the Matrix Language cannot change within a CP.

Myers-Scotton also asserted that the notion of a Matrix Language is necessary in the study of codeswitching; otherwise, there will be an over-prediction of possible codeswitched utterances. Without a Matrix Language, it will be possible for “the source of system morphemes to change from one constituent to another within the same CP” (Myers-Scotton 1997a:247). Another shortcoming regarding the notion of a Matrix Language was addressed by Muysken and De Rooij (1995), who found fault with the notion of a Matrix Language assuming the possibility of an alternation of the Matrix Languages in a conversation. As an argument against this criticism, Myers-Scotton (1997a) contended that this possibility should be permissible because discourse dominance in codeswitching is dynamic; it is subject to change when social values attached to each linguistic code change and when the proficiency of the speakers changes.

2.3.1.1 The roles of Matrix Language

With the refinement of the Matrix Language Frame model, the Matrix Language takes on a different role from that in the prior version of the model. In Myers-Scotton (1993a), the Matrix Language sets the grammatical frame for two of the three types of constituents: mixed constituents and Matrix Language islands. The third type, Embedded Language islands, was excluded from the previous model. Muysken and De Rooji (1995) problematized Myers-Scotton’s notion of embedding in part due to its failure to address
Embedded Language islands, which the more general claim now includes. Accordingly, a Moroccan Arabic/French sentence such as the following is not ruled out in the Matrix Language Frame model (adapted from Myers-Scotton 1997a:245):

(1) Tajziw tajdiru dak la regulation [sic] djal les naissances...
  T-A/IMPF/come/3P T-A/IMPF/take/3P that the regulation of the births
  ‘They come and do that-the limitation of the birth’

In the above example, the French Embedded Language island, *la regulation*, is well-formed in French. However, since this Embedded Language island is part of a larger Arabic NP constituent in the Matrix Language, it is preceded by the definite article, *dak*, in Arabic. In Arabic, demonstrative NPs must be preceded by definite articles (Myers-Scotton 1997a).

In addition, Myers-Scotton (1995) pointed out that the notion of the Matrix Language has a close relationship with the unmarked choice in her markedness model, reviewed in § 2.1. She stated that there is a correlation between the Matrix Language and the unmarked choice:

The sociolinguistic basis of the [Matrix Language] is clear since the [Matrix Language] is the language which is the more unmarked (in terms of its socio-psychological associations) for the interaction type in which the [codeswitching] occurs. Often the [Matrix Language] will be the language most associated with solidarity-building functions for the speakers (Myers-Scotton 1995:237).

2.3.1.2 The Embedded Language Island

Muysken and De Rooij (1995) criticized Myers-Scotton’s treatment of peripheral elements and questioned the ability to maintain her notion of Embedded Language islands. Further, they also questioned her explanation of peripheral elements. According to the early version of Matrix Language Frame model, Embedded Language islands involved two hypotheses: the Embedded Language Island Trigger Hypothesis and the Embedded Language Hierarchy Hypothesis. The Embedded Language Island Trigger
Hypothesis stated that Embedded Language non-congruent words could enter the Matrix Language in the form of an Embedded Language island, or a multi-word codeswitch. The Embedded Language Hierarchy Hypothesis set up a hierarchy for likely optional Embedded Language islands. It stated that (Myers-Scotton 1993a:144):

1. The more peripheral a constituent is to the theta-grid of the sentence (to its main arguments), the freer it is to appear as an [Embedded Language] island.
2. The more formulaic in structure a constituent is, the more likely it is to appear as an [Embedded Language] island. Stated more strongly, choice of (any) part of an idiomatic expression will result in an [Embedded Language] island.

To cope with the imperfect application of these two hypotheses, Myers-Scotton (1997a:250) proposes a new Embedded Language Island Hypothesis to replace both of the earlier hypotheses: all Embedded Language islands are now obligatory.

Embedded Language Island Hypothesis: When there is insufficient congruence between the lemma underlying an [Embedded Language] content morpheme and its [Matrix Language] counterpart at one or more of the three levels of lexical structure, the only way to access the [Embedded Language] element is in an EL island.

Take the following Spanish/English sentence for example (Myers-Scotton 1997a:251):

(2) No va-n a bring it up
   No go-3P to bring it up
   'They aren't going to bring it up'

According to Myers-Scotton, Spanish equivalence for *bring up* would be *no lo van a sacar* (translated morpheme by morpheme into English as *no it go-3P to bring*). In this Spanish equivalent sentence, versus the English *bring it up*, the Spanish verb appears without a satellite particle, and the nominal object precedes a finite verb. Assuming that *bring up* best conveys the speaker's intention, the only way to maintain both Spanish and *bring up* is to have an Embedded Language island in the utterance.

53
2.3.2 Content-system morpheme distinction

There is an issue of terminology confusion regarding the content-system morpheme distinction. Different researchers refer to these notions in different terms, such as closed-class items versus open-class items, and functional elements versus thematic elements. However, these terms do not encompass the same generalizations. According to Myers-Scotton (1997a), the distinction between closed-class and open-class items hinges on whether categories accept new members, and the distinction between functional and thematic elements refers to lexical category defining. However, not all members in these categories behave identically cross-linguistically in codeswitching. Therefore, the terms CONTENT and SYSTEM morphemes are more appropriate in the study of codeswitching as this distinction not only represents well the differential roles of morpheme types but also the relationship of these two morpheme types to one another. In her own words,

... in the composition of linguistic structures, content morphemes contribute to whatever semantic and pragmatic messages the speaker wishes to convey and, in this initiating role, they project directions selecting the system morphemes which are the nuts and bolts—the grammatical system—which give the constituent its form (Myers-Scotton 1997a:255).

Features define content and system morphemes; content morphemes are [+thematic role assigner/receiver, -quantification]; system morphemes are [-thematic role assigner/receiver, +quantification]. Accordingly, content morphemes include nouns, verbs, and other constituents that can occur in Comp position in a CP as well as elements that can occur in the spec of Comp. System morphemes, then, include quantifiers, determiners, possessive adjectives, degree adverbs, and tense and aspect (Myers-Scotton 1997a, 1997 b, 2002).

Additionally, the notion of thematic role is expanded to include discourse thematic roles. Thus, discourse markers assigning discourse thematic roles are classified
as content morphemes at the discourse level. Myers-Scotton and Jake (1995:984) describe content morphemes as follows:

Nouns, adjectives, time adverbials, and most verbs and prepositions are prototypical content morphemes. The feature shared by these lexical categories is that they constitute the predicate-argument structure, by either receiving or assigning thematic roles. Nouns and descriptive adjectives receive thematic roles; verbs, predicate adjectives, and some prepositions assign thematic roles. Discourse markers (e.g. well, because) are content morphemes because they assign thematic roles at the discourse level.

This expansion subsequently results in a revision of the characteristics of complementizers in codeswitching, rendering invalid their characterization as system morphemes in the previous version of the Matrix Language Frame model. The characteristics of complementizers now cover content morphemes from either Matrix or Embedded Language, provided congruence between these languages in term of complementizers exists.

2.3.3 CP as the Unit of Analysis

One of the major changes made to the Matrix Language Frame model is the change of the unit of analysis. In the previous version of the Matrix Language Frame model, the unit of analysis was the constituents within a sentence. It is now the CP, maximal projection of COMP. In Myers-Scotton’s own words, “A CP is the highest unit projected by lexical elements. It can be defined unambiguously in terms of phrase structure as a complementizer or an element in Specifier (Spec) position followed by an IP” (Myers-Scotton 2002:55). Myers-Scotton (1997a) claimed that the CP is a proper unit of analysis because: (1) it is used in syntactic theory and it is a point at which parameter settings and universal principles apply; (2) it allows more precision in testing the constraints of the Matrix Language Frame model because what counts as switching in a CP is clearer than in a sentence; and (3) it accounts for the status of constituents with
nulls because null elements can be in this particular position; accordingly all exclamations, such as *What!?* or *Never!*, are monolingual CPs containing null elements (Myers-Scotton 1997a, 2002). Additionally, the CP can become a bilingual constituent when it contains one or more mixed constituents or one or more CP-dominated Embedded Language Islands.

Furthermore, Myers-Scotton (1997a) added that the constraints presented in the Matrix Language Frame model would not be invalidated, because the constituents within a sentence, the former unit of analysis, are maximal projections within a CP of a sentence. The following are Swahili/English examples illustrating the divisions within sentences by CP, adopted from Myers-Scotton 2002:56-57:

(3) 

{Ndio wa-zungu wa-na-sem-a} [old habits die hard] 

Yes CL2-European CL2-NONPST-say-FV

‘Yes [as] Europeans say, old habits die hard.’

(4) 

{U-na-wez-a ku-m-pat-a a-me-va-a} 

2S-NONPST-able-FV INF-OBJ-find-FV 3S-PERF-wear-FV

nguo ny-ingine bright kama color y-a red] 

clothes CL9-other bright as color CL9-ASSOC red

‘You can find her (she is) wearing other bright colors [such] as red [ones].

‘You can find her [that] she is wearing other bright clothes [such] as red [ones].’

(5) 

{Lakini sasa wewe angalia profit} 

[amba-yo a-li-end-a ku-make] 

But now you look at profit rel-cl9 PST-go-FV INF-make

‘But now you look at [the] profit that he went [ahead] to make.’

(6) 

{Lakini a-na so many problems, mtu} 

[a-me-repeat mara ny-ingi] 

but 3S-with so many problems person 3S-PERF-repeat time CL9-many

‘But he has so many problems, [that] [he is] a person [who] has repeated many times.’

Example (3) is simply one bilingual sentence containing two monolingual CPs. Examples (4), (5), and (6) contain bilingual CPs. Example (4) contains one bilingual CP with mixed constituents. Example (5) contains two bilingual CPs with the second CP embedded in the first. Example (6) also contains two bilingual CPs, the first of which contains an
Embedded Language Island, *so many problems*. The second CP in this example contains a null constituent, referring to *problems*, in COMP.

Myers-Scotton previously defined codeswitching as "... the selection by bilinguals or multilinguals of forms from an embedded variety (or varieties) in utterances of a matrix variety during the same conversation" (Myers-Scotton 1993a:3). Given that the unit of analysis has been changed to the CP, intrasentential and intersentential codeswitching is redefined as follows (Myers-Scotton 1997b:223):

A CP shows intrasentential CS if it contains at least one constituent with morphemes from Language X and Language Y (a mixed constituent). This CP may also contain other constituents which are monolingual (i.e., [Matrix Language] or [Embedded Language] islands). Also, intersentential CS is now best defined as switching between monolingual CPs which are in different languages.

In addition to the revision of definitional issues and the unit of analysis, Myers-Scotton and Jake elaborate on psycholinguistic aspects of the Matrix Language Frame model in their recent publications (Myers-Scotton and Jake 1995, 2000, 2001, Myers-Scotton 2002). The model is largely influenced by psycholinguistic theories: it assumes the differential activation of two relevant languages (Grosjean 1988 cited in Myers-Scotton 1993a), different retrieval process of lexicon (Garret 1975 cited in Myers-Scotton 1993a), as well as the concept of lemma linking conceptual information and grammatical function (Levelt 1989). With these underlying assumptions, Myers-Scotton (2002) proposes two sub-models, the 4-M model and the Abstract level model, to support the content-system morpheme opposition and the concept of congruency in the Matrix Language Frame model.

### 2.3.4 The 4-M Model

The 4-M model is developed in support of the content-system morpheme oppositions. Not only does the model add precision to the Matrix Language Frame model,
it connects grammar with language production and processing. A quote from Myers-Scotton (2002:85) is informative:

... the [Matrix Language Frame] model implies that the Matrix Language - Embedded Language opposition and the content-system morpheme distinction are universal features underlying language production when classic codeswitching data is involved. However, the 4-M model offers indirect evidence about how language production actually works. This shows one way in which competence and performance are linked. It does this—its major virtue—by showing how surface morpheme distributions can be explained if they are related to abstract entries in the mental lexicon.

(Original italics)

The oppositions on which the 4-M model is based are as follows (Myers-Scotton 2002:73):

[+/- conceptually activated]
[+/- thematic role receiver/ assigner]
[+/- looks outside its immediate maximal projection for information about its form]

In accordance with these oppositions, there are four types of morphemes: (1) content morphemes: i.e., all nouns, adjectives, and most verbs, (2) early system morphemes: i.e., plural, and determiner, including articles and possessive adjectives, (3) bridge system morphemes: i.e., possessive of and 's, and (4) outsider system morphemes: i.e., subject-verb agreement, case affixes, and clitics/affixes. Both bridge and outsider system morphemes belong to the same class: late system morphemes. Figure 2 represents feature-based classification of these morphemes.
The assumption behind this model is that these four types of morphemes are related to production in different ways and are accessed differently in the process. When speakers intend to convey certain messages, their intentions activate language-specific semantic/pragmatic feature bundles that select best-fit content morphemes conveying the intended conceptual information. The content morphemes’ lemmas are directly related to speakers’ intentions in this case, and they are salient at the mental lexicon level. The lemmas of the early system morphemes are also salient at this level. However, these lemmas are not activated by speakers’ intentions, as in the content morphemes; they are activated by lemmas supporting content morphemes. In other words, lemmas supporting content morphemes select appropriate early system morphemes. An English noun phrase, *the book*, is an example of content and early system morphemes. The noun *book* is the content morpheme, and the determiner *the* is the early system morpheme, adding
definiteness to its head noun. Myers-Scotton (2002) claims that the early system morphemes are more likely to appear in bilingual clauses framed by the Matrix Language than the other two types of system morphemes, which are activated at the formulator level.¹⁸ That is, their lemmas are salient in the formulator level; thus their name—late system morphemes. The occurrence of these two late system morphemes in bilingual clauses produces larger linguistic units, such as CPs and IPs.¹⁹ An example of a bridge system morpheme is the preposition of as in ball of Lena (Myers-Scotton 2002:18). Subject-verb agreement is a good example of outside system morphemes. They are called “outside” system morphemes because these types of late morphemes “... look outside their immediate maximal projection for information about their form” (Myers-Scotton 2002:18).

To sum up, there are four different types of morphemes; each of which involves different activations in the production process. In general, content morphemes and early system morphemes are conceptually activated, whereas bridge system morphemes and outsider system morphemes are structurally activated. Finally, Myers-Scotton (2002) claims that the oppositions underlying the 4-M model are universal; however, the classification of morphemes may differ cross-linguistically.

2.3.5 The Abstract Level Model

The Abstract Level model yields further support to the Matrix Language Frame model—the element of congruence checking in particular. The Abstract Level model is concerned with what counts as “sufficient congruency.”

---

¹⁸ There are levels of activation in the bilingual production process: (1) the conceptual level, (2) the lemma level, and (3) the formulator level. The formulator level is the last level for lemma activation (Myers-Scotton 2002).
¹⁹ IP stands for “inflection phrase” in x-bar theory and Government and Binding theory (Crystal 1997).
Similarly to the 4-M model, the Abstract Level model is based on the nature of mental lexicon. According to Myers-Scotton and Jake (1995, 2001) and Myers-Scotton (2002), all three of her models, including the Matrix Language Frame model, the 4-M model, and the Abstract level model, presuppose the model of language production illustrated below.

**Conceptual level**

![Conceptual level diagram](image)

**Lemma level**

![Lemma level diagram](image)

**Functional level**

![Functional level diagram](image)

**Positional level**

![Positional level diagram](image)

Figure 2. Production process diagram: lemma activation (adapted from Myers-Scotton 2002:24)

According to the diagram, the production procedures begin at the conceptual level, where speakers map their intentions onto language. Bilingual speakers decide whether to produce bilingual speech at this level, taking into consideration social value and attitude toward codeswitching in a given community. The behavior of language switching could
be acceptable in some communities, while denigrated in others. If the speakers decide to produce bilingual speech, they will have to select a Matrix Language that sets a morphosyntactic frame for their utterances. At this point, speakers’ intentions would choose language-specific semantic/pragmatic feature bundles interfacing with language-specific lemma in the mental lexicon. These bundles are what mediate between the conceptual level and the lemma level or mental lexicon. According to Myers-Scotton (2002:24):

Lemmas are tagged for specific languages in the speaker’s repertoire. They contain all necessary information that will result in surface-level morphosyntactic structures. This information is what is referred to in the Abstract Level model as the three abstract grammatical levels of lexical-conceptual structure, predicate-argument structure, and morphological realization patterns.

As stated in the quote above, the Abstract Level model (Myers-Scotton 2002) assumes three levels of abstract lexical/grammatical structure: (1) lexical-conceptual structure, (2) predicate-argument structure, and (3) morphological realization patterns. Congruence checking occurs at these three levels. Level (1) refers to the matching between semantic/pragmatic feature bundles and speakers’ intentions at the conceptual level. Level (2) refers to the mapping relations between thematic role assigners and the relevant arguments. Level (3) refers to requirements by grammatical constraints for surface-level realizations in terms of elements and constituent orders. According to Myers-Scotton (2002), singly occurring Embedded Language constituents that appear in the Matrix Language frame have to pass the congruence checking at all three of these levels. To explicate, congruence checking begins when speakers’ intentions implicate lemma selections of the Embedded Language elements. If there is sufficient congruence between the Embedded Language elements and the Matrix Language counterpart, the Embedded Language elements can appear at the surface structure within the grammatical frame
provided by the Matrix Language. Sufficient congruence means congruence to the extent that the encoded features satisfy the requirements of the Matrix Language Frame model.

In case of insufficient incongruence, the Embedded Language content morphemes will be blocked: the Blocking Hypothesis. According to Myers-Scotton and Jake (1995:986-987), “[i]ncongruence in pragmatic messages in cross-linguistic lemmas that are otherwise nearly identical often motivates mixed constituents in [codeswitching]; other types of congruence problems motivate other [codeswitching] patterns, such as [Embedded Language] islands ...” Other compromise strategies can also be employed when there is insufficient incongruence: (1) producing bare forms of Embedded Language elements without relevant Matrix Language system morphemes, and (2) producing do constructions instead of Embedded Language verbs with inflections from the Matrix Language (see Myers-Scotton 2002).

To sum up, a central assumption behind the Matrix Language Frame model is that there is an asymmetrical relationship between two participating languages: the Matrix Language and the Embedded Language. The Matrix Language is the language that provides a grammatical frame in mixed constituents. The Embedded Language is the other language that is used in codeswitching. The revised Matrix Language Frame model rests on two principles, recapped below (Myers-Scotton 2002:59):

The Morpheme Order Principle: in Matrix Language + Embedded Language constituents consisting of singly occurring Embedded Language lexemes and any number of Matrix Language morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the Matrix Language.

The System Morpheme Principle: in Matrix Language + Embedded Language constituents, all system morphemes which have grammatical relations external to their head constituent (i.e. which participate in the sentence’s thematic role grid) will come from the Matrix Language.

The concept of congruence between the Matrix Language and the Embedded Language is also vital to the model. The Blocking Hypothesis (Myers-Scotton 1993a:120) states that
"[i]n [Matrix Language] + [Embedded Language] constituents, a blocking filter blocks any [Embedded Language] content morpheme which is not congruent with the [Matrix Language] with respect to three levels of abstraction regarding subcategorization." Also, in the revised Matrix Language Frame model, Myers-Scotton proposed two psycholinguistic based models: the 4-M model and the Abstract Level model to support the Matrix and Embedded Language opposition as well as the content-system morpheme opposition in the Matrix Language Frame model.
CHAPTER 3
RESEARCH DESIGN AND METHODOLOGY

This chapter addresses information regarding the participants, methodology, procedures for data collection, and languages involved in the present study.

3.1 DATA COLLECTION

For the present study, data collection involves naturalistic participant observation and recordings of conversation. This is supplemented by information gathered through interviews.

3.1.1 Participant Observation

To gather everyday naturally occurring conversational data, the researcher (myself) first asked the participants for their permission to tape-record their conversations. The participants were informed that their conversations would be recorded but they were not told when the actual recording would be. The conversations were recorded on a small digital Sony IC-recorder ICD-BP220 over the course of two years, and in different informal occasions such as group outings, get-togethers for lunch or dinner, or at temple events and festivals. The recorded conversation is approximately 80 hours. The presence of the researcher did not interfere or change the nature of the interactions of the in-group get-togethers as the researcher herself had long been an in-group person, being friends with the participants, taking part in social activities, and helping out with the temple for

---

20 Most of the previous Thai/English codeswitching studies were carried out by using the matched-guise technique, questionnaire, self-report, and interviews, e.g., Warie (1977), Rukthamying (1995), Thitiwattana (1996), and Boonkongsaen (1999). The present study differs from the previous literature in its methodology: long-term participant observation and naturally occurring data. Interview is also necessary in the present study. The purpose of the interview is to elicit speakers' opinions concerning Thai-English codeswitching.

21 To understand how language is actually used, analysts should depend on naturally occurring data as opposed to their own judgment or intuition.
over four years. The observer's paradox (Labov 1972, 1973), which is the interfering effect of one's knowing they are being observed, is minimized in this case. All instances of shift from Thai to English and vice versa were transcribed from the recordings by the researcher.

According to Li (1994), “the methodology of participant observation contrasts with survey research in that data collection and analysis are carried out consecutively in the latter but concurrently in participant observation” (original italics). That is, the methodology used in survey research is usually a linear process, beginning with research question(s) and formulating a hypothesis. A research tool is subsequently designed and the data are collected from the sampled group according to predefined procedures. The collected data are then organized or coded and analyzed. Participant observation, on the other hand, starts with some general questions in the mind of the researcher who is already out in the field collecting data and trying to make sense out of the data he or she has. Therefore, the data collecting and its analysis take place at the same time. Also, the researcher constantly defines and redefines questions according to the actual situation. The advantages of this particular methodology are stated in Milroy (1987:78, original italics):

The very high quality of the data in terms of capacity to provide a good sample of everyday language is of course the major advantage of a participation observation method. Another major advantage is the insight it is capable of yielding into the social and communicative norms of the community. Under this head is included not only information on informal social ties and organization, but also the fields of study generally described as ‘the ethnography of speaking’ (Saville-Troike 1982) and ‘interactional sociolinguistics’ (Gumperz 1982; see further 8.4.1). A third advantage, pointed out by Labov (1981:25), is that ‘by emphasizing deeper studies of groups and social networks, we gain in the possibility of explaining linguistic behavior’, in other words, we might derive insights into why a speaker's language occupies a particular position in a wider social structure.

The present study adopts this methodology because it yields simultaneous naturally occurring data in everyday life. Since one of the objectives of the present study
is to discover language choice patterns as well as its functions in naturally occurring contexts, the methodology of participant observation is the most appropriate fieldwork method for such a micro-interactional level study.

3.1.2 Interviews

In addition to the participant observation, the present study also employs another method of data collection: an interview. Two types of interviews were conducted in the present study: the attitudinal interview and the retrospective interview. After the conversations were recorded and transcribed, the participants were invited to participate in the one-on-one interviews with the researcher. The purpose of the attitudinal interviews was to elicit speakers’ opinions concerning Thai-English codeswitching. The information regarding their opinions and attitudes towards Thai/English codeswitching is necessary in the analysis of the data, e.g. whether or not the participants share the same background information regarding Thai/English codeswitching. Each participant was interviewed in a casual setting: in the participant’s room. Each interview lasted approximately fifteen minutes to half an hour, but was begun with no time limit. The interview consisted of a combination of highly structured and open-ended interview questions. During the interviews the researcher tried to hold to the sequential order of the questions as much as possible. Follow-up questions are also allowed. The list of questions used during the interviews is included in Appendix A.22

The notion of intentionality has been dealt with in different fields of study, e.g., anthropology, psychology. For anthropologists, interpretation of conversational activities relies on conventionality as much as on intentionality. It is possible and quite common

---

22 The list of the questions is in English. The participants have the choice of being interviewed in English or in Thai. All the participants in the study chose to do the interviews in Thai.
that interlocutors respond to the contextually relevant conventions and ignore the issue of speakers’ intentions. That is, the interlocutors can recognize words or actions as having a particular conventionally defined goal (Duranti 1993). In the present study, the researcher, being a Thai speaker and a Thai/English codeswitcher herself, has insights to the interpretation of Thai/English codeswitching practiced by the participants. However, to ensure and enhance the validity of the research results, the researcher also conducted a retrospective interview with all the participants, by showing the participants the transcription and asking them for their interpretations of each conversation excerpt. This is done with an assumption that the participant’s choices of linguistic varieties are potentially conscious and subject to introspection.23

The retrospective interviews in the present study require the participants to reflect on their own use of Thai/English codeswitching as well as others’. According to Fræsch and Kasper (1987) and Robinson (1992), introspective methods help researchers gain better insights into language production process. However, Poulisse et al. (1987) stated that the use of introspective data, including retrospective data, could be controversial because such data might be affected by researcher bias. In their discussion of retrospective data, Ericsson and Simon (1984) argued that retrospective data can be reliable if the data are collected under certain conditions: the data should be collected immediately after the task, contextual information should be provided to the subjects, information asked must be retrievable and specific, no leading questions should be asked, and the informants should not be informed that they will be asked for retrospective comments after the task. The present study satisfies most of the conditions except that the

23 During the interviews, all the participants reported that they are very conscious about the act of codeswitching.
retrospective interviews were not conducted immediately after each conversation. It is in fact infeasible with the nature of the data collection in the present study. The main methodology adopted in the present study is the long-term participant observation. The participants were not specifically told what the researcher is investigating. They were told that their conversations would be recorded but they were not told when the actual recording would be. Asking retrospective questions right after the recording of each conversation is thus impracticable in the present study. Also, doing so would unnecessarily draw participants' attention to their use of Thai/English codeswitching and would affect their codeswitched speech as a whole. To remedy this problem, the researcher provides all the contextual information the participants may need to activate their memories. The detailed transcription also lends itself very helpful in this case. Examples of questions asked during the retrospective interviews are: Why did you say that? Why did you switch to ... (English, Thai, or Chinese)? Why do you think Speaker ... (A, B, or C) switches? In cases in which the participants cannot recall why they switch, some evaluation questions are asked: What were your alternatives? What else could you have said? What did you notice about the situation? Would it be any different if you/they were to use another language?

3.2 BACKGROUND INFORMATION

This section gives a brief history of Thais since their first arrival in Hawai‘i, followed by a description of fieldwork conducted among the Thai participants on O‘ahu.

Information regarding the Thai population in the United States or particularly in Hawai‘i is very scarce. The information presented in this section was thus accumulated
from a few existing materials and the author's participant observation within the community. Because different linguistic varieties have different social and indexical values in different communities and because all the participants have originated from Thailand, it is essential to consider the background information including the history, people, and languages of Thailand. This section is intended to give an overview of such background information.

