

The Second Language Acquisition of Mandarin Chinese Tones
by English, Japanese and Korean Speakers

Hang Zhang

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Linguistics.

Chapel Hill
2013

Approved by
Jennifer L. Smith
Elliott Moreton
Misha Becker
Katya Pertsova
Ji-Yeon Jo

© 2013
Hang Zhang
ALL RIGHTS RESERVED

ABSTRACT

HANG ZHANG: The Second Language Acquisition of Mandarin Chinese Tones by English, Japanese and Korean Speakers
(Under the direction of Jennifer L. Smith)

This dissertation explores the second language acquisition of Mandarin Chinese tones by speakers of non-tonal languages within the framework of Optimality Theory. The effects of three L1s are analyzed: American English, a stress-accent language; Tokyo Japanese, a lexical pitch accent language; and Seoul Korean, a non-stress and non-pitch accent language. The study tests for three possible sources of L2 tonal errors; namely, 1) universal phonological constraints (i.e. the Tonal Markedness Scale (TMS), the Obligatory Contour Principle (OCP), and Tone-Position Constraints (TPC)); 2) the transfer of L1 pitch patterns; and 3) a pedagogical problem of Tone 3. The data shows that these three factors jointly shape the properties of interlanguage grammars.

This study finds that the TMS, the OCP, and TPC constrain L2 tone acquisition, but do so to varying degrees. Evidence is found that the TMS applies to both word- and sentence-level L2 productions. Some effects of the OCP are found to interact with the TMS and with L1 transfer effects. For example, patterns regarding tone pairs (more T1-T1 productions than T4-T4, and in turn more than T2-T2) can be attributed to either a case of the “emergence of the unmarked” interacting effects of the TMS and the OCP, or to local conjunction of the TMS. Learners are better at maintaining Rising (T2) at word-initial positions, but Falling (T4) at word-final positions. L2 learners often substitute other tones for target tones and the substitution patterns provide evidence for L1 transfer. For example, English speakers often use high falling tone while Japanese speakers tend to lengthen low tones to express monosyllabic narrow focus in sentences. This study found conflicting error and substitution patterns pertaining to Tone 3, as well as greater accuracy in processing Pre-T3 sandhi than the sandhi occurring elsewhere. This effect is argued to be attributed to the “T3 [214]-First” teaching method.

In light of the three factors affecting L2 tone acquisition, this study proposes a constraint re-ranking model to provide a new way of viewing positive and negative transfer. It is demonstrated that some markedness constraints are promoted while some are demoted in the acquisition of tones.

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere thanks to my advisor, Dr. Jennifer Smith, for her inspiration, guidance, support, and patience. I cannot remember how many times Jen has rescued me and put me back on track. She is strict, but very understanding, approachable and always helpful. Jen is my role model for being a linguist, a mentor, and a teacher. Words fail to express my appreciation to her! I am exceedingly grateful to Dr. Elliott Moreton for his engaging discussions. His enthusiasm and love for linguistic research is contagious. I also greatly appreciate my other committee members, Dr. Katya Pertsova, Dr. Misha Becker, and Dr. Ji-Yeon Jo, for their great support and insightful suggestions.

This project could not have been completed without the assistance of many people. This study relies heavily on the statistical analysis guided by Chris Wiesen, and I truly appreciate all of his hard work and patience. Special thanks are also due to Jia Lin for data analysis; Azusa Saito, Katsuhiko Sawamura, Jeehwa Kim, Seul Bee Lee, and Insun Lee for translating subject-recruitment documents and experiment reading materials into Japanese and Korean; the teaching staff at Zhejiang University (Hangzhou, China); and the 67 participants who agreed to be a part of this study. I especially appreciate Hong Shi, Linghong Huang, and Feiyan Wang, whose enormous amount of help with the collection of pilot data in 2011 has made this study possible.

I don't think I could have found a better fit for my grad career than UNC-Chapel Hill. This project is supported by the Center for Global Initiative (Pre-Dissertation Travel Award) and Graduate School (Dissertation Completion Fellowship). I would like to thank the entire Linguistics Department. This includes Professor Emeritus Craig Melchert, who helped me start in the Linguistics Department. I also truly appreciate the insightful comments and invaluable help given by my fellow grad cohort in our

P-side group (a UNC-CH phonetics/phonology research group)—Mary Kohn, Lúcia Fischer, Emily Moeng, Jen Griffin, Anna Powers, Amy Reynolds, Katherine Shaw, Matt Fuller, Justin Pinta, and other grads. Thank you all!

I also owe a particular debt of gratitude to the Department of Asian Studies at UNC-CH for opportunities for professional and personal development during the past eight years. I am truly grateful to Dr. Gang Yue for the help he has given me to make my grad career at UNC possible. I also deeply appreciate Dr. Wendan Li's consistent support for my full-time teaching position in the Asian Studies Department and my research. I still think fondly of my time talking with her in New West and over the phone. I would also like to thank Dr. Jan Bardsley very much for her support to my trip to China for data collection, and all my colleagues in Asian Studies for their professional support and friendship.

I cannot end without thanking my family and friends. I dedicate this dissertation to my husband Jie Cai and my daughter Emma Xiaoman Cai. I have relied so much on their love and encouragement throughout my time conducting research. For so many weekends, Jie has taken Emma to libraries and playgrounds to allow me to concentrate on my research at home. I remember being five-months pregnant when I wrote the “acknowledgements” section for my Master's thesis and now Emma is five years old. I know how long you two have been waiting for the day you can officially call me Driver (Dr.) Zhang. I finally earned the license!

TABLE OF CONTENTS

LIST OF TABLES	xii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS AND SYMBOLS	xvi
Chapter	
1. INTRODUCTION	1
1.1. Overview	1
1.2. The significance of the study	3
1.3. The theoretical context	4
1.4. Research questions and general hypotheses	6
1.5. The organization of the dissertation	9
2. PHONOLOGICAL BACKGROUND AND PREDICTIONS	10
2.1. Optimality Theory and the second language acquisition of Mandarin tones.....	10
2.1.1. Some basic concepts of Optimality Theory	10
2.1.2. The OT grammars and Second Language Acquisition	15
2.2. The tone system of Mandarin Chinese	21
2.2.1. Mandarin tonal features and representation	21
2.2.2. Mandarin Tone Sandhi	25
2.2.2.1. Mandarin T3 Sandhi and the domain	25
2.2.2.2. T3 and T3 Sandhi Pedagogy	26
2.2.3. The perception of tones	28
2.2.4. Three phonological constraints relevant to Mandarin tone acquisition	31

2.2.4.1. The Tonal Markedness Scale (TMS)	32
2.2.4.2. Obligatory Contour Principle (OCP)	35
2.2.4.3. Tone-Position Constraints (TPC)	38
2.3. Assumptions of prosodic structures and the stimulus sentence	42
2.3.1. Assumptions of prosodic structures	43
2.3.2. Factors affecting prosodic phrasing	44
2.3.2.1. Syntactic structure	44
2.3.2.2. Focus	46
2.3.3. Prosodic structure of the Mandarin stimulus sentence	49
2.4. The prosodic characteristics of the three first languages and predictions concerning each.....	51
2.4.1. Representation of intonation	52
2.4.2. English prosodic structures and predictions.....	54
2.4.2.1. Prosodic phrasing of stimulus sentences for English speakers.....	56
2.4.2.2. Tone assignment	56
2.4.3. Tokyo Japanese (Japanese) prosodic structures and predictions	60
2.4.3.1. Japanese Speakers' Prosodic Phrasing of Stimulus sentence	61
2.4.3.2. The assignment of tones to AP	63
2.4.4. Seoul Korean (Korean) prosodic structures and predictions	68
2.4.4.1. Predicted prosodic phrasing for Korean speakers	69
2.4.4.2. The tone assignment of test words for Korean Speakers	70
2.4.5. Summary of L1 transfer predictions	72
3. METHODOLOGY	74
3.1. Test materials	74
3.2. Subjects and recording procedures	78
3.3. Judgment and data transcription	80

3.4. Data Analysis	83
4. WORD-LEVEL EXPERIMENT RESULTS	84
4.1. General information of L2 production errors	85
4.2. On the test of Tonal Markedness Scale (TMS)	89
4.2.1. The statistical analysis of TMS	89
4.2.2. Substitution patterns for individual tones	90
4.2.2.1. Substitution patterns without target tone information	91
4.2.2.2. Substitution patterns with target tone information	94
4.2.2.3. A discussion on the error and substitution patterns of Tone 3	96
4.2.3. Summary on TMS tests	99
4.3. The test of Obligatory Contour Principle (OCP)	100
4.3.1. OCP test 1	101
4.3.2. OCP test 2	105
4.3.3. Summary of OCP tests	108
4.4. Test of Tone-Position Constraints (TPC)	109
4.4.1. General positional effects and statistical test of TPC	109
4.4.2. Positional error rates details: substitution patterns	112
4.4.3. The correlation between error and substitution patterns for positional effects	115
4.4.4. Summary of TPC tests	119
4.5. The test of typical L1 pitch pattern transfer	120
4.5.1. Statistical test of L1-like pitch pattern transfer	120
4.5.2. On the most frequently produced tone combinations	123
4.6. Summary of Chapter 4	126
5. SENTENCE-LEVEL EXPERIMENT RESULTS	129
5.1. General error patterns and the TMS, OCP effects	131
5.1.1. General error patterns	131

5.1.2. Individual tone errors and Tonal Markedness Scale	133
5.1.3. OCP test in Experiment 2	138
5.1.4. Summary of 5.1	140
5.2. On the most vulnerable positions at sentence level	141
5.2.1. An overview of vulnerable positions in sentence level tonal productions	141
5.2.2. Error rates of syllables inside and outside of focused constituents	145
5.2.3. The error patterns in neutral constituents	147
5.3. Different types of L1 learners' expressions of focus	152
5.3.1. The monosyllabic focused constituents (F1)	153
5.3.2. Long focused constituents (F2-F5) and the disyllabic tone patterns	158
5.3.2.1. L1 transfer of phrasal level pitch patterns	158
5.3.2.2. The transfer of typical L1 disyllabic tone patterns	161
5.4. Summary of Chapter 5	168
6. A DISCUSSION OF THEORETICAL IMPLICATIONS	169
6.1. The constraint re-ranking model for second language acquisition (SLA)	170
6.1.1. The general structure of the conceptual SLA constraint re-ranking model	170
6.1.2. Constraints used in the analysis of the SLA of Mandarin tones	172
6.1.2.1. Markedness Constraints	173
6.1.2.2. Faithfulness Constraints	175
6.1.2.3. The OT treatment of tones and intonations	175
6.1.3. The promotion and demotion of Markedness constraints	177
6.1.3.1. The postulation of L1, L2, and interlanguage grammar rankings	177
6.1.3.2. Four types of Markedness constraints and the effects of constraint re-ranking	179
6.1.3.3. A discussion on constraint re-ranking in Second Language Acquisition	182
6.2. The interlanguage tonal grammars	183

6.2.1. The summary charts of response tone combination patterns	184
6.2.2. Representative interlanguage tonal grammars	186
6.2.2.1. Japanese speakers' mismatched pairs	187
6.2.2.2. English and Korean speakers' mismatched pairs.....	190
6.2.3. Summary of partial Interlanguage rankings	192
6.3. The “Emergence of the Unmarked” cases in the SLA of tones	194
6.3.1. About Markedness and “The Emergence of the Unmarked” (TETU)	195
6.3.2. A TETU case of interacting effects of TMS and OCP	197
6.3.3. A TETU case of interacting effects of TMS and TPC	201
6.4. Summary	204
7. CONCLUSIONS AND IMPLICATIONS	205
7.1. Summary of research findings	205
7.1.1. Three universal phonological constraints in the SLA of Mandarin tones	205
7.1.2. L1 transfer effects	208
7.1.3. The constraint re-ranking model	209
7.2. Pedagogical implications and future directions	212
7.2.1. The pedagogical problem of Tone 3.....	212
7.2.2. Other pedagogical implications regarding difficult tones in context	217
7.3. Concluding remarks	220
APPENDIX 1. Tokyo-type Accent map and Recruitment Criteria Regarding Language Background for Japanese Speaking Learners	223
APPENDIX 2. Sample Language Background Survey (for Japanese Speakers).....	224
APPENDIX 3. Pre-Test Materials	225
APPENDIX 4: Sample Experiment 1 Reading Materials	226
APPENDIX 5: Sample Experiment 2 Test Materials	227
REFERENCES	228

LIST OF TABLES

Table 2.1. The four traditional lexical tones of standard Mandarin Chinese	22
Table 2.2. Predictions of Mandarin tone sequences mimicking L1 intonation patterns	72
Table 4.1. Rao-Scott Chi-Square test results for TMS hypothesis	90
Table 4.2. English speakers' substitution patterns	94
Table 4.3. Japanese speakers' substitution patterns	94
Table 4.4. Korean speakers' substitution patterns	95
Table 4.5. English speakers' OCP test 1 results	102
Table 4.6. Japanese speakers' OCP test 1 results	103
Table 4.7. Korean speakers' OCP test 1 results	103
Table 4.8. Error patterns with positional information (Statistical test results included)	111
Table 4.9. English speakers' response-tone rate details with positional information (%).....	112
Table 4.10. Japanese speakers' response-tone rate details with positional information (%).....	113
Table 4.11. Korean speakers' response-tone rate details with positional information (%).....	114
Table 4.12. Overall Substitutions with positional information (without target tone information).....	116
Table 4.13. The comparison of positional errors and substitutions in English speakers' data set	117
Table 4.14. The comparison of positional errors and substitutions in Japanese speakers' data set	117
Table 4.15. The comparison of positional errors and substitutions in Korean speakers' data set	117
Table 4.16. The test results of L1 transfer of typical F0 patterns	121
Table 4.17. Top 10 most frequently produced tone combinations	123
Table 5.1. TMS test (*T2>>*T4>>*T1) results in Experiment 2	135
Table 5.2. English speakers' substitutions in Experiment 2.....	135
Table 5.3. Japanese speakers' substitutions in Experiment 2.....	136

Table 5.4. Korean speakers' substitutions in Experiment 2	136
Table 5.5. OCP test results for English speakers in Experiment 2.....	139
Table 5.6. OCP test results for Japanese speakers in Experiment 2.....	139
Table 5.7. OCP test results for Korean speakers in Experiment 2.....	139
Table 5.8. The top ten positions with highest error rates	144
Table 5.9. Comparison of average error rates of focused and unfocused syllables in Exp.2.....	147
Table 5.10. English speakers' error rates of syllables in unfocused (neutral) substituents	148
Table 5.11. Japanese speakers' error rates of syllables in unfocused (neutral) substituents.....	148
Table 5.12. Korean speakers' error rates of syllables in unfocused (neutral) substituents.....	149
Table 5.13. The most frequently produced tone sequences for "Ft1," "Ft2" and dummy word	164
Table 5.14. Prosodic word tone patterns in F2, F3, F5 and neutral constituents	165
Table 6.1. English speakers' summary chart of disyllabic respondent tones	184
Table 6.2. Japanese speakers' summary chart of disyllabic respondent tones.....	185
Table 6.3. Korean speakers' summary chart of disyllabic respondent tones.....	185
Table 6.4: Accuracy rates of target tone pairs (cited from summary Table 6.1, 6.2 and 6.3).....	197
Table 6.5: The proportions of tone pairs in L2 tonal production (cited from Fig.4.6).....	198

LIST OF FIGURES

Figure 2.1. Diagram of an OT grammar	11
Figure 2.2. Positive transfer and negative transfer in the framework of OT	18
Figure 2.3. Schematized pitch contour of the four Mandarin lexical tones	22
Figure 2.4. Tonal representation I (Bao 1999)	23
Figure 2.5. Tone 3 Sandhi Processes from a traditional view	28
Figure 2.6. Hasse diagram of the ranking in Tableaux (20)	37
Figure 2.7. Syntactic structure of the stimulus sentence.....	46
Figure 2.8. Metrical grid for a bitonal metrical foot with one stressed syllable.....	55
Figure 2.9. Intonation Model of Korean (Jun 2005).....	69
Figure 4.1. Proportions of within-inventory and out-of-inventory errors	86
Figure 4.2. Out-of-inventory errors in Experiment 1.....	87
Figure 4.3. General error rates for TMS test in Experiment 1.....	89
Figure 4.4. Substitutions within each subset in Experiment 1.....	92
Figure 4.5. The default forms and surface forms of T3 in Experiment 1.....	97
Figure 4.6. The occurrences of tone pairs in L2 productions (OCP test 2)	107
Figure 5.1. Internal syntactic structure of focused constituents.....	130
Figure 5.2. General error rates in Experiment 2	132
Figure 5.3. Error rates of individual tones in Experiment 2 (TMS test).....	134
Figure 5.4. Accuracy rates and substitution rates for target F1-f1 in the English speakers' data set.....	155
Figure 5.5. Substitute tones for target F1-f1 in the Japanese speakers' data set	156
Figure 5.6. Substitute tones for target F1-f1 in the Korean speakers' data set	157
Figure 5.7. The distribution of T4 in long focused constituents (English speakers).....	160

Figure 6.1. Positive transfer and negative transfer in the framework of OT.....	171
Figure 6.2. Hasse diagram for Japanese Speakers' ILG	189
Figure 6.3. Hasse diagram for English and Korean speakers' ILG	191
Figure 7.1. Constraint re-ranking model for SLA	210
Figure 7.2. Japanese speakers' partial interlanguage tonal grammar (i.e., Fig. 6.2)	211
Figure 7.3. English and Korean speakers' partial interlanguage tonal grammars (i.e., Fig. 6.3).....	211
Figure 7.4. Two teaching methods of Tone 3.....	215

LIST OF ABBREVIATIONS AND SYMBOLS

AM	Autosegmental-metrical model
CFL	Chinese as a Foreign Language
CON	Constraints
Con-PT	Constraints involved in Positive Transfer
Con-NT	Constraints involved in Negative Transfer
EVAL	Evaluator
FAITH	Faithfulness Constraints
GEN	Generator
ILG	Interlanguage Grammar
ITC	Identical Tone Combination
L1	First Language
L2	Second Language
NITC	Non-Identical Tone Combination
OCP	Obligatory Contour Principle
OT	Optimality Theory
ROA	Rutgers Optimality Archive (http://roa.rutgers.edu)
SLA	Second Language Acquisition
TBU	Tone-bearing Unit
TETU	The Emergence of The Unmarked
TMS	Tonal Markedness Scale
TPC	Tone-Position Constraints
UG	Universal Grammar

Mandarin Tones

[5]=high tone, [1]=low tone

FT3 low dipping tone [214]

T0 Neutral Tone

T1 Tone 1 (level tone [55])

T2 Tone 2 (high rising tone [35])

T3 low tone [21]

T4 Tone 4 (high falling tone [51])

T5 Rising tone (alike in phonetic shape to T2) which results from the Pre-T3 Sandhi.

Symbols

ι (IP) intonational Phrase

φ (PhP) phonological Phrase

(ip) intermediate Phrase (used to describe the English prosodic structures)

(AP) accentual phrase (used to describe the Japanese and Korean prosodic structures)

ω (PW) prosodic word

ft foot

σ syllable

μ mora

H* accentual H, which associates to the stressed syllable

F0 fundamental frequency, in Hertz

[+U] high register (tone) ([+Upper])

[-U] low register (tone) ([-Upper])

CHAPTER 1: INTRODUCTION

1.1. Overview

Among the languages in the world, about half of them are tonal (Hyman 2011, Yip 2002). Mandarin Chinese is one of the best-known Asian tonal languages. There are nearly 885 million native speakers of Mandarin Chinese, making it the world's most commonly spoken language. With China's rising role in the global economy and other fields, there has been a significant increase in the population of second language (**L2**) learners of Mandarin Chinese all over the world and most of the L2 learners of Mandarin are non-tonal language speakers (Hu 2008). Because of its complex tone system, Mandarin is considered to be one of the most difficult languages to learn as an adult (Kiriloff 1969, Bluhme and Burr 1971; Shen 1989; Wang 1995; Sun 1998; Chen 2000; Wang et al. 1999; among others). The pitch of the voice, which relies on the tension of vocal folds and is quantified by fundamental frequency (F0), plays a completely different role in tonal and non-tonal languages. In Mandarin, the contours of pitch over a syllable, known as tones, functions at the lexical level and is used to distinguish the meaning of words. However, in non-tonal languages, pitch functions mainly at post-lexical levels (that is, at the phrase, sentence, and discourse levels), and is independent of the meanings of words (Gussenhoven 2004). That is, pitch patterns may be specified either mainly at the lexical level, at the phrasal/sentential level, or at both lexical and phrasal/sentential levels, resulting in more or less complex tonal specifications for different types of languages. Because of the differing roles that pitch plays in different languages, the second language acquisition (**SLA**) of Mandarin tones can be extremely difficult for adult speakers of non-tonal languages.

Mandarin tonal error patterns made by L2 learners are notorious for being "wild" in the linguistic literature since the large number of mis-productions in the L2 tone system appears to be no more than

random errors. In the past several decades, tone error analysis of the second language acquisition of Mandarin tones has been the focus of much debate and research. Researchers believe that if the sources of tonal errors made by L2 learners could be identified, educators in the field of Teaching Chinese as a Foreign Language (**CFL**) would be able to significantly improve the teaching and learning of Mandarin tones (Sun 1998; Chen, Q-H. 2000; Zhang, H. 2010; Yang 2011; among others). The majority of previous studies on tonal errors are concerned with the L2 tonal productions made by English-speaking learners (Chen G-T. 1974; Chen Q-H. 1997; Elliot 1991; Lu 1992; McGinnis 1996; Miracle 1989; Shen 1989; Sun 1998, Chen 2000; Yang 2011; among others). These studies usually attribute the source of error to the learner's first language (**L1**, here is English), to the L2 (Mandarin), or to the relationship between L1 and L2. Some studies have concluded that tonal error originates from an interference from the learner's mother tongue (referred to as "L1 negative transfer") by analyzing the differences in pitch ranges between English speakers and Mandarin speakers (Chen 1974; Miracle 1989) and by analyzing the L1 (English) suprasegmental transfer (Chiang 1979; White 1981; Lin 1985; Broselow et al. 1987; Shen 1989; Wuelser 1994; Chen, Q-H. 2000). Some studies involved with the second language acquisition of tones have focused on the complex tonal registers and contours that combine to form Mandarin tonal structures (Yip 1980; Shen 1989; Miracle 1989), while some claim that the distinction of Mandarin rising tone and low dipping tone is intrinsically harder than other Mandarin tone pairs, thus causing difficulty for learners (Hao 2012; Elliot 1991; Shen and Lin 1991; Shen, Lin and Yan 1993; Moore and Jongman 1997). Various studies have researched the perception and production of the SLA of Mandarin tones and attribute the sources of error to a complicated mapping between the L1 and L2 tone systems (Leather 1990; Elliot 1991; Hao 2012; Wang 2006; among others). However, there are very few studies which have conducted tonal error analysis from a perspective of phonological universals as sources of error. As mentioned earlier, most studies dealing with the SLA of Mandarin tones have only looked at learners who speak English as their L1; few have examined the L2 tone productions by learners with different L1 backgrounds (Wang 1995; Zhang 2010; So 2006). Cross-linguistic studies would be especially informative to the study of L2 tonal errors, since they could help researchers determine the potential

effects that L1 transfer operations and phonological universals might have on the acquisition of Mandarin tones (Major 2008; Eckman 2008). As Sun (1998) points out, a study which assumes that transfer is operative in the perception and production of L2 tones should “examine groups of learners who have demonstrably different discourse-prosodic features in their L1s. Such a study has yet to be conducted.” (p. 8) Furthermore, there are hardly any studies which account for the L2 tonal errors within the framework of Optimality Theory (OT, Prince and Smolensky 1993, McCarthy and Prince 1993, 1995) in the field of SLA of Mandarin tones, in spite of the fact that OT is one of the most influential pieces of work in linguistics within the past 25 years.

1.2. The significance of the study

In order to fill the gap in the previous studies and advance the study of L2 tone acquisition by bridging language pedagogy with linguistic theories, this dissertation conducts a phonological study of L2 tonal productions made by American English, Tokyo Japanese, and Seoul Korean speaking learners of Mandarin within the theoretical framework of Optimality Theory. American English, Tokyo Japanese, and Seoul Korean (henceforth English, Japanese, and Korean) represent three different types of non-tonal languages according to the characteristics of word prosody: stress-accent languages (represented by American English), lexical-pitch-accent languages (represented by Tokyo Japanese), and non-stress and non-lexical-pitch-accent languages (represented by Seoul Korean) (Jun 2005).¹ In addition, the speakers of English, Japanese and Korean make up the majority of the population of L2 learners of Mandarin worldwide (Hu 2008). The objective of this dissertation is to identify sources of tonal errors (L1 transfer, phonological universals, and a pedagogical problem related to the teaching of tones), and to clarify how these factors jointly shape the L2 tonal interlanguage grammars.² L2 learners often substitute other tones for target tones in productions. Both error patterns and substitution patterns are examined to test for

¹ The term “stress” and “accent” have been used in linguistic literature in diverse contexts. In this study, “accent” refers to the underlying specification for prominence whereas the term “stress” refers to the metrical head physically realized on surface.

² “transfer” here is used as a neutral term, and could be either positive or negative.

evidence of these three factors and to provide suggestions on improving the pedagogy of Mandarin. It is argued in this study that the SLA of Mandarin tones is constrained not only by L1 transfer, but also by some phonological universals and by pedagogical problems related to the traditional method used to teach Mandarin tones. The bulk of the study tests for the evidence of three phonological universals and examines how these phonological universals interact with other factors (i.e. L1 transfer and pedagogical issues) in shaping non-native tonal grammars under the framework of OT. To summarize, this study is most concerned with identifying error or substitution patterns which cannot be attributed to L2 input, nor to structure derived from learners' L1, in order to determine what role phonological universals play in the second language acquisition of Mandarin tones.

1.3. The theoretical context: Second Language Acquisition and Universal Grammar

Following Chomsky's works (1965, 1975, 1981 and 1999), the present study assumes a particular perspective on linguistic universals. That is, the grammars of all languages share a common structure, referred to as Universal Grammar (UG), which stems from a set of universal linguistic principles. It is proposed in Chomsky (1965, 1981) and Pinker (1984, 1994) that UG is part of an innate, biologically endowed language faculty which allows an infant to acquire the grammar of his or her L1 even with limited linguistic experience. UG is composed of certain constant principles that generally apply across languages, but also includes several parameters that allow for variation from language to language. There have been some changes in the way universal principles and parameters have been formalized in UG with the development of linguistic theories such as Minimalism (Chomsky 1995) and OT. For this project, UG is taken as a system of grammatical constraints pertaining to primary language (L1) development and constraints on the grammars of adult native speakers. L2 learners actually face a task similar to that of L1 acquirers in that both types of learners need to arrive at a linguistic system that accounts for linguistic exposure. The apparent similarity between primary and secondary language acquisition has given rise to much investigation and debate; the assertion that UG also governs L2 acquisition and the extent of its influence have been debated since the 1990s (White 2003).

In the study of SLA, Interlanguage Grammar (ILG) refers to non-native grammar developed by L2 learners. The concept of ILG was proposed in the late 1960s and early 1970s by researchers such as Corder (1967), Nemser (1971), Selinker (1972) and Adjemian (1976). One of the characteristics of the system is that it preserves some features of the learners' first language while creating innovations in speaking or writing the L2. Through substantial development in SLA research in the last decade, L2 phonologists found that although native language or L1 transfer has proven itself as a source of learner errors, interlanguage grammars are systematic in a way that stays within the limits of variation allowed by UG (Major 2001, Broselow, Chen and Wang 1998). Some research in SLA has discovered patterns that could not have been learned from the target language input alone, nor derived from the grammar of learners' first languages. Such patterns often reveal a preference for less marked structures (Broselow, Chen and Wang 1998; Epstein, Flynn, & Martohardjono 1996). Although it has been discovered that universal principles of markedness play a role in the SLA, how the aspects of UG are accessible to L2 learners and the extent to which these principles shape the interlanguage grammar has yet to be determined. For example, if L2 acquisition is assumed to be constrained by UG, does UG influence the ILG through L1 transfers? What happens to the aspects of UG that are not required in any particular L1; do they deteriorate from lack of use and become inaccessible, or are they still accessible for L2 learners?

Until the development of Optimality Theory (OT), L2 phonologists did not consider UG to be naturally incorporated in a formal theory of the phonology in facilitating the investigations of SLA patterns (Eckman 2004). OT is a linguistic model which proposes that the observed forms of language arise from the interaction between conflicting constraints. OT grammar consists of a language-specific ranking of universal constraints that evaluate the well-formedness of possible outputs. Constraints are generally regarded as universal, but their rankings differ from language to language, which gives rise to cross-linguistic differences. Language acquisition can be viewed as the process of adjusting the rankings of these universal constraints to match the language one is learning. A constraint ranking determined by the interaction of some factors or constraints relevant to L2 acquisition could therefore manifest the

interlanguage grammars possessed by L2 learners at a given stage of acquisition. The following factors will be considered for the development of tonal interlanguages: L1 grammars, since learners use their L1 knowledge when producing L2; target language grammars, since the target language drives the L2 learners to develop the target language's structures; and universal constraints, since these keep learners from making "wild" errors (Major 2001). By examining actual L2 productions, OT grammar can be used to work out rankings of related constraints, thereby modeling the unconscious linguistic knowledge that exists in the minds of L2 learners. More information about the basic design of OT grammar and the use of OT grammars in SLA is introduced in Chapter 2 (§2.1). OT has been very influential in phonology, generating a considerable amount of insightful analysis on a variety of issues (McCarthy & Prince 1993, 1995, Kager 1999). Yet, it has had little impact on the field of SLA, not to mention the SLA of Mandarin tones. L2 phonologists at the frontier of SLA theoretical research believe that OT "demonstrates promising directions for future research" of SLA (Hancin-Bhatt 2008; Eckman 2004).

1.4. Research questions and general hypotheses

In the theoretical context stated above, the present phonological study of English, Japanese and Korean speakers' acquisition of Mandarin tones is conducted with the goal of answering the following research questions:

- (1) Is UG accessible to L2 learners? More specifically, is the SLA of Mandarin tones by learners with different L1 backgrounds constrained by tone-related phonological universals? If so, how do the L2 learners access UG and to what extent?
- (2) How do the prosodic structures of these three types of L1s affect the SLA of Mandarin tones, especially in expressing sentence-level prominence (namely, contrastive focus and narrow focus)?
- (3) How does the current method of teaching tones influence the SLA of Mandarin Tones (especially Tone 3)?
- (4) How do phonological universals, the L1 transfer of typical tone patterns, and Tone 3 pedagogical issue interact with one another and influence the tone production of L2 Mandarin speech?

The present study hypothesizes that L1 transfers and some aspects of UG, especially those phonological universals related to tone acquisition, and current tone teaching method (namely that of

Tone 3) constrain the SLA of Mandarin lexical tones. While there may be many phonological universals affecting the L2 tone acquisition, this study will focus on three phonological principles relevant to Mandarin tonal acquisition: the Tonal Markedness Scale (**TMS**), the Obligatory Contour Principle (**OCP**) and Tone-Position Constraints (**TPC**). All three are formalized within the theoretical framework of OT. It is predicted that the L2 tonal productions are governed by these universal phonological constraints at both word and above-word level and that different types of L1 speakers will share some similarities in error and substitution patterns. It will be claimed that these phonological universals influence the L2 tonal productions either through L1s or in the form of “the Emergence of the Unmarked” (**TETU**, McCarthy and Prince 1994), which would suggest that some aspects of UG are still accessible to L2 learner even if they are not required in some L1s. However, the effects of UG may not show up clearly and regularly in error or substitution patterns of participants, since these effects often interact with the L1 transfer. This study assumes that the SLA initial state is the steady L1 final state, and therefore that the L1 transfers will lead to differences in the three ILGs. When the tone structure of the L1 and Mandarin are similar or the same, transfer is positive.³ Positive transfer leads to high accuracy rates and high substitution rates. Substitution rate here means the rate at which the target tone is substituted for another tone. For example, the “substitution rate of T1” means the rate at which T1 is substituted for some other tone. When the prosodic structures of L1 and Mandarin are very different, negative transfer may occur and lead to high error rates and corresponding low substitution rates. It is predicted that some typical L1 intonation patterns employed in the realization of phonological events in the native L1s, such as sentence-level prominence (namely, contrastive focus and narrow focus), will also be used in the L2 non-native tonal productions and thereby affect the acquisition of Mandarin lexical tones. This hypothesis will be detailed

³ A lot of SLA research on similarity, such as the Speech Learning Model (SLM) by Flege (1992, 1995), seem to come to the same conclusion: the more similar the phenomena the more likely (negative) transfer will operate (Major 2008), or, similar sounds are difficult to acquire because a speaker perceives them as equivalent to those in the L1. However, what counts as “similar” is controversial. Most studies hold that this ‘similarity causes difficult acquisition’ claim is about vowel or consonants learning and involves the sound contrasts in more than one dimension (for example, a different place and manner of articulation). This study looks at pitch patterns (mainly relying on the tension of vocal folds) for which the L2 contrast involves some dimension that is not used contrastively in the learners’ L1.

by identifying typical L1 pitch patterns which may be transferred to the L2 tonal productions in Chapter 2. The third factor affecting the interlanguage grammars is a problem of Mandarin pedagogy, especially the “Full-Tone 3 First” teaching method. This paper hypothesizes that the problematic assumption of the underlying form of Tone 3 and the associated teaching method of Tone 3 lead to high error rates of T3 in L2 tonal productions. More details are provided in Chapter 2.

To test for evidence of each of these three factors affecting L2 acquisition of Mandarin tones, a word-level phonological experiment (Experiment 1) is designed. Experiment 1 is concerned with the L2 tonal error and substitution patterns in disyllabic words. Disyllabic words will be tested since it is widely recognized that the dominant metrical structure in Mandarin is a disyllabic unit and because the disyllabic unit is the basic prosodic domain in various studies of phonology (Chen 1979; Shih 1986) and morphology (Lü 1963; Feng 1998, 2001; Duanmu 1999). Experiment 2 is a sentence-level phonological experiment containing stimuli of identical tone sequences. This experiment is designed to confirm and expand the findings obtained from the word-level experiment. The error and substitution patterns in both experiments are closely examined and statistical analyses are conducted. The results of the Experiment 1 show that some phonological universals (namely the TMS, the OCP, and TPC) constrain the disyllabic tone sequences. Systematic error and substitution patterns indicating the effects of these constraints are found in participant speech patterns, including cases of UG constraining SLA through L1 and cases of UG not required in L1s but still accessible to SLA (TETU cases). Typical L1 prosodic structures (contrastive focus structures) on disyllabic words are found transferred to L2 disyllabic tonal productions in the examination of the most frequently-produced tone sequences by different types of learners in Experiment 1. Conflicting error and substitution patterns are found and it will be argued that these conflicting patterns are caused by a pedagogical problem related to Tone 3. The effects found in Experiment 1 are confirmed in the sentence-level tonal productions (Experiment 2). In Experiment 2, some effects of the L1 transfer (English as a “head-prominence” or “deaccenting” language, Japanese and Korean as “edge-prominence” or “dephrasing” languages) regarding narrow focus expressions are identified, although a lot of

similarities among the three languages are also found. In addition, the positional effects pertaining to TPC testing are expanded to the domain of sentences and the most vulnerable positions for specific tone types are identified in Experiment 2.

To address the interacting effects of L1 language transfer, universal phonological constraints (including the TMS, the OCP, and TPC) and pedagogical problems (of Tone 3 teaching) that emerge in the development of second-language tonal grammar, this study addresses theoretical issues concerning the SLA of Mandarin tones by examining and comparing overall substitution patterns of English, Japanese and Korean learners. For a general view of the developmental trajectory of L2 tone acquisition, this study proposes an idealized constraint re-ranking model by observing the L1 and L2 prosodic structures and the interlanguage properties inferred from the experimental results, and sketches out how these learners develop their L2-like tonal grammars from their L1 grammars. Four types of constraints and their behavior patterns during the course of SLA (re-ranking of some related universal constraints) and two cases of the “Emergence of the Unmarked” are discussed. Some important features of tonal interlanguage grammars are accounted for through representative OT tableaux and it is argued that the three factors (UG, L1, and pedagogical issues) jointly shape interlanguage grammars at the intermediate stages of SLA.

1.5. The organization of the dissertation

The dissertation is organized as follows. Chapter 2 gives an overview of the phonology of each of the four languages involved in this project, i.e., Mandarin, English, Japanese and Korean, and offers detailed hypotheses and predictions regarding the L1 transfer, three universal phonological constraints, and the pedagogical problems in the SLA of Mandarin lexical tones. Chapter 3 presents the design of the two phonological experiments. In Chapter 4 and 5, experimental results pertaining to the L2 tone patterns in the domain of disyllabic words (Experiment 1) and sentences (Experiment 2) are presented. Chapter 6 discusses how UG, L1 transfer and Tone 3 pedagogical problems jointly shape the interlanguage tonal grammars and regards some theoretical issues in the SLA of Mandarin tones within an OT framework. The dissertation ends with concluding remarks and pedagogical implications in Chapter 7.

CHAPTER 2: PHONOLOGICAL BACKGROUND AND PREDICTIONS

The objective of Chapter 2 is twofold: one is to introduce the phonological background of OT and the prosodic structures of the four languages involved in this study; the other is to detail the predictions made regarding the respective roles played by phonological universal constraints, the transfer of typical L1 melodic patterns, and tone pedagogical issues in the SLA of Mandarin. This chapter is organized as follows: Section §2.1 offers a brief introduction to OT grammar and the use of OT grammar in the context of SLA. In section §2.1, an idealized constraint re-ranking model in the spirit of OT and sketch of the hypothesized rough trajectories of L2 tonal grammar development will be proposed. Section §2.2 offers an introduction of Mandarin tonal phonology including tone representation, tone sandhi rules, tone perception and the three phonological universal constraints formalized in OT which are related to the acquisition of Mandarin lexical tones. Section §2.3 introduces the prosodic structures of the stimuli sentences in this study, preparing for the predictions of typical L1 melodic pattern transfers in the following sections. In section §2.4, the prosodic structures of the three L1s (English, Japanese, and Korean) will be outlined, and typical L1 pitch patterns which may occur in the L2 tonal productions will be identified. Following the descriptions of intonational structures for each language, predictions of the error or substitution patterns regarding L1 transfers are provided.

2.1. Optimality Theory and the second language acquisition of Mandarin tones

This section introduces some basic concepts of Optimality Theory (OT) in §2.1.1. Section §2.1.2 discusses the use of OT in the context of SLA and predicts the development of interlanguage grammar within an OT framework.

2.1.1. Some basic concepts of Optimality Theory

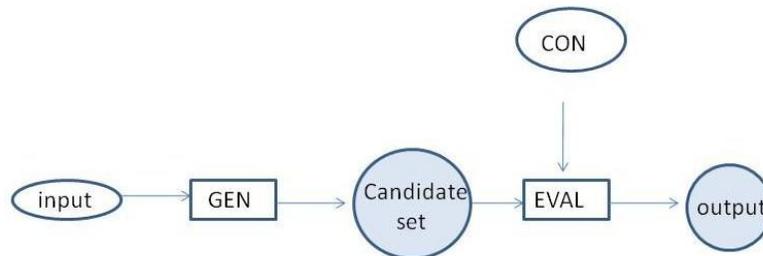
The analysis presented in this dissertation makes use of the theoretical framework provided by Optimality Theory (OT). A brief introduction of a few basic and important concepts which are related to the analysis is provided in this section. Readers who are already comfortable with OT may skip to §2.1.2.