3.2.1 The Participants: Thai People in Hawai‘i

The Thai-English speakers participating in the present study are graduate students, monks, and permanent residents\textsuperscript{24} who come from different parts of Thailand and speak different dialects. The permanent residents include graduate students who are married to Americans; these students intend to stay in the United States after their graduation. All of the participants speak Standard Thai. However, some participants can speak the northern dialect, some the northeastern dialect, and others of Chinese ancestry can speak a dialect of Chinese. The participants in this study were selected for their verbal fluency in their spoken languages and dialects. The two main languages involved in this study are standard Thai and English; however, some Thai dialectal utterances and Chaozhou, a dialect of Chinese, are observed. There are a total of thirty Thai speakers in the present study, fifteen male and fifteen female speakers. Among these, there are four speakers who are monks, five speakers who are of Chinese ancestry, four speakers who can understand Lao, and two who can understand the northern dialect of Thai. On arrival in the state of Hawai‘i, Thai people must adjust their speech repertoires. English is no

\textsuperscript{24} The term “permanent residents” is used to refer to Thai emigrants. The participants prefer to be referred to by this term, rather than “emigrant” or “immigrant.”
longer their private code, as it could be in Thailand; Thai, on the other hand, has gained that function. It serves as an ethnic symbol (Surawan 1975).

With regard to the number of Thai population in the state, the Hawai‘i State Data Center has a record of a total of 819,914 residents who are five years old and over, in Honolulu County. Of this number, there are 1,504 people who were born in Thailand. In addition, according to the Census 2000 Sample Data File, there are a total of 1,496 Thai speakers in the state of Hawai‘i. That accounts for 0.1% of the whole state census (a total of 1,134,351 people in the state of Hawai‘i). Among the 819,914 residents on the island of O‘ahu, there are 1,106 Thai speakers, who account for 0.1% of the whole population on the island (http://www.census.gov/population/cen2000). Detailed information regarding the Thai population in the state of Hawai‘i is provided in Appendix B.

Furthermore, there are two Thai temples on the island of O‘ahu: the one that is the focus of the present study is the Wat Buddhajakramongkolvararam at Pearl City. This temple is the fifteenth Thai temple of The Council of Thai Bhikkhus in the United States. It was established in June 19, 1986 as a Non-Profit Organization, permitted by the U.S. government. The temple was planned and built with the patronage of Thai, Laotian, and Cambodian Buddhism followers. At present, there are four monks at the temple. According to the data-base of the Thai temple, there are approximately 262 people who support as well as participate in the temple events. This number includes both Thais and non-Thais (Wat Buddhajakramongkolvararam, Personal Communication, March 2005).

According to the Spring 2004 Fact Sheet provided by the International Student Services (ISS), University of Hawai‘i at Mānoa, there are 1,562 international students
registered for Spring 2004 under a non-immigrant student visa. Among these, there are a total of 51 Thai students (http://www.hawaii.edu/issmanoa).

This group of participants can be considered as "the functionally bilingual speaker" in the typology of immigrant communities, proposed by Li (1994). The key factor in the typology is the different social network structures of speakers and their relationships with language choice. According to him, there are six types of immigrant communities: (1) the monolingual community language speaker, (2) the functionally monolingual community language speaker, (3) the functionally bilingual speaker, (4) the "mixed" bilingual speaker, (5) the functionally bilingual 'host' language speaker, and (6) the functionally monolingual "host" language speaker (see Li 1994). The Thai participants in the present study belong to the third category. The first language of this type of bilingual speakers is the ethnic language which they use in many key social contexts. They can codeswitch, but usually they keep the two languages separated. They have network ties with both Thais and non-Thais. They can be considered as a bridge between the monolingual and the wider community where the use of English is necessary.

3.2.2 English in Thai Society

As far as English is concerned in Thailand, it is a foreign language, mainly used in international situations. The English language was first introduced into Thailand as early as 1612 through contact with British people. In the beginning, it was taught only to members of the royal family (Ministry of Education, Thailand 1981:1-5). During the process of westernization and modernization in Thailand (1851-1960), a number of selected young people were sent to many foreign countries to extend their studies for the

25 Even though this typology was originally designed for immigrants, the linguistic behavior and background of the graduate students participating in this study also fit with this type.
purpose of the country's development. As a result, there was an influx of English words during that time and we find the first attestations of switching and borrowing between Thai and English (Warie 1977:25). At that time, when English was virtually restricted to the royal family and the upper class, the ability to communicate in English had become the mark of the educated and the elite. It symbolized knowledge, power, sophistication, and modernism. It was also a way to "show one's ostentation" as Warie (1977:25) states:

In those days, when foreign advisors flourished and the students who returned from Europe wanted to display their superior knowledge of things, Europeanism, particularly English words and phrases, were imported wholesale...

English is now formally taught in school as a foreign language from the upper elementary level to the sophomore year of college. Among Thai people, English is best regarded for its value and prestige in business situations and society as a whole, as well as for its value in academic fields. Knowledge of English has been referred to as "a stepping stone by which Thai people can get from one level of society to another, higher one" by Surawan (1975:2). Such knowledge can help one attain a better job, a higher salary, and higher regard in society. At present, those holding degrees from foreign countries are an exclusive group of people assured of securing better jobs at many well-established companies. Many "classifieds" these days target only those who have such academic exposure.

Today, English is used widely among Thai people, but it is still the language of the educated, especially with other English speakers and more noticeably among educated Thais. This group includes those who have been exposed to English-speaking countries, as well as those who do not have such exposure but have proficiency in English, such as students majoring in English and those who often engage in English conversations in the workplace. English items ranging from single words to whole
expressions appear in Thai discourse, including loanwords and technical terms. According to Surawan (1975), the mixture of English and Thai is mostly used in informal situations among in-groups. When used with strangers or in formal situations, it is likely to be critiqued as an attempt by new graduates from abroad to show off. In addition, it is deemed a "flaw" or a "corruption" of the language (Thonglor 1970 cited in Warie 1977) and as such considered "improper" speech. Since the reign of King Rama VI (1910-1925) the linguistic "purism" movement has arisen and is felt in attempts to purify the Thai language of non-native vocabulary. These efforts have been successful in certain milieus (Warie 1977:26, 35).

3.2.3 Thailand and Its Language

The meaning of its name, Thailand, is literally the "land of freedom." Thailand has never been colonized. Its national language is the Central Thai or standard Thai. Generally, there are four regional dialects in Thailand: (1) Central Thai, (2) Northern Thai, (3) Northeastern Thai, and (4) Southern Thai. These dialects are different in tones and word choices. Central Thai is the language of urban people as opposed to rural. The other Thai dialects are usually used among peers who share a similar background.

In addition, Thailand has a concept of behavior appropriate to rank that extends throughout the Thai society. This behavior by all means includes linguistic behavior. Social rank in Thai society is divided by relationship between speakers, drawing on social position, age, gender, politeness, and so on. There are different lines of vocabulary to be used with different people with different status. There is a set of vocabulary to be used with the King and the royal family. There is also one to be used with monks. As for the common people, politeness and familiarity between the speakers play an important role in
determining one’s word choice. For examples, the words ‘to eat’ in Thai are [sawej] to be used with the king, [chan] to be used with monks, [rapprat ha:n] to be used among commoners who wish to express politeness in formal situation, [tha:n] to be used among commoners who wish to express moderate politeness, [kin] to be used among close friends and family members.

Kinship terms are another characteristic of the Thai language. It is customary among the Thai people to address family members and each other by the appropriate kinship term, e.g., grandmother, father, mother, mother’s elder brother, mother’s younger sister, father’s elder brother, older sister, older brother, younger sister, younger brother, and so on and so forth. For people who are employed, their professional positions can also be used as address terms, e.g. [phu:catka:n] ‘manager’, [sa:tsatra:ca:n] ‘professor’, and [aca:n] ‘teacher, master’. These address terms can be used alone or together with persons’ names. The system of address terms is influenced by the authority inherent within the family as well as the workplace.

Additionally, kinship terms can be used as first, second, or third person pronouns in Thai. It can occur alone or with a person’s name. For examples:

(1) 1st Person Pronoun: na: (p^at) paj sm: kho:o rna:
Auntie Pat go buy stuff come
‘Auntie (Pat) went shopping.’ (‘I went shopping.’)

(2) 2nd Person Pronoun: na: (p^at) ca paj rmplaw k^a
Auntie Pat Asp go Q. Part.
‘Are you (Auntie Pat) coming?’

(3) 3rd Pronoun: na: (p^at) maj ma:
Auntie Pat Neg. come
‘Auntie (Pat) didn’t come’

26 The transcription of Thai in the present study follows Burusphat et al. (1999). See Appendix C.
Also, a person’s name can be used as first, second, or third person pronouns.

(4) 1st Person Pronoun: \[p^h\text{at } k^h\text{it } wa: \ p^h\text{at } ju: \ b\text{a:n } di:k^*a:\]
Pat think that Pat stay home better
‘Pat (I) thinks that Pat (I) had better stay home.’

(5) 2nd Person Pronoun: \[p^h\text{at } ca \ p\text{aj } ro:rig\text{ian } maj\]
Pat Asp go school Q.
‘Will Pat (you) go to school?’

(6) 3rd Pronoun: \[p^h\text{at } majdaj \ k\text{i}n \ k^h\text{aw } wanni:\]
Pat Neg. eat rice today
‘Pat (she) didn’t eat.’

The Thai language also has a set of sentence-final particles. Different dialects make use of different sets of these particles. The sentence-final particles are mostly related to the expression of politeness and to the emotions of the speakers, e.g., \[k^h\dot{\alpha}\] to express politeness by female speakers, \[k^h\dot{r}\text{áp}\] to express politeness by male speakers, and \[wa\] to express dissatisfaction or anger in central Thai. Similar particles for the Northern Thai dialect are \[c\dot{a}w\], \[k^h\dot{\text{áp}}\], and \[b\dot{\text{á}}\], respectively (Wimonkasem 1995).

Additionally, \[ja\] indicates dissatisfaction and sometimes sarcasm. It is used among close friends. \[na\] or \[na:\]\:
indicate request, persuasion, and sometimes surprise.

Thai belongs to the Kam-Tai language family. It is an isolating language; thus, the most obvious difference between English and Thai is one of inflectional morphology: English has inflectional morphemes while Thai has none. Information about number, tense, and gender of a word in question is given by the addition of extra words. Examples are:

(7) \[s\text{unak } laj \ tua \ ji: \ b\text{on } t^h\text{anon}\]
dog many Cl be on street
‘Many dogs are on the street.’

(8) \[nok \ s\text{o:n } tua \ bi\text{n } ji: \ b\text{on } fa:\]
bird two Cl fly Asp on sky
‘Two birds are flying in the sky.’
The above examples show that information on number can be given to a word in question by adding a Numeral and a relevant Classifier to a noun, or adding a Quantifier.

(9) khaw klap ba:n mma chaw ni:
 3rd Pro return home when morning this
  'He went home this morning.'

(10)dm:ŋ paj sm: kʰɔŋ tʰi: tala:t lam:w
  Dang go buy things at market Asp
  'Dang went to buy things at the market (already).'

(11)mːw tua mia kamlaj kin pla:
cat Cl female Asp eat fish
  'The female cat is eating fish.'

(12)chan ca kin kʰanom
  1st Pro will eat snack
  'I will eat a snack.'

(13)khaw paj phrujni:
  3rd Pro go tomorrow
  'He will go tomorrow.'

Past tense is indicated by a temporal adverbial clause, muachawniː, in (9); an aspect word, lamːw, in (10). Progressive tense is marked by aspect words, kamlaj, in (11); and juː in (8) above. Future tense is indicated by an aspect word, ca, in (12); and a temporal adverb, phrujniː, in (13). Female gender is indicated by a noun phrase meaning 'female', tua mia, in (12).

Regarding syntax, English and Thai share the same basic word order: SVO. However, the sequential order of some grammatical categories in Thai differs from that in English. The structure of the simple sentence in Thai is formulated as follows:

(14) S → NP + VP

For example (adapted from Pankhueankhat 1998:222),

(15) fon tok
    rain fall
The adverbial phrase can appear in front of the noun phrase or after it. It can also appear at sentence-final position.

\[(16) S \rightarrow (\text{AdvP.}) \text{NP (AdvP.)} + \text{VP (AdvP.)}\]

Below are some examples:

\[(17) [\text{mùawanni:}] \text{AdvP} [\text{fon}] \text{NP} [\text{tʰiː:\text{chiammaj}]} \text{AdvP} [\text{tok}] \text{VP} [\text{nak maːk}] \text{AdvP} \ni \text{rain} \ni \text{at Chiangmai} \ni \text{fall} \ni \text{heavy very}\]

\[(18) [\text{mùawanni:}] \text{AdvP} [\text{fon}] \text{NP} [\text{tok}] \text{VP} [\text{nak maːk}] \text{AdvP} [tʰiː:\text{chiammaj}] \text{AdvP} \ni \text{yesterday} \ni \text{rain} \ni \text{fall} \ni \text{heavy very} \ni \text{at Chiangmai}\]

\[(19) [\text{fon}] \text{NP} [\text{tok}] \text{VP} [\text{nak maːk}] \text{AdvP} [tʰiː:\text{chiammaj}] \text{AdvP} [\text{mùawanni:}] \text{AdvP} \ni \text{rain} \ni \text{fall} \ni \text{heavy very} \ni \text{at Chiangmai} \ni \text{yesterday}\]

\[(20) [tʰiː:\text{chiammaj}] \text{AdvP} [\text{fon}] \text{NP} [\text{tok}] \text{VP} [\text{nak maːk}] \text{AdvP} [\text{mùawanni:}] \text{AdvP} \ni \text{at Chiangmai} \ni \text{rain} \ni \text{fall} \ni \text{heavy very} \ni \text{yesterday}\]

Below are more examples illustrating Thai sequential order for some grammatical categories (adapted from Burusphat 1999:8-9):

\[(21) \text{Verb} + \text{Adverb} + \text{Degree word} \ni \text{e.g.} \ni \text{kin chaː maːk} \ni \text{eat slow very}\]

\[(22) \text{Noun} + \text{Adjective} \ni \text{e.g.} \ni kʰon diː \ni \text{person good}\]

\[(23) \text{Noun} + \text{Possessive} \ni \text{e.g.} \ni kradaːt kʰoːŋ kʰaw \ni \text{paper of 3rd pro}\]

\[(24) \text{Noun} + \text{Numeral} + \text{Classifier} \ni \text{e.g.} \ni miː saːm tua \ni \text{bear three Cl.}\]

Examples (21) and (22) show that degree word and adjective occur after the constituents they modify. Examples (23) and (24) illustrate the word order of possessive noun phrase and a noun phrase containing a classifier, respectively.

In addition, there are no articles in Thai. Also, the structure of affirmative and interrogative sentences is the same. However, an interrogative word will be added at the
end of the interrogative sentence. Additionally, Thai is a pro-drop language. Example (25) shows all of these characteristics.

(25) hen pʰuːchaj kʰon tʰi jɯːŋ jɯː troŋ nam maj
see man Cl Rel stand Asp at there Q
'Do you see the man who is standing there?'

3.3 CONVENTIONS AND ABBREVIATIONS USED FOR PRESENTATION OF DATA

Generally, the data presented in the study, including the sample Thai/English data presenting syntactic aspects of the Thai/English codeswitching in Chapter 5, are in the following format:

(26)²⁷ hɪl man maj hɪgh
3rd Pro Neg.
'The hill is not high.'

The first line is a transcription of the recorded utterances; the second line is a morpheme-by-morpheme translation of the first line, followed in line three with a free translation into English. A codeswitch is indicated with underlined italic text, as is standard in the literature. The underlined italic text also indicates a switch from a Matrix Language to an Embedded Language.

As for the sociolinguistic/functional analysis in Chapter 4, the recorded Thai/English data are transcribed according to the conversation analysis custom of putting all the actual conversational details, such as length of pause, in the presentation of data. Note that all the data will be in the lower case. Capital letters represent loudness of a given word in utterances. More information regarding conventions and abbreviations used in the study is listed in Table 1. Below is an example of the data presentation:

²⁷ The numbering of examples starts over at the beginning of each Chapter.
EXAMPLE (27) (The switch into English is in underlined italics.)

1 A: ləː w pʰiː kə səːwəːr flower thəː səəˈfiːwər
   and older siblings then buy at safeway

2 B: ʔəː ræː təː toːnːiː: <X> ((NAME: B)) məː niːːoː cai waː:
oh yeah but now Neg. sure that

3 kʰaw miː:
   3rdPro have

Translation:
A: and older sibling (I) bought the flower at safeway
B: oh yeah but <X> (I) now is not sure if they have it

Below is a list of conventions and abbreviations used in the present study.

Table 1. A list of conventions and abbreviations

<table>
<thead>
<tr>
<th>Abbreviations and conventions</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1°Pro</td>
<td>first person pronoun</td>
</tr>
<tr>
<td>2°°Pro</td>
<td>second person pronoun</td>
</tr>
<tr>
<td>3°°Pro</td>
<td>third person pronoun</td>
</tr>
<tr>
<td>Asp</td>
<td>aspect</td>
</tr>
<tr>
<td>Cl</td>
<td>classifier</td>
</tr>
<tr>
<td>Det</td>
<td>determiner</td>
</tr>
<tr>
<td>Lnk</td>
<td>linker</td>
</tr>
<tr>
<td>Part</td>
<td>particle</td>
</tr>
<tr>
<td>Prog</td>
<td>progressive</td>
</tr>
<tr>
<td>Pst</td>
<td>past</td>
</tr>
<tr>
<td>Q</td>
<td>question word</td>
</tr>
<tr>
<td>Rel</td>
<td>relative clause marker</td>
</tr>
<tr>
<td>Int</td>
<td>interjection</td>
</tr>
<tr>
<td>(.)</td>
<td>micro-pause</td>
</tr>
<tr>
<td>{</td>
<td>simultaneous talk</td>
</tr>
<tr>
<td>(2.0)</td>
<td>length of silence in seconds</td>
</tr>
<tr>
<td>&lt;&lt;X&gt;&gt;</td>
<td>((NAME)), names mentioned in the conversation</td>
</tr>
<tr>
<td>((0))</td>
<td>extra textual information and transcriber's comments</td>
</tr>
<tr>
<td>=</td>
<td>latching (i.e. no interval between adjacent turns)</td>
</tr>
<tr>
<td>(.....)</td>
<td>omitted sections</td>
</tr>
<tr>
<td>(  )</td>
<td>unintelligible stretches of talk</td>
</tr>
<tr>
<td>?</td>
<td>question</td>
</tr>
<tr>
<td>,</td>
<td>rising intonation</td>
</tr>
<tr>
<td>!</td>
<td>exclamation mark</td>
</tr>
<tr>
<td>:</td>
<td>lengthened vowel</td>
</tr>
<tr>
<td>(laugh)</td>
<td>laugh</td>
</tr>
<tr>
<td>P</td>
<td>piano/softly</td>
</tr>
<tr>
<td>pp</td>
<td>very quietly</td>
</tr>
<tr>
<td>f</td>
<td>forte/loudly</td>
</tr>
<tr>
<td>ff</td>
<td>very loudly</td>
</tr>
<tr>
<td>CAPITALS</td>
<td>loud</td>
</tr>
<tr>
<td>A, B, C, ...</td>
<td>speakers are randomly labeled as Speaker A, B, and so on in each conversation excerpt</td>
</tr>
<tr>
<td>MA, MB, ...</td>
<td>speakers who are Buddhist monks, randomly labeled as Speaker MA, MB, and so on in each conversation excerpt</td>
</tr>
<tr>
<td>((1.42.30))</td>
<td>record time ((hr.min.sec))</td>
</tr>
</tbody>
</table>
CHAPTER 4
FUNCTIONS OF THAI/ENGLISH CODESWITCHING

This chapter explores the practice of Thai/English codeswitching in relation to two frameworks: (1) the markedness model and (2) the conversational codeswitching approach.

4.1 A SUMMARY OF THE RESULTS FROM THE INTERVIEWS

Despite the linguistic purism movement, most participants in this study are reported to have positive attitudes toward Thai-English codeswitching. During the interviews, the participants stated that Thai-English codeswitching can reflect the socio-economic status, and educational background of a person who codeswitches. In particular, the person who is codeswitching is generally perceived by his/her interlocutors to have a "good image" in term of socio-economic status; however, this image may not match the actual socio-economic status and educational background of the speaker. The use of codeswitching can help him/her project the image of a well-educated, middle- or upper-class person. Presumably, this finding supports the other result found in the interviews--every participant thinks codeswitching is a way to help one gain acceptance or even a way to "display one's ostentation" or to elevate one's status (i.e., to show off, or to display accomplishment) when used with out-group members.

In addition, the participants also pay attention to the proper use and the pronunciation of the codeswitched words. Particularly, they have a higher opinion of those with exposure to English-speaking countries and those who codeswitch with the right words in the right context and with native-like pronunciation. Most of the
participants put the highest value on pronunciation. Some of the participants even mentioned in the interview that codeswitching is prestigious because being able to speak English with proper pronunciation and having experience abroad are well-regarded.

Finally, most of the participants in the present study said that they are quite conscious and selective in terms of whom they codeswitch with and where they do it. They tend not to codeswitch in situations in which the use of English is marked, e.g., with out-group members or in Thailand. The participants said that they codeswitch, but not with everybody. They resist codeswitching with those who do not share a similar bilingual background and with those who are not their peer members. This implies that codeswitching is an in-group mode of talk and that rational choice is not, or at least, not always unconscious.

Even though the results from the interviews show that the act of codeswitching is considered common among all the Thai participants in the present study, their use of codeswitching is quite minimal. The community language norm is still very much oriented towards monolingualism. Table 2 illustrates the Thai/English codeswitching ratio.

Table 2. Ratio of Thai/English codeswitching Frequency among the participants (Every 100 word, English:Thai).

<table>
<thead>
<tr>
<th>Speakers</th>
<th>Ratio of Thai/English codeswitching Frequency (every 100 words, English:Thai)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permanent residents</td>
<td>3.3:100&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td>Graduate students</td>
<td>1.7:100</td>
</tr>
<tr>
<td>Monks&lt;sup&gt;29&lt;/sup&gt;</td>
<td>0:100</td>
</tr>
<tr>
<td>Average frequency</td>
<td>2.5:100</td>
</tr>
</tbody>
</table>

<sup>28</sup> This ratio reads as follows: there are 3.3 English words in every 100 Thai words.
<sup>29</sup> Personal exposure to communicating with the monks reveals that they do codeswitch, albeit very infrequently. The data collected in the present study do not contain any examples of their Thai/English codeswitching.
4.2 APPLICATION OF THE MARKEDNESS MODEL

Codeswitching is common in the speech of the Thai people who participated in the study. They switched back and forth between Thai and English to display that they are well-educated, bilingual, middle-class people. The interview participants associated the use of English with a good social and educational status. Accordingly, the switch into English may be associated with their identities as well-educated, middle-class people.

As stated in §2.1, the markedness model distinguishes between marked and unmarked codes. The linguistic codes that appeared in the Thai/English codeswitching data are Standard Thai, Northeastern Thai, Northern Thai, English, and a dialect of Chinese (Chaozhou). In many instances, Thai-English codeswitching functions as the "overall codeswitching as the unmarked choice." The fact that every participant in the study codeswitches implies that in this linguistically diverse environment codeswitching is an unmarked way to speak. The result from the interview gives full support to this finding. Every participant mentioned that it is common to codeswitch in a place where there is linguistic diversity, as in the United States. They also added that they are quite conscious and selective in terms of whom they codeswitch with and where they do it. They resist codeswitching with those who are not their close friends as well as those who do not share the same bilingual background. They also avoid codeswitching in a context in which English is marked, like in Thailand or in situations where out-group members are present. The markedness of Thai and English can be put on a continuum, illustrated as follows:

\[
\begin{array}{c}
\text{unmarked} \\
\text{Thai} & \text{English with in-group members} & \text{English with out-group members} \\
\end{array}
\]
Following are examples of Thai/English codeswitching as unmarked choice. According to Myers-Scotton (1993b), unmarked codeswitching can be categorized into “sequential unmarked choice” and “overall codeswitching as the unmarked choice,” each of which will be discussed respectively.

4.2.1 Sequential Unmarked Choice

“Sequential unmarked choice” is a switch from one unmarked choice to another, resulting in different RO balances when there is a redefinition of the exchange between participants. The switch of linguistic choice encodes the emerging unmarked relationship between participants and their recognition of this relationship. Instances of such switching are found in the Thai/English data. Examples are below.

Example (1) is a conversation among friends. Speakers A and B are Thai. Speaker C is male Filipino-American. He is a student and a member of the National Guard. The conversation took place in a get-together of Thai students and international students for a Thai cooking demonstration. The conversation is about Speaker B’s husband, who is also a military officer.

Example (1) (The switch into English is in underlined Italics.)

1 → A: (......) <X> ((NAME: B)) (. ) t bammaj <X> ((NAME: B’s husband))
why
2 maj ku: thi: ba:n k haj ne:wi: not stay at house Pos. navy
3 B: ha:, what
4 A: ba:n th i:ju: tron lan ne:wi: sksch e:q (. ) ha: ba:n house that stay at behind navy exchange look for house
5 ju: majchaj re:? asp Neg. Q.
6 B: k haw ju: maj djaj (. ) k haw maj ku: phrowa: 3rdPro stay Neg. can 3rdPro Neg. stay because
7 → k haw ca maj djaj nen he will not get the money (. )
3rdPro will Neg. get money
8 → because if he (. ) <X> ((NAME: B’s husband)) stays at the
9 hundred for his apartment but he gets nine hundred from the navy (. ) if he
10 stays in the navy housing he will not get paid (. ) he can live there for free
11 but he will not get paid (. ) so it’s better if we live outside
12 A: [oh okay

84
14 → C:  [yeah (. ) that’s true

Note: The first line of the transcription represents what was actually said in the conversation. The second line is the word-by-word translation (provided when necessary).

Translation:
A:  <X> ((NAME, one syllable: B)) (. ) how come <X> ((NAME: B’s husband)) is not living in the navy housing?
B:  what,
A:  those houses behind navy exchange (. ) you are looking for a house aren’t you?
B:  he can’t stay there (. ) he does not want to stay there because he will not get the money he will not get the money (. ) because if he (. ) <X>
((NAME: B’s husband)) stays at the housing he will not get the extra money (. ) he is now paying seven hundred for his apartment but he gets nine hundred from the navy (. ) if he stays in the navy housing he will not get paid (. ) he can live there for free but he will not get paid (. ) so it’s better if we live outside
A:  [oh okay
C:  [yeah (. ) that’s true

The markedness model assumes that speakers have a sense of what is marked and unmarked in the community, and this sense is called the markedness evaluator. The linguistic repertoires of Speaker A and B include Thai and English. Speaker A starts off this conversation in Thai even though the setting is “international.” From the retrospective interview, Speaker A said that she chooses Thai partly because of the topic, and partly because it is common for Thai speakers to talk in Thai in almost all occasions, regardless of the presence of a non-Thai speaker. Speaker B first responds to Speaker A in Thai. She then switches to English in line 8 to engage Speaker C, who knows about the military and its housing systems and does not speak Thai. By doing so, Speaker B acknowledges a change in a situational factor (i.e., that she now wants to include Speaker C), the shared language and group membership or a new RO set of all three speakers. To explain this in terms of the markedness model, situational factors belong to the first filter.
of the markedness model. The first filter is followed by the markedness evaluator, which is the second filter. Even though English is less unmarked, Speaker B switches to English in order to engage Speaker C in the conversation. In line 8, she mentions her husband’s name right after the pronoun, but it is mutually understood by both Thai speakers whom she is referring to. Speaker B said that she mentions her husband’s name in order to give Speaker C information he needs to understand what is going on. This can also be explained in terms of costs and rewards: even though English is less unmarked between two Thai speakers, it allows her to include Speaker C who can support her statement, as shown in line 14. Speaker B succeeds. Additionally, the switch into English indexes Speaker B’s identity as a cooperative, flexible, and considerate bilingual communicator. This finding echoes that in Finlayson et al. (1998). The participants in his study are perceived to be cooperative and flexible when they codeswitch to accommodate their interlocutors.

The switching to engage another speaker in a conversation is very common among the Thai/English codeswitchers. Example (2) below is an instance of an encounter in which a Thai codeswitcher switches into English to include a non-Thai speaker in the conversation.

Example (2) is a conversation among Speaker A, B, and C. Speaker A and B are Thai. Speaker C is American. Speaker B and C are husband and wife. The conversation took place at Speaker B’s house while all the speakers are having lunch. The conversation is about movies.

Example (2) (The switch into English is in underlined Italics.)

(.....)

1 A:  wan ko:n pʰom wa: ca paj du: nia
day before 1st Pro think Asp go watch Part.

2 B:  lo:
yeah

30 These filters are discussed earlier in § 2.1.
A: soŋ rwan maj lat samuraj ko (1.0) [kbo mawthən
two story Neg. last samurai then cold mountain
B: [paj du: sa (.). paj du:
go watch Part. go watch

lej jen nia (1.0) naŋ di:
already evening this movie good
A: tæ: lat samuraj kəna:t naŋ kaw læ:w kəaw pe:t soŋ
but last samurai Even movie already 3rd pro open two
roŋ lej na:
Cl. Part. auntie
B: ?um=
yeah
A: = kbo mawθən naŋ maj (.). roŋ diaw
cold mountain . movie new Cl. one
B: naː paj duː wan raː k lej na (.). ʔoː ho: kəon nəːn ((f))
auntie go watch day first Part. Part. Int. people crowded
A: ?um=
yeah
B: hajcaj maj ʔoː k han paj naj miː tæː huə kəon ʔa ((f))
breath Neg. out turn go where have only head people Part.
A: (laugh)
B: la last samurai () the day we went. ((f))
C: uh huh
B: the whole place was packed huh, nothing but people’s heads every where()
kə̞ ʔun <X>(NAME: C) bʊː k waː after the movie’s
Mr. say that
B: over he went () in the bathroom () even guys cried ((f))
A: ?um
yeah
B: I said they must be Japanese () they’re sissy ((f))
((everybody laughs))

Translation:
(.....)
A: I was going to watch it the other day
B: yeah
A: two movies either last samurai or (1.0) [cold mountain
B: [go watch . go watch this evening (1.0)
good movie
A: but for the last samurai even it’s already old it’s shown in two theatres auntie
B: yeah=
A: = cold mountain is new but is shown in one theatre
B: auntie went to watch on the first day (.). Oho very crowded ((f))
A: yeah
B: I couldn’t breathe wherever I turned to it was all people’s heads ((f))
A: (laugh)
B: la last samurai () the day we went. ((f))
C: *uh huh*
B: *the whole place was packed huh, nothing but people's heads everywhere.*
Mr. <X>((NAME: C)) said that *after the movie's over he went in the bathroom even guys cried* ((f))
A: *yeah*
B: *I said they must be Japanese they're sissy* ((f))
((everybody laughs))

As mentioned above, Thai is the neutral unmarked code, or the "safe" code for all Thai speakers who participated in this research. In fact, the community language choice norm is very much oriented towards monolingualism. A conversation among Thai codeswitchers usually starts off in Thai regardless of the presence of a non-Thai speaker.

Example (3) below shows that even though a conversation is in English, it will first switch into Thai when another Thai speaker joins in before it switches back to English.

Example (3) is a conversation among close friends, Speaker A, B, C, and D. All of the speakers except Speaker B are graduate students at UH. Speaker A and B are husband and wife. The conversation takes place in a car. Speaker A, B, and C are picking up Speaker D.

**EXAMPLE (3) (The switch into Thai is in underlined Italics.)**

(....)

1 A: I *(1.0)* we *(.)* we think that the subway is easier than the train
2 B: uh hm=
3 A: *the train is all like=
4 B: *the train is a whole lot confusing *(.)*
5 subway is all=
6 A: *Because sometime they have no English at all,*
7 C: uhm
8 A: *( )* ((Speaker D comes in))
9 → C: *p: tua ni: re:
I see Cl. this Q.*
10 D: *tua ni: raj e: tua ni:*
Cl. this Q. yeah Cl. this
11 C: *tæ: mo:n maj hen diaw kʰ op daj hen*
but see Neg see a moment maybe can see
12 A: uhm
13 B: uhm
14 D: *ko di: tæ: kʰ æ:n man *(.)* man jau*
also good but sleeve 3rd Pro 3rd Pro long
15  *paj naja kmap tʰwŋ nǐw Ḹa*
   over a bit almost reach finger Part.
16  C:  *leːː*
   yeah
17  A:  *pʰiː <X> ((NAME: D)) Ḹa (.) tōŋ pᵃj*
   older sibling  Int.  must go
18  *kʰaːŋ nɔ̀n (1.0) pʰiː <X> ((NAME: D))*
   side there  older sibling
19  *tōŋ pʰːm pᵃj ʔiː k kʰaːŋ nɔ̀n kʰa*
   (laugh)
   must detour go another side one Part.
20  C:  *cheːːn kʰap ((smile voice))*
   please  Part.
21  D:  long time no see how have you been doing? ((f))
22  B:  I’m trying to make it easier for you (.....)
   ((The conversation continues in English))

Translation:

(.....)
A:  I (1.0) we (.) we think that the subway is easier than the train
B:  uh hm=
A:  =the train is all like=
B:  =the train is a whole lot confusing (.)
   subway is all=
A:  =Because sometime they have no English at all,
C:  uhm
A:  ( ) ((Speaker D comes in))
C:  *ahh is it this one?*
D:  *this one what yeah this one*
C:  *I can’t see I might be able to see it later*
A:  uhm
B:  uhm
D:  *it’s good but the sleeves (.) they are too long almost cover my fingers*
C:  *yea,h*
A:  *sister <X> ((NAME: D)) (.) you have to go to the other side*
   (1.0) older sibling <X> ((NAME: D)) you have to go over to the other side (laugh)
C:  *please ((smile voice))*
D:  long time no see how have you been doing? ((f))
B:  I’m trying to make it easier for you (.....)
   ((The conversation continues in English))

As seen from line 1 on, the conversation among Speaker A, B, and C continues in English until Speaker D joins the conversation. In line 8, Speaker C greets Speaker D by inquiring about her new shirt. The conversation then continues in Thai for a few minutes.
until Speaker D switches to English to greet Speaker B, a non-Thai speaker, in line 20. The conversation then continues on in English for half an hour about their outing to karaoke.

In all three of these examples, both English and Thai are the unmarked codes. Codeswitching is unmarked here, as it is a conversation among bilingual in-group members. The speakers in these conversation strips begin the conversation with Thai and switch to English, or vice versa, when there is a change in the situation—that is, when a speaker acknowledges a new RO set among the participants in the conversation.

Additionally, the “sequential unmarked choice” also includes the most common activity in codeswitching literature: the reported speech. According to the markedness model, the code used in reported speech is unmarked for the RO set of the event being quoted and the code used elsewhere is unmarked for the RO set of the present conversation. The use of codeswitching in reported speech makes the narration more vivid and real.

EXAMPLE (4) (The switch into English is in underlined Italics.)

1 A:  
   ko bep phun th'am aj nu: ko bo:k wa: th'am aj
   then like friend that go together 1pro then say that why

2 →  k'haw maj truat bat th'am aj (;) k'haw ko bo:k wa: my face is my
   3pro Neg check I.D. 2pro 3pro then say that

3  

4 B: [(laugh)

5 C: [(laugh) 2ə: k'haw na: kæ: əd i
   yeah 3pro face old Part.

Translation:

A:  
   like my friend who went with me I said to him why didn’t they check
   your I.D. he then said that my face is my I.D.

B: [(laugh)

C: [(laugh) he looks old then

In this example, Speaker A is telling a story about what happened to her when she went out with her foreign friends. Speaker A uses Thai to report her own utterance (line 1). She
then switches to English when she quotes her foreign friend’s utterance (line 2). English is the unmarked rights and obligation set at the quoted event. However, when Speaker A quotes herself, she chooses Thai even though it is not the language she actually spoke at the time of the quoted event. This is very common in reported speech found in the Thai/English data.

EXAMPLE (5) (The switch into English is in underlined Italics.)

1 A: თქვა: თო:ნტ h i: რა ჭან კუ დ i: ნა
   but when go Cambodia 3rd pro then good Part.
2 ქ '<x> ((NAME)) წარ ჰ e's weird this time (laugh) then like

Translation:
A: but he was good when we went to Cambodia <x> ((NAME)) was like he’s weird this time (laugh)

The conversation in example (5) is about the unfriendly behavior of a third person, the person being spoken about. Speaker A switches to English when she starts the quotation. By so doing, she distances herself from the comments made by another foreign friend whom she quotes. This function of reported speech is also found in other works, e.g., Besnier 1993, and Irvine 1993. According to Hill and Irvine (1993:7), “… the function of the reported speech is not so much to enhance the utterance’s consequentiality, as to distance the speaker from an utterance deemed somehow reprehensible … Here speakers seek to evade responsibility by invoking additional sources for an utterance, rather than seeking to augment their own utterances’ force”. An investigation of reported speech in Nukulaelae Atoll, a Polynesian community, shows that reported speech is a strategy employed by speakers to “… present themselves as minimally accountable for the content of their talk” (Bernier 1993:161). Quotative forms of speech are also used by speakers to dissociate themselves from reprehensible comments in insults and verbal defamation (Irvine 1993).
Example (6)

A: <NAME> (NAME: B)) cam daj maj thi: phi: remember can Q. that older sibling
bok wa: phi: paj ce: tu tu: num thi: say that older sibling go meet cabinet cabinet one that
phi: chomp ma:k older sibling like much
cam daj remember can
A: wan ko:n phi: paj la:w phi: paj ce: cawk phi: rain day before older sibling go then older sibling go meet owner shop
khaw wa: no: phi: chomp cin cin re: tha: no: phi: chomp 3rd pro say younger sibling like really really Q. if younger sibling like
cin phi: ca lot haj sa:msip pesen really older sibling will discount give thirty percent
khaw pen khon thaj re: Int. 3rd pro be person Thai Q.
A: plaw () majchaj (.....)
B: no no

Translation:
A: <NAME> (NAME: B)) do you remember that older sibling said older sibling saw a book shelf that older sibling really likes
B: I remember
A: the other day older sibling went and older sibling met the shop owner she asked does younger sibling really like it? if younger sibling really likes it older sibling will give you 30% discount
B: Oh is she Thai?
A: no (.) no (.....)

Example (6) is a conversation between two close friends, Speakers A and B. Speaker A is a permanent resident. She moved from Thailand to the United States to live with her husband, an American-born Thai, almost ten years ago. Both speakers have many non-Thai friends in their social network. Speaker B is a graduate student. Both speakers have good proficiency in Thai and English. In this excerpt, Speaker A reports the utterances of the furniture store owner, a non-Thai speaker, in Thai and causes a little misunderstanding. Speaker A overuses the community language norm, in this case. This

In the quoted conversation between the shop owner and Speaker A, the owner of the shop refers to herself as older sibling and Speaker A as younger sibling.

92
can be seen as evidence for the observation that "the community" is very much oriented towards monolingualism and that Thai is the "safe" code for the participants in the present study. The participants opt to speak Thai rather than English in most cases. Most of the quotations are actually reported in Thai regardless of the language of the actual event. It is reported in its original language, English, mainly for special effects, e.g., telling a joke, climax, or disassociation from the quoted utterances. Example (4) above is an example of the joke effect and the disassociation from the quoted utterances. By switching to English, she disassociates herself with the quoted utterance; it is not she who said that he looks old. This is also commonly found in the Thai/English data. When a quoted utterance is a negative comment about someone, it is usually reported in its original language, English.

4.2.2 Overall Codeswitching as the Unmarked Choice

"Overall switching as the unmarked choice" is the most common type of codeswitching found in the Thai/English data. This type of codeswitching is a switch to another unmarked choice with no changes at all in the situation. This type of switching usually occurs between bilingual peers when the participants want more than one social identity to be salient in the current exchange. According to the markedness model, unlike other types of codeswitching, each switch in this type of codeswitching is unmarked and does not necessarily have a special indexicality; it carries communicative intention. This type of codeswitching also contains a good amount of switching within sentence boundaries.

Examples below are excerpts from conversation among in-group members. All the interactions are informal and involve only bilingual in-group members, whose socio-
economic status is equal. Most of the participants are graduate students at UH. They all have positive attitudes towards the act of codeswitching among in-group peers in a multilingual setting.

Examples (7) and (8) are from conversations between two Thais. Speaker A is a U.S. permanent resident from Thailand. She went to a university in the state of Illinois. She is now a housewife. She has a good command of English. Most of her friends are non-Thai speakers. Speaker B is a graduate student. They have been friends for over seven years. In Example (7), Speaker A talks about her Taiwanese realtor. In Example (8) Speaker A and B talk about the benefits of emigrants’ having another family in their country of origin and how beneficial that is for the younger generations to know that they have other family members back in Thailand or other countries.

EXAMPLE (7) (The switch into English is in underlined Italics.)

1 → A: \(khaw\) \(pen\) \(khon\) \(tajwan\) (\(ke\)\(n\) \(na\) \(ambitious\) \(ma\):\(k\) \(khaw\) \(bo\):\(k\) \(3^\text{rd Pro}\) \(be\) \(person\) Taiwan \(smart\) \(Part\) \(very\) \(3^\text{rd Pro}\) \(say\)

2 → \(wa: \) \(re\):\(mont\) \(maj\) \(maj\) \(nia\) (\(co\):\(nju\) \(nain\) \(nine\) \(to\) \(five\)
that \(start\) \(new\) \(Part\) \(true\) \(work\)

3 → \(lae: \) \(ko\) \(whatever\) \(t\):\(a\) \(khaw\) \(pok\) \(paj\) \(thuk\) \(wan\) (\(khaw\) \(ca\) and then \(but\) \(3^\text{rd Pro}\) \(out\) \(go\) \(every\) \(day\) \(3^\text{rd Pro}\) \(will\)

4 → \(pok\) \(ca\):\(k\) \(bain\) \(t\):\(a\) \(chaw\) \(thuk\) \(wan\) \(somm\)\(ut\) \(wa: \) \(khaw\) \(out\) \(from\) \(home\) \(since\) \(morning\) \(every\) \(day\) \(suppose\) \(that\) \(3^\text{rd Pro}\)

5 → \(ja\):\(k\) \(ca\) \(specialize\) \(t\):\(h\):\(w\) \(kh\):\(aj\)\(lua\) \(mi\):\(cha\)\(n\) \(ti\):\(i\):\(h\)\(aw\) \(want\) \(will\) \(area\) \(kailua\) \(have\) \(period\) \(that\) \(3^\text{rd Pro}\)

6 → \(re\):\(m\) \(khaw\) \(ko\) \(d\):\(z\)\(a\) \(b\):\(a\)\(p\) \(t\):\(h\):\(w\) \(kh\):\(aj\)\(lua\) (\(ph\):\(o\) \(t\):\(jaw\) \(khaw\) \(start\) \(3^\text{rd Pro}\) \(then\) \(will\) \(like\) \(rea\) \(kailua\) \(when\) \(morning\) \(3^\text{rd Pro}\)

7 → \(ko\) \(d\):\(z\)\(a\) \(k\):\(hap\) \(rot\) \(pok\) \(paj\) \(le\)\(n\)\(na\) \(then\) \(will\) \(drive\) \(car\) \(out\) \(go\) \(Part\) \(Part\).

Translation:
A: she’s taiwanese . smart \(very\) \(ambitious\) (\(she\) said \(when\) \(she\) \(first\) \(started\) \(it’s\) \(true\) \(that\) \(her\) \(work\) \(is\) \(nine\) \(to\) \(five\) \(or\) \(whatever\) \(but\) \(she\) \(goes\) \(out\) \(Every\) \(day\) \(she\) \(would\) \(leave\) \(early\) \(in\) \(the\) \(morning\) \(every\) \(day\) \(suppose\) \(that\) \(she\) \(wanted\) \(to\) \(specialize\) \(in\) \(Kailua\) \(area\) \(when\) \(she\) \(started\) \(she\) \(would\) \(like\) \(Kailua\) \(area\) \(in\) \(the\) \(morning\) \(she\) \(would\) \(drive\) \(out\) \(there\)

EXAMPLE (8) (The switch into English is in underlined Italics.)

1 → B: \(ko\) \(di: \) \(khaw\) \(ca\) \(daj\) \(ru: \)\(suk\) \(belong\) \(wa: \) \(tua?e \):\(n\) \(then\) \(good\) \(3^\text{rd Pro}\) \(will\) \(able\) \(feel\) \(that\) \(self\)

2 → \(ko\) \(pen\) \(suan\) \(num\) \(kh\):\(on\) \(thi\):\(nan\) \(also\) \(be\) \(part\) \(one\) \(of\) \(there\)

3 → A: \(chaj\) \(lae: \) \(ko\) \(make\) \(connection\) \(kap\) \(kh\):\(on\) \(ru?n\) \(yeah\) \(and\) \(also\) \(with\) \(people\) \(other\)
Translation:
A: that's good so they could feel that they belong that they are a part of there
B: yeah and also make connection with other people (.....)

Example (9) is a conversation between two Thai friends. Both of the speakers are graduate students at University of Hawai'i. The conversation takes place in a car when all of them are going to have dinner at a Japanese restaurant. The conversation is about another Thai student, who recently moved to the state.

EXAMPLE (9) (The switch into English is in underlined Italics.)
1 A: tæ: sɔŋsaj nɔːŋ <X> ((NAME)) kʰaw chɔːp .
   but suspect younger sibling 3rdPro like
2 → hang out kap pʰiː: ma:kṣwaː pa
   with older brother more Q.
3 B: ḷːmːm[ yeah
4 → A: [ kwaː tʰi kʰaw ca paj hang kap kʰon ḷːmːn pa
   more to 3rdPro will go with person other Q.
5 B: chaj pʰiː: kæː læːw nəj () pʰuak nan () pʰuak nan
   yeah older brother old already Part. group that group that
6 → man dek kʰaw ləj maj ruːsmuk comfortable (.....)
   3rdPro young 3rdPro then Neg. feel

Translation:
A: but I suspect that <X> ((NAME)) likes to hang out with you more yeah?
B: yeah[
A: [more than to hang with other people yeah?
B: yeah I'm already old . those people () those people they are young he might not
   feel comfortable (.....)

Example (10) is a conversation between two Thai friends who have known each other for over six years. Speaker A is an American-born Thai. He is very fluent in both Thai and English. Speaker B is a graduate student at University of Hawai'i. The conversation is about parking rules and parking tickets.

EXAMPLE (10) (The switch into English is in underlined Italics.)
1 → A: kʰaw ko pratʰuəŋ doːj kəːn tʰiː ca () walk () ?ə:
   3rdPro then rebel by prefix to will Int
2 doːn tɑːm miteː tʰuk ?an læːw ko jɔːt təŋ haj kʰon
   walk along meter every Cl. and then feed money give people
3 tʰiː maː cɔːt tʰiː nan ?ə
   that come park at there Part.
4 B: uh huh
5 A: man ko pen kaːn tat rajdaj kʰɔːŋ muaŋ
   it then be Prefix cut revenue of city
6 B: uh huh
7 A: læːw muaŋ ko cap kʰaw læːw ko prap kʰaw duaj
   and city then arrest 3rdPro and then fine 3rdPro also
A: he rebelled by walk. uh. walking and feeding every single meter for people who parked there
B: uh huh
A: it cut down the revenue of the city
B: uh huh
A: and the city arrested him and fined him for cutting down the revenue of the city
B: Yeah
A: in fact it’s their rule. the rule said not to engage in meter-feeding
B: not to [meter-feeding
A: [yeah you’re not supposed to feed the meter. he broke the rule (.....)

Example (11) is a conversation between Speaker A and Speaker B, both of whom are Thai. Speaker A is a wife of a West Point officer. Speaker B is a graduate student at University of Hawai‘i. In this example, Speaker A and Speaker B are talking about the educational background of Speaker A’s husband.

EXAMPLE (11): (The switch into English is in underlined Italics.)

(.....)

1 A: bæpwa: si: pi:jaŋ məj cop ləj () kʰaw ko bæpwa: like four year yet not finish already 3rdPro then like
2 → credit man wa: credit man wa: () service ma:la:w pi: məŋ () 3rdPro say 3rdPro say come already year one
3 thəməŋqni: chajma () thəm wə:la: thə: paj ce: kap something like that yeah arrive time if go meet with
4 nakriannajrōj th:ı: cop mahaːlaj()hen wa: lieutenant that finish university see that
5 → nakriannajrōj duajkan nia kʰaw ca mi: creditability lieutenant together Part. 3rdPro will have
6 → ma:kkʷa: ruːwa: advance sommut wa: cop pi diawkən() more or suppose that finish year same
sommut wa: cop pi: pæ:t si: pi: diawkan tæ: suppose that finish year eight four year same but
we:la: ca(.) th:u:k () haj () pʰidʒa:rana haj time will Pass give consider give
lu:manjot nia kʰaw dʒa lumanjot pʰuak west pʰo:nt ko:n promotion Part. 3ʰPro will promote 3ʰPro West Point first
B: o: kʰaw pen west pʰo:nt re:
yeah 3ʰPro be West Point Q
A: ha kʰaw pen west pʰo:nt kʰaw pen bærwa: kʰaw haj yeah 3ʰPro be West Point 3ʰPro be like 3ʰPro give
creditability ko:n advance one year th:æ:n th:hi: ca du: wa:
first instead of will see that
pi: pæ:tsi:pi: kʰaw bo:k wa: year eighty five () in
year eighty-six 3ʰPro say that
service te:wa: graduated year eighty four
but
B: ʔw:m th: i: ho: kʰo:ŋ u: kʰa hale manuako mi:
yeah at dorm of Part. Hale Manoa also have
dek west pʰo:nt tem lej
student West Point full Part.
A: lə:
yeah
B: rian keŋ[ study good
19 → A: [yeah rian keŋ pʰuak ni: kʰaw rian keŋ() fæ:n study good 3ʰPro this 3ʰPro study good boyfriend auntie Part like
na: ni: bærwa:() three point eight. three point five ʔaraj
20 → na: ni: bærwa:() three point eight. three point five ʔaraj
21 → pʰuak ni:() lə:w kʰaw all().pʰuak ni: kʰaw all around na
like this and 3ʰPro 3ʰPro this 3ʰPro Part.
majchaj kʰɛ: rian keŋ jaŋdjaw ni: na:
Neg. just study good only this auntie
majaŋ kʰu:j pʰro fœn rau pen west pʰo:nt na:
Neg. brag because boyfriend 1ʰPro be West Point auntie
kʰɛ: date west pʰo:nt ko:nnə: fœn na:
used to West Point before boyfriend auntie
ni: na() pʰuak ni: kʰaw all around kind of person
26 → ni: na() pʰuak ni: kʰaw all around kind of person
this Part. 3ʰPro. this 3ʰPro
kʰaw majdaj keŋ rwaŋ diaw majdaj rian keŋ
3ʰPro Neg. good thing one Neg. study good
jaŋdjaw() krachakrachæŋ pajmotlej
only enthusiastic altogether
wa: ʔaraj wa raw niana kʰit rwaŋ nun tʰam rwaŋ
that that that 1ʰPro Part. think thing one do thing
nun ʔw:m la [sadut pup ʔw:m lej ca wa: tʰa:m ʔaraj
one forget already trip as soon as forget already will that ask what
Translation:

A: like they couldn’t finish in four years (.) they’re like the credit said the credit said (.) already in service for one year (.) something like that yeah (.) if he meets with a lieutenant who graduated from college (.) they see that they are all lieutenants he will have more creditability or more advance suppose that they graduated the same year (.) suppose that they graduated in year eighty-four same year (.) but when they are being considered for promotion they will promote west point people first

B: yeah he’s a west point grad?

A: yeah he is a west point grad he is like they give him creditability first advance one year instead of looking at the graduation year eighty-four they said year eighty five . in service but graduated year eighty-four

B: yeah at the uh dorm hale manoa there’re lots of west point grads

A: yeah

B: good at study=

A: =yeah good at study they are good at study (.) my boyfriend he’s like (.) three point eight , three point five something like that (.) and he’s all (.) they’re all around not just good at study hey I didn’t say it because my boyfriend is a west point grad I used to date west point grad before my boyfriend . they are all around kind of person they are not good at only one thing not only good at study (.) enthusiastic altogether I’m all stunned like what on earth like I’m thinking of one thing then asking another thing forgot already [as soon as I trip I forget all of what I would or was doing

B: [yeah

Example (12) is a conversation among three Thai friends. They are talking about Speaker A’s surprise birthday party, arranged by Speaker A’s husband.

EXAMPLE (12) (The switch into English is in underlined Italics.)
(....)