OT is a linguistic model proposed by the linguists Alan Prince and Paul Smolensky in 1993, and expanded by John J. McCarthy and Alan Prince in 1993 and 1995. It was first applied in the area of phonology and it shares its focus with the investigation of universal principles, linguistic typology and language acquisition. OT shares with *The Sound Pattern of English* (SPE, Chomsky and Halle 1968) the notion of an underlying form, or input, as well as the notion of a surface form, or output. However, SPE derives the output step by step from an input, whereas OT simply selects the optimal output from many candidates.

The OT grammar assumes an input-output design in which a linguistic input is parsed, not by rules, but by a universal set of constraints on the well-formedness of linguistic structures. OT is constraint-based, where constraints are generally regarded as universal. The ranking of these constraints differs from language to language, which allows for cross-linguistic differences. The central idea of OT is that the outputs (surface forms) of the language arise from the resolution of conflicts between grammatical constraints. A surface form is “optimal” in the sense that it incurs the least-serious violations of a set of violable constraints, ranked in a language-specific hierarchy. The basic design of the OT grammar is illustrated by the diagram below.

Figure 2.1.

Diagram of an OT Grammar (Archangeli 1999)



A linguistic input enters the grammar and a set of candidates for that input is automatically generated through a function known as the Generator, or GEN. These candidates are then evaluated according to how well they conform to the set of constraints (CON) on linguistic structures. The harmony evaluations are determined by the evaluator, or EVAL, which is guided by a hierarchical ranking of the universal constraints. The candidate description which is evaluated to be the most harmonic is the optimal output for the given input. Two types of constraints are used in this process of evaluation, namely Markedness constraints and Faithfulness constraints.

Markedness constraints reflect generalizations of linguistic structures that commonly occur in natural languages (unmarked) and those that do not commonly occur (marked). The notion of *markedness* embodies linguistic universality in a “soft” sense. The idea is that all types of linguistic structure have two values: “marked” and “unmarked”. Unmarked values are cross-linguistically preferred and basic in all grammars, while marked values are cross-linguistically avoided and used by grammars only to create contrast (Kager 1999). For example, all languages have open syllables (CV, V), but only a subset of language allow closed syllables (CVC, VC). Hence the unmarked value for syllable structure is “open.” Markedness may also involve scales. For example, the simpler a tone structure the more likely its occurrence in languages (for example, level tones are more unmarked than contour tones). OT turns markedness statements into the actual substance of grammars since the structural well-formedness of the output relies on Markedness constraints. Examples of Markedness constraints are given in (1)-(3):

- (1) a. Onset (Prince and Smolensky 1993): Syllables must have onsets.
b. NoCoda (McCarthy & Prince 1994): Syllables must not have codas.
- (2) a. *Contour >>*Level (Yip 2002): Contour tones are more disfavored than level tones.
b. *Contour: No contour tone (two or more pitch targets on a tone-bearing unit) is allowed.
c. *Level: No level tone (one pitch target on a tone-bearing unit) is allowed.⁴
- (3) OCP (tone) (Leben 1973): Identical tones on adjacent positions are prohibited.

⁴ Level tone has the lowest tonal complexity (Ohala 1978, Zhang, J.2004).

(1) includes the unmarkedness values of syllable structures. (2) addresses the tone markedness scale concerning two general categories of tones, namely contour tones such as rising (LH), falling (HL), and even convex tones, versus level tones, such as high level (H) and low level tones (L).⁵ The constraint in (3) indicates that identical tone sequences, such as H-H and L-L are disfavored.

Faithfulness constraints ensure similarity between the input and the output. That is, Faithfulness constraints require the output (surface form) to be identical, or faithful, to the input (the underlying, lexical form) in some particular way. The requirement to be faithful to the lexicon is the major force counterbalancing markedness. Following McCarthy and Prince (1995), some examples of Faithfulness constraints relevant for the current study are given in (4).

(4) Example Faithfulness constraints:

- a. Max (Tone): every tone of the input has a correspondent in the output (that is, don't delete tones).
- b. Dep (Tone): every tone of the output has a correspondent in the input (that is, don't insert tones).
- c. Ident (Tone): let α be a tone-bearing unit (TBU) in the input, and β be any TBU corresponding to α in the output; if α has tone T, then β has tone T.

Constraints are violable. Markedness and Faithfulness constraints are inherently in conflict with one another, so every logically possible output of any grammar will necessarily violate at least some constraints. Violation of a constraint is not a direct cause of ungrammaticality. Instead what determines the best output of a grammar is the least costly violation of the constraints. In a grammar, the constraints are ranked relative to one another, with the principle that the higher the ranking of a constraint, the more serious its violation or, alternatively, the more strongly it holds in a language. Constraint rankings are language-specific, which accounts for the variation in structures observed cross-linguistically. A brief illustration of a tableau representing harmony evaluation is offered in (5). The “*” denotes a violation of the constraint. “*!” denotes a fatal violation, signifying that that particular constraint violation prohibits the candidate from being the optimal one in the evaluation. A shaded cell indicates that the cells have no

⁵ Convex tones such as the full third tone (a low dipping tone) in Mandarin Chinese.

effect on the outcome. Cells in both loser rows and winner rows can be shaded, and shaded cells can contain violation marks or not. The “→” signifies the winner or the most “harmonic” output form.

(5) Sample Tableaux

a.

/input/	Constraint 1	Constraint 2
→Candidate A		*
Candidate B	*!	

b.

/LH.L/	*Contour	Ident (Tone)
→H.L		*
LH.L	*!	

Ranking: *Contour >> Max (Tone)

(5b) shows that the Markedness constraint *Contour dominates the Faithfulness constraint Ident (Tone). The input /LH.L/ generates multiple candidate structure descriptions, and, for the purpose of illustration, only two potential candidates will be considered here, namely a tone sequence that deletes a pitch target of the initial contour tone (decreases the tone complexity) and one that is faithful to the input. The candidate “LH.L” fails because it crucially violate the highest ranked Markedness constraint *Contour, even though it does not violate any Faithfulness constraints. Thus, the optimal candidate, H.L, has the less serious constraint violation and is picked as the surface form.

To clearly illustrate the evaluation process, this paper also adopts the format of *comparative tableaux* introduced by Prince (2002) and uses *combination tableaux* (McCarthy2008) which include information about violations as well as W and L annotations of the comparative tableau. While the original tableaux focus on constraint violations, comparative tableaux help illustrate why a specific candidate is picked. For each losing candidate in the tableau, a W or an L is placed under each constraint to specify whether that constraint favors the winning candidate over that particular losing candidate (W), favors that particular losing candidate over the winning candidate (L), or favors neither (blank). The W and L symbols are limited to loser rows because they represent how a loser compares with the winners for each constraint. Every loser row is expected to have a W and that W is to the left of (dominates) every L. The tableaux (6) below illustrate the new format of combination tableaux:

(6) Sample Combination Tableaux

a.

/input/	Constraint 1	Constraint 2
→Candidate A		*
Candidate B	*W	L

b.

/LH.L/	*Contour	Ident (Tone)
→H.L		*
LH.L	*W	L

This dissertation takes the target forms of tones or tone sequences of Mandarin as the input, or underlying forms, and the L2 tonal productions, or, the non-native productions made by English, Japanese and Korean-speaking learners of Mandarin as the outputs, or the surface forms.

2.1.2. The OT grammars and Second Language Acquisition

In recent years, OT has been developed to account for a range of phonological alternations and has generated a considerable amount of insightful analysis (McCarthy & Prince 1993, 1995, Kager 1999). This theory of the knowledge of language and its acquisition is especially intriguing for researchers in language acquisition studies because of its proposal of language representations and of how language develops over time. Although there is increasing interest in adopting OT to understand L2 data, few studies have actually done so, and even fewer have analyzed the L2 acquisition of tones in an OT framework. Of the few studies which have adopted OT to analyze L2 data, most focus on L2 syllable structures (e.g., Hancin-Bhatt and Bhatt 1997; Hancin-Bhatt 2000; Broselow et al. 1998; Hayes 1999; Villafana 2000). As discussed in Hancin-Bhatt (2008), viewing L2 data within an OT framework has generated more and more interest for a number of reasons. For example, OT does not require different basic assumptions of how the linguistic input types are parsed. Therefore, an OT grammar does not have to assume special rules or representations in order to accommodate L2 input that cannot be passed within an L1 grammar. In other words, the OT grammar removes the need to design specific rules for L2 learners. In addition, OT begins to acknowledge and encode how language variation, a very important topic (variant L2 productions and variable competence) in SLA studies, is a natural consequence of a dynamic

system, particularly during acquisition. In OT, variation results simply from differences in the ranking of the given constraints in different grammars. Furthermore, since OT naturally incorporates UG and accommodates language variation, it not only provides the studies of SLA a new perspective of research but also supplies researchers with very practical tools to account for interlanguage properties, especially those focusing on or involving UG and markedness.

This section will account for the properties of tonal interlanguage grammars in an OT framework and will propose an SLA constraint re-ranking model. Following the introduction of this constraint re-ranking model, the proposal about the three sources of error for L2 tonal productions will be viewed from the perspective of L1-L2 constraint ranking similarities and differences. The central concept behind an OT grammar is that a grammar consists not of a set of rules, but rather of a set of ranked constraints which are presumed to be universal. Given materials from the target language (i.e., Mandarin in this study), what a second language learner (i.e., the English, Japanese, Korean speaking learners in this study) must infer is the ranking of these universal constraints in L2. That is, within an OT framework, language acquisition can be roughly described as the process of adjusting the L1 ranking of relative constraints to match the L2 grammar one is learning.

This study assumes a Full Transfer/Full Access Model (Schwartz & Sprouse 1996) of L2 learning. “Full Transfer” refers to a full instantiation of the L1 constraint ranking into the initial state of the interlanguage grammars.⁶ Therefore the SLA of Mandarin tones can be taken as a process of adjusting the tonal grammar rankings from the L1 prosodic grammars toward the target L2 tonal grammar. During this process, the L1 prosodic properties inevitably influence the acquisition of the L2. In other words, the

⁶ Hancin-Bhatt (1998) suggests a specific development in complex coda acquisition and proposes that during initial stages of language acquisition, the constraint hierarchy is instantiated onto the L2, with a simultaneous demotion of the Faithfulness constraints. The Hancin-Bhatt (1998) proposal is an idealized and simplified scenario regarding Faithfulness constraints. As Hancin-Bhatt (2008) state, “the beginnings of constraint re-rankings as a reflection of interlanguage grammar....are underdeveloped in the L2 phonology literature” (p.136). The present study is concerned more with the behavior of Markedness constraints. Therefore the discussion of constraint re-ranking here focuses on Markedness constraints only.

L1-specific rankings of universal constraints are transferred to the interlanguage grammars because L2 acquirers begin with the knowledge of L1 properties before exposure to the target language input (L2).

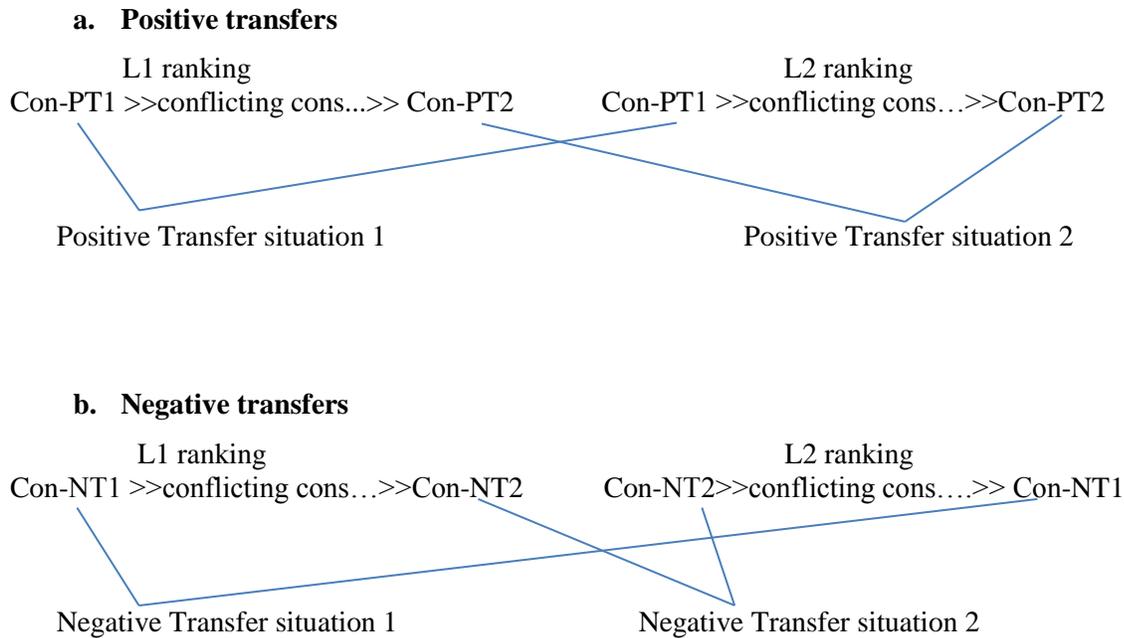
With respect to constraint re-ranking during language acquisition, Tesar and Smolensky (2000) assume that Markedness constraints are demoted in the course of L1 acquisition. Boersma and Levelt (2000) discuss the L1 acquisition order of Dutch syllable types and assume both Markedness constraint demotion and Faithfulness constraint promotion. Boersma and Levelt (2000) hold that, in the initial state of the child's grammar, all Markedness constraints are ranked above Faithfulness constraints (Gnanadesikan 1995). Faithfulness then gradually rises in the hierarchy, one by one overtaking Markedness constraints. In the end, faithfulness is ranked on top and the child masters all syllable structures. The constraint re-ranking in second language acquisition differs from the constraint re-ranking in first language acquisition because of the different initial states. Following the "Full Transfer" hypothesis (Schwartz & Sprouse 1996), we would expect a full instantiation of the L1 constraint ranking into the initial state of the ILGs. Since this study focuses on the effects of phonological universals, we are particularly interested in the re-ranking of Markedness constraints. We assume that, during the SLA process, some Markedness constraints will be demoted while others will be promoted depending on the specific L1 and L2 grammars.

Previous studies reveal that "L2 structures that are similar or the same as their counterparts in the L1 can have a generally facilitative effect in learning, while L2 structures that are not present in the L1 grammar provide a substantial challenge" (Hancin-Bhatt 2008). Thus, we assume that L1 grammar rankings similar to or the same as the L2 structures will be transferred to ILG easily and quickly which can be taken as cases of Positive Transfer. However, if the L1 grammar rankings are very different from that of L2's and re-ranking is required, Negative Transfers would occur.⁷ The actual SLA process and the interlanguage properties are very complicated and it is even considered "extralinguistic phenomena" (Hancin-Bhatt 2008: 121). In order to sketch out the rough trajectories of L2 tonal grammar development

⁷ In some situations, the L1 ranking of some constraints is too low to be visible in L2 grammar.

in the spirit of OT, this study proposes a simplified and idealized SLA constraint re-ranking model, as displayed in Fig. 2.2 below. Error and substitution patterns found in non-native tonal productions will be accounted for using this constraint re-ranking model in Chapter 6.

Figure 2.2. Positive transfer and negative transfer in the framework of OT



The constraints involved with the Positive Transfer situations are either ranked high or low relative to conflicting constraint in both L1 grammar (L2 initial ranking) and L2 grammar (L2 final ranking). There is no need to re-rank constraints during the course of acquisition and therefore the acquisition of the rankings is quick and easy. In this dissertation, the two situations of Positive Transfer will be referred to as PT1 and PT2 and the constraints involved with these two situations will be referred to as Con-PT1 and Con-PT2 respectively. As displayed in Fig. 2.2a, “Con-PT1” constraints are ranked high (Positive Transfer situation 1) while “Con-PT2” constraints are ranked low (Positive Transfer situation 2) in both L1 and L2. However, if the constraints are ranked very differently in the L2 initial ranking and the L2 grammar ranking (i.e., low in the L2 initial state ranking but high in the target language grammar, or vice versa), new rankings must be learned and the process of re-ranking would

result in slow acquisition (Negative Transfer) of some structures. The two situations of negative transfer will be referred to as NT1 and NT2 and the constraints involved with them will be referred to as “Con-NT1” and “Con-NT2” respectively. As seen in Fig. 2.2b, “Con-NT1” is ranked high in L1 but low in L2, while “Con-NT2” is ranked low in L1 but high in L2. The re-ranking of Con-NT1 means L2 learners are trying to acquire a greater range of linguistic structures or trying to acquire a new inventory of linguistic structures. However, the re-ranking of Con-NT2 means L2 learners are trying to delimit the range of possible L2 structures.

In this dissertation, the study of L1 transfer focuses on the transfer of typical L1 pitch patterns (especially in disyllabic words), which varies among the three L1s. The typical L1 pitch patterns are determined based on the L1 prosodic grammars and the structure of stimuli sentences. Both positive and negative transfer may occur regarding these typical L1 pitch patterns, and it depends on whether the properties, P, of the typical L1 pitch patterns are associated with or carried by Positive Transfer-type constraints or Negative Transfer-type constraints. If the constraints bearing the property P, of the L1-specific prominent intonation patterns undergo positive transfer, we expect quick and easy acquisition. However, when the constraints bearing P of the typical L1 intonation patterns undergo negative transfer, it may cause a relatively slow acquisition. We expect that when the target tone sequences are similar to the melodic patterns typical of the L1 (that is, if the target tone sequences share similar rankings of relevant constraints), the error rates will be low and substitution rates of these patterns will be high. However, if the target L2 tone sequences are unlike the typical L1 pitch patterns (i.e., rankings of relevant constraints are very different), learners may experience difficulties in the acquisition process, which may lead to high error rates and low substitution rates for specific tone type or tone sequences.

As mentioned in Chapter 1, we predict not only that L1 transfers constrain the SLA of Mandarin tones, but also that some universal phonological constraints constrain the SLA of Mandarin tones, since we assume a “Full Access” model. The “Full Access” of UG is reflected in the UG-constrained interlanguage grammars (ILG). As we know, some constraints reflecting universally-preferred structures

may already be contained in the constraint-ranking of the L1, and these rankings could be transferred into ILGs via typical L1 pitch pattern transfer. In this way UG is accessed by L2 learners through knowledge of their L1. However, it may also be the case that UG is accessible to L2 learners without the help of L1. If this is the case, it should be reflected in the innovations of ILG during SLA. If some Markedness constraints reflecting universally-preferred structures are neither required by the learner's L1 nor found in the learner's L2 input but instead emerge during the development of the learner's interlanguage grammar during SLA, we can infer that aspects of UG which are not in use for the speaker's L1 do not decay with lack of use, but instead continue to be accessible to the speaker. If evidence of this is found, it would be an example of "the Emergence of the Unmarked" (TETU, McCarthy and Prince 1994) in SLA.

In the spirit of OT, some constraints are "inactive" in L1 because they are not necessary in determining the actual outputs. These "inactive" constraints may be ranked very low and masked in the learners' native-language grammar by the effects of higher ranking constraints, or masked by some Faithfulness constraints in the correct productions. Thus we cannot identify these constraints by only examining L1 transfer effects. However, these "inactive" constraints are assumed to be present in the grammar. In some cases, these hidden rankings have been uncovered through examination of loan-word phonology or reduplication and truncation phenomena (Ito and Mester, 1995, McCarthy and Prince, 1995, among others). Studying second language acquisition may be another way of revealing hidden rankings since some of these "inactive" constraints may become visible in the interlanguage grammars reveal their ranking relative to other constraints. These patterns frequently show a preference for less marked forms and, as Epstein, Flynn & Martohardjono put it, this is "generally described as an effect of universal principles of markedness, often conceived of as part of the innate endowment provided by Universal Grammar" (1996). This study uncovers some tonal patterns that appear to be independent of both the L1 and L2 grammars to some extent, and are thought to be example of TETU. The experiments in this study test for the evidence of the access of UG by L2 Mandarin tone learners by examining the emerging effects of three tone-relevant universals: the TMS, the OCP, and TPC. The definitions of these constraints are

introduced in the section of §2.2 and a discussion of TETU as it relates to Con-PT2 and Con-NT1 is offered in Chapter 6. It is found that in these cases of TETU, some Markedness constraints, such as the OCP (LH) constraint, are not simply demoted (if they are Con-NT1) or remaining low or inactive (if they are Con-PT2) and dominated by other constraints (such as Faithfulness constraints), but sometimes rear their heads in interlanguage grammars, playing a decisive role in picking the surface forms.

The third error source this study identifies is a pedagogical problem of a target tone type, Tone 3. The low T3 is phonetically easy (unmarked) and assumed to be acquired by L2 learners easily. That is, the rankings of T3-related constraints (such as “No dipping tones” or *FT3) in L1s and L2 are similar, so dramatic re-rankings are not required.⁸ However, it will be claimed that the assumption that the underlying appearance of Tone 3 is a dipping tone, rather than a low falling tone, and the teaching method associated with this assumption, interferes with the acquisition of Tone 3. This traditional teaching method associated with Tone 3 may cause an unnecessary re-ranking of some specific constraints (specifically, learners may be assigning *FT3 a low rank in L2, but a high one in L1), leading to high error rates of Tone 3 and conflicting error and substitution patterns. This is detailed in the following section §2.2, which gives an overview of T3 Sandhi in Mandarin.

2.2. The tone system of Mandarin Chinese

This section offers an introduction to the tone system of Mandarin, the target language in this study. Sub-section §2.2.1 gives an overview of the inventory of Mandarin tones and their representations; §2.2.2 outlines tone sandhi processes in Mandarin; §2.2.3 gives an overview of the perception of Mandarin tones; and §2.2.3 gives background information on the three tone-related phonological constraints used in OT: the TMS, the OCP, and TPC.

2.2.1. Mandarin tonal features and representation

⁸ Since FT3 is not allowed in the surface forms of test words in the experiment 1, we assume the constraint *FT3 is ranked high in L2. L1s in this study, such as Korean and Japanese, also do not allow FT3; therefore *FT3 is also ranked high in L1 grammars.

In Mandarin, morphemes are almost exclusively monosyllabic. The pitch contour over a syllable can distinguish word meaning, yet the tone each syllabic morpheme takes is entirely arbitrary and unpredictable. There are four types of tones on full syllables, as well as a neutral tone, which is carried by unstressed syllables (such as function words) and whose pitch value is determined by the preceding tone. The neutral tone will not be discussed in this study. An example of a stressed syllable in each of the four (non-neutral) tones is provided in table 2.1 below. The numbers of contour pitch value indicate the relative starting and ending pitch of each tone on a 1-5 scale, where 5 represents the highest pitch, and 1 represents the lowest (Chao 1930, 1948).

Table 2.1. The four traditional lexical tones of standard Mandarin Chinese

Wade-Giles name	Pinyin Spelling	gloss	Shape of tone	contour pitch value
Tone 1	mā	“mother”	(high) level	55
Tone 2	má	“hemp”	(high) rising	35
Tone 3	mǎ	“horse”	(low) dipping/low	214/21
Tone 4	mà	“scold”	(high) falling	51

Figure 2.3. Schematized pitch contour of the four Mandarin lexical tones

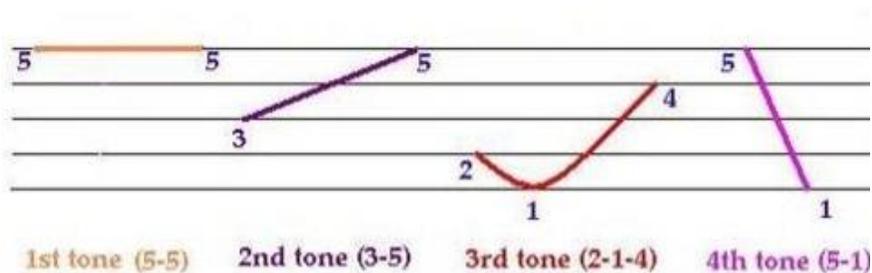


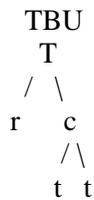
Table 2.1 illustrates each of the four traditional tones. In Chao’s (1948) description, Tone 1 is the only level tone; the other three are contour tones. The three pitch levels of Tone 3 form a complex contour. Traditionally, we assume the low dipping tone is the standard form of Tone 3. Tone 3 is realized as a full dipping contour [214] mostly in isolated syllables and in prosodic-final positions, while it is usually realized as a low tone [21] in non-final positions. However, in some word- and phrase-final positions,

[214] is not a possible realization of T3. Only [21] is a possible realization of T3 in these positions and [21] has a much wider distribution in Mandarin. In this study, we assume that Tone 3 is phonologically a low level tone, and that it can appear not only at non-word- and phrase-final positions but also at word- and phrase-final positions in Mandarin Chinese (Kratochvil 1968; Shih 1986, 1997; Yip 2002). Therefore following Yip (2002), this study assumes [21] is the default or standard form of Tone 3, and we label [21] as “T3,” while the full low dipping tone [214] will be labeled “FT3”, which stands for “Full Tone 3”.

Following Goldsmith’s (1976) theory of autosegmental phonology in which segments and tones appear on separate “tiers,” we assume that tones are linked with Tone-Bearing Units (TBU). We assume the TBU in Mandarin are syllables.⁹ In the large body of previous work on Chinese Linguistics, tonal categories are consistently classified by two sets of descriptive terms: one denoting pitch height (high/low or “yin”/ “yang”), the other pitch movement (rising, falling, dipping, etc.) This practice implies that tone consists of two independent orthogonal dimensions: register (i.e. pitch height) and pitch contour (Chen 2000). This paper follows the model of tone proposed by Bao (1999), which is displayed in Fig. 2.4.

Figure 2.4. Tonal representation I (Bao 1999)

TBU: tone bearing unit; T: tone root; r: register; c: contour; t: terminal tone segment



All of these features, including register and contour, are borne by a syllable. This model suggests that a contour tone behaves like a single unit since the contour node dominates the terminal tone segments. In this model, register is a sister of contour but does not dominate it. It predicts that register and contour could behave independently.

⁹ This is a general claim for contour tones, including contour tones in Mandarin. According to Zhang, J. (2004b), contour tones are properties of the syllable and the sonorous portion of the rime is the Tone-Bearing Unit.

In register feature analysis, a binary feature [+/- upper] is proposed (Yip 1980). This register feature “situates the beginning of the tone in either the upper or the lower part of the speaker’s register, since contours have only an initial tonal specification, and the phonetic contour is the result of drift away from that initial target. Level tones have two targets, firmly anchoring both ends and keeping the pitch stable” (Yip 2001). This paper employs the binary feature, i.e., [+/- upper], henceforth [+/- U]. Following the model of Bao (1999), this paper employs “h” and “l” to refer to high and low component tones in the terminal tone positions of a contour tone here, where two terminal tone segments make up a whole tone borne by a syllable. Following this claim, the citation forms of four lexical tones can be represented as (7):

(7) Mandarin Tonal representation II

tone	pitch value	register + component tones
Tone 1 (T1)	55	[+U, hh]
Tone 2 (T2)	35	[+U, lh]
Tone 3 (T3)	21	[-U, ll]
(FT3)	214	[-U, llh]
Tone 4 (T4)	51	[+U, hl]

For the sake of economy, I will just use the regular H and L system, the “most parsimonious account” (Yip 2002) proposed for the African tonal system, in the general discussions of tone behaviors in Chapter 2. Following Yip (2002) who assumes that a contour tone is composed of two level tones, the four tones can be represented as H (T1), LH (T2), L (T3) and HL (T4). However I will use the Mandarin tone representations of T1, T2, T3, FT3 and T4 elsewhere because in most cases the constraints, such as the OCP, refer to Mandarin whole tones in the present study.

It’s clear that Mandarin is a tone language due to the nature of the lexical tones stated above. However, there are very few studies concerning the stress system in Mandarin. Of the few studies which have been conducted, there have been two main findings. One is a prevailing view of the stress system in Mandarin which holds that stressed sounds or syllables carry full lexical tones while unstressed syllables carry neutral tones. The other is about the relative prosodic strength pattern in Mandarin. The other finding (see Kochanski et al. (2003) and Lai et al. (2010)) is that there is a robust strong-weak alternation

with respect to fundamental frequency (F0), the quantitative representation of pitch, in Mandarin. The prosodic strength pattern of this strong-weak alternation in Mandarin will be considered in the analysis of universal high-low registered tone pattern found in Experiment 2.

2.2.2. Mandarin Tone Sandhi

2.2.2.1. Mandarin T3 Sandhi and the domain

This section considers Tone Sandhi in Mandarin and the teaching of T3 Sandhi in the field of Chinese as a Foreign Language (CFL). In connected speech, adjacent tones undergo dramatic alternations to their phonetic shape. This tonal alternation is referred to as Tone Sandhi (Chen, M. 2000). Among the few existing Tone Sandhi processes in Mandarin, Tone 3 Sandhi, is the most common. T3 Sandhi is illustrated in (8). The T3 Sandhi rule turns a T3 on the first full syllable of a two-T3 sequence into a rising tone which is alike in phonetic shape to T2. This rising tone which results from the T3 Sandhi rule will be referred to as **T5** in this study.

(8) Tone 3 Sandhi rule (or, Pre-T3 Rule)

- a. *mai* “buy”
L citation form (T3)

- b. *mai jiu* “to buy wine”
L.L citation form (T3-T3)
HL.L sandhi form (T5-T3)

This dissertation follows Shih’s (1986) proposal about the Sandhi domain. Shih (1986) proposes that the domain of Mandarin Tone Sandhi is prosodically-defined and that it is the foot that constitutes the domain of Mandarin sandhi. The Foot Formation Rule (FFR) proposed in Shih (1986) to derive prosodic feet is provided in (9) and (10):

(9). Foot (ft) Construction

- a. Link immediate constituents into disyllabic feet.
- b. Scanning from left to right, string together unpaired syllables into binary feet, unless they branch to the opposite direction.

(10). Super-foot (ft') Construction

Join any leftover monosyllable to a neighboring binary foot according to the direction of syntactic branching.

The Foot Formation Rule correctly predicts the sandhi domains for tri-syllabic phrases. Two examples of tri-syllabic phrases (Shih 1986) are given below.

(11) a. “The umbrella is small”

Yu san xiao
rain umbrella small
() ft
() ft'
LH LH L

b. “small umbrella”

xiao yu san
small rain umbrella
() ft
() ft'
L LH L

That the foot is the domain for Tone Sandhi is also evidenced in constructions involving *yi* “one” (一) and *bu* “no” (不). These are two T4 morphemes which, when strung together, undergo a tone sandhi rule similar to the T3 Sandhi rule. When these two T4 morphemes are followed by another T4 morpheme, they become rising tones. This particular T4 Sandhi rule only undergoes with these two morphemes, and not with any other T4 morphemes.

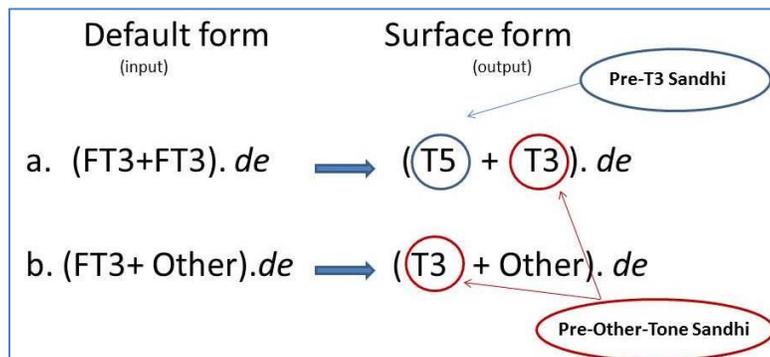
The FFR will be referred to in the target tone sequences for Experiment 2. Note that although the Sandhi rules described above (i.e. the T3 Sandhi rule and the T4 Sandhi rule) are both cases of dissimilation when there are two adjacent similar tones, in native Mandarin, these rules do not apply to other identical tone sequences. Therefore L2 learners should acquire sandhi rules for T3-T3 sequences, but not for T1-T1, T2-T2, or T4-T4 sequences.

2.2.2.2. T3 and T3 Sandhi Pedagogy

In the field of Teaching Chinese as a Foreign Language (CFL), almost all Chinese language schools or programs in the world employ the prevailing teaching method of “Full Tone 3 (FT3)- First,” including the the universities from which subjects for this study were recruited (the University of North Carolina at Chapel Hill (**UNC-CH**) and Zhejiang University). The traditional “FT3 First” method follows Chao (1948,1968), who lists the citation tone shape as “low dipping” for Tone 3 and further notes that in non-citation context, Tone 3 surfaces as “high rising” before another third tone in environments where Tone Sandhi applies, and as “low (falling)” elsewhere. That is, the L2 learners first study the citation form of the third tone as FT3, a low dipping tone, and then subsequently learn two T3 Sandhi rules: 1) the **Pre-T3 Sandhi**, which changes FT3 into T2 (referred to as T5 in this paper) when followed by another FT3, and 2) a rule which drops the final rise of the dipping FT3, changing FT3 into T3, the so-called “half-third tone,” when FT3 is followed by other tones (i.e., T1, T2, T4 and neutral tone). I call the second one the **Pre-Other-Tone Sandhi**. These two sandhi rules make the difficult task of tone acquisition even more difficult when it comes to T3, because within sentences, even at word- or phrase-final positions, T3 Sandhi rules should be applied nearly everywhere (Kuang and Wang 2006). L2 learners must learn to change FT3 into either the rising tone T5 before FT3 (Pre-T3 Sandhi) or to T3 before all other tones (Pre-Other-Tone Sandhi), except in utterance final positions.

In this paper, we examine the behaviors of T3 and T5 separately. In Experiment 1, the test disyllabic words are followed by a structural particle *de* (的) bearing neutral tones. Therefore, if there is a Tone 3 at the final position of a disyllabic word, it should surface as a T3 [21] because of the Pre-Other-Tone Sandhi. For example, when the test disyllabic word *shuǐ guō*, a FT3-FT3 tone sequence, is followed by particle *de*, a neutral tone, the first morpheme *shuǐ* should surface as a rising tone because of Pre-T3 Sandhi, while the second morpheme *guō* should surface as T3 [21] according to Pre-Other-Tone Sandhi. This means no FT3 will appear in the surface forms in the test materials. Below, the diagram of the surface forms of Tone 3 displays the Sandhi processes Tone 3 undergoes at specific positions following the traditional assumption of FT3:

Figure 2.5. Tone 3 Sandhi Processes from a traditional view



The diagram shows that T5 experiences Pre-T3 Sandhi while T3 at both word-final position in (a) and at word-initial position in (b) undergoes Pre-Other-Tone Sandhi. Therefore in the L2 tonal productions, if target T5 has a high error rate, this means L2 learners have difficulty with Pre-T3 Sandhi. If target T3 has a very high error rate, this means L2 learners have difficulty with Pre-Other-Tone Sandhi. This paper predicts that the “FT3-First” teaching method leads to high error rates of T3 and T5 because L2 learners tend to overuse the “standard” form they have been taught, FT3, and under-value the importance of or forget to apply the complicated Sandhi rules.. The experimental results regarding T3 sandhi are reported in Chapter 4. The pedagogical implication about T3 is discussed in Chapter 7. This study advocates the assumption of the low level tone (T3) as the default form for Tone 3 and supports a “T3 First” teaching method, as opposed to the current widely-used “FT3 First” method, which will not only simplify the rule application process, but will also significantly reduce the interference of Tone 2 and Tone 3 with one another in tone perception and production.

2.2.3. The perception of tones

In this section, we briefly review previous studies done on the perception of tones by L2 learners and justify the assumption made that the three groups of L1 speakers in this study can correctly perceive the Mandarin tones used in the experiments detailed in Chapter 3.

Previous studies suggest that, in addition to the primary acoustic parameters of F0 height and F0 contour, other parameters such as amplitude and temporal properties (overall duration and Turning Point) are also effective phonetic correlates of the tones (Wang et al. 2006, Moore & Jongman 1997) There are two broad categories of findings about non-native tone perception by L1 tonal language speakers and non-tonal language speakers. Many studies reveal that tone language speaking learners and non-tonal language speaking learners may perceive tones in different ways, while some studies suggest speakers of tonal L1s and non-tonal L1s do not differ in the overall accuracy of their perception but may have different specific confusion patterns.

Among the first category of studies, Gandour & Harshman (1978) compared the discrimination of paired synthetic tones by speakers of Thai, Yoruba, and English. Auditory dimensions such as pitch height and length of the stimuli were important for all three groups of subjects. However, English speakers did not attach much importance to the contour, while speakers of tonal languages, Thai and Yoruba, did. It seems that speakers of tone and non-tone languages may attach differing degrees of importance to different perceptual cues. Another finding regarding the perception of tone is that speakers of tone languages tend to exhibit more categorical perception of pitch patterns than speakers of non-tone languages. Stager & Downs (1993) found that speakers of tone languages are not sensitive to slight frequency changes because they make more categorical judgments of pitch.

Some studies claim that L1 tone background facilitates non-native tone discriminations. Wayland and Guion (2004) trained L1 English speakers and Mandarin speakers to discriminate a pair of Thai tones and found that speakers of tonal L1s (Mandarin) might have an advantage over speakers of non-tonal L1 (English) in learning to perceive tones in an L2 (Thai). Lee, Vakoč, and Wurm (1996) make a similar observation, finding that Cantonese speakers' tonal background made it easier for them to discriminate both Mandarin and Cantonese tone pairs, as compared to English speakers. This study concludes that a tonal L1 might only give an advantage in discriminating L2 tones if the L1 tonal system is more complex than that of the L2.

However, there are also some studies which reveal that speaking a tonal L1 negatively affects the perception of L2 tones. Wang (2006) tested speakers of Hmong (a tonal language), Japanese (a pitch accent language), and English (a stress language) on how well they could identify Mandarin tones. The results showed that English and Japanese speakers performed equally well, but that Hmong speakers performed significantly worse than the other two groups. The author concludes that the lack of exact mapping of L2 tones onto L1 tones may interfere with the acquisition of nonnative tones, especially at the initial stage of learning. So (2006) examines the effects of native (L1) prosodic background and audiovisual (AV) training on the perception of the four Mandarin tones by groups of naïve listeners. So (2006) found that the Cantonese tonal system hindered the learning of Mandarin tones, while the Japanese pitch-accent system facilitated the establishment of a new tonal system. The English stress-accent system neither helped nor hindered tone learning.

In the second category of studies, Francis et al. (2003) assessed Mandarin Chinese and English speakers' identification of Cantonese tones before and after training. There was no significant difference between these two groups after training, i.e. their stronger tones and weaker tones are the same before and after training.

Among the studies on the perception of L2 Mandarin tones, there are additional findings regarding the confusion patterns and positional effects in perception. Various studies of perception of Mandarin tones (see Gandour (1978) for a review) confirm that both native speakers and learners of Mandarin find T1 and T4 easiest to identify, while T2 and FT3 [214], which are acoustically most similar to one another, are harder to differentiate. Kiriloff (1969) and Chen (1997) found that English-speaking learners often confused Tone 2 [35] and Full Tone 3 [214] in perception. Hao (2012) investigates the perception and production of Mandarin tones by English and Cantonese (a tonal language) speakers. Hao (2012) found that the Cantonese group did not perform significantly better than the English group in perceiving and producing Mandarin tones, and that both groups had significant difficulty in distinguishing Mandarin Tone 2 and Full Tone 3.

Some positional effects in Mandarin tone perception by L2 learners are also found. Broselow et al. (1987) found that the perception of T4 is affected by position. T4 is significantly easier to perceive when it occurs at the end of an utterance. This positional effect is likely due to the fact that an utterance-final falling tone (T4) is phonetically similar to the English pitch contour. The falling tone may be more difficult for English speakers to perceive when it appears in an unfamiliar context, such as in a disyllabic word-initial position, or in the middle position of a triplet. Furthermore, English speakers sometimes misidentify the target T4 as T1 at the ends of strings. Chen (2013) investigates possible innate perceptual biases that may shape the phonological rule in Mandarin T3 Sandhi. Chen (2013) found that Mandarin listeners perceive the Mandarin T2-FT3 contrast categorically while Dutch listeners perceive this contrast in a psycho-acoustical manner. Nevertheless, both groups were more accurate in discriminating Tone 2 and Full Tone 3 if Full Tone 3 occurred first in the to-be-discriminated pair.

The subjects in most of the studies reviewed have either had no specific L2 learning experience of the stimuli (are new to the target lexical tone systems) or have received only limited training. The subjects in the current study have received at least 5 month of intensive face-to-face classroom instruction of Mandarin tones. In addition, in this dissertation project, all Chinese morphemes/words/sentences used in the two phonological experiments are picked from elementary-level textbooks, and it is confirmed with all subjects that they are very familiar with these morphemes in the pre-test. The proficiency level of stimuli morphemes and words are also confirmed with the handbook of Outline of Graded Vocabulary for HSK (1992), a standard book on the subject. Thus we assume all subjects correctly perceive the target tones used in the experiments in a similar way. This study is concerned only with the production of Mandarin tones during SLA, but the findings on the perception of Mandarin tones reviewed above are considered in accounting for the L2 tone patterns.

2.2.4. Three phonological constraints relevant to Mandarin tone acquisition

The Markedness constraints relevant to SLA acquisition of Mandarin tones assumed by this study include the Tonal Markedness Scale (TMS), the Obligatory Contour Principle (OCP), and Tone-Position Constraints (TPC). These three phonological principles are generalized based on a survey of natural languages and formalized within the OT framework. These constraints are ranked relatively low in the target language, Mandarin, because generally speaking Mandarin allows all the tone types (levels and contours) and identical tone sequences (except for T3-T3), and there is no restriction on the distribution of tone types in terms of position (such as word-initial or final positions) in Mandarin. However, the rankings of these Markedness constraints are relatively high in English, Korean, and Japanese (see details in §2.4). The re-ranking of these constraints during the course of SLA may bring up interesting error or substitution patterns, thus lead to some innovations in the interlanguage grammars. This study hypothesizes that these three phonological principles govern the L2 tonal productions. That is, these constraints may not jump from their high L1 rankings directly to the target L2's low rankings, and may instead show up over the course of SLA, playing decisive roles in picking surface forms which may appear to be SLA errors. We expect that some evidence of the effects of the TMS, the OCP, and TPC will be detectable in the L2 learners' tonal productions in this study. This section will introduce these three phonetically-grounded principles and their relevance to the SLA of Mandarin tones.

2.2.4.1. The Tonal Markedness Scale (TMS)

The Tonal Markedness Scale (TMS) is a universal and phonetically-grounded constraint scale usually presented as *Rising >> *Falling >> *Level (Ohala 1978, Hyman & VanBik 2004). In other words, rising tones are more complex than falling tones, which in turn are more complex than level tones. Or, rising tones are the hardest to produce, falling tones are the second hardest, and levels are the easiest.

This tone complexity scale is a fixed ranking based on a cross-linguistic survey of the tonal inventories in the world's languages and phonological alternations. Clearly, contour tones have greater complexity than level tones because contour tones contain more pitch targets within a tone-bearing unit

than level tones (Ohaha 1978). Therefore *Contour >>*Level is proposed to address this observation in the present study. Within the group of contour tones, there are some reports in the tonal literature that falling tones behave differently than rising tones. According to Cheng (1973) and Ohala (1978), falling tones are more numerous in tonal inventories of languages because they are both easier to produce and perceive than rising tones. It is reported that the pitch interval between the tonal sequence low-high is more likely to be reduced than the interval between a high-low sequence (Hyman 1973) and that falling contours must cover a greater pitch range than a rising contour in order to be perceived with equal levels of “prominence.” Ohala and Ewan (1973) and Sundberge (1979) reveal that speakers are able to produce a falling pitch over a given pitch interval much faster than a rising pitch over the same interval, which may indicate that falling tones make better tonal contrasts and thus are perceptually more salient. For this reason, falling tones are found in a greater number in languages. Based on Zhang, J.’s (2002) survey of 187 genetically diverse tone languages, three implicational tendencies were established to address the tonal markedness scale:

(12) The TMS effect reflected in the tone distributions in natural language (Zhang, J. 2004a)

- a. If a language has contour tones, then it also has level tones.
- b. If a language has complex contour tones, then it also has simple contour tones.
- c. If a language has rising tones, then it also has falling tones.

According to the TMS (*Rising>>*Falling>>*Level), we would predict that “*T2>>*T4>>*T1” in the context of Mandarin lexical tones. As noted earlier, FT3 contains three tone targets, represented as [214] in the system of Chao (1948), and is regarded as the most complicated tone type. So we predict that FT3 is more complicated than other tones, which contain at most two tone targets. In terms of level tones, we predict that T1, a high level tone, is more marked than T3, a low level tone, following the universal tone complexity scale *High>>*Low discussed in Yip (2002). Therefore, the predicted complete tonal markedness scale as it pertains to Mandarin, including FT3 and T3 could be proposed as (13).

(13) *FT3>>*T2>>*T4>>*T1>>*T3: FT3 is more disfavored than T2; T2 is more disfavored than T4; T4 is more disfavored than T1; T1 is more disfavored than T3.

The TMS is reflected in the first language (L1) acquisition of Mandarin tones. Most studies of L1 acquisition of Mandarin tones reveal that Mandarin tones are acquired very early, long before the inventory of segmental sounds is mastered. Li and Thompson (1977) looked at seventeen Mandarin-speaking children aged 1; 6 to 3 years. They find that at the earliest one-word stage, high level tones (T1) are produced first, followed by high falling tones. Rising (T2) and dipping tones (FT3) come later, and syllables with such tones are either avoided or changed to T1 or T4. When these last two tones T2 and FT3 are acquired, children often confuse the two, and this confusion continues on into the two- to three-word stage. Zhu & Dodd (2000) confirm the above findings examining the error patterns made by 129 monolingual Mandarin speaking children, aged 1;6 to 4;6 years. The findings above indicate that the high level tone (T1) is easier to acquire than contour tones (T2, FT3), and that the simpler contour tone (T4) is easier to acquire than complex contours (T2, FT3). Therefore it seems that the tone complexity scale (14) correctly predicts the order in which tones are acquired by Mandarin-speaking children.

Studies which have focused on the order of acquisition of Mandarin lexical tones in second language acquisition have reported different findings (Kiriloff 1969; Leather 1990; Lu 1992; McGinnis 1996; G-T.Chen 1974; Miracle 1989; Shen 1989; Zhao 1988; Q-H.Chen 1997; Elliot 1991; Guo 1993; Sun 1998; Zhang, H.2010).¹⁰ However, the majority of them have reported that L2 learners acquire the rising and/or dipping tones last (Kiriloff, 1969; Miracle, 1989; Leather, 1990; Elliot, 1991; Guo, 1993; Chen, Q-H, 1997; Zhang 2010). This is compatible with the prediction made by the Tonal Markedness Scale regarding rising and dipping tones.

The present study predicts that L2 learners' non-native tonal productions are constrained by the TMS. As noted earlier, two T3 sandhi rules are applied to all tone sequences containing T3 and the

¹⁰ As Sun (1998) summarized, the orders are "T1<T4<T3<T2" (T1 is acquired first, T2 is acquired last) as reported in Miracle (1989); "T2<T3<T1<T4" as stated in Shen (1989); "T1<T4<T2=T3" as in Leather (1990), "T1<T4<T2<T3" as in Elliot (1991) and Chen,Q-H (1997); "T4<T1<T3<T2".

current pedagogy regarding the “FT3-First” is problematic (as discussed in §2.2.2.2). Since these issues may interfere with the error and substitution rates of Tone 3, this study will only predict the ranking of “*T2>>*T4>>*T1”. The performance of the two allophones of Tone 3 will be discussed whenever relevant though. This study predicts that T2 will be more difficult for L2 learners to acquire than T4, which will be more difficult than T1. The effects of the TMS (*T2>>*T4>>*T1) may be manifested by error and substitution rates. That is, it is predicted that T2 will have the highest error rates, T4 will have the second error rates, and T1 will have the lowest error rates. On the other hand, it is predicted that the T1 will have the highest substitution rate (the rate at which T1 is used as a substitute), followed by T4, followed by T2 with the lowest substitution rate.

2.2.4.2. Obligatory Contour Principle (OCP)

The Obligatory Contour Principle (OCP) (Leben 1973, Goldsmith 1976, McCarthy 1986) has played an important role in formal phonology and has been used to explain various dissimilation effects. The OCP is used in describing the distribution of phonological units like tones and place features, and is also used to characterize many phonological processes, both in motivating the application of a process, and restricting its output.

(14) OCP (Obligatory Contour Principle) (McCarthy 1986:208)

Adjacent identical autosegments are prohibited.

The OCP is supposed to govern tone behaviors universally. A strong case of the application of the OCP in the native Mandarin tonal system is the T3 Sandhi process wherein a T3 changes to T2 if placed before another T3. Here I cite a representative account by Yip (2002), who also takes Tone 3 to be underlyingly specified as L and analyzes T3 Sandhi as a dissimilation process which is motivated by the OCP. As mentioned in §2.2.1, Yip (2002) assumes that a contour tone is composed of two or more level tones. For example, a rising tone is composed of a low component tone and a high component tone as it has been proposed for the African tonal system. The OCP as it refers to these component or constituent

tones will be referred to OCP (constituent). Yip (2002) refers to the dominating OCP constraint as the OCP (whole tone), meaning two identical whole contour tones are prohibited. The whole tone refers to an entire tone borne by a full syllable, a contour tone's component tones. Yip (2002) proposes a high-ranked positional Faithfulness constraint (FaithPrWdHead) which resists changes made to the second syllable of a word. For the sake of simplicity, the analysis (19) will not include this constraint, and will restrict itself to candidates which do not change the second syllable of the word. The constraints are defined below:

- (15) I-dent (Tone): let α be a tone-bearing unit (TBU) in the input, and β be any TBU corresponding to α in the output; if α has tone T, then β has tone T.
- (16) OCP (Whole tone) (Yip, 2002) : Identical whole tones cannot occur on adjacent syllables.
- (17) OCP (L): Identical whole Low (T3) tones cannot occur on adjacent syllables
- (18) *Rise (*T2): no whole rising tones (or component tones HL associated to a TBU) is allowed.

Below is a tableau for the input of T3 followed by T3, which is taken to be /L.L/ underlyingly.

(19) Tableau for the grammar of T3 Sandhi

/L.L/	OCP(L)	Ident (Tone)	*Rising	OCP(whole tone)
→LH.L		*	*	
L.L	*!W		L	*W

In this tableau, the candidate L.L is completely faithful to the input, but violates both the OCP (L) and the OCP (whole tone). The tone on the first syllable in the candidate LH.L is different from the input, violating Ident (Tone). However, the specific OCP constraint OCP (L) outranks Ident (Tone), picking the candidate LH.L as the optimal output because the violation of Ident (Tone) is less serious than the violation of OCP (L). As a result, the grammar of T3 sandhi could be generally represented by the following constraint ranking: OCP (L) >> Ident(Tone), *Rise, OCP (Whole tone).

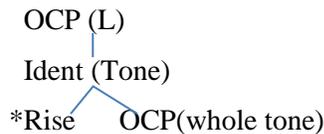
Other combinations of two identical tones adjacent to one another (i.e. T1-T1, T2-T2, T4-T4) remain faithful to the input. Therefore some Faithfulness constraint(s) must dominate some Markedness

constraints. The tableaux (20) below illustrate this point. The *Rise and OCP (whole tone) constraints are violated by the chosen surface forms (like LH.LH and H.H), so these constraints must be ranked low relative to Faithfulness constraints. Since H.H, LH.LH, and HL.HL all satisfy the high-ranked Faithfulness constraint Ident (Tone), they are chosen as the surface forms.

(20) Tableaux for T1-T1, T2-T2, and T4-T4, all of which keep their underlying tones as outputs in native Mandarin

/H.H/	OCP(L)	Ident (Tone)	*Rise	OCP(whole)
→H.H				*
L.H		*!W		L
/LH.LH/	OCP(L)	Ident (Tone)	*Rise	OCP(whole)
→LH.LH			**	*
L.LH		*! W	*L	L
/HL.HL/	OCP(L)	Ident (Tone)	*Rise	OCP(whole)
→HL.HL				*
L.HL		*! W		L

Figure 2.6. Hasse diagram of the ranking in Tableaux (20)



In this case, the effect of Markedness constraints such as the specific OCP (L) is visible in T3-T3 combinations, but the general OCP (whole tone) is masked by the Faithfulness constraints in other identical-tone combinations (T1-T1, T2-T2, T4-T4). Therefore the OCP (whole tone) is “inactive” in the grammar of Mandarin. The OCP is widespread in African languages, but in Chinese, it is relatively rare (low ranked) and only occasionally appears in the form of the OCP (L) in T3 Sandhi situations.

We predict that the general OCP (whole tone) is relevant in all L2 tonal production data sets, which would mean that the OCP (whole tone) is ranked relatively high in interlanguage grammars. It

would indicate that, although the tone pairs T1-T1, T2-T2, and T4-T4 are equally good in Mandarin, L2 learners disfavor all identical-tone-sequences, including T1-T1, T2-T2, and T4-T4. This preference may be manifested both in error and substitution patterns, i.e., significantly higher error rates and lower substitution rate of identical-tone combinations than expected in the L2 tonal productions. The OCP effects will be tested in both word-level and sentence-level experiments, the results of which will be reported in Chapters 4 and 5.

2.2.4.3. Tone-Position Constraints (TPC)

As we know, contour tones (pitch changes within a syllable) are more marked than level tones, implying that level tones are more widely distributed in world languages than contour tones. Because contour tones are more complex in nature than level tones, articulation and perception of contour tones require the Tone-Bearing Unit to have greater tone-bearing ability. While other complex segments, such as obstruents, allow overlapping oral constrictions, the articulation of contour tones relies on the state of a single articulator, the vocal folds. Because of the lack of overlapping articulators, contour tones require duration in order to be implemented. Therefore, contour tones with higher complexity are disfavored since they place a higher demand on the duration and sonority of the rhyme (Zhang, J. 2004).¹¹ The phonological distribution of contour tones has been of much theoretical interest in the literature (Goldsmith 1976; Bao 1990; Duanmu 1990; Yip 1989, 1995; Zhang, J. 1998, 2002, 2004a, 2004b).

Research focusing on positional licensing of contour tones conducted by Zhang, J. (2004a) claim that “phrase-final syllables and syllables in shorter words are the preferred bearers of contour tones, even though they are usually not privileged for other phonological contrasts.” Word-initial syllables, which have been shown to selectively license many other phonological contrasts (Steriade 1993,1995; Beckman 1997), “do not show up on the list of privileged contour tone bearers” because “the positional licensing

¹¹ According to Zhang, J. (2004a), the main perceptual correlate of tone is the fundamental frequency (F0), especially the 2nd, 3rd and 4th harmonics. Since sonorants possess richer harmonic structures than obstruents including the crucial 2nd, 3rd and 4th harmonics, sonorants are better tone bearers than obstruents.

behavior of contour tones is ... sensitive to the phonetic properties that are crucial to contour tones *per se*, namely duration and sonority. ” Final lengthening is the basis for considering the prosodic-final position as a relevant parameter influencing tonal-bearing ability and the phonetic literature has shown that the final syllable of a prosodic unit is subject to lengthening (Klatt 1975; Wightman et al. 1992). So, within a prosodic unit of the disyllabic word, we predict that word-final positions are better bearers of contours.

In this study we test the positional effects in the realization of L2 tonal productions. In Mandarin, contour tones (rising and falling) are phonologically contrastive and they cause lexical meaning different. Contour tones are freely distributed in Mandarin. That is, rising and falling tones are freely distributed in disyllabic word-initial and word-final positions. However, based on Zhang, J.’s (2004a) finding, we hypothesize that the L2 learners may be affected by the contour licensing constraint. Therefore, contour tones (T2, T4) may be easier for L2 learners to produce in word-final positions, compared to word-initial positions. We propose some constraints concerning the positional effects of contour tones:

- (21) a. * Contour-I >> * Contour-F (word-final positions (F) are better bearers of contour tones than word-initial positions (I))
b. * Contour-I: contour tones are not allowed at word-initial positions.
c. * Contour-F: contour tones are not allowed at word-final positions

(21a) can be expanded into two sub-rankings to fit this study, shown in (22):

- (22) a. * Rise-I >> * Rise-F:
Rising tones are more disfavored at word-initial positions than at word-final positions.
b. * Fall-I >> * Fall-F:
Falling tones are more disfavored at word-initial positions than at word-final positions.

Since the constraints used in the following analysis (and in Chapter 6) refer to whole tones, the labels T1, T2, T3, T4 and FT3 will be used in the tableaux hereafter to keep the Mandarin tone representations simple. Below is an example tableau illustrating how the ranking of (22b) works to picks

the surface form T1-T4 for the underlying target form T4-T4, a sequence found in many English speakers' L2 tonal productions in the experiments described later. For these participants, it seems that the OCP (whole tone) is ranked low in the ILG. Because of this, the OCP (whole tone) cannot be the reason why the target T4-T4 is mis-produced as T1-T4 by L2 learners. The tableau (23) shows how the TPC constraint “*Fall-I” interacts with other constraints, such as the TMS constraint “*Rise” (*T2) and the Faithfulness constraint, to determine the surface form T1-T4. Only candidates which either end or begin with T4 are included in the tableau. No candidate that keeps the initial T4 will win because of the high ranking of *Fall-I. The Faithfulness constraint “Id-Register” indicates that the register feature of the output tone should be identical to the register feature of the correspondent input tone. This Faithfulness constraint eliminates those tone sequences with different register features, such as T3-T4 ([-U] [+U]). Without Id-Register, T3-T4 would win out because *T3 is ranked lower than *T1 according to the TMS.

(23) Tableau for output T1-T4 (input T4-T4)

Input: T4-T4	*Rise	*Fall-I	Id-Register	OCP(whole)	*Fall-F
→ T1-T4					*
T2-T4	*!W				*
T3-T4			*!W		*
T4-T4		*!W		*W	*
T4-T1		*!W			L
T4-T2	*!W	*W			L
T4-T3		*!W	*W		L

The above tableau shows that the violation of *Fall-I is more serious than the violation of *Fall-F. Although T1-T4 violates *Fall-F, it satisfies the higher-ranked *Rise, *Fall-I and Id-Register, making it the winning candidate.

As discussed in section §2.2.4.1, even among the contour tones there is fall-rise asymmetry (Ohala 1978; Zhang, J. 2004b). It is predicted that the universal fall-rise asymmetry may involve some positional effects in the realization of L2 contour tones. Most studies on the phonetics of the fall-rise asymmetry are concerned with finding phonetic reasons, whether articulation-based or perception-based,

which might lead to this asymmetry. Survey studies (such as Ourso 1989 and Zhang, J. 2002, 2004b) reveal that falling tones are allowed in many more languages than rising tones, and even within a language, rising tones are often more restricted in distribution than falling tones. Although these studies provide phonetic reasons for the fall-rise asymmetry, very few explicit claims have been made about whether there is a fall-rise asymmetry specifically for prosodic-final positions. Zhang, J. (2004b) reports a case of the asymmetrical distributional property of rising and falling tones with regard to word-final positions in Mende, the major language of Sierra Leone. In Mende, a falling tone can occur on a short vowel in the final position of a polysyllabic word, while a rising tone cannot. The present SLA study regarding the production of Mandarin tones by non-tonal language speakers will not only observe the surface distribution of all tone types at all positions (both word-initial and word-final) in interlanguages, but will also pay particular attention to the realization of contours at word-final positions by L2 learners. Based on the previous studies regarding fall-rise asymmetry, we predict that this kind of asymmetry may also affect L2 tonal productions. That is, while the L2 participants produced both the rising and falling contour tones better in word-final positions than in word-initial positions in general, it may be the case that the falling tones are produced better than rising tones in word-final positions.¹² The positional effects are formalized as a Markedness constraint, seen in (24):

(24) * Rise-F >> * Fall-F:

Rising tones are more disfavored than Falling tones at word-final positions.

This study closely examines these positional effects for all tone types and tests for evidence of these fixed rankings regarding the positional effects of contours mainly in Experiment 1, an experiment focusing on disyllabic words. If the rankings in (21) are working in the L2 interlanguage grammars, we expect that the error rates of contour tones (both T2 and T4) would be significantly higher at word-final

¹² According to previous studies (Dreher and Lee, 1966; Kratochvil, 1971; Ting, 1971; Chuang et al., 1972; Howie, 1976; Nordenhake and Svantesson, 1983; Jongman et al. 2006), FT3 and T2 are the longest in duration whereas T4 is the shortest in native Mandarin. Here, temporal properties of native Mandarin tones are not taken into consideration because it's not clear whether the temporal characteristics of lexical tones are also valid in L2 learners' tonal productions.

positions than at word-initial positions. If the ranking in (24) is working in L2 tonal productions, the error rates of T4 at word-final positions would be significantly lower than those of T2. Although these rankings do not inform us of the behavior level tones will exhibit, we still predict that the level tones, T1 and T3, will be performed better than contour tones at all positions due to its lower tonal complexity. TPC effects will be confirmed in Experiment 2, a sentence-level test. Since this project also attempts to provide pedagogical implications to the field of Teaching Chinese as a Foreign Language, we will examine the L2 performance of all tone types at both word-initial and word-final positions (Experiment 1), and all positions or syllables in sentences (Experiment 2), to get a complete view of positional effects. By doing so, we can identify those positions which are most vulnerable (i.e. have the highest error rates) for specific tones and provide valuable pedagogical tools for Mandarin educators and learners. This study will allow us to determine which tones are most difficult to implement at which positions (syllables) in prosodic units (words, phrases and sentences). The error rates of all target tones will be examined and reported in Chapter 4 and 5.

This section has introduced three phonological universals which are predicted to constrain L2 tonal productions. Because of these constraints, we predict that the three interlanguage grammars possessed by English, Japanese and Korean speakers may share some similar error and substitution patterns regarding the TMS, the OCP, and TPC. However, we also predict that specific L1 prosodic structures may affect L2 tonal productions and lead to varying L1-specific interlanguage properties. The following two sections introduce the assumptions concerning the prosodic structures of three L1s (§2.3) and detail the L1 transfer predictions based on the prosodic structures of the three L1s (§2.4).

2.3. Assumptions of prosodic structures and the stimulus sentence

The goal of this section is to clarify the phonological assumptions of prosodic structures. This will inform us as to how these factors affect prosodic phrasing of stimulus sentences, thereby influence the tone assignments for the test materials by L1 speakers. The clarification is to prepare us for making

detailed predictions about L1 speakers' possible tone assignments to stimulus sentences in the next section §2.4.

The stimulus sentences used in the two phonological experiments are declarative sentences in Mandarin, all of which share similar syntactic and phonological structures. The three L1s used in this experiment differ in prosodic structure: American English) is a stress language in terms of word prosody, and a “head-prominence” language in terms of focus expression characteristics; Tokyo Japanese is a pitch-accent language in terms of word prosody and an “edge prominence” language in terms of focus structure; and Seoul Korean) is a non-stress non-pitch accent language in terms of word prosody characteristics and an “edge prominence” language regarding focus expressions. It is predicted that the L1-specific prosodic structures will affect the interlanguage grammars. We predict that L1 transfer will occur, so typical L1-like tone patterns will likely be found in the L2 tonal productions. Section §2.3.1 provides a brief introduction to the autosegmental-metrical model (**AM**) of intonation; section §2.3.2 outlines the factors affecting prosodic phrasing and the phonological representation of tone in sentences; and §2.3.3 gives information concerning the stimulus sentence (exemplified with the carrier sentence in Experiment 1) and its tone assignments by native Mandarin speakers.

2.3.1. Assumptions of prosodic structures

It is widely assumed that a prosodic structure consisting of prosodic constituency and prosodic prominence form the basis of phonological representation (Selkirk 2009). That is, both prosodic phrasing and prosodic prominence contribute to determining the phonological representation of tone in the sentence.

Following Selkirk (1978, 1984, 1986, 2009, 2011) and the autosegmental-metrical model (AM) of intonation developed by Pierrehumbert and her colleagues (Pierrehumbert 1980; Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988; see Ladd 1996, 2008 for extensive review), prosodic structure has a fundamental property: the *prosodic hierarchy*. The *prosodic hierarchy* refers to

an ordered set of prosodic category types which constitute possible node labels for prosodic structures. Since the representations of prosodic units vary in the literature on intonation, this study assumes that prosodic constituents include six levels of prosodic units, as displayed in (25). The prosodic units will assumed to be the same for all four languages involved in this study:

- (25) Prosodic constituents (prosodic hierarchy)
Intonational Phrase (ι , IP)
Phonological Phrase (φ , PhP)
Prosodic Word (ω , PW)
Foot (ft)
Syllable (σ)
Mora (μ)

The level of Intonational Phrase (ι , IP) corresponds to the breath group (BG) and prosodic group (PG) in the literature, especially those on Chinese prosodic structures. The level of Phonological Phrase or prosodic phrase (φ , PhP) consists of a wide range which includes *Intermediate Phrase* (ip, which is often used to describe the English intonation system) and *Accental Phrase* (AP, which is usually used to describe the Japanese and Korean prosodic structures). AP is the smallest rhythmic/prosodic unit above the word (Selkirk 2009, Jun 2005). In this study, I will use the same set of letters in parentheses to annotate the prosodic phrasings. However, in some discussions, I will refer to the prevailing names of the prosodic units for specific languages, such as “ip” for English, and “AP” for Japanese and Korean.

2.3.2. Factors affecting prosodic phrasing

The prosodic phrasing of a sentence is affected by a lot of factors, such as syntactic structure, semantic properties, focus structures and the length of the constituents of a sentence, etc. Here we discuss two major factors we believe closely relate to the prosodic phrasing of the stimulus sentences used in this study--the role of syntactic structure and focus structure (or information structure).

2.3.2.1. Syntactic structure

Syntax mediates between phonology and semantics. Previous studies show that, in the context of generative grammar, the presence or absence of various types of phonological phenomena at different

locations within a sentence correlates with differences in syntactic structure (Chomsky and Halle 1968, Selkirk 2011). Selkirk (2009, 2011) proposes an initial sketch of a Match Theory of syntactic-prosodic constituency correspondence. This theory defines the simplest possible correspondence between three levels of constituents of syntactic structure and constituents of prosodic structure, namely a match between clauses, phrases, and words. It is worth noting that, Selkirk (2009, 2011) also states that some other factors, such as phonological Markedness constraints, may influence the syntactic-prosodic correspondence. That is, in the spirit of OT, if some phonological Markedness constraints on prosodic structure are ranked high, it may lead to violations of Match constraints and produce instances of non-isomorphism between syntactic constituency and phonological domain structure. This interactive effect is due to the idea that “influences on the phonological domain structure of a sentence are highly modular; it cannot be accounted for by the theory of syntax alone” (Selkirk 2011, p.478).

Among the three Match constraints, the match constraint on clause has a strong motivation and is supported by the data in Selkirk (2009). Selkirk (2009) provides evidence from the syntactic grounding of prosodic constituent structure (particularly, that of clauses) and claims that “there is indeed an effect of clause structure on sentence prosody.....clauses in syntactic representation are mirrored in phonological representation by the presence of corresponding prosodic domains”. These clause-grounded prosodic constituents are referred to as intonational phrases (IP), as is common practice in the literature (Ladd 1996, Nespor and Vogel 1986). Following this research, I assume the phrasing of Intonational Phrase for the stimulus Chinese sentence used in this study is heavily affected by the syntactic structure. Since the carrier sentence in Experiment 1 and test sentences in Experiment 2 share very similar syntactic and focus structures, I only take the Experiment 1 carrier sentence as an example to discuss the prosodic phrasing issues. Figure 2.7 shows the syntactic structure and (26) shows that the clause forms a separate Intonational Phrase from the subject and the following verb. The disyllabic test word is labeled as XX.

Figure 2.7. Syntactic structure of the stimulus sentence

exception of the test words. Thus we take the test words as the location of contrastive focus and predict they would affect the L2 tone implementation. In Experiment 2, all test sentences are designed to contain narrow focuses, which we predict will affect phonological phrasing and tonal productions.

I follow Selkirk (2007)'s definition of "contrastive focus," which is used to designate the status of a constituent in sentences like "I only give one to Sarah". The meaning of the sentence includes a specification that there exist alternatives to the proposition expressed by the sentence which are identical to that proposition except for different substitutions for the contrastively focused constituent. In this sentence, "Sarah," carries contrastive focus.¹³ Some researchers of the focus-prosody interface have claimed that the principles of grammar do not assign contrastive focus any distinctive prominence (Ladd 1996, Gussenhoven 2004), while others have proposed that contrastive focus is subject to a special grammatical principle for the assignment of phrase stress which can lead to a grammatically represented prominence distinction between contrastive focus and non-contrastive constituents (Truckenbrodt 1995, Rooth 1996b, Selkirk 2002, 2006, Féry and Samek-Lodovici 2006, Büring 2006). This study predicts that the test words in Experiment 1 are assigned prosodic prominence. The narrow focuses in the Experiment 2 stimulus sentences demonstrate stronger prominence, which will affect phonological phrasing and tonal implementations by L1 speakers. That is, the focus structure will potentially violate the Match correspondence constraints mentioned above.

For the pitch expressions for prosodic prominence (such as phrase accent in English and sentence-level prominence, focus), we adopt and recast the idea of tone and prominence combinations discussed in De Lacy (2002). De Lacy (2002) proposed a family of constraints in a fixed ranking concerning the relationship between tone and prosodic positions. He assumes a Tonal Prominence scale in which higher tones are more prominent than lower tones; that is, H>M>L. This scale combines with structural scales to form two sets of constraints in a universally fixed ranking. One set regulates the

¹³ This type of focus is referred to variously as contrastive focus, identificational focus, alternatives focus, or simply focus (Jackendoff 1972; Jacobs 1988; Krifka 1992; Rooth 1992; Rooth 1996a, Kiss 1998, Kratzer 2004, as cited in Selkirk 2007).

relation of tone to heads: *HD/L>>*HD/M>>*HD/H, which expresses the empirical claim that foot heads and higher tone have an affinity for one another; the other set deals with tone on non-heads: *NON-HD/H>>*NON-HD/M>>*NON-HD/L, which means non-head feet and lower tone have an affinity for one another. In this study, we adopt the Tonal Prominence scale and the idea of stress-driven tone. We recast the idea in the present study and propose a fixed ranking: *HD/[-U]>>* HD/[+U] in which high registered ([+U]) tones such as T1, T2, and T4 are attracted to metrically prominent positions, such as the stressed syllables in words/phrases and the focused or stressed syllables in sentences. That is, we predict in the L2 tonal productions, L1 speakers may use high registered tones ([+Upper] or [+U]) at those “stressed” or “focused” syllables to substitute for the target Mandarin tones, thus making tonal errors. If English speakers transfer their dominating L1 stress patterns to their interlanguage, for example trochaic foot, then we will see a lot of [+Upper] [-Upper] (or High-Low registered) tone sequences in disyllabic words. At the sentence level (as in Experiment 2), this study predicts that the narrow focused constituents will attract high registered tones in the L2 tonal productions. For example, the mono-syllabic focuses and the syllables with “Stress Focus” or the strongest prominence within multiple-syllable focused constituents in test sentences may be prone to bearing [+Upper] tones.¹⁴

Following is a discussion about each of the three L1 groups’ phrasing of stimulus sentences. However, the concrete prosodic realization of focus varies depending on language-specific phonological constraints, therefore the intonational patterns on the test words may be different from one another. Native speakers of the target language, Mandarin, usually mark sentence prominence by manipulating pitch range (i.e, expand pitch range on focus, narrow post-focal pitch range) and durations, but do not change the value of lexical tones (Ladd 2008, Jun 2011, Flemming2008, Chen 2004). However, non-tonal language speakers mark focus in different ways from tonal language speakers. For example, head-

¹⁴ “Stress Focus” is proposed by Selkirk (2007, 2009) and assumes that a Focus-marked constituent is required to contain the greatest stress prominence within some relevant domain. This assumption is related to the long focused constituents with more than two syllables in Experiment 2. In order to satisfy “Stress Focus”, the phrasal stress that is produced will induce the presence of a phonological phrase edge adjacent to that stress, and therefore may introduce a phonological phrase domain structure that does not correspond to a syntactic phrase (see Selkirk 2002, 2009).

prominence languages such as English, German, and Greek, mark prominence through the use of pitch accent (i.e., salient pitch movement on the stressed syllable), and the focused word receives a nuclear pitch accent while ‘deaccenting’ the following word(s).¹⁵ However, for some edge-prominence languages, such as Korean and Japanese, which are also called ‘dephrasing’ languages, mark prominence by manipulating the phrasing (i.e., grouping) of words. Because of this, the focused constituents may be assigned L1-specific phrasal intonation patterns. So we expect slightly different prosodic phrasings of the stimulus sentence and tone assignments on the test words by different types of learners, which in turn will affect the L2 tonal production and cause different tonal errors. The following section §2.3.3 displays the proposed prosodic phrasing of the target stimulus sentence by native Mandarin speakers. The predictions of L1-specific phrasing and tone assignments are discussed in the section §2.4.

2.3.3. Prosodic structure of the Mandarin stimulus sentence

The interaction of lexical tone and utterance-level prosody (intonation) in Mandarin is a recurrent theme in the study of Chinese phonology (Shen 1985; Lee 2005; Yuan 2004, among others). Traditionally, people follow the “overlay” models of F0 (Ladd 1988; Bolinger 1964) in the study of tone-intonation interactions. In an “overlay” or “superposition” model, we treat the relationship between tones and intonation in Mandarin as “small ripples riding on large waves” (Chao 1933). Within an AM model of intonation, the interaction between intonation and lexical tone should be described in terms of local phonological events (Ladd 2008). In Mandarin, as mentioned above, the sentence-level prominence (focus) or phrasal boundaries are believed to be expressed mainly by expanding or narrowing pitch range, but not changing the Mandarin lexical tone identity (Chen 2004; Ladd 2008; Flemming 2008). That is, generally speaking, no matter what prosodic events undergo with the sentences, the lexical tone identity should be preserved in Mandarin. Due to the high density of pitch specifications at the lexical level in Mandarin, this study assumes the original lexical tones remain in declarative sentences. Therefore, we

¹⁵ The term “pitch accent” is used in the literature of English phonology very often. It refers to the phrasal level accent or prominence. It is different from the lexical “pitch accent” in Japanese, which is a property stored in the lexicon.

assume the target tone identities in our study, including the test words embedded at sentence-medial positions in Experiment 1 and those in the declarative sentences in Experiment 2, will be preserved.

According to the Mandarin ToBI prosodic notation system (as in Jun 2005) and the discussion about the nature of prosodic structure above, we assume Mandarin has the following prosodic constituents: the breath group (BG) or IP, the prosodic phrase (PP, φ), the prosodic word (PW), the foot, the syllable, and the mora (Peng et al. 2005; Tseng and Chou 1999; and Tseng et al. 2005).

Tseng (2003) argues that phrase-level intonation in Mandarin is not as significant as that in non-tonal languages, such as English. Thus, Peng et al. (2005), Tseng and Chou (1999), and Tseng et al. (2005) treat the prosodic constituent above the prosodic phrase as a breath group (BG). A BG is what a speaker utters with one full breath. It is analogous to the IP discussed in §2.3.1.

The prosodic phrase is the prosodic constituent immediately above the prosodic word in the prosodic hierarchy. A prosodic phrase is usually composed of two or three prosodic words, even though sometimes a single prosodic word may constitute a prosodic phrase. This constituent approximately corresponds to other terms in different theories, such as the intermediate phrase (Beckman & Pierrehumbert 1986), or the phonological phrase (Nespor & Vogel 1986). It will be labeled as PhP as discussed in (25).

The prosodic word is the minimal unit of articulation in normal communication and it usually consists of a single lexical word together with associated unstressed functional words (i.e., auxiliaries, determiners, conjunctions, or prepositions) (Wheeldon 2000). In Mandarin, prosodic words are mostly disyllabic or trisyllabic, and occasionally monosyllabic when the morpheme is uttered in isolation as an imperative. Prosodic words may or may not coincide with lexical words (Wang 2003). Wang (2003) offers an example that *tā měitiān* ('he everyday...'), consisting of two lexical words *tā* and *měitiān*, can be uttered as one or two prosodic words in actual speech. Following this practice, I assume the monosyllabic subject and the disyllabic verb *wǒ juéde* ('I feel ...') in the stimulus sentence of this study

form a prosodic word; the test word and the following monosyllabic particle form a prosodic word; and the compacted adjective phrase containing a monosyllabic adverb *hěn* “very” and monosyllabic adjective *hǎo* “good” also form a single prosodic word.

Based on our discussion in the section above, (27) is the phonological phrasing of the stimulus sentence for Mandarin speakers. Following Shih (1997) and Chen (2004) Mandarin focus inserts an intonational phrase boundary at the left edge of a focused constituent, so there is an IP boundary before the test words. This is compatible to the Match Theory in Selkirk (2011) since the test words also begin a clause in the sentence. The test word, a modifier, and the following noun *dōngxi* “thing” form an NP which serves as the subject of the clause, and is followed by the AP (adjective phrase) *hěn hǎo* “very good”. According to the Match Theory (Selkirk 2011), I propose a PhP boundary between the noun phrase *XX de dōngxi* and adjective phrase *hén hǎo* in the clause.

(27) Phonological phrasing of the Mandarin stimulus sentence for Mandarin speakers

Chinese pinyin:	Wǒ	Jué de	XX de	DōngXi	Hén Hǎo.
Chinese characters	我	觉得	XX 的	东西	很好
Gloss:	I	feel	XX -particle	things	very good.
Phrasing:	ɿ{ φ[ω]} ɿ{ φ[ω	ω] φ[ω] }

The sentence consists of two Intonation Phrases with the boundary aligned to the major syntactic boundary. This boundary is also aligned with the left edge of the focus word. This major prosodic break is shared by all L1 speakers.