1 A: tæ: ciŋ ciŋ na wannı: phı: thäm no:n thäm ni: (.)
   but true true Part. today older sibling do that do this

2 ➔ phı: ko wa: (.) you’d better bring something home (.) klap ma:
   older sister then that return come

3 thäm kapkʰaw thı: ba:n ko toŋ du: lu:k ko nuaj ʔik (.)
   make food at home also must watch offspring then tired also

4 ➔ phı: wa: (.) you’d better buy something (laugh)
   older sibling that

5 B: [(laugh)

6 C: [(laugh)
but actually today when I was doing this and that (.) I was like (.) you'd better bring something home (.) you cook some food I still have to watch the kids still tired I was like (.) you'd better buy something (laugh)

B: [(laugh)]
C: [(laugh)]

In Examples (7)-(12), the RO sets associated with Thai and English are of internationally well-educated bilingual peers. Thai allows them to maintain their solidarity with the group. It is also associated with being conservative, original, and sincere. English, on the other hand, is associated with being modern, authoritative, and well-educated. The participants in these examples are graduate students and permanent residents. The graduate students are very accustomed to using English in school; English is the medium of their course work as well as a language to speak with classmates. For other participants, English is the language that is used at home. Accordingly, every participant is very well accustomed to using English with the people, friends and family members, in their social network. In addition, they have plentiful opportunities to speak Thai. All of them have Thai friends in their social network. The act of Thai/English codeswitching allows them to proclaim the dual indexical values associated with the two codes. Examples (7) – (12) show that the Thai participants switch to English here and there in the overall discourse. There is no change in the situation, and thus the shared RO set of the participants remains the same: the RO set of the internationally well-educated bilingual Thai/English speakers. In sum, they use Thai/English codeswitching to signal their multiple identities between their peers and maintain their ethnic balance. This is in compliance with the analytic framework that the “overall codeswitching as the unmarked choice” is predicted to occur among bilingual peers who want to make salient two or
more positively evaluated identities or to involve dual memberships signified through
codeswitching in a conversation.

The Thai/English codeswitching data also show that “the overall codeswitching as
the unmarked choice,” perform/signify conversational functions and forms.

EXAMPLE (13) (The switch into English is in underlined Italics.)
1 \(\rightarrow\) A: rian 605 man pen le:k le:k \(\rightarrow\) quantitative (.) ?a:ca:n bo:k
study 605 3\(^{th}\)pro is math math teacher say
2 wa: we:la:thi: samo:n \(\rightarrow\) process...
that when brain

Translation:
A: I'm taking 605 it's mathematics (.) quantitative (.) the teacher said that
when the brain processes...

In this example, Speaker A switches to English to clarify what she meant by
[le:k le:k]. That is she self-repairs her former ambiguous choice of word by switching
to English. The word [le:k] itself means ‘mathematics,’ and by reduplicating, it can be
interpreted as a type of mathematics, or something that is related to statistics or merely
something involving calculation. The meaning of the word is therefore ambiguous. By
coding her reiteration in English, Speaker A provides a clarification to what she meant by
[le:k le:k] in this particular context.

Example (14) is a conversation between two Thai about gambling. Speaker A is a
permanent resident who is married to an American soldier. Speaker B is a graduate
student at University of Hawai‘i.

EXAMPLE (14)
1 A: na: ko jaŋ bo:k lej (. ) jaŋ bo:k wa: (. ) th^mp maj
auntie also still tell Part. still say that though Neg.
2 kro:ț k^on rau mo:ho: th:i: sia p^r^wa: (. )
mad people 1\(^\text{st}\)Pro angry that lose because
3 \(\rightarrow\) ba:nk^r^raŋ (.) cannot afford to lose you know
sometimes
4 B: ðum:n
yeah
5 \(\rightarrow\) A: cannot afford to lose and you (.) you go lose it (.) you get more frustration
6  yeah.
7  B:  ʔɯ:\m
    yeah
8  →  A:  become no(.) no more fun already (.) tʃai\ma
    Q.
9  B:  ʔɯ:\m
    yeah

Translation:  (The switch into English is in underlined Italics.)
A:  auntie also said (.). also said that (.). even though they’re not mad people are
    angry because (.). sometimes (.). cannot afford to lose you know
B:  yeah
A:  cannot afford to lose and you (.). you go lose it (.). you get more frustration yeah.
B:  yeah
A:  become no. no more fun already (.).yeah?
B:  yeah

In Example (14), the topic of this example is gambling. When Speaker A makes a
comment about those who gamble, she switches to English. In Example (12) above,
Speaker B switches to English to mark that the words are words she only thought to
herself (inner monologue). By switching into English, the speakers create a contrast that
allows them to contextualize their conversation. Codeswitching contextualizes the former
example such that it sets apart a topic from her opinion: what the speaker thinks of people
who gamble. In the latter example, it sets apart the speaker’s narrative and her inner
monologue.

4.2.3 Codeswitching as a Marked Choice

“Codeswitching as a marked choice” is a switch to the marked choice to negotiate
different RO sets between participants. A general motivation to make marked choices is
to increase or decrease the expected social distance between participants. The switch can
be either positive or negative depending on the situation. In addition, “codeswitching as a
marked choice” can be seen as a strategy to exclude other ethnic out-group members. In
multi-ethnic cities, when speakers use their ethnic languages as a marked choice, they
risk offending those who are excluded. According to the markedness model, the decision
to switch or not switch is a matter of costs and rewards.

Example (15) is a conversation taking place at a Thai temple. It is a birthday party for one
of the monks as well as a thank-you party for all the disciples who help out at a temple
fundraising event. There are many speakers, all of who participate in this conversation at
different times. The conversation consists of Standard Thai, English, and Chinese. All of
the speakers are Thai. Only three speakers: A, B, and C understand Chaozhou Chinese.

EXAMPLE (15) (The switch into English is in underlined Italics, the switch into Chinese is in Bold.)

1  \( \rightarrow \)  A:  \( \text{cacan lu:ksit ca klap ba:n la:w na} \)  
   master disciple will return home already Part.  
   \((1.35.10)\)

2  MA:  ( )=

3  \( \rightarrow \)  A:  \( =\text{no (f))t\text{on paj so}\text{p} \text{pi: } \text{khaw ko:n} \)  
   must go send older sibling 3rd Pro first
   \((1.36.27)\)

4  MA:  \( \text{paj } \text{he (.) paj du na\text{nsuw: so:p}} \)  
   go Part. go look book take an exam
   \((1)\)

5  B:  \( \text{?e: paj du: na\text{nsuw: so:p}} \)  
   yeah go look book take an exam
   \((1)\)

6  A:  \( =\text{no (f))nu: maj mi: so:p} \)  
   l^\text{Pro} Neg. have exam
   \((1)\)

((The conversation continues about who else has an exam coming up))

((m.))

19 \( \rightarrow \)  A:  \( \text{paj la cacan sawatdi: } \text{khrap } \text{happy birthday (f)} \)  
   go Part. master good-bye Part.
   \((1.36.27)\)

20  MA:  \( \text{watdi: na khrap=} \)  
   bye Part. Part.

21  B:  \( =\text{sawatdi: khrap} \)  
   good-bye Part.

22  F:  ( )

23  B:  \( \text{waj waj waj waj waj lu:ksit th\text{e lu:k (ff)}} \)  
   salute salute salute salute salute disciple Part. child
   ((laughter continues for 5.0 seconds))

24  MA:  \( \text{majchaj waj p\text{hra th\text{e lu:k (ff) (laughter)}}} \)  
   Neg. salute monk Part. child

25  F:  \( \text{?a la:w maj waj p\text{hi: } ?a (f)} \)  
   Exc. then Neg. salute older sibling Q.

26  B:  \( \text{th\text{ammaj ton waj}} \)  
   why must salute

27  F:  \( <X> ((\text{NAME: A})) (1.0) \text{waj p\text{hi: rew}} \)  
   salute older sibling quick

28  F:  \( ((\text{laughter (2.0)}) \text{diaw do:n top}} \)  
   moment Pass. slap

29  B:  \( \text{diaw do:n top (laugh}} \)  
   moment Pass. slap
30 → C: ʔacaːn watdiː: (1:36:50)
master good-bye

31 → A: caːn bajbaj ({melodious tone}) see you tomorrow
master bye bye

32 → E: baːːːːːj
bye

33 → A: pʰruŋːiː nɯŋ kʰawniːaw [waj na toːn jen
tomorrow steam sticky rice already Part. when evening

34 → C: [jaː lumːm tʰaːm noːŋ kʰaw na ( )
don't forget ask younger sibling 3rd Pro Part.

35 → F: thank you na nuː tʰiː maː duː (.) maː chuaj duː:
Part. child to come watch come help watch

36 → F: luːːk pʰiːː a na luːːk pʰiːː paj naj maː baːŋ ha
daughter older sibling Part. Part. daughter older sibling go where come some Part.

37 → C: <X> ((NAME: C)) pʰɯŋ ma
just come

38 → F: ʔa lə
Exc. Q.

39 → C: wːm
yeah

40 → F: majdaj paj dəːn kap kʰaw lə
Neg. go walk with 3rd Pro Q.

41 → C: majdaj paj kʰaw paj kap <X> ((NAME: E))
Neg. go 3rd Pro go with

42 → MA: paj səːm paj set ko paj dajʔeː tɔː:
go Sam go finish then go Daiei continue

43 → C: paj dajʔeː:
go Daiei
((unintelligible for 6.0 seconds))

44 → A: caːn diaː pʰruŋːiː tʰam namcim ʔaraj ʔaraj ((1.37.18))
moment tomorrow make dipping sauce delicious delicious

45 → F: namcim ʔaraj
dipping sauce what

46 → G: ok ləːw ceː kan
then meet together

47 → A: pʰruŋːiː tɔn jen ma kin
tomorrow when evening come eat

48 → F: ha luŋpʰiː namcim ʔaraj((ff)) (6.0)tʰam namcim reː:
Exc. elder brother monk dipping sauce what make dipping sauce Q.

49 → C: hej <X> ((NAME: B)) raw waː ca tʰaːm
Exc. 1st Pro think will ask

50 → F: waː miː ʔaraj kʰon paj tʰiːː kʰaw bap ca proofread
that have foreigner person which that 3rd Pro. like will

51 → B: ca ʔaraj na
will what Q.
((The conversation continues on this topic for a few minutes, then the topic changes to eating vegetarian))

98 C: raw kin maj daj la phrun: raw kin ce: 1°Pro eat Neg. can Part. tomorrow 1°Pro eat vegetarian

99 B: uhu kin ce: kan mot ((ff))
Exc. eat vegetarian together all

100 MB: sipsi majchaj re: re:m sipsi majchaj re:
fourteen Neg. Q. start fourtens Neg. Q.

101 F: phruni:
tomorrow

102 C: wanni: la: tθ: η
today clean stomach

103 F: nu: nu: phruni:=
child child tomorrow

104 MB: =sipsi thmq ji:psο:η
fourteen until twenty-two

105 F: phruni: nε phrowa: mmanθaj rau man wanthi:
Tomorrow Part. because Thailand 1°Pro it date

106 sipsi læ:w
fourteen already

107 B: chaj chaj
yeah yeah

((The conversation continues about eating vegetarian))

147 → B: a: khaw ca klap kan læ:w khaw ca klap kan
Int. 3°Pro will return together already 3°Pro will return together

148 læ:w () waj mi: ?araj khο: be: kan khuj kan[ di:kwa:
then when have what ask number each other talk each other better
((laughter))

149 F: 
[la: sa:m wan
leave three day

150 læ:w ph:i: ko kin to: nia wanthi: sipsi: nia
then older sibling then eat continue Part. date fourteen Part.

151 C: uh huh

152 B: θ: (!) haj e-mail di:kwa:
yeah give better

153 → A: ba:n ju: klap
home locate far

154 E: chat char

155 B: haj e-mail læ:w ko chat kan ((laughter))
give and then each other

156 → F: haj ((high tone))?ah tua lek man kin ce: ph:i: ca daj
Exc. Part. body small 3°Pro eat vegetarian older sister will can

32 The word chat is reduplicated. It is used in a Thai manner.
Translation:
A: master I'm going home
MA: ( )=
A: =NO I have to send older sibling back first
MA: go(.) go study for the exam
B: yeah go study for the exam
A: NO I don't have an exam
((The conversation continues about who else has an exam coming up))
T: I'm going master good bye happy birthday
MA: bye=
B: =good-bye
F: ( )
B: let let let let let the respect you pay be paid to the disciple (laughter (5.0)) not to the Buddha
F: hey why you don't pay respect to older sibling (2.0)
B: why do they?
F: <X> ((NAME: A)) (.1) pay respect to older sibling quick
(laughter (2.0)) I’ll slap you
B: I’ll slap you (laugh))
C: master good-bye ((1.36.50))
A: master bye bye ((melodious tone)) see you tomorrow
E: bye:::::
A: tomorrow you prepare the sticky rice [in the evening
C: [don’t forget to ask her please
F: thank you child for taking (. ) helping me taking care of older sibling’s daughter
which places did she go
C: I just came
F: oh yeah
C: yeah
F: you didn’t go walk with them
C: no I didn’t she went with <X> ((NAME: A))
MA: they went to sam, then daiei
C: they went to daiei
((unintelligible for 6.0 seconds))
A: master tomorrow you make a delicious dipping sauce
F: what dipping sauce?
G: ok see you
A: I’ll come eat tomorrow
F: HEY elder sibling monk what dipping sauce? (6.0) make dipping sauce yeah
C: Hey I thought I was going to ask if there’s any foreigner who is willing to
proofread my work
A: willing to what?
((The conversation continues on this topic for a few minutes, then the topic changes to
eating vegetarian))
(.....)
C: I cannot eat it any more tomorrow I’ll eat vegetarian
B: UHUUM everybody is eating vegetarian
MB: fourteenth isn’t it starting on the fourteenth isn’t it?
F: tomorrow
C: today is the cleaning stomach day
F: kid kid tomorrow=
MB: =fourteenth to twenty-second
F: tomorrow yeah because in Thailand it’s already fourteenth
B: yeah yeah
((The conversation continues about eating vegetarian))
(.....)
B: Ah they are going home they are going home (. ) if you have anything
exchange your numbers so you can talk [that’s better (laughter) ((1.41.24))
F: stopped eating for three days and older
sister will continue eating on
this fourteenth
C: uh huh
B: Uh better give the e-mail address
A: her home is far
E: chat chat
B: give e-mail address and then chat together (laughter)
F: UHH the small one will eat vegetarian older sister wants to share the food what's wrong with that?
B: [Older Sibling Is Fatter
A: older sibling is short
F: yeah older sister is short ((softer voice))
((laughter continues for 5.0 seconds))
F: hmm hmm hmm hmm when he teased he ((angry tone))
MC: ( ) are happy (laughter)
A: what time eh go home ((1.41.57))
D: did you ask everything you wanted to?
C: gotta go hurry (.....)

There are more than ten native Thai speakers present at this particular event. Speakers A, B, and C are the only speakers who have a Chaozhou Chinese origin and they have some knowledge of the language. Chaozhou Chinese is marked in this conversation because it is an ethnic language that is intelligible only to the three speakers. Speaker A takes a gamble and switches to Chaozhou Chinese to negotiate a new RO set with Speaker C. The language makes salient the shared ethnic identity of Speaker A and C, but alienates the rest. Yet, again, such a choice is a matter of costs and rewards. As shown in line 1 and 3, Speaker A is trying to leave, but he has to drive Speaker C home first. In line 19, approximately one minute after the first attempt, he expresses his valediction again to the monk. In line 30, Speaker C also bids farewell to the monk. Nevertheless, in line 35, Speaker F comes in to thank Speaker C for taking care of her daughter. In line 44, Speaker A attempts to leave and he bids farewell to the monk again by telling the monk to make a good BBQ dipping sauce tomorrow. It is quite common for Thai speakers to end a talk by mentioning a future plan whether it is tomorrow, or the next time they will meet or talk again. As shown in line 46, Speaker G takes it as a valediction token and
bids farewell to Speaker A. However, the conversation still goes on with Speaker F coming in and inquiring about the sauce, and Speaker C asking Speaker B about a proofreader. The conversation topic shifts to eating vegetarian. It involves many speakers, but mainly Speaker C and F, and continues for a while. In line 147, Speaker B interrupts the conversation by telling Speaker C and F to exchange their contact information so that they can talk later. In line 153, Speaker A mentioned that Speaker C’s house is far from the temple. Yet, the conversation still continues, as Speaker F is dissatisfied with Speaker B’s action. Finally in line 167, approximately seven minutes after Speaker A’s first attempt to leave, Speaker A switches to Chaozhou Chinese asking Speaker C to go home. Speaker A succeeds.

Example (16) is a conversation between Speaker A and B, who are friends. Both of them are graduate students at UH. Speaker A was trying to explain to Speaker B what B had to do in order to carry out a sociolinguistic interview.

EXAMPLE (16) (The switch into English is in underlined Italics.)

(.....)
1 A:  pʰɔ cop læ:w kɔ bɔk kʰaw ʔaːn niː haj ʔX> ((NAME: B))
      when end already then tell 3rd pro read this to
2  ⇒  faŋ (2.0) ʔaːn haj ʔX> ((NAME: B)) faŋ (1.0) read this for me
      listen read to listen
3  B:  ʔuːm
      yes

Translation:
(.....)
A:  when you’re done, you can tell them to read this for ʔX> ((NAME: B)) (2.0) read this for ʔX> ((NAME: B)) (1.0) Read this for me
B:  yes

Example (17) is a conversation among a group of Thai friends who are about to go out for dinner together. They are waiting for everyone to come so that they can leave the dorm.

EXAMPLE (17) (The switch into English is in underlined Italics.)

1 A:  maj miː leːw kʰrap faŋ niːtɔŋ roː kʰraːj ʔiːk maj maj
      no have already Part. side this must wait who more Q Neg

108
2 ru: wa:
know Part
3 B: maj mi: lœ:w ni:
no have already Part.
4 C: aw rom kan paj rœplaw?
bring umbrella together go Q
5 D: mi: nuŋ an (laugh) (2.0) tukata:mi: na:rak
have one Cl. teddybear cute
6 C: 2m:m
yeah
7 → E: ṇan raw paj kan læj maj khrap (3.0) shall we?
then 1stPro go together already Q Part.
8 F: paj paj paj
go go go

Translation:
A: nobody else do we have to wait for anybody else I don’t know
B: nobody else
C: did you guys bring any umbrellas?
D: I have one (laugh) (2.0) the teddy bear is cute
C: yeah
E: then shall we go now? (3.0) shall we?
F: go go go

Examples (16) and (17) are similar to each other: a Thai speaker switches to a marked choice, English, in order to receive a response from the interlocutor(s).\textsuperscript{33} English is associated with being formal and authoritative. Even though the conversation excerpts in these two examples are casual, Speaker A in Example (16) and Speaker E in Example (17) decide to switch to a marked code to accomplish their goal.

\textit{Quotation: Standard Thai/Northern Thai}

\textbf{EXAMPLE (18) (The switch into northern Thai is in underlined Italics.)}
1 A: lã:w pʰman pʰi: kʰon ni: man chw: ʔaj
and friend older sibling person this 3rdPro name Part.
2 <X> ((NAME)) pʰon lampaŋ (.). man
be person Lampang 3rdPro
3 pʰuːt pʰroː pʰuːt chaː (.). ((slow high pitch))
speak beautiful speak slow

\footnote{\textsuperscript{33} Previous work in conversation analysis have shown that the 2.0 seconds of pause (lack of response from the interlocutor) is adequate to trigger a repetition of the first pair part, e.g. a question, a request, etc. (see Li 1994, 1998)}
man ca pʰuːt jaŋ ȵia
3rd pro will speak like this

B: ?uːm
yeah

A: sommut wa: man ca chuan <X> ((NAME:B))
suppose that 3rd pro will ask

7 → paj kin kʰaw man ko ca bo:k wa: <X> ((NAME:B))
go eat rice 3rd pro then will say that

8 paj kin kʰaw pa ((slow high pitch))
go eat rice Q.

9 B: (laugh)

A: læ:w thːiːni: kʰon kʰap rot pʰajaba:n rot rot rot. oj
and now person drive ambulance car car car int.

rotnakrian ko pʰa:n paj chajma
student bus then pass go Q.

12 B: ?uːm
yeah

13 A: thːiːni: ?ai <X> ((NAME)) man ko
now Part. 3rd pro then

14 bo:k wa: () man ko bo:k () thːiːni: man lej ba:n
say that 3rd pro then say now 3rd pro pass home

15 man pai læ:w ḥai
3rd pro go already Part.

B: ?ːe:
yeah

3rd pro then say that 2nd pro Part. 2nd pro Part. 2nd pro seem will

18 → pʰa:n ba:n <X> ((NAME)) maː: ((slow high pitch))
pass home come

19 læw kwaː man ca pʰut cop rot ko (laugh)
and until 3rd pro will speak finish car then

20 B: (laugh)

A: ?a:j thːaː ca pʰa:n ba:n <X> ((NAME))
2nd pro seem will pass home

22 → maː: 1a ((slow high pitch)) (laugh) ?oːm klap (laugh) (.2)
come already detour return

23 pen kʰon naː rak pen pʰ man pʰiːː maː tɑŋtæː: (
be person lovable be friend older sibling come since

24 B: ?uːm
yeah

25 A: weːlaː ca waː na hēː ?aj <X> ((NAME:A))
time will scold Part. Int. Part.

26 → ?a (.2) ?aj aŋ ko bo huː ((slow high pitch)) (laugh)
Part. what then Neg. know

27 → ?aɾaj jaŋ ȵia ka
what like this Part.
Translation:
A: ...and this friend of mine is called \(<x>\) (person's name: one syllable) she is from Lampang. she speaks politely speaks slowly. (slow high pitch)) she would speak like this
B: yeah
A: suppose that she'd want to ask \(<x>\) (NAME: B) to go have lunch together she'd say. \(<x>\) (NAME: B) want to go eat? (slow high pitch))
B: (laugh)
A: and then the driver of the ambulance bus bus bus . eh student bus already passed yeah
B: yeah
A: then \(<x>\) (NAME) she said . she said . then it already passed the house
B: yeah
A: then she said. you you you probably already passed \(<x>\) (NAME))'s house (slow high pitch) and by the time she finished saying the bus then (laugh)
B: (laugh)
A: you probably already passed \(<x>\) (NAME))'s house (slow high pitch)) (laugh) they had to take a detour back (laugh) (.2) she's lovable she's been my friend since ***
B: yeah
A: when she's gonna scold huh \(<x>\) (NAME: A)) I don't know what you're doing (slow high pitch)) (laugh) something like that

Example (18) is a conversation among close friends. Speaker A is from the northern part of Thailand and she can speak the northern dialect. Speaker B is from the central part of Thailand and cannot speak the northern Thai dialect. Speaker A switches to the northern Thai dialect when talking about her friend, who is the topic of this conversation. Speaker A’s switching to the northern Thai dialect is marked. The switching adds aesthetic effects to her story about her friend’s personality.

4.3 APPLICATION OF THE CONVERSATIONAL CODESWITCHING APPROACH

As stated in §2.2, the conversational codeswitching approach focuses on the micro-analysis of interactions that take place naturally in small groups. The premise
underlying this interpretative framework is that codeswitching signals contextual information. The conversational codeswitching approach requires two things. First, the transcription of naturally occurring data has to include conversational details like pauses, overlaps, latches, and so on. Second, the interpretation of any relevant social meanings has to be demonstrated by participants in the sequential organization of the actual conversation. According to Auer (1995), the act of codeswitching typically involves certain activities, e.g., change of addressee selection, reported speech, reiterations, and so on. This section is organized according to these activity types.

4.3.1 Reiteration

Example (19) is a conversation between Speaker A and B in a restaurant. The conversation is about what kind of food Speaker C (absent in the conversation) wants Speaker B to order for her.

EXAMPLE (19) (The switch into Thai is in underlined italics.)

1 A: she said anything with avocado and onion ring (.) wasn’t that all she
2      said? (2.0) kʰaw boːk waː ʔaw ʔarajkoːdaj thːiː miː
          3rdPro    say    that    take    whatever    that    has
3    ʔavokʰado kap ʔonienriŋ chaŋpa?
        avocado    and    onion    ring
4    B:    chaŋ kʰa
        yes    part

Translation:
A: she said anything with avocado and onion ring (.) wasn’t that all she
said? (2.0) she said anything with avocado and onion ring, didn’t she?
B: yes ((in Thai))

As shown above, Speaker A uses Thai to code her reiteration when her request for confirmation is responded to with silence. What can be explained in this context is that by using different linguistic codes, the Speaker A builds up a contrast which helps her index herself as a confirmation seeker and helps her get attention from the interlocutor. The confirmation seeker here is defined as a person who tries to elicit a confirmation or a
response from his/her interlocutor. As shown above, Speaker A has successfully elicited a response from Speaker B.

A pattern that emerges from these examples is that codeswitching is used to contextualize turn transition. By building up a structural contrast, codeswitching is used here primarily to signal turn handovers, similarly to other contextualization cues, e.g., prosodic or gestural cues.

4.3.2 The Addressee Selection

The example below is a sample conversation from §4.2.1. It is a conversation between three students, two of whom are Thai (J and NV) and one of whom is a male Filipino-American (B). Speaker B is a student as well as a member of the National Guard. The conversation took place in a get-together of Thai students and other international students for a Thai cooking demonstration. The conversation was about NV’s husband who is also in the military. Speaker NV sat across from Speaker J and B, who sat next to each other. When NV answered J’s question, she first answered it in Thai and then switched to English to include speaker B in the conversation.

EXAMPLE (1) (The switch into English is in underlined Italics.)

1 → A:  

   (. . . .) <X> ((NAME: B)) (. . . .) thammaj <X> ((NAME: B’s husband))
   why

2  

   maj ju: th:i: ba:n kh:ɔŋ ne:wi: 
   not stay at house Pos. navy

3  

   B:  
   ha:, what

4  

   A:  
   ba:n th:i: ju: troŋ lαŋ ne:wi: ekʃɛŋ (. . . .) ha: ba:n 
   house that stay at behind navy exchange look for house

5  

   ju: majchaj re:? 
   Asp Neg. Q.

6  

   B:  
   kh:aw ju: maj daj (. . . .) kh:aw maj ju: pʰrowa: 
   3rdPro stay Neg. can 3rdPro Neg. stay because

7 →  

   kh:aw ca maj daj nɛn  he will not get the money (. . . .) 
   3rdPro will Neg. get money

8 →  

   because if he (. . . .) <X> ((NAME: B’s husband)) stays at the
   housing he will not get the extra money (. . . .) he is now paying seven
   hundred for his apartment but he gets nine hundred from the Navy (. . . .) if he
   stays in the navy housing he will not get paid (. . . .) he can live there for free
   but he will not get paid (. . . .) so it’s better if we live outside

9  

10  

11  

12  

13  

14 → C:  

[oh okay]

[yeah (. . . .) that’s true]
Translation:
A: \(<X> ((\text{NAME, one syllable: B})) .) \text{ how come } <X> ((\text{NAME:B's husband})) \text{ is not living in the navy housing?} \\
B: \text{ what,} \\
A: \text{ those houses behind navy exchange (.) you are looking for a house aren’t you?} \\
B: \text{ he can’t stay there (.) he does not want to stay there because he will not get the money } \text{ he will not get the money (.) because if he (.) } <X> ((\text{NAME: B's husband}) \text{ stays at the housing he will not get the extra money (.) he is now paying seven hundred for his apartment but he gets nine hundred from the navy (.) if he stays in the navy housing he will not get paid (.) he can live there for free but he will not get paid (.) so it's better if we live outside} \\
A: \text{ oh okay} \\
C: \text{ yeah (.) that's true}

In line 1, Speaker A uses Thai to open the conversation with Speaker B. Speaker A subsequently receives a response from B in line 6. In line 7, Speaker A interprets Speaker B’s response as a request for clarification. She then formulates her question in the next turn by specifying the location of the Navy Housing to which she was referring. In line 8, Speaker B responds to A in Thai. She then switches to English in line 9. The switch in line 9 is a self-repair of her utterances for the purpose of including Speaker C in the conversation. As shown in the following turn, Speaker C gives a response to Speaker B by commenting on Speaker B’s explanation.

As for the repair in line 9, Speaker B does not repair all her previous utterances in the turn, but instead translates back to a certain point. This can be analyzed in terms of membership category. Speaker B’s husband is in the military. Speaker C is also in the military and is not living in military housing for similar reasons to Speaker B and her husband – they will not get extra money. It is not necessary for her to repair the whole of her utterances, as both speakers share the same background. This also allows her to put
emphasis on the reason why they do not want to move into military housing. By switching to include Speaker C (who is in the military), Speaker B indexes her identity as someone who is also related to the military, such that she knows about military housing and the financial coverage for housing provided by the US government, knowledge which C also shares.

EXAMPLE (20) (The switch into English is in underlined Italics and the switch into North-eastern Thai is in Bold)

1 A: hiw læ:w kʰrap
hungry already Part.
2 B: ?aw
here
3 A: mi: raj kin ʔa () maj mi: ʔaraj kin lɛj ʔa
have what eat Part. Neg. have what eat at all Part.
4 → C: som ju: ni:
fork be here
5 → A: buaŋ buaŋ
spoon spoon
6 D: pʰak
vegetable
7 A: pʰak
vegetable
8 D: pʰak
vegetable
9 A: pʰak
vegetable
10 D: malako:
papaya
11 A: pʰat mi: ʔa
fried noodle Part.
12 D: maj mi: ʔa () maj ju: troŋni:
Neg. have Part. Neg. be here
13 A: ()
14 D: kʰawniaw ʔa kʰa
sticky rice Q. Part.
15 B: kʰawniaw kʰaw pack jaːŋdiː: ((melodious tone, unpacking the sticky rice 3rdPro well container))
16 (.5) ?um (.1) ?oho: rɔ:n lɛj (.1) ?aw
uhm Exc. hot Part. here
17 → A: buaŋ
spoon
A: I’m hungry already
B: here
A: anything to eat (.) nothing to eat at all
C: here’re some forks
A: spoon spoon
D: vegetable
A: vegetable
D: vegetable
A: vegetable
D: papaya
A: stir-fried noodle?
D: don’t have it (.) not here
A: ( )
D: is that sticky rice?
B: they well-packed the sticky rice ((melodious tone)) (.5) ah (.1) hot (.1) here
A: spoon
C: here
B: please help yourself
A: what can I eat?
B: here
C: the chicken feet aren’t cook yet
A: oh aren’t cook yet
C: over there already cook

This conversation takes place at a fundraising concert for the Thai temple. Speakers A, B, and D are native Thai speakers who are students at University of Hawai‘i. They are regular volunteers for the temple fundraising events. Speaker C is a Laotian permanent
resident who has been in Hawai‘i for over ten years. His native language is Lao. He has been working with the temple bringing in singers from Thailand to Hawai‘i for fundraising. As for his full-time job, he is a middle person between local farms and vendors in Chinatown. In line 4, Speaker C uses Thai to offer Speaker A a fork to use with the food. Speaker A refuses the offer by switching into Lao, an ethnic language of Speaker C.

Example 21 is a conversation among three Thai speakers, Speakers A, B, and C. Speakers A and C are female. Speaker C is male. Speaker C is the oldest among the three speakers. Speaker C is the youngest. As for their linguistic repertoires, all of the speakers can speak both standard Thai and English very well. Additionally, Speaker A and B can speak the Northern Thai dialect. Both of them lived in the northern part of Thailand before coming to Hawai‘i. Speaker A is a permanent resident; she has lived in the United States for almost ten years. Speaker B and C are graduate students at University of Hawai‘i. There are other standard Thai speakers present but they do not participate in this conversation. The conversation is about Speaker A offering Speaker B’s children a ride to a party. Speaker B herself and her husband have to go to another party as well.

EXAMPLE (21) (The switch into English is in underlined Italics and the switch into Northern Thai is in Bolds)

1 A: diaw haj dek dek ma: naŋ rot <X> ((NAME: A))
   moment let child child come sit car
2     læ:w haj pʰi: <X> ((NAME: A's husband)) kʰap paj drop ko daj
     then let older sibling drive go also can
3  -> B: majpənraj <X> ((NAME: A)) paj tʰə diaw saj
     it's ok go Part. moment late
4  -> A: naːn jaŋ maj rəːm rək
     event yet Neg. start Part.
5  B: majpənraj <X> ((NAME: A)) paj tʰə
     it's ok go Part.
6  C: pʰi: <X> ((NAME: C)) waː paj duajkan
     older sibling think go together
7     thːiː diaw leːj diːkwaː:
     time one already better
8  -> B: majpənraj <X> ((NAME: A)) paj tʰə
     it's ok go Part.
9  -> A: pʰi: <X> ((NAME: B)) (. .) <X> ((NAME: A)) waː kʰap
     older sibling that drive
10    paj pʰrəːmkan diːkwaː:
     go together better
11 -> B: <X> ((NAME: A)) klap paj hə
     return go Part.
A: majpenraj ((NAME: A)) paj daj cin cin
it's ok go can true true
dek dek ca daj maj ton ro: duaj
child child will can Neg. must wait too

B: majpenraj ph:i: khap ma: rap 2i:k ro:p daj
it's ok older sibling drive come pick up another round can

A: <X> ((NAME: A)) wa: paj duajkan boday
that go together Part.

B: 2aw gan ko daj () la:w to:n kh:a: klap la
take that also can and when time return Part.

A: diaw kh:a: klap <X> ((NAME: A's husband))
moment time return

ko ma: rap dek dek t'h:i: ba:n ph:i: (.....)
then come pick up child child at home older sibling

Translation:
A: let the children come to sit in <X> ((NAME: A)) 's car and then let older sibling
((NAME: A's husband)) drive to drop them off
B: it’s ok <X> ((NAME: A)) will be late
A: the event hasn’t started yet
B: it’s ok <X> ((NAME: A)) go already
C: older sibling <X> ((NAME: C)) think it’s better to go together
B: it’s ok <X> ((NAME: A)) go already
A: it’s ok <X> ((NAME: A)) can go really so the children wouldn’t have to wait too
B: it’s ok older sibling can drive back and pick them up again
A: <X> ((NAME: A)) thinks go together is good also
B: that’s also good (.) what about when we come back?
A: when you come back older brother ((NAME: A’s husband)) will come and pick up
the children (.....)

Employing Auer’s distinction of codeswitching (1995), the skeleton of this conversation
can be outlined as follows:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A1:</td>
<td>offer</td>
</tr>
<tr>
<td>B1:</td>
<td>dispreferred second pair part</td>
</tr>
<tr>
<td>A1:</td>
<td>indirect offer (2\textsuperscript{nd} attempt)</td>
</tr>
<tr>
<td>B1:</td>
<td>refuse (2\textsuperscript{nd} attempt)</td>
</tr>
<tr>
<td>C1:</td>
<td>comment in favor of the offer</td>
</tr>
<tr>
<td>B2:</td>
<td>refuse (3\textsuperscript{rd} attempt in northern dialect)</td>
</tr>
<tr>
<td>A1:</td>
<td>re-offer</td>
</tr>
<tr>
<td>B2:</td>
<td>refuse (4\textsuperscript{th} attempt in northern dialect)</td>
</tr>
<tr>
<td>A2:</td>
<td>offer (in northern dialect)</td>
</tr>
<tr>
<td>B2:</td>
<td>accept the offer (in northern dialect)</td>
</tr>
</tbody>
</table>

\footnote{\textsuperscript{34} 'A' refers to Speaker A and '1' is Language 1 (Thai), '2' is Language 2 (the northern Thai dialect)}
This conversation starts out with Speaker A offering a ride to Speaker B's children in standard Thai. The utterance in line 1 is the first pair part, which sets up conditional relevance and expectation to be fulfilled by a second pair part. Her turn is responded to by Speaker B in standard Thai. It is a dispreferred second or a dispreferred response, as it refuses the offer. In line 6, Speaker C supports the offer. The following turn, in line 11, Speaker B switches to the northern Thai dialect to mark a dispreferred second when his first two attempts in standard Thai failed in line 4 and 9. In line 15, Speaker A switches to the northern Thai dialect and insists on the offer. Speaker B eventually takes the offer.

These three conversation excerpts serve to highlight the role of codeswitching as a contextualization cue, by setting up a structural contrast. It is a discourse strategy whereby bilingual speakers accommodate and collaborate with each other. Example (1) demonstrates a switching to engage a non-Thai speaker into a conversation. Example (20) and (21) demonstrate a switching to an ethnic code to mark dispreferred second, or to reject the offer. This finding echoes Li's finding that conversational participants use the ethnic language to mark dispreferred response (1994).

4.3.3 Reported Speech

Codeswitching is a frequent correlate of reported speech in conversation among Thai students. The direction of the switch is not always significant, and the code used for the quotation is not necessarily that of the actual speech being reported as shown by the example below.

EXAMPLE (22)

(.....)

1 A:  muawain <X> (NAME) tʰʰaw m kʰaw wa:()
yesterday ask 3rd Pro that
2  cho:phu:jin ji:pun rwp law () ʔaraj jaːŋ ni:()  
like woman Japan Q something like that
A: ...yesterday <X> (NAME) asked him (. ) if he likes Japanese women (. ) something like that (. ) he said yeah (. ) same ethnic background speaking the same language but if that’s not possible it’s okay too =

B =No he talked to me (. ) he said this (. ) he said that (. ) as for him <X> ((NAME)) the one who came last semester [He said a lot of people were giving him a full support for his relationship with her

A: [yeah yeah people were giving him support

B: [but he said=

A: =yeah yeah
B: <X> ((NAME)) is out he said it’s hard to grasp (smile voice) what she’s thinking (laugh)
A: [yeah
C: [yeah
D: [yeah
B: He said that (.) like (.) he’s not looking for a short-term relationship (.....)

Example (22) is another example of the use of Thai-English codeswitching as a contextualization cue in quotation. As shown above, Speaker A uses Thai to code her reported speech, and Speaker B uses both Thai and English. The conversation is about a male friend and his personal relationship. He is a Japanese male graduate student who has been in the United States for over six years. Speaker A is telling other participants what she knows about this Japanese friend.

As shown above, as soon as Speaker A finishes her turn in line 4, B starts hers in line 5 by disputing the information given by Speaker A. Note that there is no pause between Speaker A and Speaker B’s turns. Then, Speaker B starts to quote the Japanese friend’s utterances in Thai. Beginning with line 5, there are a number of overlaps between Speaker B and A; specifically there are overlaps in line 7 and 9-10; and 11 and 12. Speaker B then switches to English in her last quote of the Japanese friend’s speech. This switch indexes her identity as a bilingual speaker of Thai and English. Also, by switching into English, Speaker B builds up a contrast to negotiate her position as someone who has a direct interaction with this Japanese friend and, as such, she is someone who knows better in this matter. Being able to quote in the original language shows authority in a sense that she has a direct interaction with the third person being talked about, and she does not only know what he said, she also remembers from direct experience. As a matter of fact, her negotiation is successful: every speaker in the interaction in the following turns accepts her information as true.
Additionally, switching into English also serves to show Speaker B's identity as less disparaging than could have been thought of had she used Thai. In fact, other participants do the same thing when they are giving somewhat derogatory information about a third person not present in the conversation. By switching to the code of the actual speech, they are not responsible for the comments made.
CHAPTER 5
SYNTAX OF THAI/ENGLISH INTRASENTENTIAL CODESWITCHING

This chapter accounts for the Thai/English codeswitching data from a grammatical perspective, employing the Matrix Language Frame model.

5.1 TYPES OF MIXED CONSTITUENTS

Similar to other codeswitching data reported in previous literature, noun, verb, and adjective are the predominant categories at which switches take place in the speech of the Thai participants. This supports the Matrix Language Principle and the 4-M model that content morphemes are more likely to be switched than other types of morphemes.

There are switches of some English system morphemes; however, they are quite infrequent. When they occur, they often occur as a part of an Embedded Language island.

In the present study, there are only two examples of singly occurring system morphemes on Thai base forms. This is because Thai is not an inflectional language. Information on number, tense, and gender of a word is given by the addition of extra words. Below are some examples of Thai/English codeswitching found in the present study.

Examples of noun switching:

(1) man pen ka:n sa:n awareness haj pʰuːjiŋ
   it is Prefix build to woman
   'It was an awareness promoting among women.'

(2) land man pa:j tʰwŋ naj
   3rdPro go reach where
   'How far does the land reach?'

(3) miː loft tʰiː either pen office ruː pen playroom
   have Rel. be or be
   'There is a loft that could either be an office or a playroom.'
(4) phrowa: man thw: pen contract paj la:w
because 3rd pro count be go already
'Because it's already counted as a contract.'

(5) tæ: man pen walk up
but 3rd pro be
'But it's a walk-up.'

(6) khaw ca bæp we:la: tham nat paj du: ca bæp chaphɔ
3rd pro Asp like35 when make appointment go look will like only
serious buyer
'They are like, when making an appointment to look at the house, it will be only for serious buyers.'

(7) tæ: ha: wajkaj pen first priority
but be
'But Hawaii Kai is the first priority.'

(8) tɔŋ mi: mystery lek lek
must have little little
'They must have a little mystery.'

(9) maj mi: kʰraj ru: ra:jla?iat kiawkap organization ni:
Neg. have who know detail about this
'Nobody knows any detail about this organization.'

Examples of verb switching:

(10) tæ: ko:n th:i: kʰaj offer <X>((NAME)) kʰaw maj kin na
before Rel. used to 3rd pro Neg. eat Part.
'I offered <X> (the sushi) before but she didn’t want to eat.'

(11) ru: tɔŋ ru: sunk offended wa: pʰwan paj sw: swa ma:
or must feel that friend go buy shirt come
same
'...or feel offended that your friends bought the same shirts.'

(12) tæ: ko mi: kʰon mention
but also have person
'But there are people who mention that...'

(13)pʰɔ: mæ: ko push kʰaw ke:n paj
father mother then 3rd pro over
'His parents push him too much.'

35 The Thai, [bæp] functions similarly to the English, like (Kangkun 2000).
(14) ʔaːtca **afford** majdaj ka  
maybe  Neg.  Part.  
'Ve may not be able to afford it.'

(15) pʰuːjìŋ **aware** kap rwan ṭareŋ  
woman  with  story  cancer  
'Women are aware of cancer.'

(16) raw ko **enjoy** thːiː ca len  
1stPro then  to  will  play  
'We then can enjoy playing.'

(17) kʰawniaw  kʰaw **pack** jaːŋdiː:  
sticky  rice  3rdPro  well  
'They packed the sticky rice well.'

(18) tiːn kaj niː jɑŋ bɔdaj **cook**  
foot  chicken  here  yet  Neg.  
'They have not cooked the chicken feet.'

(19) kʰit waː nɔːŋ nɔːŋ nia kʰon ca **rest**  
think  that  younger  sister  younger  sister  Part  probably  Asp.  
'(I) thought that you guys would probably rest'

(20) kʰaw ko maj maː **hire** pʰiː rʰoːk  
3rdPro  then  Neg.  come  older  sibling  Part.  
'They're not going to hire you.'

(21) kʰaw ca paj **hang** kap kʰon ʔɯːn  
3rdPro  will  go  with  person  other  
'He will go hang with other people.'

(22) ko maj kʰawcaj waː tʰammaj ca **post** maj daj  
then  Neg  understand  that  why  will  Neg.  can  
'I don’t understand why we cannot post it.'

(23) pen kʰon **add** ʔεːŋ  
be  person  self  
'I am the person who adds it.'

(24) **you** ko tɔŋ **need** kʰaw  
then  must  3rdPro  
'You then need them.'

(25) jaːŋ weːlaː lwanjot kʰaw ko **require** sen rwan ṭareŋ **qualification**  
like  when  promote  3rdPro  then  connection  issue  
'Like when it’s promotion time, they require connections regarding qualifications.'
(26) <X> ((NAME)) ko forget it daj laj
then can already
'You then can forget it already.'

Examples of adjective switching:

(27) kʰaw chɔːp kʰon tʰiː bǣp mysterious
3rd Pro like person Rel. like
'He likes the person who is like mysterious.'

(28) rwaŋ kʰon sɔːŋ kʰon nan dramatic maːk
story Pos. two person that very
'The story of both of them is very dramatic.'

(29) man mɯn kap waː man unknown
3rd Pro like with that 3rd Pro
'It's like it's unknown.'

(30) <X> ((NAME)) pen dek tʰi gentle maːk
be child Rel. very
'<X> is a very gentle kid.'

(31) kʰaw ko nɯŋ jɔːnlɔŋ waː kʰon ca bǣp miserable maːk
3rd Pro then think back that probably will like very
'She thought back and said that she probably would be very miserable.'

(32) <X> ((NAME)) majdaj kʰit waː man critical
Neg. think that 3rd Pro
'<X> did not think that it's critical.'

(33) tɔː: chɔːp tʰam naː gloomy gloomy duaj
but like make face also
'But she likes to make gloomy face.'

(34) pretty nia cheerful tɔː: beautiful nia maj campen toŋ cheerful
Part but Part Neg. necessary must
'Pretty is cheerful but beautiful is not necessarily cheerful.'

(35) ko kʰwː: man paranoid paj læːw
then be 3rd Pro go already
'It's like he's already paranoid.'

(36) kʰɔːŋlen innovative thiː mwaŋtʰaj maj kʰoŋ miː
toy at Thailand Neg. quite have
'Thailand doesn't quite have innovative toys.'
Examples of adverb and conjunction switching:

(37) raw ko cam maj daj exactly
  1st Pro then remember Neg. can
  'I don’t remember exactly.'

(38) <X> ((NAME)) ko either support
    then or say that
  '<X> then either support it or say that...'

(39) no tomar pah sen pha: khaw ko:n
    must go send older sibling 3rd pro first
  'No, I have to drop older sister off first.'

(40) yeah rian kenh
    study well
  'yeah, they are good at study.'

5.2 THE 4-M MODEL: CONTENT-SYSTEM MORPHEME OPPOSITIONS

The 4-M model assumes that each type of morpheme is activated and accessed differently in the production process. Morphemes are categorized according to their different activations at different levels in the production process, drawing on how early or late they are retrieved from the mental lexicon in the process of language production. According to the model, there are four types of morphemes: (1) content morphemes; e.g., all nouns, adjectives, and most verbs, (2) early system morphemes; e.g., plural, and determiner including articles and possessive adjectives, (3) bridge system morphemes; e.g., possessive of and 's, and (4) outsider system morphemes; e.g., subject-verb

---

36 This is a switch of a conjunction together with a verb.
37 According to Bock and Levelt (1994), the lexical selection involves three different levels: conceptual, lemma, and lexeme or sound level. The assumed lexical retrieval process in Myers-Scotton’s model is very similar to the network model proposed by Bock and Levelt (1994). This is not surprising, as Myers-Scotton’s model of language production is based on Levelt 1989. She, however, added the formulator level, at which she assumed the late system morphemes are assigned to a content morpheme. Her assumption that speakers’ intentions are directly linked to the lemma of a content morpheme, which in turn searches for a relevant system morpheme is consistent with studies conducted in psycholinguistics. The content and system morphemes are accessed differently. Nevertheless, the distinction of the four types of morphemes is from her evidence from aphasia and second language acquisition (Myers-Scotton and Jake 2000).
agreement, case affixes, and clitics/affixes. Both bridge and outsider system morphemes belong to the same class: late system morphemes.

The 4-M model assumes that these morphemes are differently activated. They are either plus (+) or minus (-) conceptually activated at the conceptual level. The content morpheme and the early system morphemes are conceptually activated, unlike the other two late system morphemes. The assumption is that speaker's intention selects an appropriate content morpheme: the one that best fits his/her intention. Then, at the lemma level, relevant system morphemes are selected by the lemma supporting the content morpheme. The lemmas of the other two types of system morphemes are salient at the formulator level.

In conformity with the model, content morphemes are the predominantly switched category, as shown in §5.1. The early system morpheme is scarcely observed in the data.

For example,

(41)
\[ \text{mre: } \text{k}^b\text{o:̀} \text{d} \text{k}^b\text{i: } \text{ce: } \text{b}^i \text{k}^b\text{o} \text{n} \text{n}^b\text{o} \text{t}^i \text{i: } \text{p}^e\text{n} \text{lawyer} \]
mother Pos. child Rel. meet another person one Rel. be
'The mother of another kid that I met is a lawyer.'

(42)
\[ \text{boss p}^b\text{i: } \text{<X> ((NAME)) ko mention wa:} \]
older sibling then that
'Older brother <X>'s boss also mentioned that...'

In example (41), the noun lawyer does not have any article preceding it. In example (42), the verb mention does not have an inflectional morpheme attached to it. That is, these switched words occur in Thai utterances without any early system morphemes. As for the two late system morphemes, their occurrence is quite infrequent. They usually appear as a part of larger linguistic units, like an embedded language island. An example of a bridge system morpheme is the preposition at as in one at a time.
This is by no means surprising, as Thai is an isolating language. The concept of "congruence" accounts for the non-occurrence of the early system morphemes or the overall infrequency of the system morphemes as a whole.

5.3 THE ABSTRACT LEVEL MODEL AND CONGRUENCE CHECKING

Central to the Abstract Level model is the congruence checking that takes place at three different levels of abstract lexical/grammatical structure: (1) the lexical-conceptual structure, (2) the predicate-argument structure, and (3) morphological realization patterns. The first level involves congruence checking regarding the matching between semantic/pragmatic feature bundles and speakers' intentions at the conceptual level. The second level involves the congruence checking between the two linguistic codes in terms of the mapping relations between thematic role assigners and the relevant arguments. The third level refers to requirements by grammatical constraints for surface-level realizations in terms of elements and constituents order. Congruence checking occurs at these three levels. If there is sufficient congruence between the Embedded Language elements and the Matrix Language counterparts, the Embedded Language elements can appear at the surface structure within the grammatical frame provided by the Matrix Language. In Examples (41) and (42) above, the switched elements do not require any system morpheme, be it an article, a tense marker, or a subject-verb agreement morpheme, as the grammatical frame provided by the Matrix Language, Thai in this case, does not call for such morphemes. In other words, the switched items, lawyer, boss, mention, from English
are perfectly well-formed in the language of the Matrix Language: there are no determiners as integral parts of NPs in Thai.

It is worth mentioning here why speakers switch when there is a perfect congruence at the lexical-conceptual level, e.g., lawyer and boss. Certainly Thai has the exact equivalents of these two English words, [tʰanŋ:i] and [nə:i]. A rational explanation is provided by the markedness model: the use of two languages in discourse reflects the communicative intentions of speakers. On the other hand, when there is insufficient congruence at this level, speakers would select the form that best represents their intentions.

(44) kʰaw huan wa: kʰaw ca tɔŋ handle kʰon lek
3rdPro worry that 3rdPro will must person little
'She is worried that she will have to handle the little one.'

This sentence is an utterance that is a part of a conversation between two close friends; one of them is trying to help the other to find a baby-sitter for her two children, age 1 and 3. The Thai pronoun, [kʰaw], in the utterance refers to a female third person who has never babysat, but would like to help with babysitting if she does not have to handle the younger child. In this context, the English word is selected over the Thai counterparts [duː], 'watch', [duː lɛː], 'take care', or [ca tɔː:n] 'manage', because it better conveys the speaker's intentions, the word 'handle' has a sense of dealing with a difficult situation or problem. The word [ca tɔː:n] 'manage' is similar to handle in this sense; however, [ca tɔː:n] would have a negative connotation if used in this context. When [ca tɔː:n] is used with animate entities, as opposed to inanimate ones, its possible meanings are 'to tackle, to punish, or to kill'.

130
5.4 GRAMMATICAL CONSTRAINTS ON THAI/ENGLISH CODESWITCHING

This section addresses Thai/English codeswitching in terms of its grammatical constraints, particularly when there is insufficient congruence between the two languages at the surface level. This will be discussed together with the Matrix Language Principle, the Blocking Hypothesis, and the Embedded Island Hypothesis. The unit of analysis is the CP, the Matrix Language Principle states that a Matrix Language cannot change within a bilingual CP.

5.4.1 The Matrix Language Principle

The Matrix Language Principle claims that the occurrence of all switched elements, including the mixed constituents, Matrix Language islands, and Embedded Language islands, has to conform to the grammatical frame provided by the Matrix Language. As mentioned in §3.5, Thai is an isolating language; it does not have inflectional morphemes. Information on the number, tense, and gender of a word is given by the addition of extra words. Below are some examples from the Thai/English codeswitching data.

(45) [man ca mi: spot num next to 3rd pro have one 3rd pro handicap]cp
   'There's one spot next to (the) handicap'

(46) [she ca maj k^o j pay attention]cp
   Asp. Neg. fairly
   'She doesn’t really pay attention.’

(47) [taw: k^aw get along kap p^uak ha:wa:j]cp
   but 3rd pro with group Hawaiian
   'but he got along with the Hawaiian people.’

(48) [k^aw tho: ma: at a bad time cin cin]cp
   3rd pro call come true true
   'he called at a really bad time.'
(49) [mæ:n naː kʰaw on the board na]cp
boyfriend aunt 3rdPro Part.
‘My boyfriend, he is on the board.’

(50) [tɔŋ chaj rubber hammer]cp
must use
‘(You) have to use (the) rubber hammer.’

(51) [nuː boːk læːw waː [large organization man diːkwaː]cp]cp
1stPro tell Asp. that 3rdPro better
‘I told (you) that (the) large organization is better’

(52) [pʰrowaː baːŋkʰ ræŋ cannot afford to lose you know]cp
because sometimes
‘because sometimes (they) cannot afford to lose you know’

The Matrix Language in these examples is Thai. Example (45) shows that the mixed constituents comply with the Thai grammar. The switched noun, *spot*, occurs in front of the numeral, [nɯŋ] ‘one’. The predicative prepositional phrase, *next to*, follows its subject, *spot nɯŋ*. The object of the prepositional phrase, *Pai handicap*, occurs in the sentence-final position. The Matrix Language island is well-formed in Example (46). Also, the Embedded Language islands in Examples (47) – (52) do not violate the structural frame provided by the Matrix Language. It is noticeable that articles, subject-verb agreement morphemes, or tense morphemes are not permitted in the Thai structure. If these morphemes ever appear in the Thai Matrix Language frame, they always occur as a part of the Embedded Language islands, as in Examples (48) and (49).