2.4. The prosodic characteristics of the three first languages and predictions concerning each

To address the hypothesis regarding L1 transfer of typical L1 melodic patterns, this section introduces the prosodic characteristics of the three L1s: English, Japanese, and Korean, and makes specific predictions concerning tonal patterns on the test words and sentences according to the prosodic structures of these L1s. I do not intend to provide a detailed and complete description of the prosodic phonology of each language but instead focus on the aspects related to Mandarin tone acquisition. We

discuss the intonational structures in each section from a cross-linguistic perspective, in order to make comparisons for similar phenomena if necessary. The first subsection §2.4.1 considers the issue of “representation of intonation,” followed by brief introductions of the lexical and post-lexical level prosodic structures of English, Japanese, and Korean in the following three subsections. The prosodic structure and the phrasing of stimulus sentences used in Experiment 1 are included in sections § 2.4.2 to §2.4.4 for these three L1 speaker groups. Since the syntactic structure of the Experiment 2 stimulus sentences is very similar to that of the stimulus sentence in Experiment 1, the three L1-specific prosodic phrasings of Experiment 2 stimulus sentences will be shown together in Chapter 5.

2.4.1. Representation of intonation

The question of whether pitch contours should be represented in terms of contours, such as rising and falling, or whether these surface contours are a series of level pitches, such as high and low, is highly debated in linguistic literature. Many linguists have treated intonation contours as gestalts (i.e., the contour as a unit) (see Bolinger (1951), Jones (1972), Cooper and Sorensen (1981), Hirst and Di Cristo (1998), Grabe et al. (2003), and Xu (2005)). These researchers claim that intonational contours should be seen as holistic and directly reflecting of certain functional or structural aspects of speech, such as the distinction between statements and questions or the distinction between levels of phrasing. However, it has been claimed that the gestalt approach “cannot account in a satisfactory manner for either intonational meaning or intonational form.” Also, “empirical evidence as well as considerations of representational parsimony strongly suggests that these linguistically-relevant aspects of F0 contours are best represented as levels rather than movements.” (Arvaniti 2010, p.774-775) A number of convincing analyses of contour tones in terms of level pitches have been presented, such as in Goldsmith (1976). Analyses of English intonation contours in terms of level pitches have been proposed by Pike (1945), Trager and Smith (1951), and later by Liberman (1975) and Pierrehumbert (1980). In all these analyses, the phonetic pitch contours of English intonation represent the surface realization of an underlying pattern of level pitches, with pitch changes merely acting as the necessary transitions from one level pitch to the next.

This paper thus assumes that the contours in all intonational languages can be decomposed into level tones (or tone features).

Pitch is significant in all languages, but as we know, it plays very different roles in specific languages; pitch patterns may be specified either at the lexical-level or at the phrasal/sentential-level, or at both the lexical and phrasal/sentential levels, resulting in more or less dense tonal specifications. As discussed in §2.2, pitch is mainly used to specify lexical meaning in the target tone language Mandarin, although tonal languages also have intonation structures. That is, in tonal languages, most syllables are lexically-specified because tone affects word meaning. In typical intonational languages like English and many other European languages, pitch features are mainly post-lexical or intonational. That is, the F0 contour is mainly specified at the post-lexical level by means of a complex interplay between metrical structure, prosodic phrasing, syntax, and pragmatics. At the word level, accents exist in English and are phonetically manifested by “stress,” but are not automatically specified with fixed pitch types. Because of this, English is called a “stress-accent language.” The detailed prediction of typical English melodic pattern borne by the test words is provided in §2.4.2. In languages often referred to as lexical pitch accent languages, such as Japanese, Swedish, and Serbian, pitch features operate on both lexical and phrasal/sentential levels. These languages share features of tonal and intonational languages in that pitch is to some degree lexically associated, and the F0 contour (edge tones) can also be specified at the phrase and sentence level. Differently from tonal languages, words in lexical pitch accent languages have at most one syllable which is lexically specified (Arvaniti 2011). That is, there is only one fixed tone type associated with accented syllables in Japanese. On the other hand, Japanese is different from English in that the lexical accent in words is marked only by pitch movement and not at all by stress, whereas English lexical stress “uses to a greater extent material other than pitch” (Beckman, 1986). Thus the surface intonation pattern in Japanese is an interaction of lexical level pitch accent and post-lexical level intonation. We detail the predictions of the typical melodic pattern over the test words in §2.4.3. In languages such as Korean, pitch patterns do not distinguish lexical meanings, so they are not tonal

languages. At the same time, according to Jun (1993, 2005), neither stress nor pitch accent exists at the word level, making Korean different from other non-tonal languages, such as English and Japanese. Some researchers claim that Korean is a stress language featuring an iambic foot (Ko 2010, 2013). We follow Jun (1993, 2005) in predicting the tone assignments on the test words in section §2.4.4 and consider the claim that Korean is a stress language in discussions in Chapter 4.

We believe the various densities of pitch specifications in different L1s will lead to different error patterns and substitution patterns in the SLA of Mandarin tones. Following are the introductions to L1 prosodic structures and the prediction of typical L1 tone patterns. Each introduction to the L1 prosodic structure includes roughly two parts. One is the prosodic phrasing of stimulus sentence by the L1 speakers depending on the L1-specific prosodic characteristics. The other is tone assignments for the test words which leads to the most likely tone patterns L1 speakers might employ in reading the test materials if L1 transfer occurs.

2.4.2. English prosodic structures and predictions

English and Mandarin share similar prosodic constituents. Following Nespor & Vogel (1986) and Selkirk (1984), English has the following prosodic constituents: the Intonation Phrase (IP), the Intermediate Phrase (ip) which corresponds to the PhP in (25), the Clitic Group or prosodic word as in (25), the foot, the syllable, and the mora.

English is known as a “stress-accent language” in terms of word prosody. This means that the location of metrically strong syllables in a word is determined at the lexical level, and that the lexical accent in English regularly appears as stress, a cluster of phonetic properties that includes increased intensity and duration as well as various spectral correlations such as the adjustment of pitch shapes. Liberman’s notion of metrical phonology suggests that linguistic prominence consists of a relation between strong and weak nodes in a binary-branching tree structure (Liberman 1975; Liberman and Prince 1977). There are two kinds of stress patterns: weak-strong (iambic) and strong-weak (trochaic).

Diagram 2.3 illustrates the two kinds of stress patterns as cited in Ladd (2008, p.55.) The two kinds of stress patterns in English are contrastive. However, according to Delattre (1965), most English words (74%) are typically trochaic.

Figure 2.8. Metrical grid for a bitonal metrical foot with one stressed syllable



In English the major pitch movements often accompany stressed syllables. In other words, these word-level metrically strong syllables (or ‘primary-stress’ syllables) often serve as docking sites to which phrasal-level pitch accent is associated at a post-lexical level.¹⁶

In the mainstream American English ToBI transcription system (Beckman, Hirschberg, and Shattuck-Hufnagel 2005), the equivalent of the prosodic phrase or phonological phrase (PhP) is the *intermediate phrase* (ip), which is the domain of phrasal stress in English. An English ip (or PhP) usually includes two content words on average, ranging from one to four content words (Jun 2005). This agrees with Ueyama (1998)’s finding that there are 5-6 syllables for every English ip (or PP). A well-formed ip in English consists of one or more pitch accents, which are aligned with stressed syllables (syllable alignment is usually indicated by “*”) on the basis of the metrical pattern of the text, plus a simple H or L which characterizes the phrase accent (indicated by a high “-”). The phrase accent spreads over the material between the last pitch accent of the current intermediate phrase and the beginning of the next, or the end of the utterance.

Intonational phrases (IP) in English are composed of one or more such intermediate phrases plus a boundary tone. It falls exactly at the phrase boundary (Hirschberg et al. 1987a), or, “in some unmarked sense lines up with the clause” (Selkirk 1978, 1984). The H or L boundary tones (labeled as H%, L%)

¹⁶ The stronger degree of stress in a word is called “primary stress” while “secondary stress” refers to the weaker of two degrees of stress in the pronunciation of a word.

indicate the relationship between the current utterance and the following one. The H boundary tone indicates that the speaker wishes the hearer to interpret an utterance with reference to the following one, whereas the L boundary tone does not convey such directionality (Pierrehumbert & Hirschberg 1990).

2.4.2.1. Prosodic phrasing of stimulus sentences for English speakers

English and Mandarin share very similar major syntactic structures, such as SVO word order. Similar to Mandarin, English has “Modifier + NP head” (such as “a new book/a Math book/ a used book”) structure. Following Selkirk (2009), I assume the clause in the stimulus sentence forms an IP. Since English speakers read the same stimulus sentences as the other L1 speakers, I assume the test words in Experiment 1 have contrastive focus for English speakers. Thus, we predict that English subjects are very likely to put metrical prominence or pitch accent on the disyllabic test words.¹⁷ According to Selkirk (2002), the pitch accent of contrastive focus in English is followed by a PhP break, marked by both an L⁻ phrase accent and temporal disjuncture. Therefore, a PhP boundary is assigned after the contrastive focus in the stimulus sentence. The test word and the following monosyllabic particle thus form a PhP with the pitch accent at the left edge. The diagram (28) shows the prosodic phrasing of Experiment 1 sentence for English speakers. The phrasing is very similar to that of Mandarin speakers except for the location of PhP boundary, which, for English speakers, immediately follow test words.

(28) The prediction of English speakers’ phonological phrasing of the stimulus sentence:

Chinese pinyin:	Wǒ	Jué de	XX de	DōngXi	Hén Hǎo.	
Chinese characters	我	觉得	XX 的	东西	很好	
Gloss:	I	feel	XX -particle	things	very good.	
Phrasing:	↑{ φ[ω]} ↑{ φ[ω] φ[ω	ω] }

2.4.2.2. Tone assignment

Intonation is a pitch contour aligned with a linguistic text. This study assumes that intonation is independent from the linguistic text in English. This is because the same sentence, with the same stress

¹⁷ This is a phrase level phonological event, which should be distinguished from the lexical pitch accent in Japanese.

pattern, can be said with many different melodies, and the same melody can occur on many different texts. It is widely accepted that in English, the tones (H, L) are loosely associated to TBUs. Sometimes, one, or two or even three tones could be linked to one TBU (see examples including tri-tonal contours on a single syllable in Gussenhoven 2004, p. 274), sometimes one tone may be linked to one or two or even more syllables, and sometimes, tones can even “exist that do not have an association to a TBU” (Gussenhoven 2004). In addition, in many or most cases, alternating Hs and Ls are found in English sentence intonation patterns. So I predict that if L1 transfer occurs, not only H and L levels but also many rising and falling contour tones where the HL or LH tones associated to one TBU could be found in English speaker’s L2 substitutions when errors occur. That is, unlike Japanese and Korean speakers who usually have single tones associated to TBUs, English speakers may produce more contour tones in the L2 production. I will make a comparison between the proportion of contour substitute tones in the English speaker data and those of the Japanese and Korean speakers’ datasets. I predict that English speakers’ proportion of contour substitute tones is very high. In the spirit of OT, this means the constraint ‘*contour’ is ranked very low for English speakers.

With the background information of tone-text association in mind, we predict the most typical tone realization of the disyllabic test words (as in Experiment 1) by English speakers in this section. As discussed above, the test word and following particle form a prosodic phrase (contrastive focus). Therefore the test word is very likely to be the site of phrase level accent or prominence. As indicated in Beckman and Pierrehumbert (1986), there are six types of pitch accent shapes in English, namely H*, L*, H*+L, H+L*, L*+H, and L+H*. It is widely accepted that the melody a speaker employs can communicate some semantic or pragmatic information (Hirschberg 1987b). However, the stimulus sentences in the reading lists of Experiment 1 are all discrete sentences without any context, and only contrast in the test words used. We thus assume there is no specific semantic or pragmatic information to be communicated and the speakers will use the most common accent to express simple contrastive focus. According to Hirschberg & Pierrehumbert (1987), the most frequently-used accent in English is a high

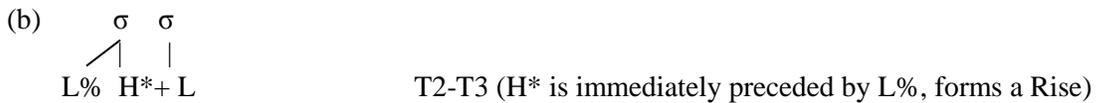
tone, which comes out as a peak on the accented syllable and is represented as H*. (This is compatible to De Lacy's (2004) claim regarding the tone prominence scale.) Therefore, we narrow down the possible pitch accent types to those accent shapes containing H*: (29) L+H*, (30) H*+L and (31) H*. It is noted that the L and H tones in intonation languages indicate relative highs and lows in the intonation contour, but are not strict phonetic realizations of lexical tones as they are in tonal languages. For these accent shapes, several counterpart Mandarin tone sequences are proposed below to imitate the L1 intonation patterns on test words.

For accent shape (29), I propose two tone sequence. The sequence shown in (29a) is a straightforward imitation of L+H*. (29b) predicts the situation where the accent is immediately followed by a phrasal tone of some sort, such as accent L⁻, so the second syllable may sound like a falling tone. The phrase accent in English is a tone that fills the space after the last pitch accent in a phrase (Pierrehumbert 1980, Beckman and Pierrehumbert 1986).

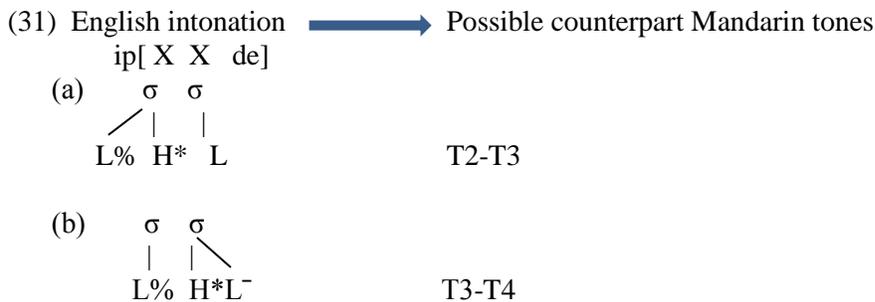
- (29) English intonation  Possible counterpart Mandarin tones
- ip[X X de]
- (a) σ σ
 | |
 L+H* T3-T1 (simple Low-High level sequence)
- (b) σ σ
 | /\
 L+H* L⁻ T3-T4 (the H* is immediately followed by L⁻, forms a falling tone)

For accent shape (30), two counterpart Mandarin tone sequences (30a) and (30b) are proposed to imitate the intonation shape. (30a) is a simple High-Low sequence. Since the test word initiates the ip, I propose an L boundary tone immediately preceding the accent which is also linked to the first syllable and forms a rising tone on the first syllable in (30b).

- (30) English intonation  Possible counterpart Mandarin tones
- ip[X X de]
- (a) σ σ
 | |
 H*+L T1-T3 (simple High-Low sequence)



For the single-tone accent case (31), two counterpart Mandarin tone sequences (31a) and (31b) are proposed, but they make very similar predictions as those made for structures (30b) and (29b). Following the description of a single-tone accent H* pitch contour in Beckman and Pierrehumbert (1986, p.258), where the penultimate syllable of the word (and also the phrase) is stressed, the F0 peak falls near the end of the associated stress syllable. Therefore the stressed syllable demonstrates a pitch rise and is followed by an L phrase accent which may sound like a Mandarin T2-T3 sequence. In the case where the single-tone accent H* is aligned with the phrase-final syllable (as in 31b), it is predicted that 1) the H* will share a phrase final syllable with an L⁻ phrase accent, 2) the accented syllable will be lengthened, 3) the peak will be earlier in the syllable, and 4) the F0 contour during the latter part of the syllable will be falling (Pierrehumbert 1980, p.72). In this case, the first syllable may be occupied by the L boundary tone while the single-tone H* and a L⁻ will share the final syllable. Therefore the word is predicted to sound like a Mandarin T3-T4 tone sequence.



Broselow et al. (1987) discussed the “the unmarked pattern for declaratives in English, traditionally called ‘rising-falling,’ involves a rising pitch on the so-called ‘tonic’ syllable, the rightmost pitch-accented syllable, followed by a fall on all material following the tonic syllable” (p.353) . I also

include this most ‘unmarked pattern’ in the prediction of English speakers’ possible L2 productions of the test words and propose that T2-T4 will be one of the typical L1 pitch pattern.

We also predict that there will not be many double stressed or double unstressed patterns in the L2 production, because English disyllables usually contain a strong and weak syllable (not both strong or both weak). That means there will be very few identical tone combinations in English speakers’ productions (this prediction overlaps with the prediction made concerning effects of the OCP). In short, when the L1 transfer occurs, there will be high occurrence rates (or bigger proportions in the L2 tonal productions than other tone sequences) of tone sequences which are similar to typical English pitch patterns over a focused disyllabic word, such as T3-T1, T3-T4, T1-T3, T2-T3 and T2-T4. That is, when the target tone sequences are these tone sequences, their error rates will be very low. If the target tone sequences are not these tone sequences, error rates will be high. When errors occur, English speakers will very likely substitute the target tones with these L1-like tone sequences.

2.4.3. Tokyo Japanese (Japanese) prosodic structures and predictions

Following Venditti’s (2005) description of Tokyo Japanese prosodic structures, we assume the following prosodic constituents in Japanese: the Intonation Phrase (IP), the Accentual Phrase (AP) which is analogous to the general Phonological Phrase (PhP), the prosodic word, the foot, the syllable, and the mora.¹⁸ Japanese is considered a pitch accent language in which words can be accented or unaccented. Pitch accent is a lexical property of a word. In Japanese, “pitch accents are the most straightforward component of an intonation contour. They have a fixed shape consisting of a sharp decline around the accented syllable and it is usually analyzed as a drop from an H tone to an L” (Beckman and Pierrehumbert 1986). The alignment between the HL sequence and the text is straightforward and the placement of the accent is lexically contrastive. The lexical pitch accent can be annotated as H*L. Both

¹⁸ Pierrehumbert and Beckman (1988) and Nagahara (1994) assume four levels of prosodic organization in Japanese: Utterance, Intermediate Phrase (ip), Accentual Phrase (AP) and Phonological words (pw). Venditti (2005) eliminates the Intermediate Phrase. Itô, J. and Mester, A.(2007) proposes that ip and AP are recursive instances of the category PP.

English and Japanese have lexical level accent, but their phonetic realization of the lexical accent is different. Japanese lexical pitch accent is mainly achieved by the changes of pitch movement (hence “pitch-accent language”), while English stress is achieved through more phonetic variables, such as increased intensity and duration (hence “stress language”) (Beckman 1986). Japanese is a mora-timed language which means the TBU is mora. However, the syllable is the bearer of accent (Kubozono 1999).

The intonation phrase (IP) serves as the domain of catathesis, in the sense that the lowering effect triggered by a pitch accent is delimited by intonation phrase breaks. The focus structure of a sentence requires a particular intermediate phrasing pattern in which every focus must be the leftmost constituent of some intermediate phrase (ip). The accentual phrase can have at most one accent, and it will be preceded by an L tone linked either to the right edge of the preceding accentual phrase or to the left edge of the utterance. It is very common that stem-particle or stem-suffix combinations result in a single accent or in no accent, forming single accentual phrases (Nagahara 1994, p.78). Many studies, such as Venditti (2005), show that Japanese AP is tonally defined by a rise to a high around the second mora, and subsequent gradual fall to a low at the right edge of the phrase. That is, the defining mark of the AP is the presence of two delimitative tones whose occurrence is determined solely by the prosodic phrase structure of the utterance. One is a high tone, the Phrasal H which is phonologically associated with the second sonorant mora, unless this conflicts with the lexical association of an accent H to the first mora. The other delimitative tone is a low boundary tone that occurs at the beginning of every utterance and at the end of every accentual phrase.

In Japanese, pitch patterns are specified at both lexical and phrasal/sentence levels. The prediction of which tone sequences are most likely to occur on the stimulus disyllabic test word is complex. We focus on Japanese speakers’ phonological phrasing of the stimulus sentence first, and then work out the Mandarin tone sequences imitating Japanese intonation patterns at the AP initial positions.

2.4.3.1. Japanese Speakers’ Prosodic Phrasing of Stimulus sentence

Nagahara (1994) formalizes the principles governing well-formed phonological phrasing. According to Nagahara (1994), there are two main organizational schemes for well-formed phrasing: with reference to the syntactic and focus structure of a sentence, and with respect to the interrelationships among prosodic categories. Japanese has the same structure of NPs as Mandarin where modifiers precede NP heads. Since all L1 learners read the same stimulus sentences, we assume Japanese speakers also take the test words as contrastive focus. I also assume Japanese speakers share the major syntactic boundaries around a clause which leads to a major phonological boundary before the clause, although native Japanese has the major syntactic structure of SOV. This would align with the boundary at the left edge of focus.

Following Nagahara (1994), “a focused constituent is always preceded by an intermediate phrase break” (P.31) and the requirements for well-formed focus phrasing can be expressed by two constraints: the FOCUS-LEFT-EDGE and FOCUS-TO-END constraints. The FOCUS-LEFT-EDGE constraint requires that the left edge of a focused constituent be aligned with the left edge of the intermediate phrase. The FOCUS-TO-END constraint requires that there be no intermediate phrase boundary intervening between any focus constituent and the end of the sentence. This is why Japanese (and Korean) are called “de-phrasing language (s)” (Ladd 2008). There is only one focus constituent in the stimulus sentence and the effect of these two requirements is such that the test words in the present study is preceded by a phrasal boundary (an AP boundary according to the model of Venditti (2005), or both an ip and an AP boundary following Nagahara (1994)). The rest of the sentence is phrased into a single AP. Actually, no matter how the rest of the clause is phrased, the test words are located at an AP-initial position. We will adopt the pitch patterns of typical Japanese APs in the following predictions.

The theoretical prosodic phrasing stated above closely follows the prosodic structure of native Japanese. However, the Mandarin stimulus sentence in this study is new to Japanese learners and we do not expect fluency, especially with the varying test words. This expectation is confirmed by noticeable pauses following the test words in actual research data recordings. Thus we also propose a possible AP boundary after the test word and consider this possible boundary in the predictions. (32) shows the

prosodic phrasing of stimulus sentence (the possible AP boundary is in gray) by Japanese speakers, which is similar to that of English speakers.

(32) The predicted phonological phrasing of the stimulus sentence for Japanese speakers:

Chinese pinyin:	Wǒ	Jué de	XX de	DōngXi	Hén Hǎo.			
Gloss:	I	feel	XX –particle	things	very good.			
Phrasing:	ι{ φ[ω]} ι{ φ[ω]} ι{ φ[ω	ω]} }

2.4.3.2. The assignment of tones to AP

According to Venditti (2005), the AP delimitative tonal pattern is “a marking of the prosodic grouping itself” which separates from the contribution of the lexical-level pitch accent. So there are two types of AP tonal patterns. One is Accented AP, in which a sequence of accented words may combine, but only the leftmost accent survives and subsequent accents in the phrase are deleted. The other is Unaccented AP, in which unaccented words combine with adjacent words to form accentual phrases. The complete tonal transcriptions of the two types of APs are as follows (Venditti 2005):

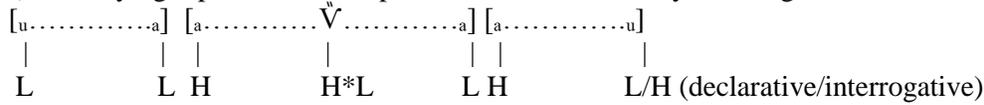
- (33) a. Accented AP %L (H-) H*L L%
 b. Unaccented AP %L H- L%

The assignment of tones on AP-initial disyllabic words follows research conducted by Nagahara (1994) and Pierrehumbert and Beckman (1988). Firstly we look into the assignment of the phrase-initial L H tone in native Japanese because it is crucial to the prediction of tone patterns on test words in this study. Since a low pitch precedes every phrase, Pierrehumbert and Beckman (1988) define the tonal pattern of an accentual phrase as follows:

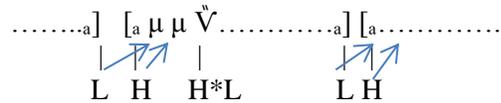
- (34) a. Define the accentual phrase as a phrase with at most one accent.
 b. Associate the pitch accent H*+L to the accented syllable.
 c. Associate H and L to the left and right edges of the accentual phrase, respectively.
 d. Finally, associate an L at the left edge of the utterance.

(35) is the underlying representation of pitch accent and boundary tone alignment. Diagrams (36)-(38) illustrate the case with two moras, the case with one mora, and the case with no moras to the left of the accented syllable (as cited in Nagahara 1994):

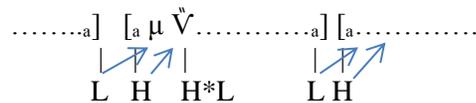
(35) Underlying representation of pitch accent and boundary tone alignment



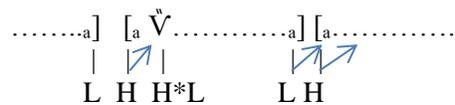
(36) When there are two moras to the left of the accented syllable, H links to the second mora of the AP in question, and L links to the initial mora if it is free.



(37) When there is only one mora to the left of the accent, the L links to the first mora and H doubly links to the pitch accent H*L in the AP in discussion:



(38) However, if the initial syllable is accented, the H doubly-links to a pitch accent H*L, and the L is linked to the right edge of the preceding AP.



According to Nagahara (1994), it's not likely that H*L can be distinguished from an HH*L when linked to a single mora, so we assume the two AP-initial moras are just linked to an H* and an L.

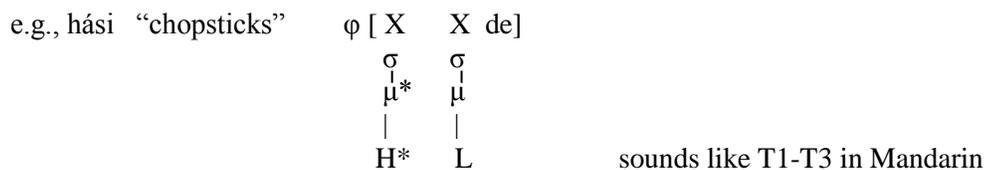
We predict that Japanese tonal structures will be transferred to the L2 production data. It is predicted that the accuracy and substitution rates of some structures (such as T1-T3, T4-T3, T3-T1, T3-T4, T2-T1, T2-T4) will be much higher than the accuracy and substitution rates of other structures because these structures are very much like the Japanese intonation pattern at AP-initial positions. The prediction

of these tone sequences is based on the examination of possible types of tone patterns that could be assigned to disyllabic words with different accent patterns. We only focus on disyllabic words in this section and expect three possible accent forms or locations for disyllabic words which are initial, final, and unaccented. We will consider accented words, and then unaccented words.

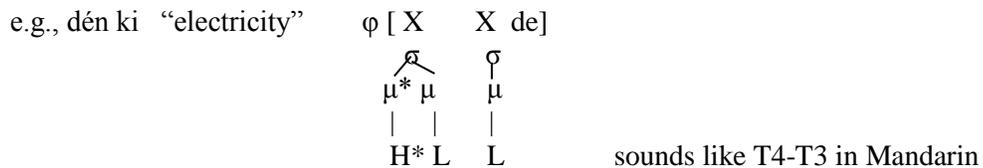
Accented words:

This section discusses three types of disyllabic words (six cases: (39)-(41) as listed below) with the accented syllables varying with mora numbers. The number of moras (1 or 2) on the accented syllable is crucial because mora numbers lead to different predictions of which Mandarin tone is most similar to it. Note that in Japanese it is syllables that are associated with accents, and it is moras that are associated with tonal segments. That is, even when a syllable is bimoraic, a pitch accent can link only to the first mora of the syllable (Nagahara 1994). The example words in (39a) (39b) and (40a) are cited from Gussenhoven (2004) and the representation of the assignment of tones to moras follows the principles utilized in Kubozono (2011). The author transcribes the example words in (40b), (41a) and (41b).¹⁹ (39a) and (39b) are cases with pitch accent on the initial syllable. Thus the assignment of tone follows the tone assignment principles illustrated in (38).

(39). a. Pitch accent is on the first syllable which has only one mora:



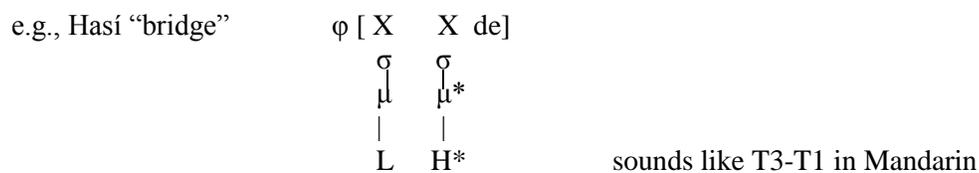
b. Pitch accent is on the first syllable which has two moras:



¹⁹ By consulting with Ms. Azusa Saito, an instructor of Japanese language in Duke University.

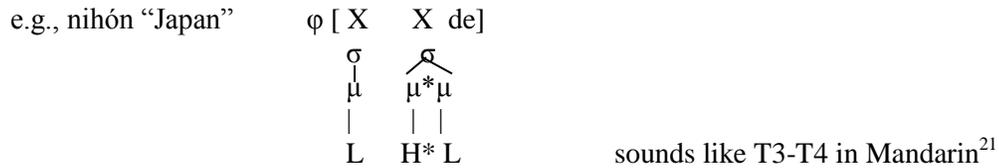
There are two reasons for assigning different Mandarin counterpart tones (T1 and T4) for single-mora and two-mora accented syllables. First, the HL tones within one syllable may follow one another faster than they do when crossing a syllable boundary (H.L). However, as we know, Japanese is a mora-timed language and tones are associated with moras, so there should be no differences in the phonetic realizations of a tone string ‘HL’ within a syllable or across two syllables. On the other hand, according to Warner (2001) who looks into the role of the mora in the timing of spontaneous Japanese speech, the timing relationships are more variable in spontaneous speech than in careful speech. The phonetic realization of mora-timing in spontaneous speech may be different in a case of HL within one syllable and H.L across syllables. At the same time, the target language Mandarin is a syllable-timed language (Lin and Wang 2007; Jun 2005). Thus, here we propose that the falling HL in a word (39a) across two syllables is analogous to the Mandarin tone string of (H)-σ (L)-σ (i.e., T1-T3), which is associated with two syllables; while the sharp falling H*L in a word (39b) within the same syllable is more like an Mandarin falling tone (HL)-σ (i.e., T4) which is associated with a single syllable. The prediction is confirmed by Japanese native speakers’ judgments, who agreed that the two H*L sequences in (39a) and (39b) sound different from one another.²⁰ Specifically, they agreed that the (39a) sequence sounded more like a Mandarin T1-T3 sequence while the H*L sequence in (39b) sounded more like a Mandarin T4. Please note that the cases with bi-moraic syllables in a word-final position for (39a) and (39b) are not included in the discussion, since both cases would lead to the same Low tone (T3) prediction in the second syllable. (40a) and (40b) are both cases of one mora which falls to the left of the accented syllable. Principle (37) is used in making predictions.

(40) a. Pitch accent is on the second syllable which has only one mora:



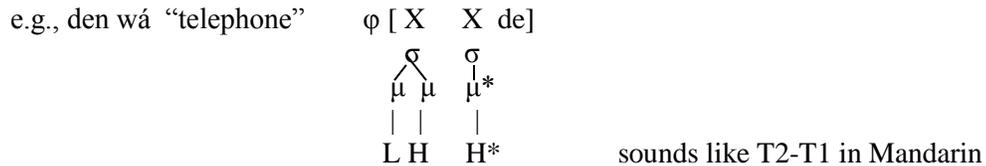
²⁰ Judgments were determined by Ms. Azusa Saito (Duke University) and Ms. Fumi Iwashita (UNC-CH).

b. Pitch accent is on the second syllable which has two moras:

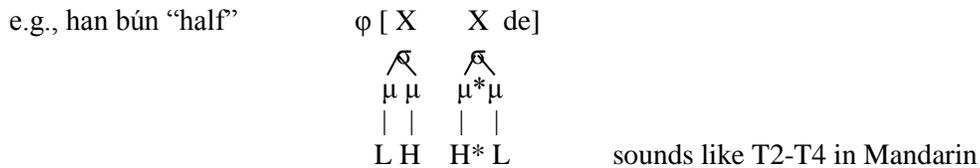


(41a) and (41b) are cases with two moras to the left of an accented syllable, I thus follow the principle stated in (36)

(41) a. Pitch accent is on the second syllable which has one mora



b. Pitch accent is on the second syllable which has two moras:



Unaccented words:

In the literature on Japanese pitch accent, it is still unresolved whether or not words that have accent on their final syllable (final-accented words) and words that have no accent (unaccented words) are acoustically different from one another. The differences between the two types are only noticeably manifested in the following particle if there is one. (Kindaichi 1947; Sugiyama 2006) For this study, since the perception of final-accented words and unaccented words are very alike, no new Mandarin tone strings mimicking the tone patterns of Japanese unaccented words will be proposed, since they would be the same as (40) and (41). That is, the cases shown in (40) and (41) cover disyllabic unaccented words.

In summary, we predict that in the L2 production data, there are many more T1-T3, T4-T3, T3-T1, T3-T4, T2-T1 and T2-T4 strings than other tone strings. That is, both the accuracy rates and the

²¹ According to Japanese speakers’ intuition, the *hon* in word (40b) could be both H and sharp falling like Chinese T4. Since (40a) already covers T3-T1, I only assign T3-T4 here.

substitution rates of these tone strings will be higher than they will be for other strings. When the target Mandarin tone sequences is the same as these tone sequences, error rates will be low, and vice versa. In addition, there are many strings of HH/HHH/HHHH and LL/LLL/LLLL in Japanese words with more than 2 syllables, (Kubozono 2011), so we predict that Japanese speakers' accuracy rates and substitution rates of T1-T1 will be higher than they will be for other tone pairs. Predictions concerning T3-T3 (LL) sequences will be difficult to test since the target language does not allow T3-T3. However, we can predict that there will be many more T1-T1 sequences than T3-T3 sequences in Japanese speakers' productions since they have been exposed to the L2 grammar for five months or longer.

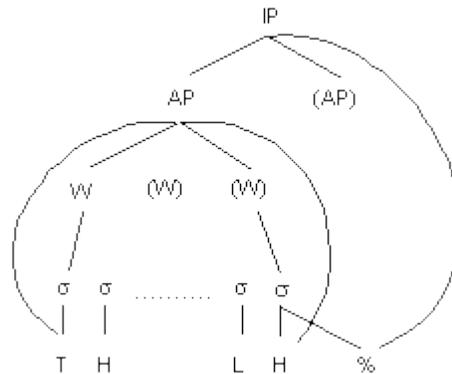
2.4.4. Seoul Korean (Korean) prosodic structures and predictions

Following Jun (1996, 1998, 2005), we assume that Korean intonations are composed of H and L tones. The TBU in Korean is the syllable. Research in Jun (1995, 1998) and Lim (2001) suggests that the fundamental frequency peaks and valleys of Korean intonation do not link to any specific syllable of a word but instead to a certain location in the phrase. That is, the pitch modulation over an utterance in Korean is a property of a sentence rather than of words. Thus Korean differs from other L1 languages in that it has neither lexical pitch accent nor lexical stress.

Jun (1990, 1993, 1998) developed a Korean Prosodic Model based on pitch contours following the Japanese prosodic models in Beckman and Pierrehumbert (1986) and Pierrehumbert and Beckman (1988). According to Jun's model (1993, 1998), the delimitation of prosodic units is primarily determined by pitch contours or tonal patterns. There are at least two prosodic levels marked by intonation in Korean: the Intonational Phrase (IP) and the Accentual Phrase (AP, which is analogous to the general PhP in (25)). The typical tonal pattern of the Korean AP is LHLH or HHLH, where the AP-initial tone is determined by the laryngeal feature of the phrase initial segment. When the segment is either aspirated or tense, having the feature [+stiff vocal cords] (Halle & Stevens 1971, as cited in Jun 1998), the AP begins with an H tone. Elsewhere, it begins with an L tone. IP may have one of several boundary tones such as L%, H%,

LH%, HL%, LHL% and HLH%. In general, IPs ending with H% often have the function of seeking information (e.g. question) and those ending with L% often have the function of making a statement. However, when the final syllable of the AP is the IP final, the final rise of the AP is replaced by the IP boundary tones. Figure 2.9 illustrates the Intonation Model in Jun (2005).

Figure 2.9. Intonation Model of Korean (Jun 2005)



IP: Intonation Phrase, AP: Accentual Phrase

w: phonological word, σ : syllable

T= H, when the syllable initial segment is aspirated/tense, otherwise, T= L

%: Intonation phrase boundary tone

2.4.4.1. Predicted prosodic phrasing for Korean speakers

The prosodic phrasing of the stimulus sentence by Korean speakers involves phrasing principles very similar to those of Japanese. In the stimulus sentence, the subject and main verb of the sentence will very likely form an IP and an AP since there are very few APs containing only one syllable in Korean. Jun (2011) suggests that a strong phrase boundary is usually inserted at the beginning of the focused argument. An additional reason for assigning a strong phrase boundary before the test word is that that location is also the location of the border of two major syntactic constituents (V and S'). Therefore a phrase boundary is assigned before the NP or the test word, meaning the test words are located at both an AP phrase-initial position, as well as an IP phrase-initial position. Similar to Japanese, Korean is known

as a “dephrasing” language (Ladd 2008) where the focused word begins an AP and all the following words were “dephrased,” meaning all phrase boundaries are removed if the sentence is not long (less than five words, following Jun 2011). However, as was discussed in the Japanese section above, the test word bearing contrastive focus may form an IP due to Korean-speaking learners’ lack of fluency in reading new L2 materials, so it is likely the case that an IP boundary will immediately follow the test word. This is confirmed by noticeable pauses following the test words in actual research data recordings, made by at least 65% of the subjects. Therefore, dephrasing likely does not occur in Korean speakers’ reading of L2 materials, and the test word is very likely ended with a phrasal low tone. Diagram (42) illustrates the assignment of the phrasal structure following the phrasing principles discussed in Jun (1996, 2011). The possible AP boundary following the test word is in gray. The phrasing of the stimulus sentence by Korean speakers is the same as that of Japanese speakers.

(42) The prediction of Korean speakers’ phonological phrasing of the stimulus sentence:

Chinese pinyin:	Wǒ	Jué de	XX de	DōngXi	Hén Hǎo.	
Gloss:	I	feel	XX –particle	things	very good.	
Phrasing:	ı{ φ[ω]}	ı{ φ[ω]}	ı{ φ[ω]}

2.4.4.2. The tone assignment of test words for Korean Speakers

Since the test word is at the beginning of an AP, we predict that the L2 tone pattern of the test words for Korean speakers will be heavily influenced by Korean AP-initial tone patterns. Lim (2001), a study on the word-level prosody in Korean, looks into the tone patterns of three-syllable test words within both IP (isolated as a single IP) and in a sentential context as the first AP in a larger IP, which is very similar to the case here. Lim (2001) found that, firstly, “there were no syllable weight effects on the fundamental frequency contours across the speakers and phrasal conditions”. Secondly, in both production and perception experiments “the second syllable position would be prominent or stressed.” That is, “the peak of F0 is shown near onset of the second syllables.” Jun (2005) also states that, when an AP has more than three syllables, the two initial tones of an AP are associated with the two initial

syllables of the AP, and the two final tones of an AP are associated with the two final syllables of the AP, where the most common tonal pattern of the AP is L/HH...LH. The syllables between the second and the penultimate syllables of the AP, if there are any, get their surface pitch values by interpolating between the H tone on the second syllable and the L tone on the penultimate syllable. (Jun 2005, Kim, J.J. et al.1997). Since the AP containing test words in the stimulus sentence is also an IP, the IP final tone L overrides the AP final high tone. (43) illustrates the predicted prosodic structure of the test words.

(43) Predicted tone assignment of the test words for Korean speakers

Phrasing	ι{φ [ω]}
Chinese pinyin	ι{φ [X	X -de]}
		σ	σ σ
Tones		L/H	H L%

If L1 transfer occurs, there may be a lot of L2 tone productions whose second syllable of the test word has a high tone H (similar to Mandarin T1), while the first syllable may be a low tone or a high tone (either a Mandarin T3 or T1). Therefore the tone strings L-H will be similar to a Mandarin T3-T1 sequence, and HH will be similar to a Mandarin T1-T1 sequence. This is illustrated in (44) below. Since the AP in question is quite short (containing three syllables only) and an L tone may immediately follows the H tone on the second syllable, a falling tone (T4) may occur. So it's also possible that the second syllable bears a HL contour for some cases. We propose the counterpart Mandarin tone sequences T3-T4 and T1-T4 for the disyllabic test words, as shown in (45). The diagrams (44)-(45) show the predictions of four counterpart Mandarin tones sequences imitating possible pitch patterns on test words by Korean speakers:

(44) a.	ι{φ [X	X	de]}	
		σ	σ		
		L	H	L%	sounds like T3-T1 in Mandarin
b.	ι{φ [X	X	de]}	
		σ	σ		
		H	H	L%	sounds like T1-T1 in Mandarin

- (45) a. $\iota\{\varphi [X \quad X \text{ de}]\}$
 $\begin{array}{ccc} \sigma & & \sigma \\ | & & | \diagdown \\ L & H & L\% \end{array}$ sounds like T3-T4 in Mandarin
- b. $\iota\{\varphi [X \quad X \text{ de}]\}$
 $\begin{array}{ccc} \sigma & & \sigma \\ | & & | \diagdown \\ H & H & L\% \end{array}$ sounds like T1-T4 in Mandarin

In this study, we will examine the numbers of the actual tone productions of these tone sequences and compare them with the occurrences of the other tonal productions. We predict that if L1 transfer occurs, the frequencies of these tone combination types will be higher than they will be in other structures, or higher than expected in the Korean speakers' L2 tonal productions. That is, the error rates of these tone sequences will be lower than they will be for other tone sequences, and these tone sequences will also be used more often as substitutes for other target tone sequences when errors occur.

2.4.5. Summary of L1 transfer predictions

The section of above introduces the prosodic structure of three L1s, predicts the phonological phrasing of stimulus sentences in Experiment 1 for the three different types of learners according to language-specific prosodic principles, and proposes the most likely counterpart Mandarin tone sequences on the test disyllabic words. The counterpart Mandarin tone sequences which are most likely to occur on the disyllabic test words are:

Table 2-2: Predictions of Mandarin tone sequences mimicking L1 intonation patterns

First languages	Predicted Mandarin tone sequences
English	T3-T1, T3-T4, T1-T3, T2-T4, T2-T3,
Japanese	T3-T1, T3-T4, T1-T3, T2-T4, T2-T1, T4-T3
Korean	T3-T1, T3-T4, T1-T1, T1-T4

It is predicted that L1 transfer will occur in the L2 tonal productions and that these counterpart Mandarin tone patterns will have a very high rate of occurrence in specific L2 data as substitute tone

sequences. The three L1s share some similarities in the predicted typical L1 pitch sequences, such as T3-T1 and T3-T4 across all three L1s, and T1-T3 and T2-T4 in both the English and Japanese accounts. We predict that when the target Mandarin tone sequences are same or similar to these typical L1 tone patterns, the error rates will be very low, meaning these learners have acquired the target Mandarin tone sequences which are similar to typical L1 intonation patterns easily and quickly. The L1 transfer of disyllabic tone sequences may be detected in both word-level (Experiment 1) and sentence-level (Experiment 2) L2 tonal productions. Since the test materials in Experiment 2 are all identical tone sequences, we predict that learners will make a large number of errors and that the actual tonal productions, especially the substitution patterns, may contain many of these disyllabic L1-like tone sequences. Some of the typical L1 pitch patterns may overlap with one another (such as T3-T1 across three L1s), or with phonological universals (such as no T2-T2, T4-T4, due to OCP effects), whereas some may be in conflict with phonological universals (such as the T1-T1 in Korean). These issues will be addressed in the data analysis in Chapters 4, 5, and 6.

The next chapter presents the designs of the two phonological experiments and the methodologies of investigation.

CHAPTER 3: METHODOLOGY

A pre-test and two phonological experiments are designed to address the research questions. The pre-test, which includes a reading task of monosyllabic morphemes, is conducted to ensure that all experimental participants are able to produce the individual lexical tones correctly. The two main experiments address the core research questions concerning word and above-word level tonal productions respectively. Experiment 1 consists of a reading task of 64 disyllabic words embedded in carrier sentences and Experiment 2 consists of a reading task of 20 sentences composed of identical tone sequences embedded in a discourse context. The experimental design of the study and methods of data collection and analysis are described below.

3.1. Test materials

The test materials for the pre-test consists of 48 monosyllabic morphemes which are used in one of the main experiments, Experiment 1 (underlined in (5) below). Chinese characters and pinyin (without tone marks) are provided in the reading list.²² All subjects had high accuracy rates (above 85%) with these tonal productions of individual monosyllabic morphemes. Because of this high accuracy rate, the pre-test will not be mentioned in Chapters 4 and 5.

The test materials for Experiment 1 are disyllabic words bearing 16 combinations of the four lexical tones. These 16 free combinations of the four Mandarin lexical tones are equally proportioned. We expect some tone combinations to be difficult to produce (such as those sequences containing tones with high complexity, identical tone sequences, and sequences with contour tones at word-initial positions, as discussed in Chapter 2), while others may be easier to produce (such as the tone sequences similar to

²² This is done to make sure the subjects are familiar with the pronunciation and tones of the morphemes. Any subject who did not know or forgot the pronunciation or tones of the individual morphemes received corresponding training before the main experiments began.

typical L1 intonation patterns). So the occurrences of these 16 combinations may differ from one another in actual L2 tonal productions.

Two words (consisting of different morphemes) for each tone combination type are used, resulting in 32 distinct words composed of 48 morphemes (underlined in (5)). The 32 words are repeated once, resulting in 64 tokens collected per speaker per trial. Neutral tones are excluded from the test words. All words are at the lowest proficiency level according to the handbook Outline of the Graded Vocabulary for HSK (1992). The morphemes or words were selected based on the following criteria:

- (1). The syllables in the test words cannot be neutralized or undergo tone Sandhi rules in spoken Mandarin (except for T3-T3).
- (2). Limit the word classes to nouns or verbs so that the test words can serve as an attributive part modifying the following nouns in sentences. No function words are used.
- (3). The use of obstruents in the test words are kept to a minimum, while the use of sonorants is kept to a maximum so that the pitch tracking can be continuous.
- (4) All syllables are CV (V) in structure without coda. The list (1) displays the 16 tone types and the 32 words used in Experiment 1, followed by Pinyin and English translations:

The test words for Experiment 1 are listed in (5)

- (5) 1. T1-T1: 书桌, 飞机 (“shū zhuō” desk, “fēi jī” airplane)
2. T1-T2: 开学, 花茶 (“kāi xué” school opening, “huā chá” jasmine tea)
3. T1-T3: 出口, 西北 (“chū kǒu” exit, “xī běi” northeast)
4. T1-T4: 开会, 书架 (“kāi huì” to attend a meeting, “shū jià” bookshelf)
5. T2-T1: 学期, 图书 (“xué qī” semester, “tú shū” picture-book)
6. T2-T2: 留学, 足球 (“liú xué” study abroad, “zú qiú” football)
7. T2-T3: 牛奶, 白水 (“niú nǎi” milk, “bái shuǐ” water)
8. T2-T4: 白菜, 学校 (“bái cài” Chinese cabbage, “xué xiào” school)
9. T3-T1: 手机, 老家 (“shǒu jī” cell phone, “lǎo jiā” hometown)
10. T3-T2: 打球, 小学 (“dǎ qiú” to play balls, “xiǎo xué” elementary school)
11. T3/5-T3: 语法, 水果 (“yǔ fǎ” grammar, “shuǐ guǒ” fruits)
12. T3-T4: 五月, 舞会 (“wǔ yuè” month of May, “wǔ huì” dance party)
13. T4-T1: 汽车, 大家 (“qì chē” car, “dà jiā” everybody)
14. T4-T2: 大学, 复习 (“dà xué” university, “fù xí” to review)
15. T4-T3: 跳舞, 下雨 (“tiào wǔ” to dance, “xià yǔ” to rain)
16. T4-T4: 睡觉, 校内 (“shuì jiào” to sleep, “xiào nèi” on-campus)

The 64 test words are embedded in sentences. Since the test words are all nouns or verbs, they are used as modifiers to modify nouns in the sentences. In order to avoid anticipatory and carry-over effects

by neighboring tones (Yi Xu 1997), the tokens are embedded in sentences where the preceding and following morphemes are both neutral tone “de”.²³ In addition, these test words were placed in a sentence-medial position, to reduce the possible interference of sentence intonation. (6) displays the carrier sentence structure and (7) lists two sample test sentences in Experiment 1.

(6). Test sentence for Experiment 1:

Chinese character:	我 觉得 X X 的 东 西 很 好.
Pinyin:	Wǒ juéde X X_de dōngxi hěn hǎo.
Gloss:	I feel X X particle things very good “I feel XX things are very good.”

(7). Two sample test sentences in Experiment 1 (more in Appendix 4):

a. 我觉得开学的东西很好。

Wo juede kaixue de dongxi hen hao.

I feel that things used for starting the semester are very good.

b. 我觉得西北的东西很不错。

Wo juede xibei de dongxi hen bucuo.

I feel that things from the northwest are very good.

The sentences are randomly ordered in the reading list. The 64 sentences were split into five groups and printed on five pages. They are transcribed in the Pinyin system of Romanization. The tonal diacritics were not used since all of the individual morphemes used in this experiment were read correctly in the pre-test by the participants. The Chinese characters of the sentences and the English, Japanese, or Korean translations are also provided to the participants in the reading lists. All of these measures were used to aid the students’ reading and to insure that they would not feel intimidated by the task.

The test materials for Experiment 2 consists of 20 sentences composed of identical tone sequences. Each sentence has a narrow focus at a fixed location (on a word or a phrase) of the sentence. This experiment is concerned with answering whether or not the prosodic events of sentence-level prominence influence the implementation of L2 lexical tones. The numbers of syllables in the word or

²³ Neutral tones also carry a pitch register which may be high or low depending on the preceding tones. However, all tokens in this study were embedded in the same sentence and the preceding tone to the first *de* is consistently Tone 2 (borne by *jue*). Thus we assume the neutral tones affect all of the tokens in the same way.

phrase bearing the focus vary, ranging from one to five syllables. Each of these will be referred to as F1, F2, F3, F4, and F5. The size of the focused constituent is varied so that we can observe how these L1 learners approach small and large focuses, and to see if they are influenced by the L1-specific focus structures (English as a “deaccenting” language vs. Japanese and Korean as “dephrasing” languages, discussed in §2.3.2.2). There will be five sentences, each with a different number of syllables in the stressed word/phrase. Each sentence contains one of the four types of tone sequences (T1/T2/T3/T4). As a result, there are 20 sentences in total: T1-F1/F2/F3/F4/F5, T2-F1/F2/F3/F4/F5, T3-F1/F2/F3/F4/F5, and T4-F1/F2/F3/F4/F5. Each of the 20 sentences is embedded in a conversational scenario. An example scenario used to elicit the utterance containing a narrow focus is given in (8). Please note that the test sentences are the answers in the conversations, and not the questions. The syllables located in the focused constituents (the morphemes bearing the new information/focus) are bolded and are capitalized in the English translations to indicate sentence stress. For all sentences, Chinese characters, pinyin (with tone marks) and L1 translations are provided.²⁴

(8). An example scenario (T1-F2: T1 sequence with a disyllabic focus)

a. Question:

Chinese: 喝 什么 的 医生 搬 书桌?
Pinyin: hē shéme de yīshēng bān shūzhuō?
Literal gloss: drink what (predicate) doctor move desks?
Translation: (English) The doctors that will move desks, what do they drink?
 (Japanese) 何を飲む医者が机を動かしますか。
 (Korean) 무엇을 마신 의사가 책상을 옮기나?

b. Answer (test sentence):

Chinese: 他 听说 喝 **鸡汤** 的 医生 搬 书桌。
Pinyin: tā tīngshuō hē **Jī Tāng** de yīshēng bān shūzhuō.
Gloss: he heard drink **chicken soup** (predicate) doctor move desks.
Translation: (English) He heard that the doctors who drink CHICKEN SOUP will move desks.
 (Japanese) 彼は**チキンスープ**を飲む医者が机を動かすと聞きました。
 (Korean) 그가 듣기로는 **닭고기 국물**을 마신 의사가 책상을 옮긴다.

²⁴ For Experiment 1 materials, no tone marks are provided in reading lists because the morphemes in Experiment 1 is already used in pre-test by which we confirmed all subjects know the correct tone marks of these morphemes. For Experiment 2, there are some new morphemes not used in Experiment 1, so tone marks are provided to ensure that subjects would not have to resort to guessing the proper tone of words which they may forgotten.

The syntactic structures of the unfocused portion (the remainder) of the sentence are the same across the five sentences with F1, F2, F3, F4, and F5 for specific tone types except for pre-focus materials. The syntactic structures of the five focused constituents are different due to the varying number of syllables and words. (6) displays the five sentences for T3 given to English speakers. Please note that the syllables in the T3 sequences are all labeled with T3 in the reading list of Experiment 2, but the sequences undergo T3 Sandhi in actual spoken Chinese and are also expected in the tonal productions of this study. Each sentence was annotated with Pinyin and translated into English. The narrow foci of the sentences are bolded and underlined.

(6). Sample test sentences for Tone 3.

T3-F1: 请你给买水的小姐写语法。

qǐng nǐ gěi mǎi **Shuǐ** de xiǎojiě xiě yǔfǎ.

Please write the grammars for the ladies who buy WATER.

T3-F2: 请你给买水果的小姐写语法。

qǐng nǐ gěi mǎi **ShuǐGuǒ** de xiǎojiě xiě yǔfǎ.

Please write the grammars for the ladies who buy FRUIT.

T3-F3: 请你给买水果的小姐写语法。

qǐng nǐ gěi **Mǎi Shuǐ Guǒ** de xiǎojiě xiě yǔfǎ.

Please write the grammars for the ladies who BUY FRUITS.

T3-F4: 请你给想买水果的小姐写语法。

qǐng nǐ gěi **Xiǎng Mǎi Shuǐ Guǒ** de xiǎojiě xiě yǔfǎ.

Please write the grammars for the ladies who WANT TO BUY FRUIT.

T3-F5: 请你给五点买水果的小姐写语法。

qǐng nǐ gěi **Wǔ Diǎn Mǎi Shuǐ Guǒ** de xiǎojiě xiě yǔfǎ.

Please write the grammars for the ladies who BUY FRUIT AT FIVE O’CLOCK.

The twenty conversations containing the test sentences are randomly ordered, split into three groups and printed on three pages. Thus, there are nine pages of reading lists including the first page containing the materials for pre-test, the following five pages of single sentences for Experiment 1 and the last three pages of conversations for Experiment 2.

3.2. Subjects and recording procedures

Sixty seven speakers participated in the pre-test. Seven were excluded from participation due to low accuracy rates in the pre-test (below 85%), leaving a total of sixty subjects to participate in Experiments 1 and 2. Twenty were native English speakers, twenty were native Japanese speakers, and twenty were native Korean speakers. All subjects had been learning Mandarin for at least 6 months, but no more than 18 months, placing them in at approximately an intermediate level.²⁵

The twenty English-speaking learners, 12 males and 8 females, were recruited from the American CFL (Chinese as a Foreign Language) students at the University of North Carolina at Chapel Hill and all spoke American English as their first language. Most were studying Mandarin in the Intermediate Chinese II (fourth semester) classes when data was collected in the spring of 2012. No participant spoke or was learning any other tonal language. (Some languages studied by participants include Spanish, Latin, French, German, Hebrew, and Russian.)

The twenty Tokyo Japanese-speaking learners (10 males and 10 females) and twenty Seoul Korean-speaking learners (8 males and 12 females) were recruited from the CFL students in the International College at Zhejiang University, Hangzhou city, China. Most Japanese and Korean speakers were studying Mandarin in the Intermediate Level classes when the data was collected in December 2011²⁶. All Japanese speakers were from areas where people speak Japanese with Tokyo-type pitch accent (Haraguchi 1997, 1988), such as Tokyo, Hokkaido, etc. (see Appendix 1 for a pitch accent map). All Korean speakers are from Seoul. Some of the Japanese speakers studied English, and only two studied other languages (French and Russian). All Korean speakers studied Mandarin and English only. All

²⁵ Studying hours per week/month vary a lot among the subjects, especially for Japanese and Korean speakers. However, the average total study hours are in the range of 384 hours to 440 hours. English speakers' average studying hours are about 384 hours (average 16 months and 24 hours per month) ; Japanese speakers' are 435 hours (average 13.5 months and 32.22 hours per month) ; Korean speakers are 440 hours (average 14 month and 31.42 hours per month). About 35% of English-speaking subjects also have experience of studying Mandarin in China.

²⁶ Due to time and budget limit, only Japanese and Korean speakers were recruited in Hangzhou, China. No qualified English speaking learners were identified during the Christmas season (December 2011, winter break for the author) in Zhejiang University. Efforts were made to lighten its influence to experiment results. In the data analysis, for example, the comparison of absolute counts of errors/substitutes between different L1 groups is avoided; instead all the comparisons are conducted between the percentages or ratios calculated within each L1 data set.

subjects claim that their Chinese is much better than English. Mandarin was the only tonal language any subject had studied. All subjects participated in this study voluntarily and each was paid for the 45-minute recording session.

The twenty English speakers' tonal productions were recorded in a soundproof recording lab. The forty Korean and Japanese speakers' tonal productions were recorded at Zhejiang University in a quiet classroom. A microphone-headset from Radio Shack and a ThinkPad laptop computer with Windows XP were used to record the productions. Version 5.2.17 of Praat (Boersma and Weenink 2011) was used. Before the recording procedure, the subjects were asked to sign a document of consent and fill out a language background survey. All Korean and Japanese speakers were provided respectively with Korean and Japanese versions of the consent forms, surveys, and recording instructions. After filling out the forms and surveys, subjects read the pre-test materials as a warm-up exercise and then read the main experiment reading lists. The participants were asked to read these sentences out loud at a normal speed. The recording was paused between pages of the reading list. Before each recording, the participant had the option of previewing the reading list for as long as he or she wanted to. If a participant misread any part of the sentence or read the sentence with a severe lack of fluency, he or she would be asked to re-read the sentence. All sound productions were saved separately on a personal computer.

3.3. Judgment and data transcription

Each subject produced 128 test syllables in Experiment 1 and 228 syllables in Experiment 2. The correctness of all L2 tonal productions was judged within sentences. The author, an instructor of Mandarin with 15 years of teaching experience and native Mandarin speaker, listened to the test syllables in sentences and judged whether or not these tonal productions were the same as the target tone productions, and marked these as "correct" or "incorrect".²⁷ To guarantee the reliability of the correctness judgments and transcriptions of incorrect tonal productions, both intra- and inter-rater agreement was

²⁷ Praat is an excellent tool to measure the accurate values of F0 and other phonetic correlates. However, it cannot judge if the L2 tonal productions are acceptable or not by native Mandarin speakers.

calculated. All tonal productions in the two experiments were judged and transcribed twice by the author, with a one month's interval in between judgments. The judgment and transcription agreement rate between them is 95.6%. For the cases of discrepancy which are mostly in the transcriptions of incorrect tonal productions, another native speaker of Mandarin who received training in linguistics was consulted in order to agree on the final transcriptions. For the inter-transcriber reliability, two other native speakers (referred to as L and C) were hired to judge and transcribe one tenth of the data independently. They are both native speakers of Mandarin and have received training in linguistics, and have taught Mandarin as a second language for 4 years and 19 years, respectively. A pairwise comparison indicates that the author and L have the highest level of agreement (93.8%), while C and L have the lowest level of agreement (91.6%). For each case of discrepancy, which were mostly in the transcriptions of incorrect tonal productions, the author discussed with L and C to arrive at a final transcription which was agreed upon by both L and C. Both the intra- and inter-rater tests show that the judgment and transcription of L2 tonal productions is reliable.

Two main parameters of judgment are the “register” and “contour” of Mandarin tones. As mentioned in Chapter 2, “register” refers to pitch height and “contour” refers to pitch movement. For register, we follow the binary feature suggested by Yip (1980), giving each tonal register either the value of upper [+U] or lower [-U]. We follow the traditional contour features of “rising,” “falling,” and “level.” A tone was considered an error if either the contour or register of the tone was wrong. That is, any register or contour features which are considered unacceptable by the judge will result in a wrong tone. For all incorrect productions, the subjects' actual tonal productions were also written down. The incorrect tones were classified as either those which do occur in Mandarin (within inventory), or those which do not (out of inventory). “Out of inventory” tones were further classified into two subcategories: one category of tones which were labeled with modified Mandarin tone names (for example, Mid-T1 (middle level tones) and L-T4 (low falling tone)). The others are transcribed using Chao's five-level system, 5 representing the

highest pitch and 1 the lowest pitch (for example a contour tone which starts out high and gradually falling down to the middle zone might be transcribed as [543]).

For the T3 and T5 productions in Experiment 1, tones were judged in the following way: at word-initial position where the target tone was followed by a T1, T2, or T4, only low tones (T3) were considered correct. T5, the T3 at a sandhi position, is phonetically a high rising tone. Therefore, any productions other than high rising were considered incorrect for target T5. Because the test words are not at phrase- or sentence-final positions and are followed by particle *de* in the test sentences, only a low T3 is acceptable for the target Tone 3 at word-final positions. In other words, no FT3 is allowed on any test syllables in Experiment 1. In Experiment 2, the surface tone targets in the T3 sequences are more complicated than those in the other tone sequences due to the applications of the T3 Sandhi rule. According to Shih (1986), the domain of Mandarin Tone 3 sandhi is prosodically defined, and it is the foot that constitutes the domain of Mandarin tone sandhi. Because word boundaries are not marked in Chinese, the three judges were asked to assign word boundaries. This, as well as consideration taken of the Foot Formation Rule proposed by Shih (1986), were used to identify the target surface tone 3 sequences. All three native speakers had a generally consistent segmentation of the text into prosodic words. Most disagreements are related to the granularity of segmentation on the pre-focus materials leading to different target tones T5 or T3 for the morpheme or syllable of *ni* “you”. Because of this, both T5 and T3 produced by the learners for the syllables of *ni* are considered correct. Both T3 and FT3 are also considered acceptable for the sentence-final syllable *fa*.

Previous studies on perceptions of Mandarin tones suggest that it’s difficult to distinguish rising (T2) and low dipping (FT3) tones, because of their acoustic similarity Both have a similar F0 onset as well as a falling-rising contour (Hao, 2012; Elliot, 1991; Moore and Jongman, 1997; Jongman et al 2006; among others). Because of this similarity, extra precautions are taken in ensuring that T2 and FT3 are judged correctly. However, it was found that the three judges in this study were consistent in their judgments of the L2 tonal productions as T2 and FT3. The acoustic measurements of the L2 productions

of T2 and FT3 as judged by the three judges conform to the acoustic differences described in acoustic studies concerning T2 and FT3. Specifically, three phonetic cues -- the timing of turning point, the $\Delta F0$ value, and duration -- were measured, and these measurements were found to match those made in previous studies (Blicher, Diehl and Cohen 1990; Shen and Lin 1991; Shen, Lin and Yan 1993; Moore and Jongman 1997; Jongman et al. 2006).²⁸ That is, the turning points are earlier in the T2 productions than they are in the FT3 productions, the $\Delta F0$ value is smaller in T2 productions than it is in FT3 productions, and lastly, the duration of T2 productions are shorter than in those of FT3 productions. Therefore, we assume that the tones transcribed as T2 and FT3 in the present study are also reliable.

3.4. Data Analysis

This study makes use of both quantitative and qualitative analytical methods. In terms of quantitative methods, statistical analyses are used to work out second language learners' general and specific tonal error patterns and substitution patterns. The hypotheses regarding the three universal constraints (the TMS, the OCP, and TPC) were tested using the SAS statistical package. For the statistical analysis, the investigator sought help from statistical analyst Dr. Christopher A Wiesen (Odum Institute for Research in Social Science, UNC-CH). Several procedures, such as the FREQ procedure, are used in the statistical analysis for the present study. The significance criterion adopted for declaring a significant difference is $p < 0.05$. More information regarding the statistical analysis used will be noted in each sub-section of data analysis in Chapters 4 and 5. The observations of error patterns and substitution patterns regarding L1 transfer and T3 pedagogical problems are interpreted within the framework of OT using a qualitative approach.

²⁸ Turning point is defined as the duration from the onset of the tone to the point of change in F0 direction. $\Delta F0$ is the decrease in F0 from the onset to the turning point.

CHAPTER 4: WORD-LEVEL EXPERIMENT RESULTS

This chapter reports test results for the evidence of our three phonological universals, L1 transfer effects and Tone 3 pedagogical problem in the word-level research data. This study hypothesizes that the L2 tonal productions are constrained by three phonological universals, namely, the Tonal Markedness Scale (TMS), Obligatory Contour Principle (OCP) and Tone-Position Constraint (TPC) and expects some similarities (such as higher error rate of rising tone than falling tones and level tones, lower occurrences of tone pairs than non-identical tone combinations and higher accuracy rate of contour tones at word-final positions, etc.) in interlanguages of English, Japanese and Korean speakers. We may not see strong effects of these constraints in native L1s and L2, but we expect SLA error and substitution patterns, as part of the innovative interlanguage grammars, would demonstrate the effects of these three constraints more clearly. For example, level tones, rising tones and falling tones are all allowed in English (L1) and Mandarin (L2), but we expect a preference for specific individual tones (ie. level tones) by L2 learners during the course of SLA of Mandarin lexical tones because of TMS. It is also believed that the native prosodic grammars affect the SLA, and we will confirm the effects of L1 transfer, especially the transfer of typical L1 disyllabic tone patterns, in the L2 tonal productions in this chapter. We predict that L2 learners could acquire those Mandarin tone sequences which sound like the typical L1 intonation pattern on the test disyllabic words much quicker than other new tone sequences, and thus we will find different error and substitution patterns among the three types of learners. The proposed counterpart Mandarin tone sequences mimicking typical L1 F0 patterns are: T3-T1, T3-T4, T1-T3, T2-T3 and T2-T4 for English speakers; T1-T3, T4-T3, T3-T1, T3-T4, T2-T1, T2-T4 for Japanese speakers; T3-T1, T1-T1, T3-T4 and T1-T4 for Korean speakers. We expect the accuracy of these tone sequences to be greater than other sequences. In this chapter, we will examine the actual occurrence frequencies of these tone sequences in specific sub-datasets, and compare the frequencies with other sequences to confirm this hypothesis. We

also predict that some of the L1 transfer effects may overlap or conflict with phonological universals because UG constrains natural languages (L2 learners access UG through L1s) but every L1 has its own unique prosodic structures. In addition to the tests concerning phonological universals and L1 transfer of disyllabic tone patterns, we also pay attention to the Tone 3 teaching issue by examining the Tone 3 error and substitution patterns in the section of TMS testing. We predict that the prevailing “FT3-First” teaching method may cause the overuse of FT3 and lead to high error rate of Tone 3 across all learners.

This chapter begins with the general error patterns and substitution patterns for the three types of L1 speakers in the section §4.1. Since this study looks at the L2 tonal pattern mainly from the UG and L1 perspectives, we focus on a statistical test and analyze the effects of three phonological constraints in the overall data set and also within each L1 group in the section of §4.2, §4.3 and §4.4. The L1 transfer effects were analyzed whenever necessary in the discussion because the universal constraints and L1 transfer effects occur together almost everywhere in the L2 productions. In order to obtain a general view of the tonal patterns, we discuss the error patterns and related substitution patterns respectively for each test. The hypothesis regarding T3 pedagogical problem is addressed in §4.2. Section §4.5 reports the test results of typical disyllabic tone pattern transfers from the three L1s, and the most frequently produced tone combinations by English, Japanese and Korean speakers respectively. This chapter ends with a summary of the Experiment 1 results in §4.6. This chapter reports the concrete test results only. The discussion of theoretical issues related to the word-level test results will be offered in Chapter 6.

It is found that sometimes L2 learners cannot successfully produce the target tones and often substitute other tones for target tones in productions. These substitutions sometimes are also called “mis-productions” and “errors.”

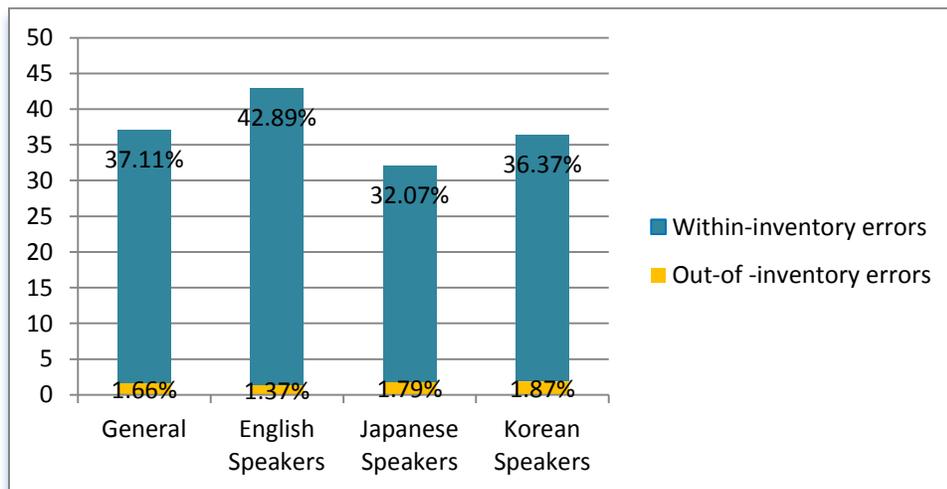
4.1. General information of L2 production errors

Before we proceed to the statistical tests of the hypotheses, this section briefly reports the general error rates of L2 tonal productions in the three L1 data sets and the mis-productions out of Chinese tone inventory to give a complete picture of the L2 tonal data. It is found that Japanese speakers possess the

highest Mandarin tonal proficiency level among the three groups of learners, and the out-of-inventory error rates are very low across the three L1 sub-datasets.

In this Experiment, 2,560 syllables were produced by each group of L1 speakers, resulting in 7,680 test syllables that were examined. The general error rates and those within each L1 group are between 30% and 45%, with the highest error rate of 42.89% falling in the English speakers' data set, and the lowest error rate of 32.07% falling in the Japanese speakers' data set. It seems Japanese speaking subjects in this study possess the highest Mandarin tonal proficiency while English speaking subjects possess the lowest. Figure 4.1 displays the overall error rates and those within each L1 group. As I mentioned in Chapter 3, two types of errors were labeled in judgments. One type consists of those errors which could be categorized into Chinese tone types, and we call them within-inventory errors, such as T1, T2, T3, T4, FT3. The other type of errors refers to those “tonal” productions that cannot be categorized into any Chinese tone types and they are called out-of-inventory errors, such as MT1 (mid-level), LT4 (low falling), [543] (high gradual falling) etc. Within such errors, most are tones within the Mandarin lexical tone inventory. Those out-of-inventory errors are at the bottom of the columns below, and the numbers are very low across the three L1 data sets.

Figure 4.1. Proportions of within-inventory and out-of-inventory errors

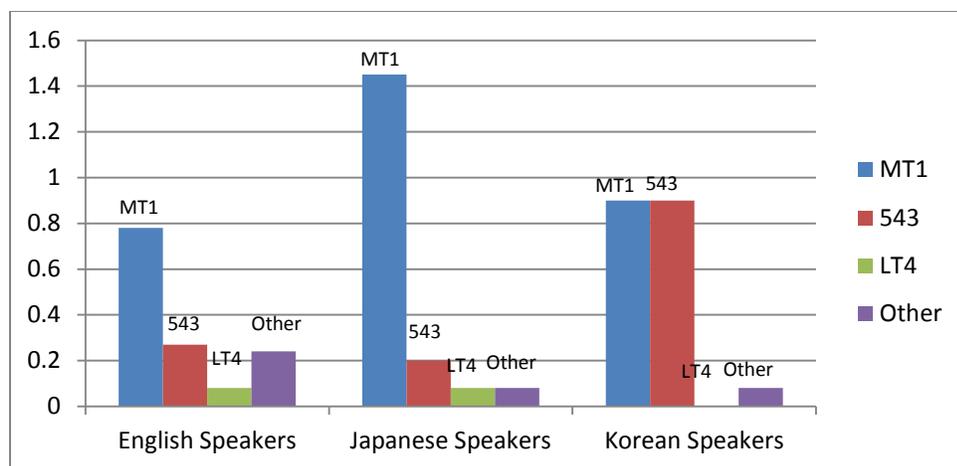


Since the error rates vary among the three L1 speakers' data sets, we will try to avoid comparisons of absolute counts in the tests in following sections. Instead, almost all of the comparisons are between percentages or ratios calculated based on the numbers within each L1 dataset.

As illustrated in the figure above, the numbers of out-of-inventory errors are very low across the three L1 data sets: 1.37%, 1.79% and 1.87% of the total productions in English, Japanese and Korean, respectively. Due to the very small numbers, these errors had no noticeable effect on the statistical analysis. However, we summarize the main types of out-of-inventory error below.

According to the judgments of the three native speakers, the most frequently produced out-of-inventory error is MT1, a middle level tone. It occurs 80 times in total, including 20 times by English speakers, 37 times by Japanese speakers and 23 times by Korean speakers. The second most frequently produced out-of-inventory error is [543] which is a high registered ([+Upper]) gradually falling tone. This tone sounds like both T1 and T4 and it occurs most often in Korean speakers' data set. It was produced by L2 learners 35 times, including 7 times made by English speakers, 5 times by Japanese speakers, and 23 times by Korean speakers. The third most frequently produced error is LT4, which is a low registered falling tone. It occurs 4 times in total, 2 times by English speakers and Japanese speakers respectively. Figure 4.2 displays the breakdown (in percentages out of total productions in each L1 data set) of four types of out-of-inventory mis-productions, i.e., MT1, [543], LT4 and others. They are all below or around 1% of the total productions. The 8 count of "Other" errors take 0.24% of English speakers' productions, but all of the eight types of "other" errors occur only once.

Figure 4.2. Out-of-inventory errors in Experiment 1



The occurrences of out-of-inventory tonal errors mean that some L2 learners don't have the correct phonological representation of individual tones. MT1 and LT4 are cases of wrong register features, while tone [543], which sounds like both a T1 and T4, is a case of incorrect representation of contour shape. Remembering the phonological representation of target Mandarin tones is as follows:

(1) Phonological representation of Mandarin tones

Tones	Pitch value	Register and Component tones
Tone 1	55	[+U, hh]
Tone 2	35	[+U, lh]
Tone 3	21	[-U, hl]
Tone 4	51	[+U, hl]

By contrast, the out-of-inventory tone MT1 can roughly be represented as [\pm U, hh], although the register feature doesn't match to any of the Mandarin tones; LT4 [-U, hl] is not a true T4 because of the wrong register feature; and tone [543] could be roughly represented as [+U, hml] where the contour feature cannot match to any of the target Mandarin tones. Due to the small amount of these errors, most of the tests in the following sections focus only on the within-inventory error patterns.

In the following three sections, we discuss the results of TMS, OCP and TPC tests. We present the quantitative study results first and discuss the findings by observing the error and substitution patterns.

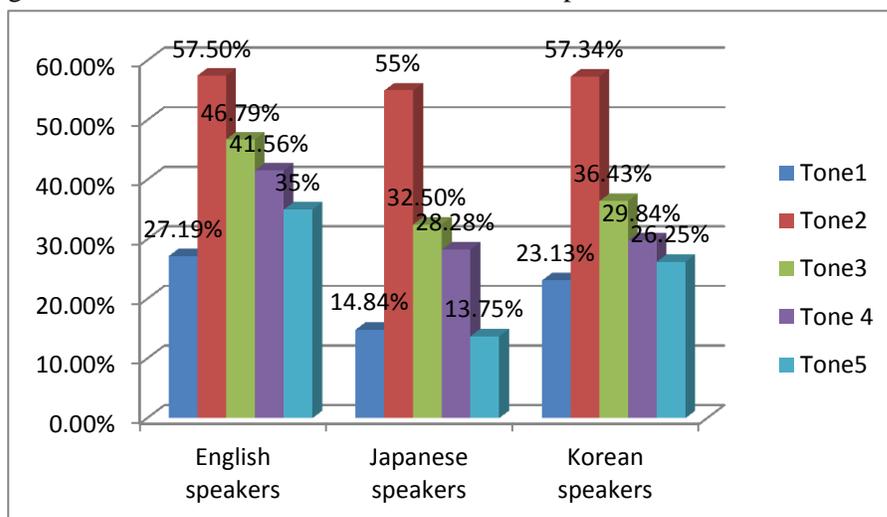
4. 2. On the test of Tonal Markedness Scale (TMS)

In this section, the related data and statistical results are closely examined to determine whether the TMS is working. As defined in Chapter 2, TMS is a universal, phonetically grounded markedness scale. It is stated as *R>>*F>>*L (Hyman & VanBik 2004) indicating that rising tones are more disfavored than falling tones, which are more disfavored than level tones. This study focuses on the three target tones of Mandarin and assumes a TMS instantiated as *T2>>*T4>>*T1 and predicts higher error rates of T2 than T4, and higher error rates of T4 than T1. With the presentation of statistical analysis of error patterns in §4.2.1 and substitution patterns in §4.2.2, we found TMS is relevant in most of the L2 production data. In addition to the test for TMS, this section also addresses two other issues. One is the influence of L1 regarding the level tone and contour tone asymmetry in substitutions. The other is the pedagogical problem of Tone 3 and its sandhi rules. Section §4.2.3 offers a summary of the TMS test.

4.2.1. The statistical analysis of TMS

The error rates for each tone in the general data set range from 21.72% (for T1) to 56.61% (for T2). The rankings of error rates for each tone within each L1 data set are very similar. The graph 4.3 shows the error rates of each tone in the three L1 groups. The statistical test of TMS below is only on T1, T2 and T4, but this graph includes the error rates of all tones (T1, T2, T3, T4 and T5) to provide a complete view of the error patterns. T5 is the rising tone (alike in phonetic shape to T2) which results from the Pre-T3 Sandhi.