Even though Thai and English are quite similar in their grammatical structures, SVO, the sequential order of some grammatical categories is different; the modifier appears after the constituents they modify. Supportive examples found in the Thai/English data involve the occurrence of English adjectives in the Thai grammatical frame.
As shown above, the order of adjective and noun complies with the Thai grammatical frame.

5.4.2 The Blocking Hypothesis and the Embedded Island Hypothesis

The Blocking Hypothesis states that any Embedded Language content morpheme that is incongruent with the Matrix Language will be blocked. Congruence is defined as a match between the Matrix Language and the Embedded Language at the lemma level. Lemmas consist of abstract pragmatic information, as well as semantic, syntactic, and morphological information. To recapitulate, the three levels of lexical structure include (1) lexical-conceptual structure, (2) predicate-argument structure, and (3) morphological realization patterns. That is, if a certain Embedded Language item is incongruent with the Matrix Language at one or more of the three levels of lexical structure, it cannot occur in the Matrix Language frame, unless it appears as an Embedded Language island. The Embedded Language island is defined as a full constituent that consists of two or more Embedded Language morphemes in a CP within a larger Matrix Language frame.

In Examples (45) to (52) above, most of the Embedded Language islands are incongruent with their Thai counterparts except for Examples (50) and (51), rubber hammer and large organization; the switch to English lexemes in these examples is due to the speakers' familiarity with using the two languages, Thai and English, together. Why this occurs is elaborately explained in §4.2.2. To recapitulate briefly, Myers-Scotton (1993b) explains that the "overall unmarked codeswitching" not only conveys referential
meaning, but also communicative intention; it reflects the dual indexical values associated with the two codes. Other switched constituents are incongruent at one or more of the three levels of the lexical structure. For example, there is not a single Thai word for the English word, *afford*. Its Thai counterparts are [mi: ŋen pʰoː] 'have enough money to', or [mi: wə:la: pʰoː] 'have enough time to'.

In addition, it is noticeable that there is no verb in Example (49); [fæːn naː kʰaw *on the board* na]cp. In Thai, subject, verb, and other constituents often times can be omitted. The following examples on the left are some examples of verbless sentences in Thai.

<table>
<thead>
<tr>
<th>dek kʰaj suːŋ</th>
<th>vs</th>
<th>dek miː kʰaj suːŋ</th>
</tr>
</thead>
<tbody>
<tr>
<td>child fever high</td>
<td></td>
<td>child have fever high</td>
</tr>
<tr>
<td>'The child has a high fever'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wanniː wanʔatʰit</td>
<td></td>
<td>wanniː pen wanʔatʰit</td>
</tr>
<tr>
<td>today Sunday</td>
<td></td>
<td>today be Sunday</td>
</tr>
<tr>
<td>'Today is Sunday.'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>baːn niː kʰoːŋ kʰaw</td>
<td></td>
<td>baːn niː pen kʰoːŋ kʰaw</td>
</tr>
<tr>
<td>house this of 3rd Pro</td>
<td></td>
<td>house this be of 3rd Pro</td>
</tr>
<tr>
<td>'This house is his.'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Accordingly, Example (49) can be analyzed as the verbless subject-predicate construction. A one-place predicate requires one argument in a sentence, while a two-place predicate requires two. *On the board* is the one-place predicate of the utterance. Therefore, Subject, [kʰaw] 'he', is the only obligatory constituent. Moreover, the omitted Thai verb [juːː] in this example functions more or less like the copula *be* in English. According to Huddleston (1988), the copula *be* has little semantic content; its primary syntactic role is to link a complement to its subject. It is, therefore, omissible in certain constructions.

Additionally, a set collocation often does not have its equivalence in another language at all three levels (Myers-Scotton 2002). Therefore, such set collocations
usually appear in the form of the Embedded Language islands. Typical Embedded Language islands also include formulaic constructs and adverbials of time or adverbia

(55) tæː fæːn  naː kʰaw stay in shape
    but  boyfriend aunt 3rdPro
    'But my boyfriend, he stays in shape.'

(56)raw maj  chóːp  <X> ((NAME)) jʊː: laːːw to begin with na
1stPro Neg. like Asp. already Part.
'I didn't like <X> to begin with.'

5.4.3 Explicating More Data: Speakers' Intentions, Economy, and Structural Constraints

The Thai/English data presented thus far is typical in codeswitching study. That is, the switched items are usually content morphemes or formulaic constructs. In the present Thai/English data, there are a few cases, however, when the early system morpheme from the Embedded Language is attached to the Thai base form.

(57) [sɔŋ sɔːɻ-m] cp
    ma  si]cp
    pass  fork-s  come Part.
    'pass the forks'

(58) [raw kɔ maː  re-faŋ da] cjcp
    1stPro then  come re-listen  can
    'I can listen to it again'

The System Morpheme Principle states that only the Matrix Language can supply the system morphemes whose grammatical relations are external to their head constituent. In other words, the late outsider system morpheme is the one and only type of morpheme obligatorily coming from the Matrix Language. Accordingly, Examples (57) and (58) above are perfectly well-formed. The plural suffix in Example (57) and the prefix re- in Example (58) do not have to look outside their immediate maximal constituents for information regarding their forms.
A plausible explanation as to why such utterances occur at all is speakers' intentions. The 4-M model claims that all the content and system morphemes in a bilingual CP are related to their activations at different levels in the production process. The speakers' intentions are directly linked to the lemmas supporting content morphemes at the conceptual or the mental lexicon level. When an appropriate content morpheme is selected, the lemma supporting the content morpheme subsequently selects an appropriate early system morpheme. In Example (57), the speaker selects the Thai word for *fork* and wishes to express plurality. The lemma supporting the content morpheme *fork* is then activated to look for an early system morpheme with the feature [+quantification]. Plurality in Thai is expressed by the addition of quantifiers and relevant classifiers. In Thai, the utterance would be something like [soŋ soːm maː saːm ?an si]; [saːm] is the quantifier ‘three’ and [?an] is the classifier for *fork*. The speaker's intentions play a key role here. That is, with the speaker’s intention to express plurality of the word *fork* and to express it as economically as possible, the speaker opts to use the English plural suffix, *-s*. A similar explanation can be offered for Example (58). To express the reiteration of an action requires another word in Thai, such as [ʔiːktʰiː] or [māi], meaning ‘again’. The speaker however chooses to use the English prefix, *re-*, to convey the meaning. In short, the speakers’ intentions call for a plural suffix of the noun in this case. It is the creativity of the speaker that seems to play a key role in these utterances. If this assumption is true, it would be in support of the idea that speaker is a rational actor; choices ultimately lie with the individual.
Nevertheless, according to Bock and Levelt (1994), “... grammatical plurality is a property of lemmas rather than of nonlinguistic concepts or messages ...” If this assumption is true, the selection of the early system morpheme cannot be activated by the speaker’s intention. In this case, the System Morpheme Principle may over-predict possible constructions of codeswitching. Also, Myers-Scotton (2002) assumes that, according to Levelt (1984), lemmas underlying content morphemes are the only lemmas that are directly linked to the speaker’s intention. The activation of relevant system morphemes is induced by lemma underlying the content morphemes. Therefore, the above explanation (speakers’ intention) is ruled out. The remaining question however is whether the lemma underlying the Thai content morpheme can select the early system morpheme from English. This question is outside of the scope of the present study. Assumptions regarding lexical retrieval in monolingual language production as well as bilingual language production are still quite controversial. Generally, the bilingual production process is assumed to be fairly similar to that of monolingual. Nevertheless, it requires additional assumptions (Costa et al. 2000). As of yet, there is no agreement as to how the monolingual or bilingual production processes actually work (see Bock and Levelt 1994, Vinson and Vigliocco 1999, Costa et al. 2000).

Another interesting utterance found in the Thai/English codeswitching data is correlative to the Matrix Language Principle and the Morpheme Order Principle, which claim that the Matrix Language is the one and the only language that supplies the surface morpheme order for the bilingual CP consisting particularly of singly occurring Embedded Language lexemes and Matrix Language morphemes.

\[(59) [k^haw bo:k jaŋ maj officially pə:t]cp\]
\[3^{rd} Pro say yet Neg. open\]
‘(They) said it’s not officially open.’
As stated earlier, Thai modifiers usually follow the constituents they modify; a typical word order of an adverb and a verb is V + Adv.

As shown above, the location of the AdvP is not fixed. Yet the adverb of manner, [nak ma:k] very heavy never appears before the VP, [tok] fall. Pankhueankhat (1998) stated that only position in which the adverb of manner can occur is after the verb or after an object (if any). I argue that the adverb of manner can also appear in front of the constituents it modifies for special effects in spoken language. For example:
This example shows that an adverb of manner can appear in front of a verb it modifies for an emphatic effect. Such constructions usually occur with a pause, the lengthening of a vowel, or the audible degree of the relevant adverb.

As far as English grammar is concerned, Quirk et al. (1985) divide adverbials (A-elements) into seven main categories according to their semantic roles: space, time, process, respect, contingency, modality, and degree. The English adverbs in Example (59), officially, and (60), actually, belong to the subclass of the adverb of manner in the process adverb, and the emphasizer in a modality adverb, respectively.

An adverb manner adjunct can also be paraphrased by in a ... manner or in a ... way. Example (59) can then be paraphrased as in an official manner or in an official way. Also, it can occur with a Negative particle; ‘not officially’ for example. As far as the emphasizers are concerned, they can co-occur with any verb or predication. They also have an emphatic effect with non-gradable verbs, e.g., he definitely saw her. Syntactically, most emphasizers precede the constituent they emphasize. Many of them tend to co-occur with clausal negation; in such a case, the emphasizer usually precedes the verb phrase. That is, they can be the focus of negation. For example, I didn’t actually see her.

Furthermore, modality adverbs and degree adverbs are similar to one another in their semantic effect. The difference is that degree adverbials are concerned with the assessment of gradable constituents in relation to an imaginary scale. To explicate briefly, modality adverbs are concerned with three aspects of a statement: emphasis,
approximation, and restriction. Degree adverbs, on the other hand, are concerned with amplification, diminution, and measure. Syntactically, they tend to occur in a medial position in English. Generally, A-elements can occur in different positions in a sentence: initial (I), medial (M), end position (E), etc. The medial position is the position between subject and verb—specifically the position immediately after the subject and the operator.

According to Quirk (1985:493), “[s]emantically, M is especially associated with modality and degree ... [g]rammatically, M is freely used for focusing and intensifying subjuncts ..., and fairly freely for optional predication adjuncts ..., as well as for some disjuncts ... and conjuncts ...” (Original italics). Accordingly, it is grammatical for the emphizer, actually, to occur at the medial position for an emphatic purpose in English.

Similarly to English degree adverbs, degree adverbs in Thai can occur in the medial position (Pankhueankhat 1998). Examples of Thai degree adverbs are [suanma:k] ‘mostly, most of the time’, [ta:mtbammada: ‘usually’, and [ta:mpokkati ‘normally’]. All of them can also occur at the initial or end position:

(69) ta:mtbammada: kha:ma: ro:nリアn chaw
    usually 3rdPro come school early

(70) kha:ta:mtbammada: ma: ro:nリアn chaw
    3rdPro usually come school early

(71) kha:ma: ro:nリアn chaw ta:mtbammada:
    3rdPro come school early usually

As for the modality adverb, the existence of this category of adverb is not discussed in Pankhueankhat (1998). However, following the semantic categorization of Quirk (1985), I suggest that such a category exists in Thai. Yet the Thai modality adverb cannot occur in the medial position. It can occur in the end position and, in some cases, in the initial position. For example,
These adverbs cannot be analyzed as adverbs of manner because they are syntactically and semantically different from the typical manner of adverb; they cannot occur with negator or be paraphrased by *in a ... manner.* In Thai, a sentence like *[kʰaw kin chaː], he eats slowly,* can be paraphrased as *[kʰaw kin jaːŋ chaː] he eats in a slow manner.* It can also occur with a negator, *[kʰaw kin maj chaː]. However, sentences like *[raw tham daj jaːŋ cincin] or *[chan henduaŋ maj naː:noː:n] are ungrammatical. The sentence, *[chan henduaŋ jaːŋ naː:noː:n] is grammatical; however, the typical adverb of manner can both occur with a negator and be paraphrased by *in a ... manner* as shown above.

Accordingly, Example (60) *[kʰaw ko jaŋ maj daj actually tham ʔaraj]cp,* seems to run counter to the Morpheme Order Principle. However, this is the only
utterance that contains modality adverb in the medial position. There are only five examples, (39), (40), (59), (60), and (61), with adverb switching in the Thai/English data—two examples of yes/no adverbs (39), (40), two examples of modality adverbs, (60) and (61), and one example of manner adverb (59). Therefore, it is also plausible to make an assumption that Example (60) is simply a speech error.

To sum up, the System Morpheme Principle in the Matrix Language Frame model can account for the occurrence of English system morphemes with the Thai base forms, Examples (57) and (58). However, the question remains whether the lemma underlying the content word in Thai can select the early system morpheme in English. Furthermore, the adverb of manner can occur in front of the constituent it modifies for special effects, particularly the emphatic effect intended by the speakers, as in Example (59). This is also analogous to the claim that codeswitching can function as a contextualization cue for emphatic effects. In the Thai construct for ‘slowly eat’, [cha: cha: (.) kin na], a pause functions as a contextualization cue for the emphatic effect. Codeswitching in Example (59) functions similarly to the pause in the Thai monolingual construct. Example (60), however, seems to run counter to the Morpheme Order Principle. The Thai language, which is the Matrix Language in this case, is not the language that supplies the morpheme order. However, the number of existing examples in the Thai/English data is too small to make any strong statement at this time. Future research is necessary to verify such a claim.
CHAPTER 6
DISCUSSION AND CONCLUSION

Chapter 6 summarizes the findings from both sociolinguistic and syntactic aspects of the Thai/English codeswitching data. This is followed by an evaluation of the theoretical frameworks employed in the study. Limitations of the present study and suggestions for future research will also be addressed.

6.1 A SUMMARY OF FINDINGS

6.1.1 Language Functions of Thai/English Codeswitching

The language functions of Thai/English codeswitching in conversation are both indexical and interactional. The markedness model provides an analytic frame by which the Thai/English codeswitching data can be analyzed in terms of their indexical and interactional functions. The markedness model takes into account not only societal norms, but also the rationality and individuality of speakers. Three types of codeswitching occur in the Thai/English discourse: (1) sequential unmarked codeswitching, (2) overall codeswitching as an unmarked choice, and (3) codeswitching as a marked choice. The Thai/English data presented in § 4.2.1, the “sequential unmarked codeswitching,” have shown that codeswitching is a convenient means for speakers to reach out to their interlocutors, both Thai and non-Thai speakers. It also allows them to make their narrations more vivid and real in reported speech. The Thai/English data that fit in the “overall codeswitching as the unmarked choice” (§ 4.2.2) reflect the dual indexical values associated with both Thai and English. Particularly, the participants use codeswitching as an unmarked choice to signal their multiple identities among their peers and maintain
their ethnic balance as Thai/English speakers. It signifies the speakers’ shared RO sets as internationally well-educated bilingual Thai/English speakers. Furthermore, codeswitching as the marked choice (§ 4.2.3) allows the participants to negotiate their RO sets with one another in order to accomplish certain communicative goals, it constitutes a resource available to participants to appeal to their group solidarity, to attract attention, and to create aesthetic effects in their narrations as a whole.

The conversational codeswitching approach demonstrates that codeswitching can be used as a contextualization cue in monolingual discourse, such as prosody, gestures, etc. Specifically, codeswitching builds up a structural contrast to contextualize turn transitions. It is used to signal turn handovers. Additionally, Thai/English codeswitching is a discourse strategy whereby bilingual speakers accommodate and collaborate with each other. The participants resort to the ethnic languages when presenting dispreferred seconds.

In terms of a wider social context, it is obvious from both the detailed analysis of the Thai/English data and long-term participant observation that Thai is the preferred language by all the speakers who participated in this study. This implies that conversations initiated by these speakers will be in Thai in most cases. When the speakers wish to build up a contrast to signal the turn-transition, they will switch to English. Also, the detailed analysis also demonstrates that ethnic languages or dialects are the languages of solidarity between peers, as shown in examples for dispreferred second pair parts.

With regard to Nivens’s claim that discourse contexts do not necessarily induce a switch from Language A to Language B (Nivens 2002), reported speech data presented
§4.2.1 and §4.3.3 in Chapter 4 show that the speakers may switch to another linguistic code in repetition or quotation for certain conversational effects. Sometimes they use Thai to quote the utterances and sometimes English, even though the utterances being quoted were originally English. The discourse contexts trigger a switch from Thai to English or vice versa for conversational effects, e.g., to attract attention, to assert one’s argument, and so on.

In addition, it is worth emphasizing that even though the results from the interview show that codeswitching is considered unmarked among all the Thai participants in the present study, their use of codeswitching is quite minimal. The community language norm is still very much oriented towards monolingualism. The most common pattern of codeswitching found in the present study is the “overall codeswitching as the unmarked choice”, and that occurs only among bilingual in-group members. The movement of linguistic purism and the connection between codeswitching and “one’s ostentation” are fairly invincible. Thai is without exception the “safe” code among the participants.

6.1.2 Syntactic Aspects of Thai/English Codeswitching

The constituents that are mixed in the Thai/English codeswitching data are noun, verb, adjective, adverb, and some affixes. The predominantly switched category is the content morpheme [+thematic role assigner/receiver, -quantification] (as opposed to system morphemes [-thematic role assigner/receiver, +quantification]). This corroborates with the Matrix Language Principle and the 4-M model that the content morphemes are more likely to be switched.
Generally, most of the data can be explained by the machinery provided by the model. Most of the hypotheses and principles are supported. Examples (52) – (54) in Chapter 5 support the Abstract Level model, the Blocking Hypothesis, and the Embedded Language Island Hypothesis.

In terms of the Matrix Language Principle and the Morpheme Order Principle, Thai, the Matrix Language, is the language that supplies the morphosyntactic frame for the singly occurring Embedded Language lexemes for all cases, except for one example, Example (20). Examples from Chapter 5 illustrate this point.

(42) \textit{boss} p^{h_i} \langle X \rangle \textit{mention} \textit{wa}: \\
older sister then that \\
‘Older brother \textit{<X>’s} boss also mentioned that...’

Example (42) shows that in the possessive noun phrase, the English noun, \textit{boss}, precedes the possessor, \[p^{h_i} \langle X \rangle\], \textit{older sibling}—the grammatical word order for a Thai possessive noun phrase.

(53) \textit{cho:p len k^{h_o}n \textit{innovative}]}cp \\
like play stuff \\
‘(They) like to play with innovative stuff.’

(54) \textit{rwa:n dramatic mi: j\textit{e:]}cp \\
story have many \\
‘There are many dramatic stories.’

Examples (53) and (54) show the Thai word order for bilingual adjective phrases. The adjective occurs after the noun it modifies in Thai.

(60) \textit{k^{haw} ko ja\textit{n maj daj actually} t^{h_am 2araj]}cp \\
3^{rd}Pro then yet Neg. can do what \\
‘He hasn’t actually done anything.’

The word order of an adverb and the constituent it modifies in Example (60) is from English, the Embedded Language. This particular example seems to counter the Morpheme Order Principle. However, a strong statement cannot be made at this time due
to the small number of data containing English adverbs. There are only five examples of Thai/English codeswitching data with English adverbs, two of which are yes/no adverbs and, therefore, are not relevant.

In addition, Thai is the language that provides the grammatical framework for utterances containing the Embedded Language islands in the Thai/English data, as shown in Examples (47) – (52). The inflectional morphemes, such as articles, subject-verb agreement morphemes, or tense morphemes, are not found in the Thai structure. If these morphemes ever appear in the Thai Matrix Language frame, they always occur as a part of the Embedded Language islands, as in Examples (48) and (49).

As for the system morpheme, the System Morpheme Principle is supported by the Thai/English data. It can account for attaching an Embedded Language early system morpheme to a Thai base form. See the examples below.

\[(57) \quad [\text{soI} \text{ pass} \text{ fork-s} \text{ ma si}]\text{cp}\]
\[\text{‘pass the forks’}\]

\[(58) \quad [\text{raw ko} \text{ re-fa} \text{ daj}]\text{cp}\]
\[1\text{Pro then come re-listen can}\]
\[\text{‘I can listen to it again’}\]

This principle suggests that the late outsider system morpheme has to come from the Matrix Language; other system morphemes are not under this obligation.

Finally, although Example (60) seems to run counter to the Matrix Language Frame model, particularly the Morpheme Order Principle, I would say that the Matrix Language Frame model is still intact at this point. The small number of examples (three examples) containing English adverbs is a limitation of the present study. This can very well be a focus of future study.
6.2 EVALUATING THE SOCIOLINGUISTIC FRAMEWORKS

This section summarizes existing criticisms of the markedness model and the conversational codeswitching approach. It also addresses arguments and modifications made to each approach. This section concludes with an evaluation of the sociolinguistic frameworks.

6.2.1 Criticisms of the Markedness Model

The markedness model is a model using socio-psychological motivations to account for speakers engaging in the act of codeswitching. Community norms are the starting-point for this model. The markedness model is speaker-oriented in that speakers are the creative rational actors making code choices by weighing costs and rewards and associating codes with the RO set within a normative framework specific to their community.

The markedness model claims that speakers have a sense of markedness regarding available linguistic codes. That is, speakers interpret all code choices in terms of their markedness. Myers-Scotton (1993b, 2002) states that, in addition to linguistic competence, speakers have a 'markedness evaluator' which is an innate (hence universal) cognitive structure, enabling them to assess or index all available linguistic choices as more or less marked or unmarked. However, the ability to assess or index all the choices in terms of their markedness is developed through social experience in interactions in a specific community. Therefore, even though the evaluator itself is universal, the code choices can be interpreted in terms of their markedness only in reference to a specific speech event in a specific community.
In addition, the markedness model views code choices as a system of oppositions: marked and unmarked. They fall along a gradient continuum as more or less unmarked. A marked choice is an unexpected or most unusual choice. An unmarked choice is a most expected one, indexing an expected interpersonal relationship. When conversing, speakers assess the potential costs and rewards of all available choices and make a decision. Speakers choose the linguistic codes based on the persona and relationships with their interlocutors which they wish to be salient. Since the unmarked is the neutral and safer choice, speakers usually choose this code. However, they sometimes choose the marked choice for certain communicative effects. In other words, speakers are considered to be active and rational actors. They act, or choose their code, with goals to minimize costs and maximize rewards.

According to this framework, code choices are understood as indexing RO sets between participants in a given interaction type. The RO set is an abstract construct. It exemplifies the speakers’ attitudes and expectations toward one another. The act of codeswitching is interpreted as negotiations of these RO sets, which derived from situational factors salient in a certain interaction type in a community. This establishes a connection between code choices and negotiations of situational-factor salience. While conversing, speakers exploit the indexical property of code choices to negotiate their identities and expectations, or their RO, with one another. Similar to many macro-societal models, the markedness model sees individuals’ language choice as being determined or constrained by macro-societal structures or norms. This perspective contrasts sharply with that of the micro-interactional one.
The markedness model has been criticized for being static and classificatory. Li (2002) commented that social-motivation-based theory is negligent in the interactional meaning of the communicative act of codeswitching.

In the recast of the markedness model, however, code choices are dynamic, and not static. Speakers are the creative rational actors such that code choices are speaker-motivated, and thus more directly associated with interpersonal relationships than the situational frame. This model can now explain all variations, both marked and unmarked choices; it is by no means an equilibrium model in that it provides an explanation for the different choices peer speakers make.

Furthermore, the amended markedness model attempts to link the social symbolism of languages on the one hand with conversational strategies of individual speakers on the other. The centerpiece of the model is a “negotiation principle” which directs speakers to choose a certain linguistic form over another by taking into consideration the RO set, costs and rewards that come with that particular code. In the recast of the markedness model, Myers-Scotton integrates the rational choice model with her original model. The amended model is ultimately speaker-based; however, it also recognizes the importance of addressees’ responses. That is, the speaker’s choice depends on such response for their success in communication—the view congruent to that of the interactional sociolinguists.

In addition, Myers-Scotton (1993b) addresses motivations for the markedness model as generated from a variety of disciplines, including sociology of language, pragmatics, social anthropology, and linguistic anthropology. According to Myers-Scotton (1993b:75), “[a] common thread in all of these approaches is that participants
'know' (at some level) that they enter into conversation with similar expectations, whether about unmarked code choices or about unmarked communicative intentions.”

These approaches also assert that speaker is a creative actor and that linguistic choices convey more than referential meaning.

All these models, including the [markedness model], rely on the notion of INTENTIONALITY in human actions; actors intend their actions to reflect goals or attitudes, and observers attribute intentions to actions. They also give at least a nod toward the notion that innate architectures coordinate readings of cost-benefit analyses of competing choices. In interpersonal contexts, such architectures can also be seen to coordinate readings of intentionality.

(Myers-Scotton 2001:12, original emphasis)

Myers-Scotton has clearly illustrated how linguistic choices are primarily driven by rationality, rather than social group membership, ethnographic background, or sequential organization of a conversation.

Unlike the distinction between high and low linguistic varieties (diglossia) in Ferguson 1959, the rational choice model does not correlate linguistic choices with activity types. Dissimilar to the interactional sociolinguists, it does not make a correlation between code choices and the sequential organization of the conversation. Rather, it makes a correlation between linguistic choices and the agent’s intent to act rationally based on available evidence.

Another drawback of the markedness model, however, lies in its interpretations of examples. An example of a conversation between a gatekeeper in Nairobi and a Luyia visitor reflects a problematic interpretation. The example is recapped below:

GUARD: Unataka kumwona nani?
‘Whom do you want to see?’

VISITOR: Ningependa kumwona Solomom I—.
‘I would like to see Solomon I—.’

‘Do you really know him? We have a Solomon A—. I think that’s the one [you mean].

151
VISITOR: Yule anayetoka Tiriki—yaani Mluyia.
   ‘That one who comes from Tiriki—that is, a Luyia person.’
GUARD: → Solomon menyu wakhumanya valuhi?
   ‘Will Solomon know you?’
VISITOR: Yivi mulole umuvolere ndi Shem L— venyanga khukhulola.
   ‘You see him and tell him Shem L— wants to see you.’
GUARD: Yikhala yalia ulindi.
   ‘Sit here and wait.’
ANOTHER VISITOR (just appearing): Bwana K— yuko hapa?
   ‘Is Mr. K here?’
   ‘Yes, he’s here—he is doing something right now. He can’t leave until he finishes. Therefore you will wait here until he comes. You will wait about five or ten minutes.’
   (Guard goes to look for Solomon A—.)

(Myers-Scotton, 1988:153-154)

Myers-Scotton stated that the gatekeeper switches to Luyia once he knows that the visitor is a Luyia person. However, the example itself does not show how the gatekeeper comes to know that the visitor shares the same ethnic background as his, except that the visitor switches to Luyia in the following turn. It is possible that Myers-Scotton did not provide adequate information in terms of the context of her examples—the visitor may have an accent of a Luyia person. Additionally, Myers-Scotton categorized this example as the “sequential unmarked choice.” However, I found that she categorized another conversation between Luyia gatekeeper and a visitor who can speak Maragoli, a Luyia variety, in another example as “codeswitching as a marked choice” to narrow social distance (Myers-Scotton 1988:169):

Gatekeeper: (To young woman stopped in the middle of the gate)
(Swahili): Ingia kwa mlango mmoja tu. (‘Enter by using only one gate.’)
Young woman (Looks behind her and sees another car pulled up so that she cannot move easily)
(Swahili): Fungua miwili. Siwezi kwenda revas! Kuna magari mengine nyuma.
   (‘Open both. I can’t reverse! There are other cars behind me.’)
(Seeing the situation, the gatekeeper very grudgingly opens both gates.)
Young woman (Driving by the gatekeeper, she says to him) (Swahili): Mbona wewe mbaya sama leo? (‘Why are you so difficult today?’)
(She says to her companions in the car – in English, ‘The man is a Luyia.’ She determines this by his pronunciation.)
(Several hours later, she drives through the gate as she leaves.)
Young woman (to gatekeeper) (Maragoli, a Luyia variety): Undindiyange vutwa. (‘You were being unkind to me.’)
Gatekeeper (Swahili; Maragoli): Pole, simbere nikhumany ta. (‘Sorry, I didn’t know it was you.’)

The setting of this example is an athletic club. According to Myers-Scotton (1988), the unmarked choice in this context is Swahili. Nevertheless, the switch to a Luyia variety is considered “marked” in this example. Apparently, there is a problem regarding the interpretations as well as the identification of marked and unmarked choices of her examples. To my knowledge, Myers-Scotton has not provided any remedy to such imperfections yet.

6.2.2 Criticisms of the Conversational Codeswitching Approach

The study of “conversational codeswitching” was developed by Gumperz (1982). This interactional/interpretative model focuses on the micro-analysis of naturally occurring data in small-group interactions. The premise underlying this interpretative framework is that codeswitching signals contextual information. There is an explicit link between monolingual stylistic cues—such as selection of prosody, intonation, and gesture—and codeswitching: both of them serve as a contextualization cue in a conversation. As such, when speakers engage in the act of codeswitching, they aspire to produce their own intentional meanings, rather than merely respond to a predetermined set of prescriptions. For interactional sociolinguists, context is not a priori factor that influences or determines linguistic variation: rather it is shaped and maintained by participants in a conversation.

Interactional sociolinguists suggested that to fully understand codeswitching practices, primary attention has to be paid to its local production in the emerging context.
They criticized other large-scale sociolinguistic models for failing to account for such local production and restricting their analysis to associating the social meaning of codeswitching phenomenon with reference to a given a priori context, e.g., participants, topic, and setting.

Even though codeswitching can function like a contextualization cue, it has a characteristic of its own that sets it apart from other cues like prosodic cues, gestural cues, and so on. The sequential organization of code choices provides a referential frame for its interpretation in terms of both meanings and functions. Auer (1995) made a distinction between discourse-related switching and participant-related switching according to sequential patterns of language choice found in his data.

The analytic approach in conversational codeswitching approach requires two things: (1) the detailed transcription of naturally occurring data, such details include things like pauses, overlaps, latches, and so on, and (2) the interpretation of any relevant social meanings demonstrated by participants in the sequential organization of the actual conversation. A closer look at conversational sequences of language negotiation can reveal a number of structural features.

The two advantages of the conversational codeswitching approach are: (1) it gives priority to speakers' linguistic choices in the turn-taking process, thereby showing the effect of choosing one code over another in the sequential organization of a conversation, and (2) it limits the scope of the interpretation of code choices essentially to the mutual understanding between the co-participants of a conversation within the sequential organization (Auer 1995).
Myers-Scotton (1993b) agreed with Gumperz that social meanings are locally negotiated. Her markedness model also exploits the idea of individual negotiation. However, in her view, it is the positions in RO balances that are being negotiated, not the communicative intention of one code choice over another, as suggested by Gumperz. Additionally, if social meanings are generated in individual interactions, what could be the explanation for the empirical fact that speakers of the same community interpret the same interaction as having the same social intention?

Myers-Scotton (1993b) and Myers-Scotton and Bolonyai (2001) further added that ultimately, data must be explained in relation to other data sets. This means there are relations among interactions, accordingly predictions can be made about unexamined data. If data are primarily individualized and interactions operate on their own terms, no predictions can be made. Conversational codeswitching, thus, has no predictive value. Generally, she stated that Li and other interactional sociolinguists have provided a thorough interpretative description in a turn-taking process in their Chinese/English data; however, they do not pay adequate attention as to why and how speakers make the actual choices, especially when those choices diverge from the expected. She also added that the conversational approach has provided a discourse device in signaling and interpreting speaker’s intentions. She commented, however, that this approach has drawbacks: this approach is primarily descriptive; it does not offer any prediction or explanation. The approach is also questionable to the extent that social meaning is locally constructed, since it is an empirical fact that speakers can effectively interpret the same interaction as conveying the same social intention. That is, speakers share a common social backdrop which helps them determine social intention. The conversational codeswitching approach
is also criticized for failing to recognize the essence of the wider social context: it fails to recognize the socio-psychological associations between code choices and speakers; accordingly, it fails to recognize social messages carried by one linguistic choice rather than another. Speaker motivations are also downgraded in this approach. The conversational codeswitching approach looks at codeswitching merely as a device to structure the interaction, neglecting the motivation for choosing one code over another and other socio-psychological messages.

Li (2002) argued that interactional sociolinguists do not deny the existence of social variables. Rather, they do not assume the existence of the variables. Interactional sociolinguists are interested in demonstrating how such categories are employed in the actual sequential patterns of a conversation. In his own words (Li 2002:163):

At the heart of the difference between CA and other sociological perspectives is a tension between language as a medium for the expression of intentions, motives, or interests, and language as a topic for uncovering the methods through which ordered activity is generated – the latter being the CA position.

Interactional sociolinguists argue that even though their unit of analysis is at the conversational level, it does not imply that macroscopic perspectives of such a phenomenon are irrelevant. Codeswitching certainly serves a conversational function, but at the same time it has a connection with macro-social facts. As a matter of fact, Li emphasized the need to integrate the macro-societal approach, the micro-interactional approach, and the social network approach in his 1994 publication; nevertheless, he failed to demonstrate such integration in the interpretation of his Chinese/English data (Canagarajah 1995).38

---

38 In addition, Canagarajah (1995) criticized Li for not addressing larger political and sociological implications of his Chinese/English data. Li's study is limited in that his conversational data come solely from intra-family interactions. The data is from only one domain (out of many different domains addressed by Fishman, e.g., work, school, friendship, religion). Li is also criticized for mainly performing a sequential
Li (2002) suggested that analysts, especially outsider-analysts, must be extremely careful in assigning social meanings to individual instances of codeswitching, for languages in different communities have a different historical development and political status. For that reason, speakers may acquire the social meanings of different codes for different purposes and reasons as well as at different rates. He further argued that interactional sociolinguists have a different concept of evidence. Their evidence is based on the interpretative and interactional procedures of conversation participants in real time. This is why the detailed transcription conventions (criticized as an obsession) are necessary in this type of analysis.

According to Li (2002:163), “[t]he ways in which conversation participants design and modify their utterances are “naturally occurring statements” made by the social actors themselves of how they make sense of each other’s contributions. This tacit, organized reasoning procedure is critical for our understanding of how social relationships are developed and higher-level social orders achieved.” This point is well taken. However, I argue that to be able to make sense of or interpret the meanings of one’s utterances, one needs to be exposed to community norms and to learn how to interpret any given utterances according to the normative orientation of that community. Primarily, people use language to communicate. When conversing, speakers have an intention or a goal to deliver messages to their interlocutors. In addition, making one word choice over another “connotes” different meanings, feelings, or ideas. The fact that speakers know the intended meaning is evidence that they have tacit knowledge of language use and normative orientation of that community. Speakers are not simply

---

analysis of his data, leaving out some traditional analyses concerning social functions of codeswitching addressed by Gumperz (1982). That is, he does not show the social meanings of the switches. The interpretation of his Chinese/English data is also problematic (see Li 1996).
passive participants who respond solely to what they hear in the turn-taking process. There are other speakers' internal factors and knowledge accumulated through experience and exposure to conversational discourse in a community. My point is that this tacit knowledge of community norms comes into play before speakers make a choice, not the other way around. Community norms impinge upon the choices speakers make; in return, the choices speakers make diachronically impinge upon community norms as well.

Furthermore, the conversational codeswitching approach considers social meanings of language use as a function of the situated contexts or the dynamics of interactions. This contrasts sharply with other large-scale studies considering linguistic variation, e.g., individual choices in particular, as derived from sociological attributes of the speaker and the situation. To explicate, conversational codeswitching views the speaker merely as a participant in an ongoing interaction, not as an identity-bearing individual. Accordingly, the conversational codeswitching approach is criticized for its inattention to the salience of sociological variables (Myers-Scotton 1993b). It neglects the existence of societal norms.

The amended markedness model "explicitly recognizes that a current speaker's choice depends on the addressee's response (on many different levels) for its 'success'" (Myers-Scotton and Bolonyai 2001:4). However, there is too much of a gap between the macro-level entities and individual choices in the conversational codeswitching approach to bilingual speech.

It is clear that language is used to communicate much more than its referential meanings. There are also intentional meanings, e.g., when utterances are phrased differently, the addressees of such utterances can recognize not only the referential
meanings, but also the intentional meanings. That is, the intention itself is also communicated linguistically. In order to account for this communicative intention and the codeswitching phenomenon, an adequate model must provide an explanation of the mechanism by which speakers interpret the intention of utterances or how code choices can convey such intentional information.

It is not possible for speakers to assume that codeswitched utterances contain communicative intentions, and it is not possible for them to interpret the utterances in terms of their intentional meanings unless there is a normative framework with which the interpretation of communicative intentions is associated. It is true that the social meaning of choices develops partly in the turn-taking process of a conversation; however, the social meanings of choices originate in the normative framework and are largely associated with societal norms. Even though contextualization cues offer a useful interactional/interpretative description of codeswitching, they do not explain codeswitching per se.

With regards to Thai/English data, the conversational codeswitching approach can account only for the data that is meaningful in the turn-by-turn analysis. Sociolinguistic background information regarding the Thai participants and their attitudes towards Thai/English codeswitching shows that the Thai participants associate the act of codeswitching with one’s trying to show off or display his/her accomplishment. As a result, the Thai participants are quite conscious about whom they codeswitch with and whether to codeswitch at all. The present study agrees with the conversational codeswitching approach to the extent that the analysis of sequential organization of the Thai/English data can reveal certain conversational functions, e.g., codeswitching is used
to build up a contrast to attract attention from the interlocutors and to mark dispreferred
seconds. However, the approach cannot account for the Thai/English codeswitching data
as a whole. Most of the Thai/English data contain examples of intrasentential
codeswitching. Accordingly, not all instances of the Thai/English codeswitching data are
accountable under the conversational codeswitching approach.

The amended markedness model, on the other hand, can account for all the
Thai/English codeswitching data in the present study. The model is more powerful than
the conversational codeswitching approach. It can both describe and explain the
codeswitching pattern. The markedness model provides an explanation of the mechanism
by which speakers interpret the intention of utterances and how code choices convey such
intentional information. It is also more interdisciplinary and multi-faceted than the
conversational codeswitching approach. There is a correlation between the premises
underlying the markedness model and theories developed in other fields, i.e. social
sciences, law, and neurobiology (Myers-Scotton 1998, Myers-Scotton and Bolonyai
2001).

The findings in the present study are in favor of the markedness model;
nevertheless, I am not claiming that the model can account for all other language pairs.
Different codeswitching situations involve different linguistic and historical factors.
Heller (1988) and Nishimura (1997) concur that sociolinguistic analyses of
codeswitching differ from one another due to the historical background of each
community, which plays an important role in determining the characteristics of
codeswitching. To explicate, the diglossia phenomenon described in Ferguson 1959 is a
good example of situational codeswitching. This particular phenomenon exemplifies the
use of metaphorical codeswitching to achieve special effect, as well as the absence of conversational codeswitching, for the use of two languages in the same conversational interaction is unlikely to occur in the diglossia situation.

Economic as well as political power relations also explain different patterns of codeswitching observed in local communities. Speakers may resist using a certain code if it is stigmatized in that community, or may even resist codeswitching as a whole. In Transylvanian villages in Romania, where German and Romanian are spoken, these two linguistic codes are rarely mixed or switched. The German-speaking minority group uses Romanian only when telling a joke, an activity considered non-respectable by Transylvanian Germans. This particular group tries to keep its own language, German, separate from another language, Romanian (McClure and McClure 1988). In contrast, Hungarians in Austria abandon their own language, Hungarian, in favor of the more prestigious linguistic code, German. Gal (1988) adds that codeswitching signifies speakers' status in certain situations, i.e., Italian children living in Germany. In this particular situation, rapid codeswitching (Poplack 1980) comes into play. Mixing two languages together within a sentence boundary helps smoothly facilitate conversational interactions between young speakers with different linguistic competencies (Gal 1988). Additionally, as shown in Myers-Scotton's African data, the markedness model can capture the effects of the African socio-political context through the use of codeswitching to reflect the identities of the Kenyan speakers (Myers-Scotton 1993b). Chinese/English codeswitching among the three generations in Li (1994) can be explained in terms of conversational codeswitching approach. Whether or not the amended markedness model can account for all other language pairs is open for future study.
6.3 CONCLUSION

The present study is in favor of the markedness model as well as the Matrix Language Frame model. Due to the ethnographic background information regarding Thai society and the Thai people, the Thai/English data are primarily comprised of codeswitching within a sentence boundary. Codeswitching at the sentence level is rare, approximately twice in an hour.\(^{39}\) Even though the act of codeswitching is reported to be unmarked by the participants, however, their frequency of codeswitching is quite low, approximately 2.5 English words to every 100 Thai words. The participants mainly switch when conversing with their close friends. There are a few times when they codeswitch with other bilingual Thai/English speakers who are not their close friends. This type of switching is infrequent. Also, it only occurs for special effects, e.g. to attract attention. Given the linguistic purification movement and the association of codeswitching with "showing off", the participants are quite selective about when and with whom to codeswitch. The Thai/English codeswitching examples in the reported speech show that the participants choose to codeswitch when it is necessary. They do not switch to the original language of the quoted utterances, except when relating themselves with the voice of authority or distancing themselves from derogatory comments about others.

Additionally, the findings in this study are partly in compliance with and partly against the claims by Poplack (1980). The present study agrees that using the conversational codeswitching approach, conversational functions cannot always be assigned to the switching data, especially intrasentential data. The conversational codeswitching approach can account for switching at turn boundary and sentence

\(^{39}\) This is the count of intersentential switches when only Thai participants are present.
boundary, but not elsewhere. As mentioned earlier, Li (1994) said that Myers-Scotton mainly focuses on intrasentential switching. The present study argues that the amended markedness model can provide an adequate account of the Thai/English codeswitching data at all levels. Poplack (1980) further said that this intrasentential switching need not be explained in terms of any social motivations, because the switching itself is a part of the repertoire of a speech community. According to her study, “situational motivations or consequences to specific intra-sentential switches ... [have] little if any pertinence for the speakers themselves” (Poplack 1980:614). The present study argues that many codeswitching studies (Myers-Scotton 1993b, Auer 1998, and Li 1994, 2002), including the present one, have shown that situational motivations and consequences to switches play a key role in determining codeswitching.

Li (1994) has attempted to integrate the concept of social network to the conversational codeswitching approach in order to explain individual as well as deviating code choices. Poplack (1980), Al-Khatib (2003), and Mayor (2004) suggest that the concept of social network has relatively little significance to the participants in their studies. Their studies have shown that speakers make their choices as individuals according to their goals in communicating with the interlocutors. This subsequently lends support to Myers-Scotton's models. Even though having speakers' intentions as machinery is very questionable if one cannot ever empirically demonstrate how intention is implemented in conversational utterances, Myers-Scotton's models are by far the most complete and promising ones available.

In terms of the syntactic aspects of the Thai/English data, the Matrix Language Frame model can generally account for the Thai/English codeswitching data, except for
one example. Its premises regarding the bilingual language production also coincide with those of other researchers (Costa et al. 2000, Grosjean 2001). In previous literature, many researchers write about the activation of the bilingual's language repertoires and language processing mechanisms, Weinreich (1953), Hasselmo (1970), Clyne (1972), and Baetens Beardsmore (1982), and Grosjean (1997, 2001). They refer to the bilingual language processing as "language modes," which is a continuum of two available languages ranging from the monolingual mode to the bilingual mode (Grosjean 1997). When speaking to monolingual or when in the monolingual mode, a bilingual would deactivate one language and be limited in term of forms from the other language. In the monolingual mode, only one language is used. On the other hand, in the bilingual mode, they would choose the base language and also have the other language ready to use in the forms of codeswitching or borrowing. This confirms the distinction between Matrix Language and the Embedded Language in the Matrix Language Frame model.

According to Grosjean (2001), any number of factors can influence a speaker or listener at one point or another on the language mode continuum. These factors have an effect on the adjustable level of activation of the speaker’s language repertoires and language processing mechanisms. These factors include (1) “participant (s)” including addressee, listener and other related factors such as language proficiency, language mixing habits and attitudes, socioeconomic status, kinship relationship, and etc., (2) “situation” including location, degree of formality, (3) “form and content of the message” including topic, languages used, amount of mixed language, (4) “function of the language act” such as to inform, to request, etc., and (5) “specific research factors” such as organization of the task used, or the aim of the research. According to Grosjean (2001),
“movement along the continuum, which can happen at any given point in time depending on factors mentioned above, is usually an unconscious behavior that takes place smoothly and effortlessly. It is probably akin to changing speech style or register based on the context and the interlocutor.”

Additionally, there are two independent factors underlying this concept: (1) “the base language chosen” and (2) “the comparative level of activation of the two languages” which falls on a continuum from monolingual to bilingual language mode (Grosjean 2001:4). “Independent factors” means that one factor can change without resulting in a change in the other factor. For example, in the bilingual mode, the base language can change from one language to the other due to a change of topic, while the comparative level of activation stays unchanged. Similarly, the comparative level of activation can change from bilingual mode to monolingual mode without a change of base language. The base language is also the language that governs language processing (Grosjean 2001). This gives support to the Matrix Language Principle; such that the Matrix Language can change in a conversation and that the Matrix Language is the language that supplies the morphosyntactic frame for the Embedded Language constituents.

As for the speakers’ intentions, it is worth pointing out that not only do speakers’ intentions play a key role in the amended markedness model, but are also influential in the Matrix Language Frame model. It seems that speakers’ intentions can be used as a convenient tool to explain any data. When there is an unexpected codeswitching construction, analysts can resort to “speakers’ intentions” as long as the construction itself does not violate any hypothesis or principle. The speaker’s intention, however, is something that is neither empirical nor falsifiable. Also, whether or not human action is
the outcome of conscious calculation is still under debate in sociology (Boztepe on-line). In my opinion, this is a huge drawback of the models. The question is, How can analysts prove what the speaker’s intention is? From a sociolinguistic perspective, the fine-grained sequential turn-by-turn analysis may be used to account for the speaker’s intentions in any actual conversation by considering the interlocutor’s reaction to what is said in the previous turn taken by the speaker. The conversational codeswitching approach is, therefore, superior to the markedness model in that aspect. However, the approach does not give a complete account of the Thai/English codeswitching data. It can only account for the data that are meaningful in the turn transition. In other words, the approach cannot provide a complete description of or explanation for the codeswitching data.

All in all, Myers-Scotton’s approach has contributed a great deal to the contemporary study of codeswitching. At the macro level, it has connected sociolinguistics, social science, psycholinguistics, and syntax together. At the micro level, the model can illustrate how negotiations of RO sets are carried out by the participants in a conversation. More importantly, the model implies that community norms impinge upon the choices speakers make; in return, the choices speakers make diachronically impinge on community norms as well, the implication that interrelates the macro and the micro level. In other words, codeswitching represents a site in which micro-approaches interrelate and intersect with macro-approaches.
6.4 LIMITATIONS OF THE STUDY AND SUGGESTIONS FOR FURTHER STUDIES

Some limitations in the present study include (1) that the focus group of the present study is mainly the first generation of Thai speakers, and (2) that a small quantity of Thai/English codeswitching data contains adverbs and other system morphemes. The present study has raised some problems for future investigation. For future research:

(1) It will be interesting to observe Thai/English codeswitching data across generations and do a comparative analysis of codeswitching patterns. This will give a test to the social network, particularly the typology of immigrant communities (Li 1994). A comparative analysis between Thai/English codeswitching practiced by Thai speakers in an English-speaking country and in Thailand will also be interesting.

(2) It will be compelling to empirically investigate codeswitching from a diachronic perspective and see how community norms shape and are reshaped by code choices.

(3) Another project that should be done is construction of a Thai/English codeswitching corpus. This will be beneficial for both sociolinguistic and syntactic studies of codeswitching. If the corpus contains conversational details, e.g., pause, it will be usable in conversational codeswitching approach as well.

(4) As mentioned earlier, there is no agreement as to how monolingual or bilingual language production actually works yet. This type of study should
be encouraged as it will provide an insight and possibly the answers to syntactic aspects of codeswitching study.
APPENDIX A

A LIST OF QUESTIONS USED IN THE INTERVIEWS

The following are the questions used during the interviews. There are sixteen questions altogether:
1. What do you think about code-switching in general?
2. What are your attitudes toward Thai people who code-switch?
3. Does it make any difference if they have native-like or non-native like pronunciation?
4. Would you feel differently if the people who code-switch are those who exposed themselves to English-speaking countries for some time?
5. How do you feel when you hear other foreigners code-switch, say Japanese speakers mix English with Japanese or French speakers mix English with French? Is it different from how you would feel if Thai people mix English with Thai? If so, why?
6. Is it appropriate for Thai people to code-switch? If yes, why and in what situations? If no, why and in what situations?
7. Do you think code-switching is also a way to “show one’s ostentation”?
8. Do you think the use of code-switching can tell something about the person who is doing it, like social status, socio-economic status, and educational background?
9. If you do, which one(s) of the three above is(are) reflected by the use of code-switching? Just one? Two? Or all three? Explain.
10. Why do people code-switch?
11. Do you think some people code-switch to gain some acceptance from others?
12. Do you think code-switching can make a person look more professional and educated? Explain.
13. Suppose that you have just come to Hawai‘i and you meet two students, one of whom does code-switching with good pronunciation and always code-switches with the right words at the right time, and the other of whom does not. What would you think of each of them? Would your attitude toward the one who code-switches be different from the one who does not? Explain.
14. Suppose there is a person who only does code-switching here in the state and does not in Thailand, and another person who code-switches all the time, no matter where s/he is. What do you think about each person or his or her practice of code-switching?
15. Do you code-switch? If yes, why and in what situations? If no, why?
16. If you do, with whom do you code-switch? Do you code-switch with everybody or do you choose to code-switch with some people and not to code-switch with others? Why? Explain and give examples.
APPENDIX B