Figure 4.3. General error rates for TMS test in Experiment 1



Within each L1 data set, the error rate of T2 ranked the highest, all at or above 55%, with the T3 as the second highest, the T4 the third highest and with T1 or T5 ranked the lowest. The sub-rankings are displayed below:

(2) Error rate rankings in Experiment 1

- a. English speakers: ***T2** (57.5%) > ***T3** (46.7%) > ***T4** (41.5%) > ***T5** (35%) > ***T1** (27.1%)
- b. Japanese speakers: ***T2** (55%) > ***T3** (32.5%) > ***T4** (28.2%) > ***T1** (14.8%) > ***T5** (13.7%)
- c. Korean speakers: ***T2** (57.3%) > ***T3** (36.4%) > ***T4** (29.8%) > ***T5** (26.2%) > ***T1** (23.1%)

Comparing the error rates of T1, T2 and T4 by statistical analysis, it is found that TMS is relevant in all L1 data sets, except for T1 vs. T4 in Korean speakers’ dataset. That means the error rates of T2 are significantly higher than T4 throughout all L1 data sets. The error rate of T4 is higher than T1 in the Korean speakers’ data, but there is no significant difference. The error rates of T4 are higher than T1 in all data sets, but the case is supported by statistical analysis only in English and Japanese speakers’ data. The SURVEYFREQ Procedure was employed to summarize the counts of errors. A statistical test of no association between tone and response using the Rao-Scott Chi-Square test was conducted to account for multiple observations within subjects. The final results of the statistical analysis, i.e. P values, are presented in Table 4-1.

Table 4.1. Rao-Scott Chi-Square test results for TMS hypothesis (Values of Pr > Chisq)

	Overall	English	Korean	Japanese
T4 vs. T1	<.0001	.0037	. 2780 (n. s.) ²⁹	.0004
T2 vs. T4	<.0001	.0044	<.0001	<.0001

Most of the sub-rankings of tone errors were supported by the statistical analysis and conform to the TMS constraint. We confirm the presence of TMS in L2 productions by observing substitution patterns in the following section.

4.2.2. Substitution patterns for individual tones

²⁹ Not significant.

In the study of SLA, substitution is not done arbitrarily, but stems from a process of avoidance and choice by L2 learners. To address the TMS rankings from the perspective of substitutions, we first look into the substitution patterns without target tone information in section §4.2.2.1, then examine both error rates and substitution together in the next section §4.2.2.2. Please note that the percentages included in this section are all out of the total production of each specific L1 data set. That is, the substitute tone percentages are the breakdowns of the total error rates within each L1 data set. Discussion on the Tone 3 pedagogical issue is provided in §4.2.2.3.

4.2.2.1. Substitution patterns without target tone information

In this section, we found that the substitution rates of individual tones are negatively correlated to the error rates. The substitution patterns in both the general data set and each L1 data set fully support the TMS constraint. In addition, by comparing the substitution patterns, we found English speakers produced more contour tones than Japanese and Korean speakers when errors occurred.

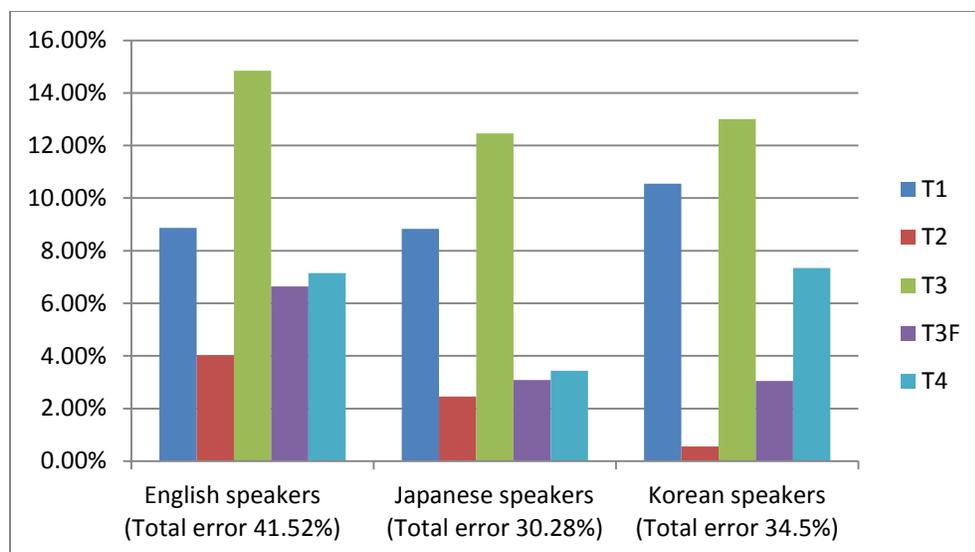
It is evident in the overall data set and each L1 subset that the most often used substitute tone is T3, the second most used is T1, the third is T4, the fourth is FT3 and the least frequently used substitute tone is T2. Ranking (3) demonstrates the frequencies of individual substitute tones in overall data set. The numbers in parentheses following each substitute tone are the frequencies. For example, T3 is used as a substitute tone (when the target tone is wrongly produced) 1,032 times throughout the L2 data set.

(3). The ranking of substitute tones in general data set

T3 (1032) > T1 (723) > T4 (459) > FT3 (327) > T2 (180)

Figure 4.4 details the substitute tones within each L1 data set. For each L1 group, the columns represent the substitute tones, with the values at the Y axis indicating the percentages of the substitute tones out of the total productions in that L1 data set.

Figure 4.4. Substitutions within each subset in Experiment 1



The rankings of individual substitute tones in each L1 data set are exactly the same, always $T3 > T1 > T4 > T2$. The substitution patterns in the general data set and all sub-data sets generally confirm the TMS ranking in §4.2.1, which was based only on error rates. There is a correlation between the error rates of the production and the substitution patterns. Wherever low error rates (high accuracy rate) occur with a target tone, the tone is often found to serve as a substitute tone at a high rate, and vice versa. This is true for T1, T2 and T4 which I have already shown to form a TMS. For example, L2 learners perform T1 the best (lowest error rate), T4 the second and T2 the worst (highest error rate). On the other hand, these learners substitute T1 for other tones most frequently while T2 the least frequently. In a more general view, a large number of level tones (T1 and T3) as substitute tones across is found L1 data sets, confirming that level tones (either low or high) are more “unmarked”.

Only the case of T3 is not compatible with this kind of pattern. For T3, we found very high error rates in all L1 data sets, and the highest substitution rate of T3 across all data sets. T3 serves as a substitute at a high rate, which means T3 is phonetically easy to produce and should have a low error rate. The reason T3 has a high error rate is likely due to the difficulty of processing T3 Sandhi. A detailed discussion regarding the acquisition of Tone 3 based on the substitution patterns is offered in §4.2.2.3.

There is another noticeable pattern regarding the use of contour tones (T2, FT3 and T4) as substitute tones among the three L1 groups. The learners use all tone types as substitute tones, but English speakers have a much bigger proportion of contour tones within its total error number than Japanese and Korean speakers. English speakers have 1,063 errors in total and 456 of them (42.9%) are contours. In contrast, Japanese speakers have 775 errors in total and only 230 errors are contours (29.68% of total errors) while Korean speakers have 883 errors and only 280 errors (31.71% of total errors) are contours. A test of no association between language and contour tones using the Rao-Scott Chi-Square test was conducted to account for multiple observations within subjects. According to the statistical test result ($P > \chi^2$ is $< .0001$), the percentage of contour tones as substitution in English speakers' data set is significantly higher than those in Japanese and Korean speakers' data sets. In other words, when errors occur, L2 learners employ all tone types to substitute target tones. However, English speakers are more likely to use contour tones than Japanese and Korean speakers. On the other hand, when errors occur, Japanese and Korean speakers are more likely to use level tones, either higher (T1) or low (T3) tones.

This contrast of preference in favor of contour tones by the English speakers is ascribed to L1 transfer. As mentioned in Chapter 2, High and Low tones are the basic component tones for intonational languages. However, the association of tones and their TBUs may be different in the three L1s. In English the tones (H, L) are loosely associated to TBUs. That means, sometimes one or two or even three tones could link to one TBU (see examples in Gossenhoven 2004, p.274), and sometimes one tone may link to one, two or even more TBUs; and sometimes some tones “exist that do not have an association to a TBU...” (Gossenhoven 2004). Therefore, as indicated in Ladd (2008), the rising and falling pitch patterns are “by no means the only possibilities in English” (Ladd 2008, p.6) because “sustained level pitch is not very common in English” (Ladd 2008, p.103). In the OT spirit, this means that in the English grammar, *Contour is ranked low. In Japanese and Korean, tones and TBU may be more tightly associated, and the correspondence between them is more stable. Overall, it seems that the different

proportions of substitute tones in the L2 tonal productions reflect L1 transfers of the surface realizations of basic pitch components.

4.2.2.2. Substitution patterns with target tone information

This section summarizes the substitution information with corresponding target tones to detail the error patterns. The analysis of the substitution patterns is organized according to the target tones. This section also includes discussion of the T3 pedagogical problem and some L1 transfer effects whenever related patterns are examined.

For each target category of Mandarin tones, the error rates and substitution rates are provided within Table 4.2, 4.3 and 4.4 for English, Japanese and Korean speakers respectively. The tones in the first line are the target tones, and those under the target tones are the substitute tones, followed by the count numbers and percentages (out of the total substitute tones for that target tone). In each chart, the high substitute percentages (higher than 25%) are highlighted in bold.

Table 4.2. English speakers' substitution patterns

Target tones	T1	T2	T3	T4	T5
Substitute tone counts& percentages	T3 85 48.8%	T3 185 50.2%	FT3 104 39.6%	T3 94 35.3%	T3 16 57.1%
	T4 52 29.8%	T1 78 21.1%	T4 65 24.8%	T1 87 32.7%	T1 10 35.7%
	T2 19 10.9%	T4 64 17.3%	T1 52 19.8%	T2 48 18.0%	T4 2 7.14%
	FT3 13 7.4%	FT3 27 7.3%	T2 36 13.7%	FT3 26 9.7%	
	MT1 4 2.2%	MT1 10 2.7%	[543] 2 0.76%	MT1 4 1.5%	
	[543] 1 0.5%	[543] 3 0.81%	[443] 1 0.38%	LT4 2 0.7%	
		[512] 1 0.27%		[355] 1 0.37%	
				[432] 1 0.37%	
				[445] 1 0.37%	
				[523] 1 0.37%	
			[543] 1 0.37%		

Table 4.3. Japanese speakers' substitution patterns

Target tones	T1	T2	T3	T4	T5
Substitute tone counts& percentages	T3 45 47.3%	T3 230 65.3%	FT3 57 31.3%	T1 107 59.1%	T1 7 63.63%
	T4 32 33.6%	T1 59 16.7%	T1 53 29.1%	T3 41 22.65%	T3 3 27.27%
	T2 6 6.3%	T4 26 7.3%	T2 35 19.23%	T2 22 12.15%	FT3 1 9.09%
	MT1 5 5.2%	MT1 23 6.5%	T4 30 16.48%	FT3 4 2.20%	
	FT3 5 5.2%	FT3 12 3.4%	MT1 6 3.29%	[543] 3 1.65%	
	[543] 1 1.0%	[345] 1 0.28%	[543] 1 0.54%	MT1 3 1.65%	
	LT4 1 1.0%	[445] 1 0.28%		LT4 1 0.55%	

Table 4. 4. Korean speakers' substitution patterns

Target Tones	T1	T2	T3	T4	T5
Substitute tone counts & percentages	T4 93 62.8%	T3 247 67.3%	T1 71 34.8%	T1 128 67.0%	T3 12 57.1%
	T3 34 22.9%	T1 66 17.9%	FT3 60 29.4%	T3 40 20.9%	T1 5 23.8%
	[543] 9 6.0%	T4 40 10.8%	T4 53 25.9%	[543] 13 6.8%	FT3 2 9.5%
	FT3 6 4.0%	MT1 9 2.4%	T2 11 5.3%	FT3 5 2.6%	T4 2 9.5%
	MT1 5 3.3%	FT3 5 1.3%	MT1 7 3.4%	T2 3 1.5%	
	53 1 0.6%		[54] 1 0.4%	MT1 2 1.0%	
			[543] 1 0.4%		

By examining the detailed substitution information, substitution patterns for each target tone types are identified and discussed below.

For target T1, all learners used a lot of T3 and T4 to substitute T1 when errors occurred. However, almost half of the substitute tones in English and Japanese speakers' data sets are T3; while in Korean speakers' data set, the use of T4 is over 60%. Korean speakers tend to confuse high level tones and high falling tones. This finding is compatible with the TMS issue of Korean speakers who don't have significant difference between the error rates of T1 and T4. This may also be related to Korean speakers' most frequent use of [543] tune (an out-of-inventory error, see related discussion in §4.1) which sounds both like a T1 and a T4. It seems Korean speakers pay more attention to the nature of tone register (High for both T1 and T4) but less with the contour shape (level or falling) when they deal with high tones.

For target T2, three L1 groups have very similar substitution pattern: in most cases (over 50%) they substitute T2 with T3. It seems that among the two level tones, T1 and T3, the low tone is even more unmarked. This confirms a general tonal markedness scale *High>>*Low (Yip 2002).

For target T3, no substitute tone takes more than 40% across the three L1 data sets. The use of FT3 as a substitute tone is noticeable, and this may due to a T3 pedagogical problem. The discussion about the Tone 3 pedagogical issue is provided in next sub-section §4.2.2.3. Other than FT3, English speakers used a lot of T4, a contour tone, while Japanese and Korean speakers used a lot of T1 to substitute T3. For Korean speakers, the use of T1 is even more frequent than FT3 and ranked the first. It

indicates that Japanese and Korean learners in this study are better in preserving the contour shape (contour or level) of the target tone than English speakers.

For target T4, all learners substitute T4 with a lot of T3 and T1. English speakers have almost the same amount of T3 and T1 (about 35%), whereas Japanese and Korean speakers substitute T4 mostly with T1 (about 60%). It's not surprising for KOR speakers because they substitute target T1 mostly with T4, and they often confuse T1 and T4. The Japanese and Korean speakers' preference of T1 substituting T4 indicates that Japanese and Korean speakers probably care more about tone registers than English speakers. It seems Japanese and Korean speakers' better performance and higher proficiency of Mandarin tones are also reflected in the substitution patterns.

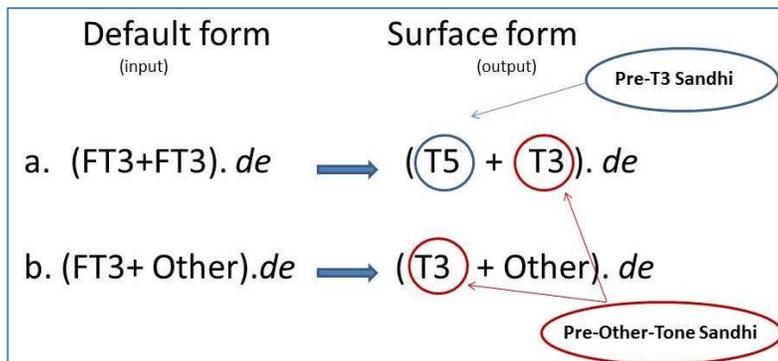
The error rates of T5, a type of T3 at Pre-T3 positions, are low, especially in the Japanese speakers' data set. When errors occur, Japanese speakers substitute T5 with T1 in most cases, while English and Korean speakers mostly substitute T5 with T3. Japanese speakers' substitute pattern of T5 (and very high accuracy rate of T3 at word-final positions) may be due to the L1 transfer of pitch accent HL which is similar to T1-T3. For English and Korean speakers, it may be because of their strong preference of T3 to substitute target T2 (as discussed above) which is also a rising tone.

4.2.2.3. A discussion on the error and substitution patterns of Tone 3

In section §4.2.2.2, we found conflicting error and substitution patterns of Tone 3. T3 serves as a substitute tone at a high rate, which means T3 is phonetically easy to produce and should have a low error rate. However, the error rates of target T3 are very high across the three L1 groups, at both word-initial and final positions (see §4.4). When we thus examine the substitution patterns for target T3, we find that all learners often substitute FT3 for target T3. In English and Japanese speakers' data, FT3 is the most often used substitute tone for target T3, while in Korean speakers' data it is the second most often used.

The reason T3 has a high error rate is very likely due to the complicated tone sandhi processes resulted from the traditional assumption of the default form FT3 for Tone 3. Following the traditional assumption of FT3 as the default form for Tone 3 and the correspondent “FT3-First” teaching method, the L2 learners have to process two Tone 3 Sandhi rules in connected speech. One is the Pre-T3 Sandhi, which is called a phonemic Sandhi, and the other is the Pre-Other-Tone Sandhi, which is also called the phonetic Sandhi rule according to Norman (1988). As discussed in Chapter 2, no FT3 should surface in Experiment 1 test materials. Below is the diagram indicating the realization of surface Tone 3 and the correspondent Sandhi these tones undergo in Experiment 1 (from a traditional view of Tone 3 Sandhi):

Figure 4.5. The default forms and surface forms of T3 in Experiment 1



The learners of this study received the “FT3-First” teaching method. Therefore, these L2 learners have to process both rules whenever Tone 3 occurs in connected speech. That is, learners have to change the FT3 into T5 [35] if it is followed by another FT3 and change all other FT3 into a T3 [21] when followed by T1, T2, T4 and neutral tone. This makes the acquisition of T3 very difficult. Because the test words are located at the first half of the noun phrase and followed by particles *de* bearing neutral tones (as shown in Fig. 4.5 above), the FT3 at word-final position in underlying form also undergo Pre-Other-Tone Sandhi and has to change into a T3.³⁰ That is, the Pre-Other-Tone Sandhi takes place both within the disyllabic test words, which may cause the high error rate of T3 at word-initial position, and also across

³⁰ When a FT3 at word-final position is followed by another FT3 in natural speech, it may be changed into either a T5 or a low T3, depending on the prosodic structure of the phrase. However, the chance of a word-final changing to a T3 is much higher than to a T5, since there are many more non-FT3 tones than FT3.

the word boundaries (but within the noun phrase) as in Pre-T3 Sandhi which may cause the high error rate of T3 at word-final positions. According to the results of Experiment 1, target T3 has high error rates at both word-initial and word-final positions (see positional information in §4.4), indicating that learners have a lot of difficulty in processing Pre-Other-Tone Sandhi, which is specifically generated based on the assumption of FT3 as the default form of Tone 3. If we assume T3 [21] as the default form, only Pre-T3 Sandhi has to be processed. Remember that in L2 learners' productions, the error rates of T5 are comparatively low across the three groups (see §4.2.1). It may be because the L2 learners process or acquire the Pre-T3 Sandhi better than Pre-Other-Tone Sandhi, or, the phonological universal constraint OCP is working in this case. This result regarding L2 learners' processing of Tone 3 Sandhi is different from the case of native Mandarin speakers. The experiment results in Zhang and Lai (2010) show that Mandarin speakers have greater accuracy with the Pre-Other-Tone Sandhi, which has a clear phonetic motivation, than the Pre-T3 Sandhi, which is with less clear phonetic motivation, in novel words. It seems the L2 learners' poor performance with Pre-Other-Tone Sandhi is not "natural" and very likely due to some "man-made reasons" such as the "FT3-First" teaching method.

This proposed account is supported by the high substitution rate of FT3 found in the current data, at both word-initial and word-final positions. While low T3 is phonetically easy as mentioned in §4.2.2.1, FT3 is phonetically difficult and more complex than other tone types in Mandarin according to the Tone Markedness Scale discussed in Chapter 2 (14). FT3 is composed of a low-registered falling and rising tone and unlikely to be picked as a substitute over some less marked tones, such as T1. Consequently, there is no reason for learners to pick FT3 as a substitute tone. However, FT3 is taught in the SLA classrooms as the basic form of Tone 3, and learners may overly produce FT3 when they process Tone 3 sandhi. That means that in the L2 learners' mind, the 'standard' phonological representation of Tone 3 is FT3 [214] instead of T3 [21]. So whenever they have to produce a target Tone 3, they are prone to map the target Tone 3 to the default or standard form of FT3. This leads to a high occurrence of FT3 in the L2

tonal productions.³¹ That is very likely the reason for the high substitution rate of FT3 for target T3 at both word-initial and word-final positions (according to the tables about positional effects in §4.4.2).

This study supports the assumption of the default form of Tone 3 as a T3 [21] instead of a FT3 [214] (Yip 2002). As introduced in Chapter 2, FT3 usually occurs at utterance-final positions or in isolation in native Mandarin. Even at these positions not all are realized as FT3 (Shi and Li 1997). It is widely acknowledged that T3 has a much wider distribution than FT3. Observing the distributions of FT3 and T3, the FT3 is more like an intonation form of Tone 3, or a lexical tone T3 [21] with some phrasal level intonation riding on it. Following the assumption of T3 as the underlying form of Tone 3, this study advocates the “T3-First” teaching method (Wang 1995, Chao, Y. 1948; Lin, T. 1979; Zhao, J-M 1988; Yue-Hashimoto, A.O. 1986) of Tone 3 and its sandhi rule. Learners under this method would take T3 as the basic form of Tone 3 and would only need to process Pre-T3 Sandhi, but not the difficult Pre-Other-Tone Sandhi, which is shown to have caused a high error rate in this study. Because T3 is also allowed at the word-final, phrase-final and utterance-final positions, a rule transforming T3 to FT3 is optional. Another advantage of the assumption of T3 as the underlying form is that it avoids the difficult task of discriminating between T2 and T3 at the beginning stage.³² Since this “T3-First” method significantly simplifies the T3 Sandhi process and reduces the difficulty in distinguishing T2 and T3, we believe it would lead to a more effective acquisition of T3, especially in producing connected speech. More pedagogical implications regarding Tone 3 acquisition is provided in Chapter 7.

4.2.3. Summary on TMS tests

³¹ Another possible explanation for the low error rate of T5 is that learners also misproduce the underlying form FT3 for target T5 (overuse of FT3), and due to the short duration of the syllables at word-initial positions, only the final rising contour of FT3 which is phonetically more salient is realized. These partially realized FT3, the rising tones, are rated as correct productions and leads to high accuracy of T5. More studies are needed to verify the proposals.

³² Many studies report that T2 and FT3 are acoustically similar, and the confusion was the prevailing major problem for learners of Mandarin Chinese at the beginning of tone acquisition (Hao, 2012; So and Best 2010; Moore and Jongman 1997; Jongman et al. 2006)

The section of 4.2.1 looks into the statistical test results of TMS and the related error and substitution patterns. It is found that, *T2>>*T4>>*T1 is statistically supported in most of the data sets, except for *T4>>*T1 in Korean speakers' data set. However, all the error and substitute tone rankings are the same across the L1 data sets and they conform to the TMS. The error rates and substitution patterns of three sub-data sets also suggest that the participants in this study, especially English speakers, have difficulty in processing T3 sandhi, especially the Pre-Other-Tone Sandhi. This pedagogical problem may cause the high error rate of T3 but also the highest substitution rates of T3 across the three L1 data sets.

Some L1 transfer effects are reflected in the substitution patterns for individual target tones. (1) When errors occur, English speakers use a higher proportion of contour tones (T2, T4 and FT3) than Japanese and Korean speakers within each error data set. (2) Korean speakers often confuse T1 and T4, and this is reflected in two facts found in this test: first, Korean speakers substitute T1 with T4, and substitute T4 with T1, both at very high rates (62.8% and 67%); second, Korean speakers employed [543], which is like a combined T1 and T4, most frequently among three groups. (3) Japanese and Korean speakers are better in maintaining the register or contour information of the target tones than English speakers, even when they make tonal errors.

4.3. The test of Obligatory Contour Principle (OCP)

This section reports the results of OCP tests. There is a “family” of OCP constraints within Optimality Theory. As mentioned in Chapter 2, the general definition of OCP was first proposed by Leben (1973) as stated in (4). Following Yip (2002), the OCP constraint in this study is presented as OCP (whole tone). The whole tone refers to an entire tone borne by a TBU instead of the component tones within a contour tone. The constraint is stated in (5).

- (4). Obligatory Contour Principle: Adjacent identical elements are prohibited.
- (5). OCP (whole tone): Two identical whole tones in adjacent syllables are prohibited.

As indicated in the Chapter 1, there are two kinds of tone combinations in Mandarin and the OCP tests. Discussions will be conducted by comparing both tone sequence types in this section. One is identical-tone combinations (**ITC**), such as T1-T1, T2-T2 and T4-T4, and the other is non-identical-tone combinations (**NITC**), such as T1-T2, T2-T4, etc. This study hypothesizes that OCP constrains the L2 tonal productions and expects L2 learners' to prefer NITC, which may be manifested in both error patterns (higher error rates of ITC than NITC) and substitution patterns (higher frequencies of NITC than ITC). In order to examine the OCP effects in the SLA data from different perspectives, two OCP tests were conducted. OCP test 1 focuses on the error rates of related tones, and the test results are reported and discussed in section §4.3.1. OCP test 2 looks into the proportion of ITC in the productions and compares the ITC proportions with the expected proportion as indicated in the stimuli. The information about Test 2 is included in § 4.3.2. Although we didn't detect much OCP effect in the Test 1, some OCP effects were supported by the statistical test in test 2. §4.3.3 summarizes the test results.

4.3.1. OCP Test 1

The OCP test 1 focuses on the error rates of the test tones at both word-initial (or “Left”; the test tone is labeled as Tone L) positions and word-final (or “Right”; the test tone is Tone R) positions. The goal of this test is to find out whether an intended tone Tx is more likely to be produced correctly when the following tone (i.e., Tx is at word-initial position, or, Tone L=Tx) or preceding tone (i.e., Tx is at word-final position, or, Tone R=Tx) is different from Tx. Since T3s go through the T3 Sandhi process in Mandarin, the test does not include T3 as a test tone. Below are the definitions of some tone types that will be used for explaining procedures comparing the error rates of test tones in different contexts.

- (6). “Tx” refers to the test tones under discussion, “Tx” could be T1, T2, T4 in Test 1.
- (7). “Ty” means any real stimuli mandarin tone other than Tx. For example, when Tx=T1, then Ty could be T2, T3 and T4.
- (8). “E” means erroneous tones for target Tx, i.e., any substitute tone for Tx; or the error rates
- (9). “N (Tx>Tx/ _Tx)” is the number of times that the learners correctly produced a target Tx as a Tx in the context _Tx. This context is labeled as Tone L=Tx in the test below.

- (10). “N (Tx>E/ _Tx)” is the number of times speakers incorrectly produced a target Tx as an E in the context _Tx (Tone L=Tx)

When Tone L=Tx, the error rates for target Tx in the two contexts of “_Tx” and “_Ty” are:

(11). $E(Tx/_Tx) = N(Tx>E/_Tx) / (N(Tx>E/_Tx) + N(Tx>Tx/_Tx)) \rightarrow$ ITC context

(12). $E(Tx/_Ty) = N(Tx>E/_Ty) / (N(Tx>E/_Ty) + N(Tx>Tx/_Ty)) \rightarrow$ NITC context

When Tone R=Tx, the error rates for target Tx in the two contexts of “Tx_” and “Ty_” are:

(13). $E(Tx/Tx_) = N(Tx>E/Tx_) / (N(Tx>E/Tx_) + N(Tx>Tx/Tx_)) \rightarrow$ ITC context

(14). $E(Tx/Ty_) = N(Tx>E/Ty_) / (N(Tx>E/Ty_) + N(Tx>Tx/Ty_)) \rightarrow$ NITC context

The test compares the error rates of Tx at ITC contexts and NITC contexts respectively. If the error rates of Tx in the ITC contexts (i.e., $E(Tx/_Tx)$ and $E(Tx/Tx_)$) are higher than those in the NITC contexts (i.e., $E(Tx/_Ty)$ and $E(Tx/Ty_)$), it means that L2 learners are better in performing the test tone under discussion in NITC than ITC, and it may imply some degree of OCP effects. The results are displayed in the three Tables 4.5, 4.6 and 4.7 for each group of L1 speakers. In the charts, the error rates of the test tones in ITC are called ITC error rates in the first line. The error rates of the test tones in NITC are called NITC error rates. We first test if there are significant difference between the NITC error rates and ITC error rates using Pearson’s Chi-Square test, and the results are listed under “Chi-Sq” in the tables below. Then we calculated the values of NITC error rates divided by ITC error rates (simplified as NITC/ITC). If OCP is working in the data sets, the error rates of NITC would be lower than ITC error rates; that is, the values of NITC/ITC are smaller than 1.

Table 4.5. English speakers’ OCP test 1 results

Test items	NITC Error rate	ITC error rate	Chi-Sq	NITC/ITC
Tone L=T1	24.17 %	25 %	0.8805	0.97
Tone L=T2	39.17 %	30 %	0.1411	1.31
Tone L=T4	53.75 %	57.5 %	0.5595	0.93
Tone R=T1	29.17 %	32.5 %	0.5731	0.9
Tone R=T2	77.08 %	81.25 %	0.4350	0.95
Tone R=T4	31.25 %	20 %	0.0534	1.56

Table 4.6. Japanese speakers' OCP test 1 results

Test items	NITC Error rate	ITC error rate	Chi-Sq	NITC/ITC
Tone L=T1	15.42 %	15 %	0.9286	1.03
Tone L=T2	29.17 %	23.75 %	0.3491	1.23
Tone L=T4	39.58 %	41.25 %	0.7921	0.96
Tone R=T1	12.5 %	20 %	0.0977	0.63
Tone R=T2	80.42 %	87.5 %	0.1516	0.92
Tone R=T4	13.75 %	25 %	0.0191	0.55

Table 4.7. Korean speakers' OCP test 1 results

Test items	NITC Error rate	ITC error rate	Chi-Sq	NITC/ITC
Tone L=T1	15 %	12.5 %	0.5810	1.2
Tone L=T2	39.17 %	23.75 %	0.0125	1.65
Tone L=T4	37.92 %	37.5 %	0.9469	1.01
Tone R=T1	32.08 %	31.25 %	0.8898	1.03
Tone R=T2	79.58 %	78.75 %	0.8732	1.01
Tone R=T4	20.83 %	25 %	0.4350	0.83

The three charts show that most of the NITC error rates are a little lower than the ITC error rates, especially in English and Japanese speakers' data sets, but there are no significant differences between the NITC and ITC error rates except for Japanese speakers' Tone R=T4. This shows very weak OCP effects, but these effects are not supported by statistical tests. All the NITC/ITC values are very close to 1, which means the error rates of NITC and ITC for each tone in all data sets are very close. There are several exceptions, or anti-OCP effects, including the cases where NITC/ITC are bigger than 1.2, namely, Tone L=T2 across all data sets and Tone R=T4 in English speakers' data set.

When Tone L=T2, all NITC/ITC values in the three data sets are higher than 1.2. This means that the error rate of T2-L is much lower when it precedes another T2 than when it precedes other tones in English, Japanese and Korean speakers' data sets. The anti-OCP effect is only with T2-L (T2 at word-initial position), not T2-R (T2 at word-final position). We looked into the breakdowns of the error rate of NITC and found it may be due to the very high error rate of T2 in the sequence of T2-T1 across all three data sets. It's always true for all three groups of speakers that the error rate of T2 preceding T1 is the highest (56.25% in English speakers' data set, 32.5% in Japanese speakers' data set and 57.5% in Korean

speakers' data set), and the T2 preceding T4 has the second highest error rates, especially in English and Korean speakers' data sets. The tone most often substituted for T2-L is T3 across the three data sets (see analysis pertaining to positional effects in §4.4) and the second most often substituted tone is T1. This conforms to the finding in Wang (1997) on a phonetic study of T2 focusing on co-articulation cases. Wang (1997) found that when the tone at word-final position is a high level or a high falling tone, the T2 at word-initial has the highest error rate and is often substituted by a low tone in an SLA study of English and Japanese speakers. As indicated in Wang (1997), this tendency may be related to anticipatory effects. Xu (1997) discovered that in native Mandarin, anticipatory effects mostly result in dissimilation: a low onset value of a tone raises the maximum F0 value of a preceding tone. According to the similar findings of this study and Wang (1997), it's very possible that a high onset value of tone may lower the F0 value of a preceding T2 and cause a misproduction of T3. That is, the high error rate of T2 preceding T1&T4 discussed above could be regarded as a case of anticipatory effect in SLA: the high onset value of a T1 and T4 may lower the F0 value or the register of the preceding T2 and leads to a substitute T3. More studies are needed to support this proposal.

Another exception is the case of Tone R=T4 in the English speakers' data set. The error rate of T4 following another T4 is expected to be higher than the T4 following other tones. However, the error rate of T4-R in the ITC context is much lower than other T4s. It should be noted that the general error rates of all sequences ending with T4 are very low, and actually are the lowest compared to other tone types at word-final position. The odd anti-OCP effect with T4-R comes from the relatively lower error rate of T4 following another T4 than T4 following other tones.

Interestingly, the two odd cases (T2-L across all data sets and T4-R in English speakers' data) occur with the tones at these positions having the lowest error rates in general (see positional effect in §4.4), and they demonstrate strong positional effects. It seems that when the error rates of the same tone types at different contexts are all low, the error rate of this tone in NITC sequences, on the contrary, is higher than in ITC.

In summary, OCP test 1 compares the error rates of specific tones (T1, T2, T4) in the contexts of ITC and those in NITC at both word-initial and word-final positions. The error rates are very close to each other in most cases; therefore, no strong OCP effects are detected. Although there is no significant difference between the error rates of tones in these two contexts, most of the NITC error rates are a little smaller than ITC error rates. This demonstrates some weak OCP effects, especially in English and Japanese speakers' data sets. Two exceptions are T2-L across all data sets and T4-R in English speakers' data, where those NITC cases possess higher error rates, and it is proposed that the "universal" T2-L case is because of the high error rates of T2-T1 sequence due to anticipatory effects.

Comparing only the error rates of these two kinds of tone combinations as in Test 1 may not be sufficient because the error count is not the only factor affecting the results of OCP testing. That is, the examination of error rates only partially demonstrates the OCP effects in SLA of Mandarin tones because the error rates do not directly correlate to the speakers' preferences of ITC or NITC. For example in test 1, when $T_x=T_2$, it has higher error rates for target T2-T2 than target T2-T1. The error rate pattern may show an OCP effect, but the substitution pattern may demonstrate very different results. For example, it's possible that the actual productions (or the substitution pattern) of the target T2-T2 is changed into T1-T1 and the target T2-T1 is also changed into T1-T1, and the substitution pattern would, on the contrary, demonstrate a strong effect of anti-OCP. Test 1 only cares about the error rates of T_x but not those of T_y or substitutions. In the following OCP test 2, more consideration is given to the substitution patterns.

4.3.2. OCP Test 2

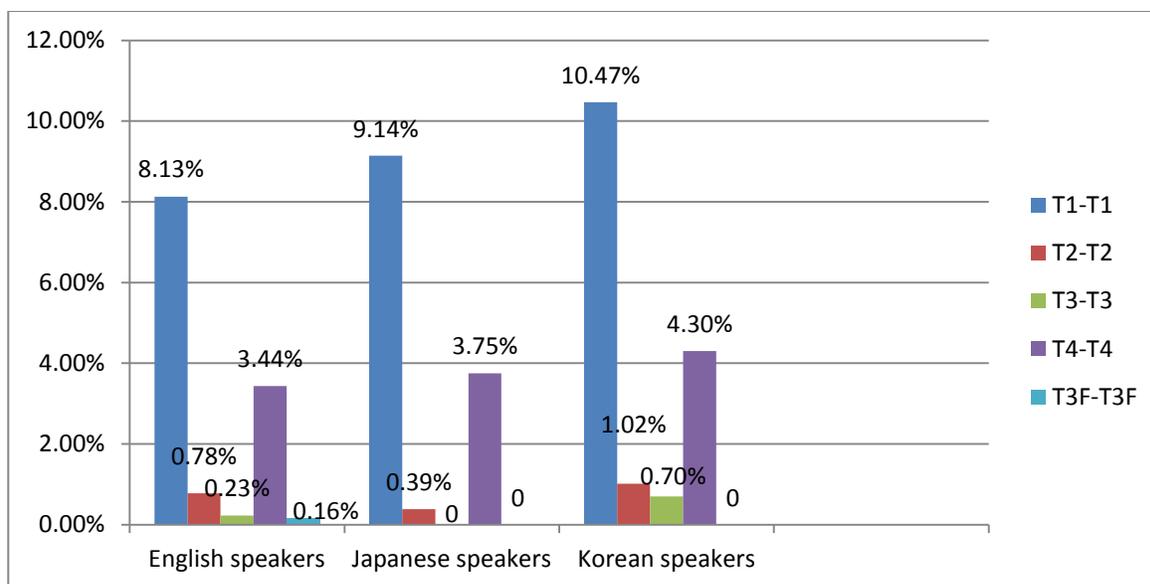
The OCP Test 2 focuses on the actual L2 tonal productions. The goal of test 2 is to find out if the proportions of ITC match those of the stimuli in each sub-data set. The 16 free tone combinations of T1, T2, T3 and T4 are evenly proportioned, and the ITC sequences as a whole constitute 3 out of 16 portions in total (T3-T3 should be produced into T5-T3). Therefore, without OCP effects, we would expect that the ITC productions would also take 3/16 (about 20%) of the total productions in each data set. Test 2

compares the observed ITC sequence percentages in the production with the expected percentages in the three sub-data sets. In each L1 data set, the general ITC proportions are significantly lower than the expected 3/16, which suggests that OCP is relevant in general. We go a step further and look into the breakdowns of each ITC sequence and find that the individual ITC sequence proportions vary a lot. The proportions of T1-T1 in all L1 data sets are significantly higher than other ITC sequences. It demonstrates some interacting effects of L1 transfer and OCP, or, the interacting effects of TMS (through L1) and OCP.

The ITC rates were compared with their expected values using the binomial test. The ITCs as a whole consist of 12.72% in English speakers' L2 tonal productions, which is significantly lower than 20% ($P < .0001$); the ITCs consist of 13.28% of the Japanese speakers' data set, which is significantly smaller than expected ($P < .0001$). Korean speakers' ITCs take 16.48%, the highest among the three groups ($P = 0.0008$; One-sided $Pr > Z$). Based on the statistical test results regarding the proportions of ITCs in general, OCP effects are obvious in all three subsets. Korean speakers produced the biggest amount of ITC sequences. This is compatible with the results of Test 1. This may be due to the L1 transfer of typical High-High (H-H) pitch pattern (like T1-T1 sequence) from Korean. The ITC breakdowns demonstrate similar patterns across the three L1s (see Fig. 4.6). All of them have T1-T1 as the highest proportion (English speakers 8.13%, Japanese speakers 9.14% and Korean speakers 10.47%), then T4-T4 (between 3% and 4%), T2-T2 (about 1%) and the T3-T3 as the smallest proportion (all about 0.5%). The occurrence rates of T1-T1 in all data sets are significantly higher than expected (both 1/16 and 1/25) and other tones are significantly lower than expected. The patterns somewhat reflect the TMS effects in tone pairs.³³

Figure 4.6. The occurrences of tone pairs in L2 productions (OCP test 2)

³³ 1/16 comes from the proportion of individual tone combination type in the stimuli (16 types of tone combination in total); 1/25 comes from the proportion of individual tone combination type in the production only (including the tone type of FT3; therefore 25 types of tone combinations in total).



We also compare the individual tone pair rates with their expected values (1/16) using the binominal test. It is found that the occurrence rates of T1-T1 across all L1 data sets are significantly higher than the expected individual proportion, whereas the T2-T2 and T4-T4 rates are significantly lower than the expected proportion. The high occurrence of T1-T1 is an anti-OCP effect. It might be attributed to both L1 transfer and the universally preferred T1, an unmarked tone. As indicated in Chapter 2, T1-T1 (H-H) is a typical L1 tone sequence in Korean and Japanese, and H-H is allowed in English. There might be a transfer of this tone sequence into the L2 production. This also explains why the T1-T1 rates are higher in Japanese and Korean speakers' data sets than in English speakers' data set, as T1-T1 is not a typical tone pattern in native English. In the data set, we also found many more T4-T4 than T2-T2, a pattern which is not motivated by L1 grammars nor L2 grammar. We will account for this pattern in Chapter 6. The highest frequency of T1-T1 across three sub-datasets involves the universal TMS effects through L1s, but the higher frequency of T4-T4's as opposed to T2-T2 is argued to represent the situation of the "Emergence of the Unmarked" (TETU) in Chapter 6.