INFORMATION REGARDING THAI POPULATION FROM U.S. CENSUS
BUREAU, CENSUS 2000

Data Set: Census 2000 Summary File 3 (SF 3) - Sample Data


<table>
<thead>
<tr>
<th></th>
<th>Hawaii</th>
<th>Hawaii County, Hawaii</th>
<th>Honolulu County, Hawaii</th>
<th>Kalawao County, Hawaii</th>
<th>Kauai County, Hawaii</th>
<th>Maui County, Hawaii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total:</td>
<td>212,229</td>
<td>15,208</td>
<td>168,246</td>
<td>30</td>
<td>7,574</td>
<td>21,171</td>
</tr>
<tr>
<td>Europe:</td>
<td>10,479</td>
<td>1,513</td>
<td>7,104</td>
<td>0</td>
<td>386</td>
<td>1,476</td>
</tr>
<tr>
<td>Northern Europe:</td>
<td>3,328</td>
<td>451</td>
<td>2,449</td>
<td>0</td>
<td>69</td>
<td>359</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2,155</td>
<td>363</td>
<td>1,445</td>
<td>0</td>
<td>56</td>
<td>291</td>
</tr>
<tr>
<td>Ireland</td>
<td>265</td>
<td>1</td>
<td>240</td>
<td>0</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>Sweden</td>
<td>505</td>
<td>48</td>
<td>444</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Other Northern Europe</td>
<td>403</td>
<td>39</td>
<td>320</td>
<td>0</td>
<td>2</td>
<td>42</td>
</tr>
<tr>
<td>Western Europe:</td>
<td>4,688</td>
<td>845</td>
<td>2,958</td>
<td>0</td>
<td>244</td>
<td>841</td>
</tr>
<tr>
<td>Austria</td>
<td>259</td>
<td>63</td>
<td>133</td>
<td>0</td>
<td>27</td>
<td>36</td>
</tr>
<tr>
<td>France</td>
<td>699</td>
<td>103</td>
<td>438</td>
<td>0</td>
<td>32</td>
<td>126</td>
</tr>
<tr>
<td>Germany</td>
<td>3,116</td>
<td>506</td>
<td>1,969</td>
<td>0</td>
<td>154</td>
<td>487</td>
</tr>
<tr>
<td>Netherlands</td>
<td>428</td>
<td>78</td>
<td>228</td>
<td>0</td>
<td>26</td>
<td>96</td>
</tr>
<tr>
<td>Other Western Europe</td>
<td>386</td>
<td>95</td>
<td>190</td>
<td>0</td>
<td>5</td>
<td>96</td>
</tr>
<tr>
<td>Southern Europe:</td>
<td>856</td>
<td>70</td>
<td>643</td>
<td>0</td>
<td>30</td>
<td>113</td>
</tr>
<tr>
<td>Greece</td>
<td>64</td>
<td>2</td>
<td>46</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Italy</td>
<td>398</td>
<td>27</td>
<td>279</td>
<td>0</td>
<td>12</td>
<td>80</td>
</tr>
<tr>
<td>Portugal</td>
<td>155</td>
<td>20</td>
<td>117</td>
<td>0</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Spain</td>
<td>239</td>
<td>21</td>
<td>201</td>
<td>0</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Other Southern Europe</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Eastern Europe:</td>
<td>1,404</td>
<td>147</td>
<td>1,051</td>
<td>0</td>
<td>43</td>
<td>163</td>
</tr>
<tr>
<td>Czechoslovakia</td>
<td>93</td>
<td>16</td>
<td>55</td>
<td>0</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>(includes Czech Republic and Slovakia)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>245</td>
<td>18</td>
<td>210</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Poland</td>
<td>291</td>
<td>36</td>
<td>226</td>
<td>0</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>Romania</td>
<td>117</td>
<td>12</td>
<td>84</td>
<td>0</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>Belarus</td>
<td>41</td>
<td>5</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Russia</td>
<td>204</td>
<td>11</td>
<td>144</td>
<td>0</td>
<td>6</td>
<td>43</td>
</tr>
<tr>
<td>Ukraine</td>
<td>89</td>
<td>0</td>
<td>75</td>
<td>0</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Bosnia and Herzegovina</td>
<td>23</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Yugoslavia</td>
<td>119</td>
<td>27</td>
<td>88</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Other Eastern Europe</td>
<td>182</td>
<td>22</td>
<td>122</td>
<td>0</td>
<td>4</td>
<td>34</td>
</tr>
<tr>
<td>Europe, n.e.c.</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Asia:</td>
<td>176,707</td>
<td>10,667</td>
<td>143,258</td>
<td>15</td>
<td>6,572</td>
<td>16,195</td>
</tr>
<tr>
<td>Eastern Asia:</td>
<td>58,399</td>
<td>2,505</td>
<td>53,689</td>
<td>0</td>
<td>516</td>
<td>1,689</td>
</tr>
<tr>
<td>China:</td>
<td>20,241</td>
<td>386</td>
<td>19,512</td>
<td>0</td>
<td>97</td>
<td>246</td>
</tr>
<tr>
<td>China, excluding Hong Kong and Taiwan</td>
<td>12,562</td>
<td>278</td>
<td>12,173</td>
<td>0</td>
<td>71</td>
<td>140</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>4,596</td>
<td>34</td>
<td>4,450</td>
<td>0</td>
<td>24</td>
<td>76</td>
</tr>
<tr>
<td>Taiwan</td>
<td>2,933</td>
<td>74</td>
<td>2,879</td>
<td>0</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>Japan</td>
<td>20,590</td>
<td>1,480</td>
<td>18,089</td>
<td>0</td>
<td>272</td>
<td>749</td>
</tr>
<tr>
<td>Korea</td>
<td>17,202</td>
<td>639</td>
<td>15,725</td>
<td>0</td>
<td>147</td>
<td>691</td>
</tr>
<tr>
<td>Other Eastern Asia</td>
<td>366</td>
<td>0</td>
<td>363</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

170
<table>
<thead>
<tr>
<th>Region</th>
<th>Country</th>
<th>Hawaii</th>
<th>Honolulu County, Hawaii</th>
<th>Kalawao County, Hawaii</th>
<th>Kauai County, Hawaii</th>
<th>Maui County, Hawaii</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Central Asia:</td>
<td>Afghanistan</td>
<td>1,240</td>
<td>106</td>
<td>979</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Bangladesh</td>
<td>65</td>
<td>5</td>
<td>60</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>India</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Iran</td>
<td>567</td>
<td>70</td>
<td>438</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Pakistan</td>
<td>229</td>
<td>27</td>
<td>134</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Other South Central Asia</td>
<td>256</td>
<td>4</td>
<td>224</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>South Eastern Asia:</td>
<td>Other South Eastern Asia</td>
<td>116,463</td>
<td>8,014</td>
<td>88,137</td>
<td>15</td>
<td>6,001</td>
</tr>
<tr>
<td></td>
<td>Cambodia</td>
<td>452</td>
<td>46</td>
<td>401</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Indonesia</td>
<td>711</td>
<td>90</td>
<td>539</td>
<td>0</td>
<td>13</td>
<td>69</td>
</tr>
<tr>
<td>Laos</td>
<td>1,592</td>
<td>24</td>
<td>1,544</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Malaysia</td>
<td>539</td>
<td>17</td>
<td>473</td>
<td>0</td>
<td>3</td>
<td>46</td>
</tr>
<tr>
<td>Philippines</td>
<td>102,063</td>
<td>7,389</td>
<td>74,908</td>
<td>15</td>
<td>5,911</td>
<td>13,840</td>
</tr>
<tr>
<td>Thailand</td>
<td>1,788</td>
<td>204</td>
<td>1,504</td>
<td>0</td>
<td>28</td>
<td>52</td>
</tr>
<tr>
<td>Vietnam</td>
<td>8,775</td>
<td>221</td>
<td>8,269</td>
<td>0</td>
<td>31</td>
<td>254</td>
</tr>
<tr>
<td>Other South Eastern Asia</td>
<td>543</td>
<td>23</td>
<td>499</td>
<td>0</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Western Asia:</td>
<td>Other Western Asia</td>
<td>532</td>
<td>42</td>
<td>392</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Iraq</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Israel</td>
<td>123</td>
<td>4</td>
<td>69</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Jordan</td>
<td>42</td>
<td>0</td>
<td>42</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Lebanon</td>
<td>91</td>
<td>10</td>
<td>59</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Syria</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Turkey</td>
<td>114</td>
<td>9</td>
<td>92</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Armenia</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other Western Asia</td>
<td>152</td>
<td>19</td>
<td>120</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Asia, n.e.c.</td>
<td>73</td>
<td>0</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Africa:</td>
<td>Africa, n.e.c.</td>
<td>1,040</td>
<td>119</td>
<td>747</td>
<td>0</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Eastern Africa:</td>
<td>355</td>
<td>47</td>
<td>277</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Ethiopia</td>
<td>37</td>
<td>14</td>
<td>23</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Other Eastern Africa</td>
<td>318</td>
<td>33</td>
<td>254</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Middle Africa</td>
<td>16</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Northern Africa:</td>
<td>294</td>
<td>45</td>
<td>195</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Egypt</td>
<td>123</td>
<td>5</td>
<td>118</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Other Northern Africa</td>
<td>171</td>
<td>40</td>
<td>77</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Southern Africa:</td>
<td>240</td>
<td>27</td>
<td>144</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>South Africa</td>
<td>213</td>
<td>21</td>
<td>123</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Other Southern Africa</td>
<td>27</td>
<td>6</td>
<td>21</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Western Africa:</td>
<td>86</td>
<td>0</td>
<td>78</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Ghana</td>
<td>38</td>
<td>0</td>
<td>38</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Nigeria</td>
<td>14</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Sierra Leone</td>
<td>11</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Other Western Africa</td>
<td>23</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Africa, n.e.c.</td>
<td>49</td>
<td>0</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oceania:</td>
<td>Oceania, n.e.c.</td>
<td>13,452</td>
<td>1,452</td>
<td>10,512</td>
<td>0</td>
<td>230</td>
</tr>
<tr>
<td></td>
<td>Australia and New Zealand Subregion:</td>
<td>1,234</td>
<td>153</td>
<td>899</td>
<td>0</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Australia</td>
<td>606</td>
<td>49</td>
<td>430</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Other Australian and New Zealand Subregion</td>
<td>628</td>
<td>104</td>
<td>469</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Melanesia</td>
<td>393</td>
<td>47</td>
<td>257</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Micronesia</td>
<td>5,846</td>
<td>969</td>
<td>4,318</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Polynesia</td>
<td>5,974</td>
<td>283</td>
<td>4,993</td>
<td>0</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Oceania, n.e.c.</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Americas:</td>
<td>Americas</td>
<td>10,551</td>
<td>1,457</td>
<td>6,625</td>
<td>15</td>
<td>337</td>
</tr>
<tr>
<td></td>
<td>Latin America:</td>
<td>6,788</td>
<td>879</td>
<td>4,338</td>
<td>0</td>
<td>135</td>
</tr>
<tr>
<td>Region</td>
<td>Hawaii County, Hawaii</td>
<td>Honolulu County, Hawaii</td>
<td>Kalawao County, Hawaii</td>
<td>Kauai County, Hawaii</td>
<td>Maui County, Hawaii</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------------</td>
<td>-------------------------</td>
<td>------------------------</td>
<td>----------------------</td>
<td>--------------------</td>
<td></td>
</tr>
<tr>
<td>Caribbean:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barbados</td>
<td>21</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Cuba</td>
<td>131</td>
<td>35</td>
<td>87</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>88</td>
<td>12</td>
<td>74</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Haiti</td>
<td>73</td>
<td>0</td>
<td>73</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Jamaica</td>
<td>418</td>
<td>43</td>
<td>336</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Trinidad and Tobago</td>
<td>176</td>
<td>18</td>
<td>154</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Other Caribbean</td>
<td>170</td>
<td>16</td>
<td>154</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Central America:</td>
<td>3,664</td>
<td>596</td>
<td>2,099</td>
<td>0</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>2,773</td>
<td>541</td>
<td>1,281</td>
<td>0</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Other Central America:</td>
<td>1,091</td>
<td>55</td>
<td>818</td>
<td>0</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Costa Rica</td>
<td>59</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>El Salvador</td>
<td>207</td>
<td>30</td>
<td>117</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>198</td>
<td>2</td>
<td>74</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Honduras</td>
<td>112</td>
<td>0</td>
<td>111</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nicaragua</td>
<td>161</td>
<td>23</td>
<td>131</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Panama</td>
<td>341</td>
<td>0</td>
<td>325</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Other Central America</td>
<td>13</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>South America:</td>
<td>1,847</td>
<td>159</td>
<td>1,340</td>
<td>0</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>128</td>
<td>0</td>
<td>66</td>
<td>0</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Bolivia</td>
<td>41</td>
<td>0</td>
<td>35</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>550</td>
<td>35</td>
<td>400</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Chile</td>
<td>126</td>
<td>43</td>
<td>76</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>238</td>
<td>39</td>
<td>197</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Ecuador</td>
<td>144</td>
<td>13</td>
<td>131</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Guyana</td>
<td>77</td>
<td>0</td>
<td>71</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Peru</td>
<td>347</td>
<td>29</td>
<td>248</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Venezuela</td>
<td>106</td>
<td>0</td>
<td>77</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Other South America</td>
<td>90</td>
<td>0</td>
<td>39</td>
<td>0</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Northern America:</td>
<td>3,763</td>
<td>578</td>
<td>2,287</td>
<td>15</td>
<td>202</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>3,743</td>
<td>578</td>
<td>2,267</td>
<td>15</td>
<td>202</td>
<td></td>
</tr>
<tr>
<td>Other Northern America</td>
<td>20</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Born at sea</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
## LANGUAGE SPOKEN AT HOME BY PERSONS FIVE YEARS OLD AND OVER, BY AGE: 2000

<table>
<thead>
<tr>
<th>Language spoken at home</th>
<th>Total</th>
<th>Less than 18 years</th>
<th>18 to 64 years</th>
<th>65 years and over</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speak English only</td>
<td>1,134,351</td>
<td>217,139</td>
<td>756,071</td>
<td>161,141</td>
</tr>
<tr>
<td>Other Pacific Islands 1/</td>
<td>832,226</td>
<td>177,581</td>
<td>556,067</td>
<td>98,578</td>
</tr>
<tr>
<td>Tagalog</td>
<td>90,111</td>
<td>16,323</td>
<td>61,478</td>
<td>12,310</td>
</tr>
<tr>
<td>Japanese</td>
<td>60,967</td>
<td>6,264</td>
<td>44,658</td>
<td>10,045</td>
</tr>
<tr>
<td>Chinese</td>
<td>56,225</td>
<td>3,999</td>
<td>25,005</td>
<td>27,221</td>
</tr>
<tr>
<td>Spanish and Spanish Creole</td>
<td>29,363</td>
<td>3,835</td>
<td>19,137</td>
<td>6,391</td>
</tr>
<tr>
<td>Korean</td>
<td>18,820</td>
<td>2,423</td>
<td>15,042</td>
<td>1,355</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>18,337</td>
<td>1,731</td>
<td>6,029</td>
<td>510</td>
</tr>
<tr>
<td>German</td>
<td>8,270</td>
<td>296</td>
<td>2,966</td>
<td>724</td>
</tr>
<tr>
<td>French (incl. Patois, Cajun)</td>
<td>3,310</td>
<td>330</td>
<td>2,517</td>
<td>463</td>
</tr>
<tr>
<td>Laotian</td>
<td>1,920</td>
<td>484</td>
<td>1,340</td>
<td>96</td>
</tr>
<tr>
<td>Thai</td>
<td>1,496</td>
<td>229</td>
<td>1,231</td>
<td>36</td>
</tr>
<tr>
<td>Other Indo-European 1/</td>
<td>1,288</td>
<td>234</td>
<td>878</td>
<td>176</td>
</tr>
<tr>
<td>Portuguese and Portuguese Creole</td>
<td>1,238</td>
<td>99</td>
<td>845</td>
<td>294</td>
</tr>
<tr>
<td>Italian</td>
<td>826</td>
<td>32</td>
<td>619</td>
<td>175</td>
</tr>
<tr>
<td>Arabic</td>
<td>729</td>
<td>124</td>
<td>560</td>
<td>45</td>
</tr>
<tr>
<td>Scandinavian</td>
<td>727</td>
<td>7</td>
<td>649</td>
<td>71</td>
</tr>
<tr>
<td>Other West Germanic</td>
<td>558</td>
<td>39</td>
<td>352</td>
<td>167</td>
</tr>
<tr>
<td>Russian</td>
<td>432</td>
<td>80</td>
<td>324</td>
<td>28</td>
</tr>
<tr>
<td>Other Indic 1/</td>
<td>421</td>
<td>46</td>
<td>342</td>
<td>33</td>
</tr>
<tr>
<td>African languages</td>
<td>359</td>
<td>43</td>
<td>292</td>
<td>24</td>
</tr>
<tr>
<td>Other Asian 1/</td>
<td>275</td>
<td>17</td>
<td>247</td>
<td>12</td>
</tr>
<tr>
<td>Polish</td>
<td>274</td>
<td>20</td>
<td>190</td>
<td>64</td>
</tr>
<tr>
<td>Hungarian</td>
<td>242</td>
<td>6</td>
<td>112</td>
<td>124</td>
</tr>
<tr>
<td>Greek</td>
<td>203</td>
<td>13</td>
<td>148</td>
<td>42</td>
</tr>
<tr>
<td>Hindi</td>
<td>200</td>
<td>18</td>
<td>174</td>
<td>8</td>
</tr>
<tr>
<td>Persian</td>
<td>192</td>
<td>15</td>
<td>153</td>
<td>24</td>
</tr>
<tr>
<td>Mon-Khmer, Cambodian</td>
<td>162</td>
<td>20</td>
<td>142</td>
<td>-</td>
</tr>
<tr>
<td>Other Native North American languages 1/</td>
<td>147</td>
<td>28</td>
<td>119</td>
<td>-</td>
</tr>
<tr>
<td>French Creole</td>
<td>128</td>
<td>7</td>
<td>100</td>
<td>21</td>
</tr>
<tr>
<td>Urdu</td>
<td>127</td>
<td>18</td>
<td>85</td>
<td>24</td>
</tr>
<tr>
<td>Other Slavic 1/</td>
<td>118</td>
<td>-</td>
<td>100</td>
<td>18</td>
</tr>
<tr>
<td>Hebrew</td>
<td>118</td>
<td>28</td>
<td>90</td>
<td>-</td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td>87</td>
<td>-</td>
<td>87</td>
<td>-</td>
</tr>
<tr>
<td>Armenian</td>
<td>59</td>
<td>-</td>
<td>43</td>
<td>16</td>
</tr>
<tr>
<td>Navajo</td>
<td>58</td>
<td>5</td>
<td>53</td>
<td>-</td>
</tr>
<tr>
<td>Yiddish</td>
<td>41</td>
<td>-</td>
<td>26</td>
<td>15</td>
</tr>
<tr>
<td>Miao, Hmong</td>
<td>30</td>
<td>-</td>
<td>30</td>
<td>-</td>
</tr>
<tr>
<td>Gujarathi</td>
<td>27</td>
<td>2</td>
<td>25</td>
<td>-</td>
</tr>
<tr>
<td>Other &amp; unspecified languages</td>
<td>253</td>
<td>24</td>
<td>196</td>
<td>33</td>
</tr>
</tbody>
</table>

1/ Excludes languages in the category which are specifically listed in this table.

Source: U.S. Census Bureau, Census 2000 Sample Data File, special tabulation by the Hawaii State Department of Business, Economic Development & Tourism, Hawaii State Data Center
### Table 13a. Hawaii – Ability to Speak English by Language Spoken at Home for the Population 5 Years and Over: 2000

<table>
<thead>
<tr>
<th>Language spoken at home</th>
<th>Total</th>
<th>Speak English</th>
<th>Speak English</th>
<th>Speak English</th>
<th>Speak English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>&quot;very well&quot;</td>
<td>&quot;well&quot;</td>
<td>&quot;not well&quot;</td>
<td>&quot;not at all&quot;</td>
</tr>
<tr>
<td>Population 5 years and over</td>
<td>1,134,350</td>
<td>(X) (X)</td>
<td>(X) (X)</td>
<td>(X) (X)</td>
<td>(X) (X)</td>
</tr>
<tr>
<td>Speak only English</td>
<td>832,225</td>
<td>(X) (X)</td>
<td>(X) (X)</td>
<td>(X) (X)</td>
<td>(X) (X)</td>
</tr>
<tr>
<td>Spanish or Spanish Creole</td>
<td>18,815</td>
<td>13,860</td>
<td>73.7</td>
<td>3,440</td>
<td>18.3</td>
</tr>
<tr>
<td>Other Indo-European languages</td>
<td>14,239</td>
<td>11,075</td>
<td>77.8</td>
<td>2,418</td>
<td>17.0</td>
</tr>
<tr>
<td>French (incl. Patois, Cajun)</td>
<td>3,310</td>
<td>2,565</td>
<td>77.5</td>
<td>485</td>
<td>14.7</td>
</tr>
<tr>
<td>French Creole</td>
<td>130</td>
<td>105</td>
<td>80.8</td>
<td>15</td>
<td>11.5</td>
</tr>
<tr>
<td>Italian</td>
<td>625</td>
<td>636</td>
<td>77.0</td>
<td>125</td>
<td>15.2</td>
</tr>
<tr>
<td>Portuguese or Portuguese Creole</td>
<td>1,240</td>
<td>920</td>
<td>74.2</td>
<td>270</td>
<td>21.8</td>
</tr>
<tr>
<td>German</td>
<td>3,065</td>
<td>3,270</td>
<td>82.1</td>
<td>565</td>
<td>14.2</td>
</tr>
<tr>
<td>Yiddish</td>
<td>40</td>
<td>40</td>
<td>100.0</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Other West Germanic languages</td>
<td>560</td>
<td>445</td>
<td>79.5</td>
<td>95</td>
<td>17.0</td>
</tr>
<tr>
<td>Scandinavian languages</td>
<td>725</td>
<td>620</td>
<td>85.5</td>
<td>90</td>
<td>12.4</td>
</tr>
<tr>
<td>Greek</td>
<td>200</td>
<td>135</td>
<td>67.5</td>
<td>45</td>
<td>22.5</td>
</tr>
<tr>
<td>Russian</td>
<td>430</td>
<td>255</td>
<td>59.3</td>
<td>150</td>
<td>34.9</td>
</tr>
<tr>
<td>Polish</td>
<td>269</td>
<td>210</td>
<td>78.1</td>
<td>45</td>
<td>16.7</td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td>69</td>
<td>70</td>
<td>78.7</td>
<td>4</td>
<td>4.5</td>
</tr>
<tr>
<td>Other Slavic languages</td>
<td>115</td>
<td>95</td>
<td>82.6</td>
<td>10</td>
<td>8.7</td>
</tr>
<tr>
<td>Armenian</td>
<td>59</td>
<td>40</td>
<td>67.8</td>
<td>15</td>
<td>25.4</td>
</tr>
<tr>
<td>Persian</td>
<td>195</td>
<td>125</td>
<td>64.1</td>
<td>70</td>
<td>35.9</td>
</tr>
<tr>
<td>Gujarathi</td>
<td>35</td>
<td>25</td>
<td>75.0</td>
<td>4</td>
<td>12.1</td>
</tr>
<tr>
<td>Hindi</td>
<td>199</td>
<td>175</td>
<td>87.9</td>
<td>20</td>
<td>10.1</td>
</tr>
<tr>
<td>Urdu</td>
<td>125</td>
<td>90</td>
<td>72.0</td>
<td>35</td>
<td>28.0</td>
</tr>
<tr>
<td>Other Indo-European languages</td>
<td>420</td>
<td>320</td>
<td>76.2</td>
<td>50</td>
<td>11.9</td>
</tr>
<tr>
<td>Other Indian languages</td>
<td>1,290</td>
<td>935</td>
<td>72.5</td>
<td>325</td>
<td>25.2</td>
</tr>
<tr>
<td>Asian and Pacific Island languages</td>
<td>267,144</td>
<td>132,375</td>
<td>49.6</td>
<td>80,805</td>
<td>30.2</td>
</tr>
<tr>
<td>Chinese</td>
<td>29,365</td>
<td>12,175</td>
<td>41.5</td>
<td>7,920</td>
<td>27.0</td>
</tr>
<tr>
<td>Japanese</td>
<td>56,220</td>
<td>29,215</td>
<td>52.0</td>
<td>16,865</td>
<td>30.0</td>
</tr>
<tr>
<td>Korean</td>
<td>16,335</td>
<td>6,795</td>
<td>41.5</td>
<td>5,710</td>
<td>35.1</td>
</tr>
<tr>
<td>Mon-Khmer, Cambodian</td>
<td>150</td>
<td>65</td>
<td>42.6</td>
<td>65</td>
<td>40.6</td>
</tr>
<tr>
<td>Miao, Hmong</td>
<td>30</td>
<td>20</td>
<td>66.7</td>
<td>10</td>
<td>33.3</td>
</tr>
<tr>
<td>Thai</td>
<td>1,494</td>
<td>725</td>
<td>48.5</td>
<td>515</td>
<td>34.5</td>
</tr>
<tr>
<td>Ladonian</td>
<td>1,915</td>
<td>810</td>
<td>42.3</td>
<td>420</td>
<td>21.9</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>8,270</td>
<td>2,260</td>
<td>27.9</td>
<td>2,675</td>
<td>32.4</td>
</tr>
<tr>
<td>Tagalog</td>
<td>60,970</td>
<td>30,715</td>
<td>50.4</td>
<td>20,810</td>
<td>34.1</td>
</tr>
<tr>
<td>Other Pacific Island languages</td>
<td>90,110</td>
<td>49,370</td>
<td>54.8</td>
<td>25,785</td>
<td>28.6</td>
</tr>
<tr>
<td>Other languages</td>
<td>1,911</td>
<td>1,310</td>
<td>68.6</td>
<td>414</td>
<td>21.7</td>
</tr>
<tr>
<td>Navajo</td>
<td>59</td>
<td>40</td>
<td>67.8</td>
<td>15</td>
<td>25.4</td>
</tr>
<tr>
<td>Other Native American languages</td>
<td>144</td>
<td>120</td>
<td>83.3</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>Hungarian</td>
<td>240</td>
<td>160</td>
<td>66.7</td>
<td>55</td>
<td>22.9</td>
</tr>
<tr>
<td>Arabic</td>
<td>725</td>
<td>480</td>
<td>66.2</td>
<td>190</td>
<td>26.2</td>
</tr>
<tr>
<td>Hebrew</td>
<td>118</td>
<td>90</td>
<td>76.6</td>
<td>25</td>
<td>21.0</td>
</tr>
<tr>
<td>African languages</td>
<td>365</td>
<td>215</td>
<td>58.9</td>
<td>85</td>
<td>23.3</td>
</tr>
<tr>
<td>Other and unspecified languages</td>
<td>269</td>
<td>205</td>
<td>79.2</td>
<td>40</td>
<td>15.4</td>
</tr>
</tbody>
</table>

(X) Not applicable.

Source: U.S. Census Bureau, Census 2000.

Internet Release Date: October 29, 2004.
APPENDIX C

THAI CONSONANTS AND VOWELS

Consonants:

<table>
<thead>
<tr>
<th>Manner of Articulation</th>
<th>Place of Articulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bilabial</td>
</tr>
<tr>
<td><strong>Stops</strong></td>
<td></td>
</tr>
<tr>
<td>Voiceless unaspirated</td>
<td>p</td>
</tr>
<tr>
<td>Voiceless aspirated</td>
<td>ph</td>
</tr>
<tr>
<td>Voiced unaspirated</td>
<td>b</td>
</tr>
<tr>
<td><strong>Affricates</strong></td>
<td></td>
</tr>
<tr>
<td>Unaspirated</td>
<td></td>
</tr>
<tr>
<td>Aspirated</td>
<td></td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
</tr>
<tr>
<td>Unaspirated</td>
<td>f</td>
</tr>
<tr>
<td>Aspirated</td>
<td></td>
</tr>
<tr>
<td><strong>Nasals</strong></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td></td>
</tr>
<tr>
<td>Tap</td>
<td>w</td>
</tr>
<tr>
<td>Semivowels</td>
<td>w</td>
</tr>
</tbody>
</table>

Vowels:

<table>
<thead>
<tr>
<th>Tongue Height</th>
<th>Position of the tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Front</td>
</tr>
<tr>
<td>High</td>
<td>i   :</td>
</tr>
<tr>
<td>Mid</td>
<td>e   :</td>
</tr>
<tr>
<td>Low</td>
<td>æ æ :</td>
</tr>
</tbody>
</table>

Diphthongs:

<table>
<thead>
<tr>
<th>Tongue height</th>
<th>Position of the tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lip-unrounded</td>
</tr>
<tr>
<td></td>
<td>Front</td>
</tr>
<tr>
<td>High</td>
<td>i</td>
</tr>
<tr>
<td>Low</td>
<td>æ æ</td>
</tr>
</tbody>
</table>

(adapted from Burusphat et al. 1999)
REFERENCES


Surawan, Porntipa. 1975. *The speech of Thai students in the U.S.A.: with reference to linguistic interference and variation.* Mater’s Thesis, Faculty of Graduate School, University of Texas at Austin.


Internet sources:


*US census*. Online: http://www.census.gov/population/cen2000

*The international student services*, University of Hawai‘i at Mānoa. Online: http://www.hawaii.edu/issmanoa