In the L2 tonal productions, extremely low occurrence rates of T3-T3 and FT3-FT3 across all data sets are found. This may be due to the L2 tonal grammars (now allowing T3-T3 and FT3-FT3), or,

strong OCP effects on T3 pairs and FT3 pairs, although T3 is the individual tone most often substituted for other tones. If the latter is the reason, it seems the conjunction of TMS and OCP are somehow contingent on tone types.

4.3.3. Summary of OCP tests

In section §4.3, the results of the two OCP tests are presented. Test 1 examines the error rates of specific tones at both word-initial and word-final positions. The error rates of the test tone in the contexts of ITC and NITC are compared. Some weak OCP effects are found in most cases, but they are not statistically supported. A few anti-OCP effects are detected when T2 is at word-initial position across all three sub-data sets and when T4 is at word-final position in English speakers' data set. The high error rate of T2-L in ITC is related to the high error rates of T2-T1, which may be due to an anticipatory effect pertaining to T2 at word-initial positions.

Test 2 compares the proportions of ITC tone sequences (identical-tone-combinations) in the L2 tonal productions and the expected portions in stimuli. It is found that the proportions of ITC sequences as a whole are significantly lower than expected, which demonstrates a strong OCP effect. Looking into the proportions of individual ITC sequences in the L2 productions, it is found there are many more T1-T1 than T4-T4, and in turn more T4-T4 than more T2-T2 across the three sub-datasets. It seems that the high occurrences of T1-T1 in Japanese and Korean are due to L1 transfer. More discussions regarding these tone pair proportions are offered in Chapter 6. We thus conclude that some OCP effects are present, especially when we examine the association of target tones and L2 productions or the distribution of productions. However, the OCP effect is not very strong in some cases because it often interacts with other factors, such as TMS effects, L1 transfers and anticipatory effects. All these tests indicate that Korean speakers are most tolerant with tone pairs and English speakers most disfavor tone pairs.

4.4. Test of Tone-Position Constraints (TPC)

This section presents the hypothesis test for TPC constraints and discussions of L1 transfer with respect to positional effects in Experiment 1. The §4.4.1 presents the statistical test results pertaining to the positional effects in the three L1 sub-data sets. We then further examine some confusion matrix charts to detail the error information within each L1 group and discuss the L1 transfer effects in §4.4.2. In §4.4.3, we again look into the substitution tones but focus more on the interaction of UG and L1 transfers, combining the findings in the preceding two sub-sections. A brief summary is provided in §4.4.4.

4.4.1. General positional effects and statistical test of TPC

Tone-Position Constraints (TPC), as mentioned in Chapter 2, are about positional effects in the L2 tone realization (especially the contour tones) and include several related constraints. Following Zhang, J.'s (2004) claim that “phrase-final syllables and syllables in shorter words are the preferred bearers of contour tones,” we propose a constraint (with two sub-constraints) to illustrate this finding:

- (15). * Contour-I >> * Contour-F (word-final positions are better bearers of contour tones)
- a. * Contour-I: Contour tones (rising and falling) are not allowed in word-initial positions.
 - b. * Contour-F: Contour tones are not allowed in word-final positions

The constraints (15) are expanded into two sub-rankings to fit for this study:

- (16). a. * Rise-I >>* Rise-F:

Rising tones are more disfavored at word-initial positions than at word-final positions.

- b. * Fall-I >>* Fall-F:

Falling tones are more disfavored at word-initial positions than at word-final positions.

Due to the findings about the fall-rise asymmetry (Zhang, J. 2004b), we hypothesize that this kind of asymmetry may also affect L2 tonal productions and propose constraint (17) below. That is, while the contour tones (both rising and falling) are performed better at word-final positions than at word-initial positions in general, the falling tones may be performed better than rising tones at word-final positions.

- (17). * Rise-F >>* Fall-F:

Rising tones are more disfavored than Falling tones at word-final positions.

This section focuses on the positional effects regarding the tone realization in L2 tonal productions at the word level. To have a complete view of the positional effects, this study examines the realization of all tone types at the word-initial and word-final positions. Both error rates and substitution rates will be considered to figure out the positional preference of each tone for different types of learners. We hypothesize that TPC is working in the L2 tonal productions and predict higher accuracy rates (or lower error rates) of contour tones, such as T2 and T4, at word-final positions than at word-initial positions within disyllabic words, and higher accuracy rates of T4 than T2 at word-final positions in Experiment 1. The results supplement the findings for the TPC principle and suggest the word-final syllables are the preferred bearers of T4, a high falling contour tone, but not for T2, a rising contour tone, in the L2 tonal productions. We confirmed the positional effects in Experiment 2 as reported in Chapter 5.

The SURVEYFREQ procedure was used to work out the numbers of clusters and observations. We also tested the null hypothesis of equal error rates (no association) between word- initial versus final positions using Rao-Scott Chi-Square Test in the general data set and also within each subset. The results show two patterns. (1) In both the general data set and all three L1 data sets, the error rates of T2 at word-initial positions are significantly lower than at word-final positions, and the error rates of T4 at word-initial positions is significantly higher than at word-final positions. It seems *Rise-F and *Fall-I apply for all three L1 speakers, or, *Rise-F and *Fall-I are ranked high in the interlanguage grammars. (2) For Korean speakers, the error rate of T1 at word-final positions is significantly higher than at word-initial positions. The Table 4.8 displays the error rates at the word-initial and word-final positions for each data set. The first column contains the tone names, and the following columns contain the error rate information of the overall data set and that of the individual language groups, with the percentages of errors at word-initial positions at left followed by errors at word-final positions. Since T5 is always at the word-initial position, this chart excludes T5. The statistical test results were displayed under each pair of

initial and final positions for each target tone. Those P values smaller than .05 indicating significant differences between the error rates at different positions are highlighted in bold and the cells are shaded.

Table 4.8. Error patterns with positional information (Statistical test results included)

Tone	General		English		Japanese		Korean	
	Initial	Final	Initial	Final	Initial	Final	Initial	Final
T1	18.02%	25.42%	24.38%	30%	15.31%	14.38%	14.38%	31.88%
	0.0567		0.4383		0.8847		0.005	
T2	33.33%	79.9%	36.88%	78.13%	27.81%	82.19%	35.31%	79.37%
	< .0001		< .0001		< .0001		< .0001	
T3	40.56%	37.08%	52.08%	42.81%	37.08%	29.06%	32.5%	39.38%
	0.4992		0.3618		0.3368		0.3939	
T4	44.17%	22.29%	54.69%	28.44%	40%	16.56%	37.81%	21.88%
	< .0001		0.0022		< .0001		0.0193	

(18) below summarizes the highlighted positional effects based on the error patterns. Each asterisk “*” represents a 5% difference between the error rates of the word-initial position and word-final position for that tone. “I” signifies word-initial positions, and “F” signifies word-final positions. For example, the “**T4-I” in English speakers’ data means English speakers’ T4 error rate at the left position is about 10% greater than at the right position, which means English speakers perform T4 poorly at word-initial positions.

(18) Positional features based on error rates

- English speakers: **T4-I || ****T2-F
- Japanese speakers: **T4-I || *****T2-F
- Korean speakers: *T4-I || *T1-F, ****T2-F

The error patterns indicate that for English and Japanese speakers, T2 at word-initial position are more resistant to change than they are at word-final position; T4 at word-final positions are more resistant to change than they are at word-initial position. For Korean speakers, T4 at word-final and T2 at word-initial position are more resistant to change than at the opposite position. In addition, T1 at word-initial position are more resistant to change than at word-final position. The *T4-I and *T2-F seems true across the three L1 groups, and we’ll confirm this finding with the substitution patterns below. It is very hard to

explain Korean’s positional preference of T1 (with lower error rate) at word-initial positions in terms of L1 transfer because only 6 out of 32 test words (about 19%) begin with aspirated consonants. Bearing the error patterns in mind, we detail the error information and explore the L1 transfer effects by considering the substitute tones with respect to positions in the next section.

The error rates of T3 at word-initial positions are much higher than those at word-final positions in English and Japanese speakers’ data sets, but lower than the T3 error rate at word-final positions in Korean speakers’ data. It indicates that English and Japanese speakers have more difficulty in processing Pre-Other-Tone Sandhi while Korean speakers may have much difficulty in processing Pre-T3 Sandhi.

4.4.2. Positional error rates details: substitution patterns

The following three confusion matrix charts demonstrate the distribution of the response tones (both correct productions and substitute tones) for specific target tones in each L1 data set. We examine these charts first with a focus on the positions with high error rates, and then compare the substitution rate patterns in terms of positions. The discussions involve L1 transfer and T3 sandhi when necessary. In each chart, the tones in the first line are the target tones at both word-initial and word-final positions. The tones in the first column are the response tones. The cells containing correct response tone rates are shaded. The highest percentages of substitutions for the target tones (within each column) are bolded and labeled with “*”. This way, we can see the most frequently used tones substituted for the target tones with very high error rates at specific positions. By examining the chart, we found some L1 transfer effects in each group.

Table 4.9: English speakers’ response-tone rate details with positional information (%)

Target Resp	T1		T2		T3		T4		T5
	Initial	Final	Initial	Final	Initial	Final	Initial	Final	Initial
T1	75.6	70	13.8	10.6	11.3	7.8	*19.7	7.6	12.5
T2	5.9	0	63.1	21.9	15	0	14.7	0	65
T3	*9.1	*17.5	*14.7	*43.1	47.9	57.2	12.5	*16.9	20
T4	5.6	10.6	5.6	14.4	5.4	16.3	45.3	71.6	2.5
FT3	2.8	1.3	1.9	6.6	*19.6	*17.8	5.9	2.2	0
OTHER	0.9	0.6	0.9	3.44	0.8	0.9	1.88	1.88	0

In the English speakers' data set, the three positions with the lowest accuracy rate (highest error rates) are T2-F, T3-I and T4-I. They are substituted by T3-F, FT3-I and T1-I most often, respectively. The substitution of T3-F and T1-I are compatible with the dominant English trochaic foot structure (Delattre 1965), which could be presented as a [+Upper] [-Upper] (or [+U] [-U]) register pattern according to De Lacy (2002)'s proposal of Tonal Prominence scale. We will detail this discussion in section §4.5.2 on the most frequently produced tone sequences.

Next, we look at the substitute tones with regards to their positional preferences. T1 is the most often used substitute tone for target T4 at word-initial position. Actually, when T1 serves as a substitute tone, it occurs more often at word-initial position than final position across all tone target tone types, in spite of the fact that T1 as a target tone has similar high accuracy rates at both word-initial and final positions. T2 is performed better at word-initial positions. This is confirmed by the substitution pattern of T2 because T2 is never substituted at word-final positions for any target tone. T3 is the tone most often substituted for other tones, especially for T1 and T2 at word-final positions. This is compatible with the higher accuracy rate of T3 at word-final positions. Opposite of T2, the accuracy rate of T4 is higher at word-final position, and T4 also serves as a substitute tone mostly at the word-final position. FT3 was substituted for T3 most often, both at word-initial and final positions. In summary, English speakers perform T1 and T2 better at word-initial positions and T3 and T4 at word-final positions. It implies that word-final syllables are preferred bearers of T3 and T4 for English speakers.

Table 4.10. Japanese speakers' response-tone rate details with positional information (%)

Target Resp	T1		T2		T3		T4		T5
	Initial	Final	Initial	Final	Initial	Final	Initial	Final	Initial
T1	84.7	85.6	10.9	7.5	10.8	8.4	*25.6	*7.8	8.8
T2	1.9	0	72.2	17.8	*13.8	0.6	6.3	0.6	86.3
T3	*9.7	4.4	*15.9	*55.9	62.9	70.9	6.3	6.6	3.8
T4	1.9	*8.1	0.9	7.2	3.3	6.9	60	83.4	0
FT3	0	0.6	0	3.8	8.8	*11.3	0.9	0.3	1.3
OTHER	1.9	1.25	0	7.8	0.4	1.88	0.9	1.25	0

Similar to English speakers, the tones which most often substitute for target tones with the top three highest error rates (i.e., T2-F, T4-I and T3-I) are T3-F, T2-I and T1-I for Japanese speakers. These substitutions are compatible with the Japanese pitch accent pattern “HL” and they have high substitute tones (T1 or T2) often at word-initial positions and low substitute tones (T3) at word-final positions.

The accuracy rates of T3 and T5 are both high and higher than other L1 speakers. It seems Japanese speakers apply T3 sandhi better than other speakers, and this may be related to their overall high Chinese proficiency level. When Japanese speakers make errors in producing T3, T2 is substituted the most often at word-initial positions, and FT3 is substituted at word-final positions. These patterns are compatible with the typical pitch accent “HL” pattern, since the T2-FT3/T3 sequence suggests a [+U] [-U] register pattern. The high substitution rate of T2 at word-initial position also reflects the phrasal level intonation structure because the AP in Japanese typically begins with a rising tone.

With respect to the substitute tones’ positional preferences, same as English speakers, T1 is substituted for other tones more frequently at word-initial positions than word-final positions, although T1 has very similar accuracy rate at the two positions. T2 is substituted for other tones more often at word-initial positions, and this conforms to its error pattern, where T2-R has very high error rate. T3 does not have a clear positional preference since it is substituted for T1 more often at word-initial positions but is substituted for T2 more often at word-final positions. T4 is often substituted for other tones at word-final positions, and it conforms to its error pattern, where T4-L always has a higher error rate than T4-R.

Table 4.11. Korean speakers’ response-tone rate details with positional information (%)

Target Resp	T1		T2		T3		T4		T5
	Initial	Final	Initial	Final	Initial	Final	Initial	Final	Initial
T1	85.6	68.1	5	15.6	10.8	*14.1	*29.4	*10.6	6.3
T2	0	0	64.7	20.6	4.2	0.3125	0.6	0.3	73.8
T3	5.6	5	*27.5	*49.7	67.5	60.6	4.4	8.1	15
T4	*6.6	*22.5	1.9	10.6	5.4	12.5	62.2	78.1	2.5
FT3	1.3	0.6	0.6	0.9	*11.7	10	0.9	0.6	2.5
OTHER	0.9	3.8	0.3	2.5	0.4	2.4	2.5	2.2	0

The tone most often substituted for T2-R, which has the lowest accuracy rate in Korean speakers' data set, is T3. This pattern is not compatible with Korean speakers' LH or HH disyllabic tone patterns, but maybe due to the unmarkedness of T3 cross-linguistically. In native Korean, if there is any syllable not mapped to the underlying tones, it will receive a Low tone by default (Jun 1996). The other substitution patterns, however, show some L1 transfer of Korean speakers' typical LH or HH tone patterns onto disyllabic words. For example, the tones most often substituted for T1 are T4 at both positions which is compatible to the [+U] [+U] register pattern like HH, for T3 are FT3-L and T1-F ([-U] [+U] registered, like LH pattern), and for T4 are T1 at both positions ([+U] [+U] registered, like HH). Interestingly, T1 and T4 substitute for each other the most often. We have seen similar patterns in the TMS test (see §4.2).

Korean speakers' substitution patterns pertaining to T3 Sandhi rules are similar to English speakers, although both the accuracy rates of T3 and T5 are a little higher. FT3 is the tone most often substituted for T3, and T3 is the tone most often substituted for T5. It again indicates that their assumption of FT3 as the standard form of Tone 3 influence L2 learners' acquisition of Tone 3 sandhi.

It is difficult to identify the positional preference of substitute tones T1 and T3, since they have conflicting positional patterns for different target tones. T2 has a very low substitution rate, and it prefers word-initial positions. T4 is substituted for other tones more frequently at word-final position.

4.4.3. The correlation between error and substitution patterns for positional effects

In the previous sections, we examined the error and substitution patterns considering positional information. The examination of the substitution patterns in the last section mainly focuses on the patterns for target tones with very high error rates. The error patterns tell us which tones the L2 learners perform poorly, and at what specific positions they do so, while the substitution patterns tell us which tones the L2 learners favor or perform well, and at what specific positions they do so. These two aspects facilitate different kinds of observations, but they often correlate in the L2 interlanguages. In this section, we

observe both the error patterns and the general substitution tones (not associated to the error rates) and discuss how the correlation between error and substitution patterns reflects the interaction of universal positional effects and L1 word-level prosody transfer.

The Table 4.12 displays the overall substitutions for the three L1 groups without error information. The first line contains the name of the language groups and the total number of substitute tones in each group. The first column lists the substitute tones employed by L2 learners when the target tones were incorrectly produced. The first five tones are the within-inventory substitute tones, and the fifth is “others” which includes those out-of-inventory substitute “tones.” In each cell, the counts of the substitute tones followed by their percentages are listed. The relatively higher substitution percentages for the two positions for each tone in each L1 data set are shaded. The percentages under each L1 are out of all substitute tones in that L1 data set. That is, the sum of all percentages of substitute tones under each first language is 100%, or, the percentages are out of total error numbers in specific L1 set.

Table 4.12. Overall Substitutions with positional information (without target tone information)

Substitute tones	English/1098		Japanese/821		Korean/931	
	initial	final	initial	final	initial	final
T1	144 (13%)	83(8%)	150(18%)	76(9%)	141(15%)	129(14%)
T2	102(9%)	1(0%)	59(7%)	4(0%)	12(1%)	2(0%)
T3	132(12%)	248(23%)	105(13%)	214(26%)	132(14%)	201(22%)
T4	51(5%)	132(12%)	17(2%)	71(9%)	42(5%)	146(16%)
FT3	81(7%)	89(8%)	28(3%)	51(6%)	39(4%)	39(4%)
Others	14 (1.3%)	21(1.9%)	7(0.6%)	39(3.6%)	13(1.2%)	35(3.2%)

Combining error patterns in Table 4.8 and substitution patterns in chart 4.12, patterns of correlation with respect to positions appear. In Table 4.13, 4.14 and 4.15, we put the error and substitution pattern data together to make comparisons. An asterisk “*” labels the tones at specific positions with a higher error rate for a target tone at either word-initial or final positions based on Table 4.8. Check signs “√” label the higher substitution rates at either positions for specific tones based on chart 4-12. If there is

no apparent preference (difference of error rates at word-initial and final positions is less than 5% in Table 4.8, or, the count difference is less than 10 in Table 4.12), both “I” and “F” will be labeled.³⁴

Table 4.13. The comparison of positional errors and substitutions in English speakers’ data set

Positional errors	*T1-F	*T2-F	*T3-I	*T4-I
Positional substitutions	$\sqrt{T1-I}$	$\sqrt{T2-I}$	$\sqrt{T3-F}$	$\sqrt{T4-F}$

Table 4.14. The comparison of positional errors and substitutions in Japanese speakers’ data set

Positional errors	*T1-I,F	*T2-F	*T3-I	*T4-I
Positional substitutions	$\sqrt{T1-I}$	$\sqrt{T2-I}$	$\sqrt{T3-F}$	$\sqrt{T4-F}$

Table 4.15. The comparison of positional errors and substitutions in Korean speakers’ data set

Positional errors	*T1-F	*T2-F	*T3-F	*T4-I
Positional substitutions	$\sqrt{T1-I}$	$\sqrt{T2-I}$	$\sqrt{T3-F}$	$\sqrt{T4-F}$

In most of the L2 productions, the positional error rates correlate negatively to the positional substitution rates; where the error rates are high, the substitution rates are low, and vice versa. It signifies that the substitution rates confirm the error rates with respect to positional effects. There are at least two evident cross-linguistic positional effects based on both positional error and substitution information: one with regard to T2 and one with regard to T4. The cross-linguistic positional effect of *T2-F is supported by the statistical analysis in §4.4.1, and confirmed by the substitution patterns ($\sqrt{T2-I}$) across the three groups of language speakers’ data. The positional effect of T4 is also very strong. T4s have much higher error rates at word-initial positions, and this is statistically supported in §4.4.1. This is confirmed by its substitution patterns, where T4 serves as a substitute tone more often at word-final positions than at word-initial positions across the L1 data sets. Therefore, this SLA study of Mandarin tones provides a supplemental finding about positional licensing of contour tones to the discussion in Zhang, J. (2004). Zhang, J. (2004) claimed that “phrase-final syllables and syllables in shorter words are the preferred bearers of contour tones, even though they are usually not privileged for other phonological contrasts.”

³⁴ The percentages in Table 4.8 are out of the total errors within the same tone type, but not out of the total errors of all tone types, different from the Table 4.12.

This study on L2 tonal productions demonstrates that the word final position is a preferred bearer of a falling contour tone (T4) rather than a rising contour tone (T2).

To summarize, in L2 learners' non-native tonal productions, we only found that the constraints *Fall-I>>*Fall-F and *Rise-F>>*Fall-F are working. Contour tones usually place a higher demand on the duration and sonority of the rhyme. Word-final position may be a better bearer of contours because word-final positions are subject to lengthening and provide more duration. We didn't detect effects of *Rise-I>>*Rise-F. On the contrary, we found very strong effects of *Rise-F based on the extremely high error rates and low substitution rate of T2 at word-final positions. So it seems there is a very clear T2 and T4 asymmetry in the L2 tonal productions: T2 is disfavored at word-final positions and T4 is disfavored at word-initial positions. The positional effect of *T4-I is predicted in this study, but *T2-F is unexpected. According to Zhang, J. (2004)'s finding that T4 has wider distribution than T2, it's not surprising that T2 is performed poorer than T4 at word-final positions, but it's unclear why T2 is better at word-initial than at word-final positions. It is possible that people often connect rising intonation with 'questions,' so they avoid T2 at word-final in order not to have a misleading semantic meaning. It may also be due to a very prevailing prosodic pattern that utterances usually begin low and then rise, but have a falling or a low tone at utterance-final position.

The only conflicting case is Korean speakers' T3 positional effect. Korean speakers have a higher error rate of T3 at word-final position, which is compatible with its typical "LH" pattern, but T3 is substituted for target tones more often at word-final positions than word-initial ones. The positional effect residing in the error patterns reflects L1 transfer. However, the high substitution rate of T3 at word-final position may be due to the high error rate of T2 at word-final position. The high error rate of T2-F thus leads to a high rate of substitution for tone T3. This reflects a cross-linguistic preference for T3, an unmarked tone, to substitute a more marked tone. We will see in a similar substitution pattern in Experiment 2 (chapter 5) that all L1 speakers, including Korean, produce a lot of trochaic footed tone patterns [+U] [-U]. More discussion will be conducted in Chapter 5 and 6.

4.4.4. Summary of TPC tests

This section examines the error and substitution patterns in the three L1 data sets focusing on positional effects. The statistical analyses show that T2 is performed better at word-initial position, and T4 is performed better at word-final position across all L1 data sets. The patterns are confirmed by the substitution patterns. This is a supplemental finding about positional licensing of contour tones to the findings discussed in Zhang, J. (2004). The L2 data suggest that the word-final position is a preferred bearer of a falling contour tone (T4) rather than a rising contour tone (T2). We also analyzed the positional patterns of T1 and T3 based on the error and substitution patterns. Although positional effects of T1 and T3 based on the error rates are not supported by the statistical tests, the error and substitution patterns reveal better performance of T1 at word-initial positions across the three L1 data sets and T3 at word-final positions in English and Japanese speakers' data sets. The conflicting case of high error rate of T3 and high substitution rate of T3 at word-final positions in Korean speakers' data set demonstrates an interaction effect of L1 transfer and universal tendency. Examination of the associated information of errors and substituted tones in the three confusion matrix charts suggest some L1 transfer effects, especially the substitution patterns for those tones with highest error rates. That is, those patterns reflect English speakers' preference for trochaic foot patterns, Japanese speakers' typical HL pitch accent pattern and Korean speakers' typical LH and HH tone patterns on disyllabic words. The detailed confusion matrix charts also confirm that all three groups of speakers processed Pre-T3 Sandhi better than Pre-Other-Tone Sandhi. The pedagogical problem of T3 may affect L2 learners' performance of T3, especially in the data sets of speakers (English and Korean) possessing lower Mandarin proficiency levels.

Sections §4.2, §4.3 and §4.4 reported the TMS, OCP and TPC test results. The TMS tests show that the TMS “*T2>>*T4>>*T1” is relevant in general and across all three L1 sub-data sets examining the error rates and substitution patterns. Only Korean speakers' *T4>>*T1 is not statistically supported. The two OCP tests examine the error and substitution patterns of ITC and NITC sequences from different perspectives. It is found that OCP is relevant in most of the cases to some degree, but it is interacting with

other factors, such as L1 transfers and TMS. The TPC tests investigated the positional effects in the L2 tonal productions and found * Fall-I >>* Fall-F and *Rise-F>>*Fall-F are relevant across all three L1 data sets. It is found that these phonological universals are working together with L1 transfer effects to some degree, and these interactions of multiple factors are also discussed in section §4.4. We then move on to conduct a test of typical L1 pitch pattern transfer effects in the next section §4.5 to address the L1 transfer hypotheses.

4. 5. The test of typical L1 pitch pattern transfer

In Chapter 2, it is hypothesized that typical L1 pitch patterns may be transferred to L2 tonal productions. That is, L2 learners may acquire those tone combination types similar to the tones often used in L1 intonation quicker than other tone patterns. If this transfer occurs, those target tone sequences mimicking L1 F0 patterns would have a higher occurrence in L2 productions than expected. The test result of the L1-like pitch pattern transfer is reported in §4.5.1. Section §4.5.2 discusses L1 transfer based on the observation of the “most frequently produced tone sequences.” The hypothesis is supported in Experiment 1, since the frequencies of the proposed counterpart Mandarin tone sequences mimicking L1 pitch patterns on disyllabic words are higher than expected.

4.5.1. Statistical test of L1-like pitch pattern transfer

Similar to OCP Test 2, the L1 transfer tests look into the proportions of tone combinations which mimic typical L1 F0 patterns in the L2 tonal productions. Since the 16 types of tone combinations are evenly proportioned in the stimuli, the individual tone combination occurrence is expected to be 1/16, and this is also the expected proportion for each individual tone combination in the L2 productions. The following test tries to find out if the tone combinations which are similar to those in their L1s are significantly higher than 1/16 (6.25%). The counterpart Mandarin tone sequences mimicking typical L1

pitch patterns are: English speakers' T1-T3, T2-T3, T3-T1, T2-T4, T3-T4; Japanese speakers' T1-T3, T2-T1, T2-T4, T3-T1, T3-T4 and T4-T3; Korean speakers' T1-T1, T1-T4, T3-T1 and T3-T4.³⁵

Tone sequences rates were compared with their expected values using the binomial test. The results in table 4.16 show that most of the tone sequences similar to the L1 tones are significantly higher than expected, except for T2-T1, T2-T4 (labeled with *) and T4-T3 in Japanese data set. The occurrence rate of T4-T3 sequence in Japanese speakers' data set is 7.5% which is higher than 6.25%, but not fully supported by the statistical test.

Table 4.16. The test results of L1 transfer of typical F0 patterns

English speakers		Japanese speakers		Korean speakers	
Productions	Pr<Z	Productions	Pr<Z	Productions	Pr<Z
T1-T3 (10.55%)	<.0001	T1-T3 (11.56%)	<.0001	T3-T4 (11.64%)	<.0001
T2-T3 (10.39%)	<.0001	T3-T1 (9.45%)	<.0001	T1-T1 (10.47%)	<.0001
T3-T4 (8.67%)	.0002	T3-T4 (8.52%)	.00004	T1-T4 (9.22%)	<.0001
T3-T1 (8.36%)	0.0009	T4-T3 (7.5%)	0.0323	T3-T1 (8.52%)	0.0004
T2-T4 (7.66%)	0.0188	T2-T4 (5.55%) *	0.1493		
		T2-T1 (4.53%) *	0.0055		

The proportion of T1-T1 in Korean speakers' data set is as high as 10.47% (and also Japanese speakers' data set with 9.14%) indicate the interaction of L1 transfer and OCP effects. This may also involve TMS. According to the OCP, two T1s at adjacent positions are not preferred and may pose difficulties for L2 learners. However, T1-T1 is an often used tone sequence in native Japanese and Korean, which may be partially due to TMS since T1 is a more unmarked tone than T2 and T4 cross-linguistically. The high occurrence rate of T1-T1 in Japanese and Korean speakers' data sets indicates that L1 transfer is stronger than OCP effects in this case.

Following is a discussion of the two cases of unexpectedly low occurrence rates (Japanese speakers' T2-T1, T2-T4 and T4-T3) in the L2 production, where I identify the reasons why their

³⁵ T5-T3 and T2-T3 are separately examined in this study. That is, whenever the target is T5-T3, the correct or true productions will be labeled as T5-T3 but not T2-T3, although these two tone sequences have the same phonetic tone values.

occurrence rates are not as high as expected. Re-examining the hypotheses I made in Chapter 2 regarding the typical Japanese tone sequences on disyllabic words, the three cases (T2-T1, T2-T4, T4-T3) share a feature that the first syllable of the disyllabic words are all contour tones mimicking Japanese disyllabic words whose first syllables are bimoraic. By contrast the other cases have only one mora in the first syllable. The T2-T1 and T2-T4 sequences imitate the tone strings with two moras to the left of an accented syllable (as of cases (41a) and (41b) in section §2.4.3). The T4-T3 sequence imitates the tone sequences with pitch accent at the initial syllable containing two moras (as in case (40b)). In the experiment stimuli (see (19) below), however, in order to keep the syllable structures as short as possible, there are no bimoraic syllables ended with codas at the word-initial positions. That may be the reason why Japanese speakers didn't produce as many T2-T1, T2-T4 and T4-T3 as expected. In the stimuli, all of the first syllables in the test disyllabic words do not have codas (nasal codas) although some of them have diphthongs (15 out of 32, the pinyin are highlighted in bold) and triphthongs (3 out of 32, pinyin are highlighted in bold).

(19) Disyllabic test words in Experiment 1:

1. T1+T1: 书(shu)桌, 飞(wei)机
2. T1+T2: 开(kai)学, 花(hua)茶
3. T1+T3: 出(chu)口, 西(xi)北
4. T1+T4: 开(kai)会, 书(shu)架
5. T2+T1: 学(xue)期, 图(tu)书
6. T2+T2: 留(liu)学, 足(zu)球
7. T2+T3: 牛(niu)奶, 白(bai)水
8. T2+T4: 白(bai)菜, 学(xue)校
9. T3+T1: 手(shou)机, 老(lao)家
10. T3+T2: 打(da)球, 小(xiao)学
11. T3+T3: 语(yu)法, 水(shui)果
12. T3+T4: 五(wu)月, 舞(wu)会
13. T4+T1: 汽(qi)车, 大(da)家
14. T4+T2: 大(da)学, 复(fu)习
15. T4+T3: 跳(tiao)舞, 下(xia)雨
16. T4+T4: 睡(shui)觉, 校(xiao)内

Japanese speakers' tonal performance may be affected by the syllable structures (ie. the number of moras) especially when the syllables are at the beginning of disyllabic words, or at the beginning of

APs since all the test words head APs in the test sentences. Japanese is a mora-timed language. It seems that the mora, the TBU of Japanese, also affects Japanese speakers' tonal performance when they learn a tone language. Another possible reason for the low frequencies of T2-T1 and T2-T4 is the anticipatory effects of T2 as we discussed in the section of §4.3.1. Due to a dissimilation process caused by this anticipatory effect, the error rates of T2 in the two sequences are very high, leading to the low accuracy rate of the tone sequences.

4.5.2. On the most frequently produced tone combinations

In this section, we examine the top 10 most frequently produced tonal combinations of the three groups of speakers, mainly from the perspective of L1 transfer. The 10 most frequently produced tone combinations (including both correct and incorrect productions) in the three data sets are displayed in chart 4.16. In each column, the leftmost is the substitute tone combinations, followed by the frequency counts and the percentages they take out of the total productions in each L1 sub-data set. The tone combinations mimicking typical L1 tone patterns are underlined.

Table 4.17. Top 10 most frequently produced tone combinations

	General			English			Japanese			Korean		
1	T1T3	395	10%	<u>T1T3</u>	135	10.55%	<u>T1T3</u>	148	11.56%	<u>T3T4</u>	149	11.64%
2	T3T4	369	9.6%	<u>T2T3</u>	133	10.39%	T2T3	126	9.84%	<u>T1T1</u>	134	10.47%
3	T1T1	355	9%	<u>T3T4</u>	111	8.67%	<u>T3T1</u>	121	9.45%	<u>T1T4</u>	118	9.2188
4	T2T3	351	9%	<u>T3T1</u>	107	8.36%	T1T1	117	9.14%	T4T3	114	8.91%
5	T3T1	337	8.7%	T1T1	104	8.13%	<u>T3T4</u>	109	8.52%	T1T3	112	8.75%
6	T1T4	302	7.8%	<u>T2T4</u>	98	7.66%	T1T4	97	7.58%	<u>T3T1</u>	109	8.52%
7	T4T3	274	7%	T1T4	87	6.80%	<u>T4T3</u>	96	7.50%	T2T3	92	7.19%
8	T2T4	228	5.9%	T4T3	64	5.00%	<u>T2T4</u>	71	5.55%	T2T4	59	4.61%
9	T5T3	151	3.9%	FT3T3	56	4.38%0	T5T3	65	5.08%	T4T4	55	4.30%
10	T4T4	147	3.8%	T4T4	44	3.44%	<u>T2T1</u>	58	4.53%	T5T3	49	3.83%

By reading the charts, we obtain the following findings.

First, the L1 transfer effects are strong. All the tone sequences imitating those typical L1 tone sequences are all in the top 10 lists as expected. The top three most frequently produced tone

combinations across the English, Japanese and Korean speakers' data sets are all predicted tone sequences in the L1 transfer hypothesis, except Japanese speakers' T2-T3.

Second, the most frequently produced tone sequences for English-speaking learners are T1-T3 and T2-T3, which have very similar frequencies.³⁶ Remember that we hypothesized in chapter 2 that typical L1 tone sequence patterns L+H*, H*+L and H* may influence the L2 acquisition of Mandarin disyllabic tones. We proposed counterpart Mandarin tone sequences T3-T1, T3-T4, T2-T4, T1-T3 and T2-T3 imitating L1 F0 patterns. According to the experiment results, the H*L⁻ affects the L2 tone acquisition stronger than L⁻H*. This shows a L2 production preference for trochees (or a strong-weak accentual pattern) which may be manifested by [+U] [-U] pitch pattern over iambs (or weak-strong accentual pattern) in adult English speakers' acquisition of Mandarin tones. It is very likely due to the "trochaic bias" in native English, and also in the study of L1 acquisition (Delattre 1965; Vihman et al. 1998). It seems the L2 tone acquisition by English speakers may be related to its native foot structures. That is, the foot structure bias at the lexical level leads to the high frequencies of T1-T3 and T2-T3, although some iambic-footed tone sequences, such as T3-T4, T3-T1, ranked the third and fourth below the trochaic-like tone sequences. Actually, the whole carrier phrase containing the test word also shows an L1 transfer of metrical structure. According to Hulst et al (1999, p.426), the accent system for English words can be described as follows: "a quantity-sensitive system with bounded, trochaic feet which are assigned from the right word edge, final extrametrical syllables, and main accent on the head of the rightmost foot...." The test word resides in a noun phrase where the test word precedes *de dongxi* "...(particle) things" as a modifier. Examining the recordings made by English speakers, there are a vast majority of words *dongxi* "things" produced as T4-T0 (High falling tone and a neutral tone), which also sounds like a trochaic foot, and the word-initial syllable "dong" bears an even stronger high falling tone. So, the whole NP phrase under discussion contains two trochaic disyllabic words, the test word and the head noun *dongxi* with the strong high-falling tone at the first syllable of *dongxi*.

³⁶ The results here are different from Zhang, H. (2007, 2010) because this project includes T3 as one of the test tones, whereas Zhang, H. (2007, 2010) only examine tone sequences combined with T1, T2 and T4.

As discussed above, the default accent pattern in English is related to its trochaic foot structure and affects the L2 tonal productions. Trochaic accentual pattern was emphasized in many areas of language acquisition studies, such as studies of speech perception and segmentation in the prelinguistic period, early word production, and patterns of function word omission in early syntax (Vihman et al 1998). There has even been posited a universal trochaic bias. In an OT spirit, this kind of prominent prosodic structure can be seen to embody its constraint ranking within the L1 grammar, and this can be seen as an UG involvement in interlanguage through L1 grammars.

The third finding is that Japanese speakers' most frequent produced tone sequences, especially the first five, are very similar to those on the English speakers' list. Both the first and the second most frequently produced tone sequences made by Japanese speakers are also T1-T3 and T2-T3. This may be due to the L1 transfer of typical Japanese pitch accent pattern HL. The T2-T3 sequence is not predicted to be a typical Japanese pitch pattern on the test words.³⁷ However, it seems the sequence T2-T3 involves the Japanese phrasal level tone patterns. The F0 contour of T2-T3 contains a rising tone at the beginning and a low tone at the end, without a sharp fall after the F0 peak. It is very like a miniature version of the typical Accentual Phrase (AP) tone pattern with an unaccented word. Therefore, it seems the L1 transfer of typical Japanese tone patterns may be involved in both lexical and phrase levels of tone patterns, although the lexical level looks very strong in affecting the L2 tonal production of disyllabic words.

The fourth finding is that Korean speakers' two most frequently produced tone sequences are T3-T4 and T1-T1. This is very likely an L1 transfer effect of typical Korean tone patterns of L-H or H-H. T3-T4 is [-U] [+U] registered, which is same as the L-H pattern. The falling tone at the word-final position may be due to the Korean learners' disfluency in reading L2 materials. That is, because the L2 learners are unable to read the whole test sentence fluently, dephrasing does not occur. Thus, there is a phrase final L% tone following the H tone at the end of the test word and forming a high falling tone. The T1-T1 tone sequence is exactly same as the typical Korean HH pattern over disyllabic words. Whether the Korean

³⁷ May be because of the different TBU in Japanese and Mandarin.

word-level stress pattern or phrase-level tone pattern lead to the most frequently produced tone sequences in L2 productions is not clear since the nature of the prosodic prominence system in Korean has been controversial. Many previous studies argue that Korean has a word-stress system, although the specific proposals did not always agree with each other (Polivanov 1936, Lee 1973, Jun 1994, Davis & Lee 1996, as cited in Ko 1999). The traditional and most influential view of the Korean prosodic prominence system analyzes it as an iambic stress system (Lee 1976). Following this view, the high frequency of T3-T4 is very likely due to the Korean word level stress system, which could be manifested by [-U] [+U] according to De Lacy (2002). On the other hand, the most recent and widely accepted proposal of Accentual Phrase theory for Korean (Jun 1993, 2005) claims that there is no word-level stress in Korean; instead, Korean prosody is best characterized by tonal patterns at the phrasal level. The typical Korean AP pitch patterns are LHLH or HHLH and a disyllabic word pattern LH. Only 6 out of 32 test words begin with aspirated consonants. So it is reasonable that a LH-like tone sequence, T3-T4, is the most frequently produced tone sequence due to the phrasal level transfer of L1 pitch patterns.

As for Korean speakers' high frequency of T1-T1 as discussed in the previous sub-section, it is a case of L1 transfer conquering phonological universals. According to L1 transfer, H-H is a very frequent pitch pattern in Korean; thus the L1 transfer predicts a high frequency of T1-T1 in L2 tonal productions. However, according to the phonological universal OCP, we predict OCP to constrain L2 tone grammars, and there not to be many occurrences of T1-T1 because identical tone sequences are not favored. The results show that T1-T1 takes 10.47% of the total Korean speakers' productions, significantly higher than expected. It signifies the L1 transfer effect is stronger than OCP effects in this case. In the spirit of OT, the OCP (H) constraint is ranked very low in Korean grammar, and there is no need to re-rank this constraint during SLA of Mandarin tones since this constraint is also very low in the target language.

4.6. Summary of Chapter 4

Chapter 4 reports the word-level experiment 1 results, including the tests of phonological universals TMS, OCP and TPC and the test of typical L1 pitch pattern transfer in sections §4.1 through §4.5. It is found that phonological universals, L1 transfers, and T3 teaching problems jointly affect the L2 tonal productions of disyllabic words. Both similarities (mostly due to phonological universals and pedagogy problem) and differences (mostly due to L1 transfer) in error and substitution patterns are found in the English, Japanese and Korean speakers' data sets, which indicates that the non-tonal language learners share similarities and differences in developing L2 acquisition paths. It is discovered that TMS, OCP and TPC are all relevant in most parts of the L2 production data set, although they interact with other factors and sometimes don't demonstrate strong effects.

There are a lot of similarities found in the L2 tonal productions. The TMS test demonstrates that all learners have the highest error rates with the rising tone (T2) and higher error rates of the high falling tone (T4) than the high level tone (T1). The error patterns are negatively correlated with the substitution patterns, which confirm the relevance of TMS in the research data. The two OCP tests investigate the OCP effects in the error patterns and the L2 tonal productions. Some OCP effects are detected and found to interact with other factors, such as TMS and L1 transfer. It is found that, generally, L2 learners favor non-identical tone combinations over identical tone combinations, especially in the substitutions. However, in the L2 tonal productions across three types of learners, T1-T1 productions are found significantly more than T4-T4, and more T4-T4 productions than T2-T2. In addition, Korean speakers have more identical tone combination productions than Japanese speakers, and in turn than ENG speakers. Since all native L1s allow H-H intonation (and H-H is a typical Korean tone pattern), the high T1-T1 is taken as a case of stronger L1 transfer effect than OCP effect. TPC is tested, and very systematic error and substitution patterns in the L2 tonal productions suggest that word-final syllables are preferred bearers of T4, and T2 is performed well at word-initial syllables by L2 learners. Some different error and substitution patterns are also found during the tests. For example, the foot and stress structure affect English speakers' acquisition of Mandarin lexical tones. When they make mistakes, English speakers tend

to employ more contour tones as substitutions than other speakers. Japanese lexical pitch accent pattern “HL” and the typical AP pitch pattern affect Japanese speakers’ interlanguage and the substitution pattern reflect this preference. Korean speakers are better at producing tone sequences with high registered tone such as T1 and T4 at the word-finals because they are similar to Korean pitch pattern.

The acquisition problem of T3 is identified in the study based on the conflicting error patterns and substitution patterns related to T3. T3 is the most frequent substitute tone when errors occur, and it implies that the low T3 is phonetically easy. However T3, as a target tone, has a very high error rate and the most frequent substitute tone for target T3 is FT3. It signifies an over-use of FT3 by L2 learners because of the problematic assumption of standard form of Tone 3. The assumption of FT3 as the standard form and the related “FT3 first” teaching method is argued to influence the acquisition of Tone 3 and the Sandhi, especially the Pre-Other-Tone Sandhi.

Section §4.5 looked into the L1 transfer of typical tone patterns on disyllabic words. The hypothesis is supported statistically by comparing the actual proportions of English speakers’ T1-T3, T2-T3, T3-T1, T2-T4, T3-T4, Japanese speakers’ T1-T1, T1-T3, T3-T1 and T3-T4, Korean speakers’ T1-T1, T1-T4, T3-T1 and T3-T4 in the L2 productions and the expected proportions. The proposed L1-like tone sequences in the L2 tone productions are all within the lists of top 10 most frequently produced tone sequences in each L1 data set, and this reflects strong L1 transfer effects. The low frequencies of Japanese speakers’ T2-T1, T2-T4 and T4-T3 may be due to the mismatch in syllable structures (mora numbers) of the stimuli words with the typical L1 tone sequences in the hypothesis. It seems English and Japanese speakers’ most frequently produced [+U] [-U] registered tone sequences are related to the dominant trochaic foot structure in native English and HL pitch accent pattern in Japanese. Japanese and Korean speakers’ most frequently produced tone sequences suggest that not only word-level but also phrasal level F0 patterns affect L2 disyllabic tonal productions.

CHAPTER 5: SENTENCE-LEVEL EXPERIMENT RESULTS

While Experiment 1 focuses on the L2 productions of disyllabic words, Experiment 2 is concerned more with above-word level L2 tonal productions made by the three L1 groups. The purpose of Experiment 2 is to confirm the results of Experiment 1 and extend the Experiment 1 findings, especially on the positional effects, by observing related phenomena in sentence productions. To confirm the TMS and OCP effects, some simple statistical tests are used, and the analysis focuses on the comparison between the word-level and the above-word-level results in section §5.1. For the positional effects, section §5.2 looks into the vulnerable positions with high error rates at the phrase and sentence level and provides a general view of the positional effects for pedagogical purposes. Section §5.3 offers analysis of narrow focus expressions (monosyllabic focus and long focus) and the transfer of L1 pitch patterns. By examining the substitution patterns for longer-focused constituents, it is argued in this section that the L2 tonal productions of sentences are affected by the binary foot constraints. The investigation of substitution patterns in prosodic words and the examination of positional effect at phrase and sentence levels sketch out a big picture of sentence-level tone grammars which are built up and framed by the lower level prosodic structures. The chapter ends with a summary in §5.4.

As mentioned in Chapter 3, the test materials of Experiment 2 are 20 sentences composed of identical tone sequences. Therefore, there are five sentences for each tone type (T1, T2, T3, T4) and they all have very similar syntactic structures. The five sentences of specific tone type vary only in the length of focused constituents, with one, two, three, four or five syllables in the focused constituents F1, F2, F3, F4 and F5. The focused constituents that are the same size across tone types share the same syntactic structures, but differ according to length across the constituents. Although the syntactic structures of pre-focused materials vary a little bit across the tone types (see appendix 5), the 20 sentences share very

similar syntactic structures overall. (1) displays example T1 sentences containing focused constituents in varying size, and Figure 5.1 presents the internal syntactic structure of the focused constituents.

(1) Example test sentences in Experiment 2:

a. T1-F1:

Chinese characters: 他听说吃 鸡 的医生搬书桌。
 Pinyin: tā tīngshuō chī jī de yīshēng bān shūzhuō.
 English: “He heard that the doctors who eat CHICKEN will move desks.”

b. T1-F2:

Chinese characters: 他听说喝 鸡汤 的医生搬书桌。
 Pinyin: tā tīngshuō hē jī tāng de yīshēng bān shūzhuō.
 English: “He heard that the doctors who drink CHICKEN SOUP will move desks.”

c. T1-F3:

Chinese characters: 他听说 喝鸡汤 的医生搬书桌。
 Pinyin: tā tīngshuō hē jī tāng de yīshēng bān shūzhuō.
 English: “He heard that the doctors who DRINK CHICKEN SOUP will move desks.”

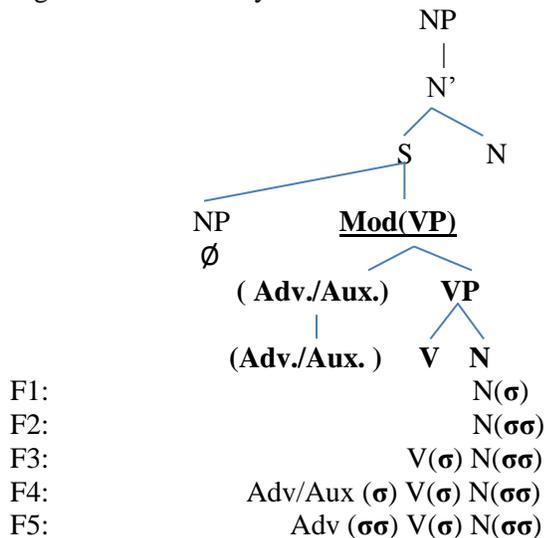
d. T1-F4:

Chinese characters: 他听说 多喝鸡汤 的医生搬书桌。
 Pinyin: tā tīngshuō duō hē jī tāng de yīshēng bān shūzhuō.
 English: “He heard that the doctors who DRINK MORE CHICKEN SOUP will move desks.”

e. T1-F5:

Chinese characters: 他听说 今天喝鸡汤 的医生搬书桌。
 Pinyin: tā tīngshuō jīntiān hē jī tāng de yīshēng bān shūzhuō.
 English: “He heard that the doctors who DRINK CHICKEN SOUP TODAY will move desks.”

Figure 5.1. Internal syntactic structure of focused constituents:



We predict that the general error rates in Experiment 2 may be higher than those in Experiment 1 because the running speech may pose more difficulties for the learners. It is hypothesized that interacting effects of phonological universals (such as TMS, OCP and TPC) and L1 transfer of similar pitch patterns could also be found in the above-word level tonal productions. The “head prominent” or “deaccenting” language (English) and “edge prominence” or “dephrasing language” (Japanese and Korean) language speakers may realize sentence level prominence in different ways, and these differences may be manifested in the error and substitution patterns in the L2 tonal productions. For example, the learners would tend to change the pitch contour in expressing narrow focus. English speakers may use L1-like pitch accent, usually high tones, in short narrow focus and place the strongest prominent at the rightmost word in long focused constituents. Japanese and Korean speakers may transfer their specific Accentual Phrase (AP) intonation patterns to focused constituents. We also predict that the disyllabic level pitch patterns found in Experiment 1 would also be reflected in the sentence level tone patterns.

5.1. General error patterns and the TMS, OCP effects

This section reports the general error patterns of Experiment 2 in §5.1.1, the tests of TMS effects in §5.1.2 and OCP effects in §5.1.3. The tests on TMS and OCP effects confirm the findings in Experiment 1. The results of Experiment 2 are briefly reported and discussions are provided only when there are significant differences between the results of the two experiments. Section §5.1.4 summarizes the findings in §5.1.

5.1.1. General error patterns

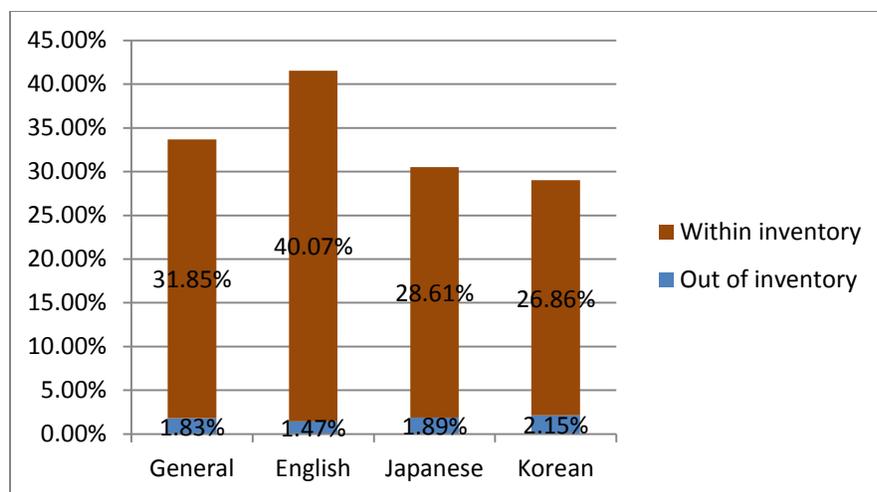
It is predicted that the same groups of Mandarin learners may make more tonal mistakes in running speech than in the tonal production of disyllabic words.³⁸ However, that’s not true according to the general error rates of Experiment 2. The general error rate of Experiment 2 is very similar and even a little bit lower than that of Experiment 1. 13,680 test syllables were examined where 4,560 syllables were

³⁸ Guo (1993) has reported that the accuracy of tone production decreases as the number of syllables in a word increases.

produced by each group of L1 speakers in Experiment 2. The general error rate of Experiment 2 (34%) is a little bit lower than Experiment 1 (37.11%). Same as Experiment 1, the English speakers' error rate (41.54%) is the highest among the three L1 groups. The Japanese and Korean speakers' general error rates are very similar. It is compatible to the finding of general tonal proficiency level in Experiment 1 that English speakers possess the lowest Mandarin tonal proficiency level. The general error rates of English and Japanese speakers in Experiment 1 and 2 are very close, but the Korean speakers' error rate of Experiment 2 (29.01%) is much lower than that of Experiment 1 (36.37%). Therefore, in Experiment 2, the Korean speaker's general error rate is very close to the Japanese speakers' (30.5%). Examining the Korean speakers' individual tone performance, the error rates of T1, T2, and T4 decreased about 5-10% from Experiment 1, while T3 and T5 show similar error rates to Experiment 1. It seems Korean speakers are better than expected at producing identical tone sequences, especially those of T1, T2 and T4. However, Korean speakers' application of T3 sandhi in sentences is not better than in disyllabic words. Another minor difference is that the English speakers' T3 error rate in Experiment 2 is much lower than in Experiment 1, while others tones preserve very similar error rates.

Figure 5.2 displays the overall error rate and those within each L1 group. Within-inventory error rates and out-of-inventory error rates are labeled. Within the errors, most of the errors are tones within the Mandarin lexical tone inventory. The out-of-inventory error rates are at the bottom of the columns. Similar to Experiment 1, the out-of-inventory errors are very few, and we assume they do not influence the results in the Experiment 2. Therefore, most of the analysis and discussions in the following sections focuses on the within-inventory errors.

Figure 5.2. General error rates in Experiment 2



The generally better tonal performance of running speech in Experiment 2 than Experiment 1 may be due to the simpler tone combinations in the test materials than those in Experiment 1. It is found in the recordings that about 30% of speakers produced very unnatural and robot-like tonal production for Experiment 2. It's very possible that the speakers learned only one tone type used in a sentence so they simply repeated the full tones of T1, T2 and T4 to successfully produce the whole sentence. That is, when speakers start speaking the first several tones in the sentences, they can predict the tones of the following materials. However, they can only take advantage of this with T1, T2 and T4 sequences without sandhi, but not with T3 sequences where tone sandhi occurs. Korean speakers took advantage of the simple tone combinations in Experiment 2 and improved their performance of T1, T2 and T4 sequences, but not T3 sequences due to the sandhi rules. We will discuss the T3 issue further in the next section by examining the substitution patterns.

Overall, the general error rates and the ratios of within-inventory and out-of-inventory errors of Experiment 2 confirm the results of Experiment 1. In the next section, details of individual error rates by different speakers are examined, and the TMS effect is confirmed.

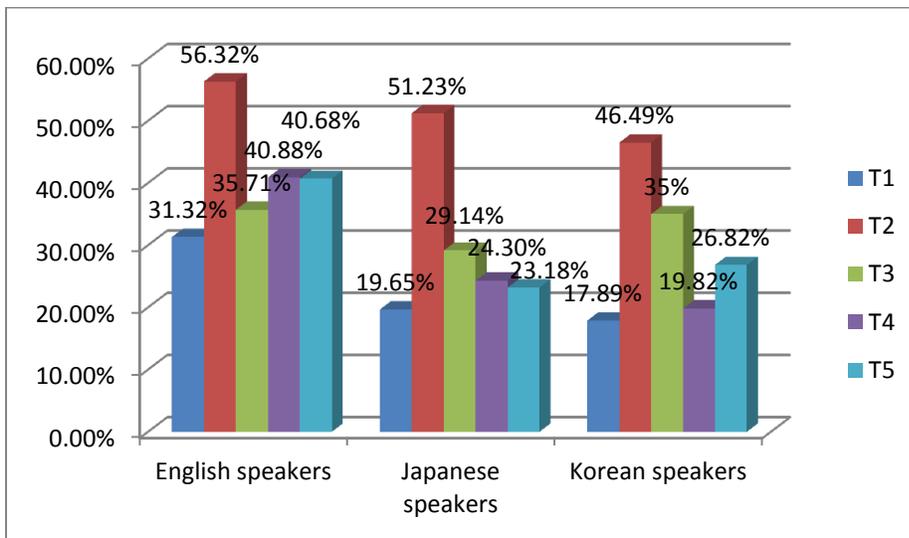
5.1.2. Individual tone errors and Tonal Markedness Scale

This section presents the error rates of individual tones, the statistical test of TMS in Experiment 2 and compares the error rankings of individual tones and substitution patterns in the two experiments. It

is found that, although the ranking of error rates between T3 and T4 in Experiment 2 is a little different from Experiment 1 across the three sub-datasets, the TMS is relevant in Experiment 2.

Same as Experiment 1, the rankings of error rates for T2, T4 and T1 within each L1 data set are the same. All L1 groups of speakers have the highest error rate with T2, then T4, and the lowest with T1, which indicates strong effects of TMS. The only two differences from Experiment 1 are (1) the rank of T3 in English speakers' data set, where the error rate of T3 is lower than T4 and T5; and (2) the rank of T4 in the Korean speakers' data set, where the error rate of T4 is lower than T5. Figure 5.3 displays the individual tone error rates and (2) lists the error rankings.

Figure 5.3. Error rates of individual tones in Experiment 2 (TMS test)



(2) Error rankings for TMS test in Experiment 2

- a. English speakers: **T2**(56.32%) > **T4**(40.88%), T5(40.68%) > T3 (35.71%) > **T1**(31.32%)
- b. Japanese speakers: **T2** (51.23%) > T3 (29.14%) > **T4** (24.3%) > T5(23.18%) > **T1** (19.65%)
- c. Korean speakers: **T2**(46.49%) > T3(35%) > T5(26.82%) > **T4**(19.82%) > **T1**(17.89%)

Tests of no association between tone and response using the Rao-Scott Chi-Square test were conducted to account for multiple observations within subjects. It is found that the error rate of T2 is significantly higher than those of T4 across the L1 groups, and the error rate of T4 is significantly higher than T1 in the English speakers' data set. The error rate of T4 is higher than T1 in Japanese and Korean speakers' data sets, but there were no significant differences. The statistical test results of Experiment 2 are very similar to that of Experiment 1 except for Japanese speakers' T1 vs. T4. The table 5.1 displays the test results (P values).

Table 5.1. TMS test (*T2>>*T4>>*T1) results in Experiment 2

Speakers	T4 vs. T1	T2 vs. T4
English	0.0174	0.0007
Japanese	0.2657 (n. s.)	< .0001
Korean	0.5834 (n. s.)	< .0001

The substitution rates of individual tones are also worked out to confirm the TMS effects determined by error rates. Comparing T1, T2 and T4 as substitute tones, T1 is the most frequently substituted for target tones, T4 the second most frequently used and T2 is the least frequently substituted for target tones when errors occur. The individual tone substitution pattern in Experiment 2 also confirms the TMS effects as in Experiment 1. Table 5.2 shows the detailed substitution patterns for each L1 group with the rankings of substitute tone frequencies under each table.

Table 5.2. English speakers' substitutions in Experiment 2

Target tone	T1	T2	T3	T5	T4
Response tones	T3 (173, 48%) T4(105, 29%) T2(41, 11%) FT3(25, 7%) MT1(10, 3%) [543](2, 1%) LT4(1, 0%)	T3(295, 46%) T1 (133, 21%) T4 (114, 18%) FT3(75, 12%) MT1(19, 3%) [543](5, 1%) LT4(1, 0%)	FT3(85, 34%) T4(65, 26%) T1(49, 20%) T2(44, 18%) MT1(4, 2%) [543](2, 1%)	T3 (58, 33%) T1(56, 32%) T4(35, 20%) FT3(25, 14%) MT1(2, 1%)	T3(200, 43%) T1(144, 31%) T2(73, 16%) FT3(30, 6%) MT1(13, 3%) [543](5, 1%)

Substitution ranking: T3(726) > **T1**(382) > **T4**(319) > FT3(240) > **T2**(158) > MT1(48) > [543](14)

Table 5.3. Japanese speakers' substitutions in Experiment 2

Target tone	T1	T2	T3	T5	T4
Response tones	T3 (92, 43%) T4(63, 29%) T2(31, 14%) FT3(14, 7%) MT1(12, 6%) [543](2, 1%)	T3(322, 55%) T1 (130, 22%) FT3(65, 11%) T4 (40, 7%) MT1(24, 4%) [543](1, 0%)	T2(72, 35%) FT3(58, 28%) T1(34, 17%) T4(28, 14%) MT1(11, 5%) LT4(2, 1%)	T1(52, 57%) T3 (24, 26%) T4(14, 15%) MT1(1, 1%) [543](1, 1%)	T3(145, 69%) T1(75, 36%) T2(20, 10%) FT3(7, 3%) MT1(15, 7%) LT4 (8, 4%) [543](7, 3%)

Substitution ranking: T3(583) > **T1**(291) > **T4**(145) > FT3(144) > **T2**(123) > MT1(63) > [543](11)

Table 5.4. Korean speakers' substitutions in Experiment 2

Target tone	T1	T2	T3	T5	T4
Response tones	T4(114, 56%) T3 (52, 25%) MT1(15, 7%) FT3(13, 6%) [543](7, 3%) T2(3, 1%)	T3(266, 50%) T1 (115, 22%) FT3(78, 15%) MT1(34, 6%) T4 (33, 6%) LT4(2, 0%) [543](1, 0%) 352(1, 0%)	T2(75, 31%) FT3(73, 30%) T1(60, 24%) T4(27, 11%) MT1(10, 4%)	T1(40, 34%) T3 (36, 31%) FT3(25, 21%) T4 (15, 13%) MT1(1, 1%)	T1(129, 60%) T3(53, 25%) [543](19, 9%) FT3(6, 3%) MT1(6, 3%) LT4 (2, 1%) T2(1, 0%)

Substitution ranking: T3(407) > **T1**(227) > FT3(195) > **T4**(189) > **T2**(79) > MT1(66) > [543](27) > LT4(4) > 352(1)

When we examine the detailed substitutions, some patterns are found. Firstly, both the general and sub-substitution patterns in each L1 speakers' data set, even the frequency rankings for specific tones, are very similar to those in the Experiment 1. For example, in both Experiments 1 and 2, English and Japanese speakers substitute T1 most frequently with T3, whereas Korean speakers substitute with T4; all speakers substitute T2 most frequently with T3, then T1; the out-of-inventory error [543] is most often used by Korean speakers, etc. There are only a few minor differences regarding the substitute tone ranks in Japanese and Korean speakers' data set. The two differences are discussed below.

Secondly, both Japanese and Korean speakers substitute target T3 most frequently with T2 in Experiment 2 instead of FT3 and T1, as in the Experiment 1, although the substitution rates of FT3 are also high. It may be due to the momentum of repeating the same tone in the sentence in this experiment. Because of the occurrence of the Pre-T3 Sandhi, the T3 sequence sentences all begin with T5 (a phonetic

T2), and the speakers just repeat T2 in reading sentences. Another possible reason is that the Japanese and Korean speakers have difficulty in applying the Pre-T3 Sandhi in the correct domains. The identification of T3 Sandhi application domains is more difficult in sentences (as in Experiment 2) than in disyllabic words (as in Experiment 1) where the words provide a simple domain of the sandhi process (see chapter 2 for more information of sandhi domains). In addition, the application of the Pre-T3 Sandhi is more frequent in Experiment 2 than in the Experiment 1. According to the error rates of individual syllables in sentences (see §5.2), Japanese and Korean speakers may have had difficulty in grouping the preposition words and verbs in to three-syllable domains of tone Sandhi and mis-produced T5 or T2 in T3 sequences.

Remember that English speakers were reported in section §5.1.1 to have a lower error rate of T3 in Experiment 2. The most frequently substituted tone for target T3 is still FT3, as in Experiment 1. However, this does not mean English speakers process T3 sandhi in Experiment 2 better than in Experiment 1, or, better than other speakers. The English speakers' T5 error rate is high in Experiment 2. Although the English speakers' T3 error rate is lower than in Experiment 1, it still is the highest among the three groups of L1 speakers. The most frequent error for target T3 in English speakers' data set is still FT3, like in Experiment 1.

The other difference is that Japanese speakers substitute T4 most frequently with T3 instead of T1 like in Experiment 1. While Experiment 1 contains various tone combinations, Experiment 2 has only one type of tone combination of T4 which is T4-T4. In experiment 1, although T3 is not the most frequently used substitute tone for T4 in general, it is most often substituted for target T4 in those T4-T4 sequences. Therefore it's very possible that the higher substitute rate of T3 for target T4 than other substitute tones may be due to the high frequency of T4-T4 sequences in the Experiment 2 stimuli.

To summarize, the TMS effects are relevant in most tonal productions in Experiment 2. The TMS ranking *T2>>*T4>>*T1 is statistically supported by the sentence level data, except for Korean and

Japanese speakers' *T4>>*T1, although the error rates are compatible with the TMS rankings. The Experiment 2 results confirm the TMS findings in Experiment 1, based on error and substitution patterns.

5.1.3. OCP test in Experiment 2

This section reports the OCP test results based on the Experiment 2 data. The detailed results of an OCP test examining the substitution patterns in Experiment 2 tonal production is presented and discussed in this section.

The goal of the test is to find out if the counts of ITC (Identical Tone Combination) match the expected counts calculated based on the individual tone error numbers in each sub-data set. To detect the OCP effects in the sentences composed of identical tone sequences, the erroneous tones or the substitutions in Experiment 2 are closely examined. Since T3 Sandhi operates in the T3 sequences, the OCP test is conducted only with target tone types T1, T2 and T4. The test consists of four steps. We take target T1 sequence as an example explaining the test procedures. First, we collect all two-tone sequences in which both tones are erroneous. Second, we count the frequencies of each of the non-T1 tones. Third, we use these numbers to predict how many of the two-tone sequences should be identical tone combinations (ITC), i.e., T2-T2, T3-T3, T4-T4, FT3-FT3. Fourth, we compare the observed number of T2-T2, T3-T3, T4-T4, and FT3-FT3 with the predicted numbers of ITCs. If the OCP is working in the data set, the ratio of observed ITC and expected ITC (i.e., O/E values) should be smaller than 1. It is found in the test results that most of O/E values of individual ITCs are smaller than 1 in Experiment 2. Specifically, English speakers have eight out of twelve O/E values smaller than 1, one O/E value is equal to 1 (FT3-FT3 for target T4), and three O/E values are bigger than 1 (T4-T4 for target T1 sequences, FT3-FT3 for target T2 sequences and T1-T1 for target T4 sequences). All O/E values in the Japanese speakers' data are smaller than 1. In the Korean speakers' data set, all O/E values are also smaller than 1, except for two equal to 1 (FT3-FT3 for target T2 sequences and T1-T1 for target T4 sequences). The OCP test results in Experiment 2 suggest that the OCP is relevant in the data sets with the strongest OCP effects in Japanese speakers' data and weakest OCP effects in English speakers' data. Below are the detailed test

results. In each table, the first columns are the target tones and the total error count of that tone. The second column displays the ITC types found in the errors for that target tone. The third and fourth columns contain the observed ITC percentages and counts, while the fifth and sixth columns contain the expected ITC percentages and counts. The rightmost column displays the O/E values for each ITC.

Table 5.5. OCP test results for English speakers in Experiment 2

Target Tone	Erroneous ITC	Obs %	Obs count	Exp %	Exp count	O/E values
T1 Total error: 126	T2-T2	0.02381	3	0.03803	4.7912	0.63
	T3-T3	0.15873	20	0.2025	25.5150	0.78
	FT3FT3	0	0	0.0049	0.6174	0
	T4-T4	0.1111	14	0.08123	10.2344	1.37 NO
T2 Total error: 294	T1-T1	0.04082	12	0.06816	20.0387	0.60
	T3-T3	0.15986	47	0.20450	60.1224	0.78
	FT3FT3	0.02041	6	0.00628	1.8467	3.25 NO
	T4-T4	0.03741	11	0.04304	12.6536	0.87
T4 Total error: 199	T1-T1	0.14573	29	0.11947	23.7736	1.22 NO
	T2-T2	0.02010	4	0.03531	7.0274	0.57
	T3-T3	0.09548	19	0.15679	31.2022	0.61
	FT3FT3	0.00503	1	0.00497	0.9882	1.01

Table 5.6. OCP test results for Japanese speakers in Experiment 2

Target Tone	Erroneous ITC	Obs %	Obs count	Exp %	Exp count	O/E values
T1 Total error:60	T2-T2	0	0	0.03341	2.0049	0
	T3-T3	0.13333	8	0.20395	12.2373	0.65
	FT3FT3	0	0	0.00416	0.2497	0
	T4-T4	0.08333	5	0.09065	5.4388	0.92
T2 Total error: 234	T1-T1	0.02991	7	0.11047	25.8499	0.27
	T3-T3	0.18803	44	0.27064	63.3299	0.69
	FT3FT3	0.00427	1	0.00752	1.7592	0.57
	T4-T4	0	0	0.00368	0.8620	0
T4 Total error: 114	T1-T1	0.04386	5	0.14830	16.9058	0.3
	T2-T2	0	0	0.01115	1.2710	0
	T3-T3	0.15789	18	0.22873	26.0756	0.69
	FT3FT3	0	0	0.00096	0.1099	0

Table 5.7. OCP test results for Korean speakers in Experiment 2

Target Tone	Erroneous ITC	Obs %	Obs count	Exp %	Exp count	O/E values
T1 Total error: 42	T2-T2	0	0	0.00022	0.0091	0
	T3-T3	0.02381	1	0.07807	3.2790	0.3
	FT3FT3	0	0	0.01060	0.4451	0
	T4-T4	15.2686	13	0.36354	15.2686	0.85
T2	T1-T1	0.0797	11	0.10147	15.7283	0.7

Total error: 155	T3-T3	0.14194	22	0.19674	30.4940	0.7
	FT3FT3	0.01290	2	0.01275	1.9758	1.01
	T4-T4	0.00645	1	0.01563	2.4219	0.41
T4 Total error: 41	T1-T1	0.48780	20	0.46814	19.1939	1.04
	T2-T2	0	0	0.00017	0.0071	0
	T3-T3	0.02439	1	0.06925	2.8393	0.35
	FT3FT3	0	0	0.00156	0.0639	0

According to the results, the OCP effects are found in most of the research data in Experiment 2.

The anti-OCP cases include the observed high frequencies of English speakers' T1-T1 and T4-T4 sequences and Korean speakers' T1-T1 sequence. This result is compatible with the finding in Experiment 1, in which we found many more T1-T1 than T4-T4 and more T4-T4 than T2-T2. This reflects some interacting effects of OCP and TMS (see more discussions in Chapter 6). When speakers make tonal errors, the possibility of making ITC of relatively unmarked tones such as T1 and T4 is higher than the relatively marked T2.

5.1.4. Summary of 5.1

This section reports the general error pattern, the result of TMS test and OCP test in Experiment 2. Most of the results are the same as, similar to, or compatible with those of the correspondent results in Experiment 1. For example, the English speakers have the highest general error rate among the three groups; TMS and OCP are relevant in most L2 tonal productions; and the substitution patterns for individual tones are very similar to those in Experiment 1. There are minor differences between the two experiments, such as the following: the general error rate of Korean is lower, very close to the Japanese speakers' error rate in Experiment 2; the English speakers' T3 error rate is lower in Experiment 2 than in Experiment 1; for target tone 3, the Japanese and Korean speakers substitute T2 most frequently instead of FT3 and T1. Possible reasons explaining the differences are provided in this section. It is very likely that the L2 learners take advantage of the design of the identical tone sequence in Experiment 2, mechanically repeat the full target tones and make robot-like tonal productions. This section also argues that the lower T3 error rate of English does not necessary mean English speakers process T3 sandhi better

in Experiment 2. The OCP test results in Experiment 2 indicate the strongest OCP effects are in the Japanese speakers' data, and the weakest OCP effects are in the English speakers' data, although OCP is still relevant in most of the English speakers' data.

5.2. On the most vulnerable positions at sentence level

This section looks into the distribution of errors to extend the findings of TPC effects in the previous chapter. It emphasizes the similarities of sentence-level positional effects across the three sub-data sets rather than the differences. By examining the positions bearing syllables with the highest error rates, this section investigates the positional effects at sentence level tonal productions, especially those in neutral (non-focused) constituents. Sub-section §5.2.1 provides a general picture of the syllables with the highest error rates, including both focused and neutral syllables. Section §5.2.2 compares the error rates of syllables inside and outside of focused constituents. The syllables with the highest error rates in neutral constituents are analyzed in section §5.2.2, and the most vulnerable positions at sentence level are identified, mainly in order to provide pedagogical implications.

5.2.1. An overview of vulnerable positions in sentence level tonal productions

Different from Experiment 1, in which the test words in carrier sentences bear contrastive focus (see chapter 4), Experiment 2 designed running speeches, or sentences, containing strong, narrow focused constituents. The annotations of the individual positions bearing syllables are provided in (3) below. In Experiment 2, five sentences containing focused constituents of different lengths were designed for each target tone type. The sentence with a monosyllabic focused constituent is called F1, with a disyllabic focused constituent F2, and so forth. All the syllables located in the focused constituents are labeled with lower case "f," and a number indicating the location in the focused constituent. Other syllables located in neutral constituents or non-focused constituents are mostly labeled with "n," followed by a number indicating the locations in the sentence, and a few with "v," meaning verbs immediately preceding focused constituents. The items labeled with "n" and specific numbers refer to the same morphemes

As mentioned in Chapter 2, the prosodic phrasing of the test sentences depends on certain factors, such as the status of focused constituents, the sentence length, L2 learners' fluency of reading new L2 materials, and different types of L1 speakers' L1 transfer of prosodic phrasing methods (see the predicted prosodic phrasing in Chapter 2). Basically, long sentences and heavy VP and NP structures may cause more phrase boundaries (Jun 2003), and focus may induce the presence of phrase edges around focus, overriding syntactic phrase (Selkirk 2011). According to previous studies on prosodic phrasing involving narrow focus, phrase boundaries should be placed at the left edge of narrow focused constituents in Mandarin, Japanese and Korean speakers' phrasings (Shih 1997; Chen 2004; Jun 1996; Jun and Lee 1998; Nagahara 1994) and at the right edge of focused constituents for the English speakers (Selkirk 2002). However, since the left edge of the narrow focus coincides with the major syntactic boundary, there is also a phrase boundary at the left edge in English speakers' phrasing. As discussed in Chapter 2, phrase boundaries after the focused constituents in Japanese and Korean speakers' phrasing may not be removed because of the L2 learners' disfluency in reading new Mandarin materials. Considering all of the above factors, the prosodic phrasings across the test sentences and L2 speakers are generally similar. In discussing vulnerable positions (with high error rates of production tones) below, we will refer to the correspondent locations in prosodic units (such as prosodic word-initial, phrase-final) according to the prosodic phrasing information presented here. (4) is an example phrasing for F5 sentences representing long focused constituents.

(4). Example prosodic phrasings for F5 sentences (T1) in Experiment 2

Pinyin: Ta TingShuo JIN.TIAN. HE.JI.TANG. de YiSheng Ban ShuZhuo.
a. Mandarin speaker:
 $\uparrow \{ \varphi [(\omega)] \} \uparrow \{ \varphi [(\omega) (\omega)] \} \varphi [(\omega)] \varphi [(\omega)] \}$
b. Japanese and Korean speakers
 $\uparrow \{ \varphi [(\omega)] \} \uparrow \{ \varphi [(\omega) (\omega)] \} \uparrow \{ \varphi [(\omega)] \} \varphi [(\omega)] \}$
c. English speakers:
 $\uparrow \{ \varphi [(\omega)] \} \uparrow \{ \varphi [(\omega) (\omega)] \} \uparrow \{ \varphi [(\omega)] \} \varphi [(\omega)] \}$

The Table 5.8 below displays the 10 syllables with the highest error rates in each L1 data set. The numbers following syllable codes in each column are the error rates. For example, the English speakers' item "T3-F1-S5-f1" has an error rate of 1.00 (or, 100%) indicating none of the 20 English speakers successfully produce the target T3, which is a monosyllabic focused constituent; Japanese speakers' item "T2-F1-S2-n2" has an error of 0.9 (or, 90%) indicating that 90% of the tonal productions for the target T2 located at the second syllable of the F1 sentence made by Japanese speakers are incorrect. Those syllables inside focused constituents (labeled with "f") are bold.

Table 5.8. The top ten positions with highest error rates

English speakers			Japanese speakers			Korean speakers		
1	T3-F1-S5- f1	1.00	1	T2-F1-S2-n2	0.90	1	T2-F2-S6- f2	0.95
2	T5-F1-S4-v	1.00	2	T2-F2-S2-n2	0.90	2	T2-F5-S8- f5	0.95
3	T2-F1-S10-n8	0.85	3	T2-F3-S2-n2	0.90	3	T2-F3-S6- f3	0.90
4	T2-F2-S4-v	0.85	4	T2-F5-S10-n5	0.90	4	T2-F4-S2-n2	0.90
5	T2-F2-S8-n5	0.85	5	T2-F2-S6- f2	0.85	5	T2-F4-S9-n5	0.90
6	T2-F4-S2-n2	0.85	6	T2-F2-S8-n5	0.85	6	T2-F5-S2-n2	0.90
7	T2-F5-S10-n5	0.85	7	T2-F4-S2-n2	0.85	7	T2-F2-S8-n5	0.85
8	T3-F3-S6- f3	0.85	8	T2-F4-S9-n5	0.85	8	T2-F4-S12-n8	0.85
9	T2-F1-S2-n2	0.80	9	T2-F1-S7-n5	0.80	9	T2-F4-S7- f4	0.85
10	T2-F1-S4-v	0.80	10	T2-F3-S8-n5	0.80	10	T2-F5-S13-n8	0.85

In the English speakers' list of high error rates, most of the items are T2, and two of them are T3, where one of them has the error rate as high as 100%. Three kinds of positions in English speakers' productions tend to have the highest error rates. First is the T3 in focus (T3 monosyllabic focus and focus-final T3 in F3). It indicates that English speakers never use T3 to express monosyllabic focus, but substituted with T4 or FT3. It can be confirmed by the substitutions for monosyllabic focused constituents discussed in the section §5.2.3.1 below. The second type of position is the syllable or verb right before the focused constituents, such as "T5-F1-v," "T2-F2-v" and "T2-F1-v." It is very possible that the speakers made tonal errors when they tended to make the non-focused items be in contrast with the focused constituent. By contrast, the verbs within the focused constituents and after the focused constituents do not have such high error rates. For example, "F3-f3" is also a verb, but the error rates are not high. The third type of position is disyllabic word-final or phrase-final positions. All the other syllables within the

English speakers' list are word-final syllables, such as “T2-F1-n8,” “T2-F2-n5,” “T2-F4-n2,” “T2-F5-n5,” “T3-F3-f3.”³⁹ However, some of them are also ip-final items, such as the ones ended with “n5,” “f3,” or sentence-final items, such as the item with “n8.”

The Japanese speakers' top 10 syllables with the highest error rates are all in T2 sequences, and all the locations share the same feature of “word-final.” Of course, most of them, or those items ended with “n2” and “n5,” are also located at the AP-final positions. Only one of them is located in the focused constituent which is also at word-final position.

The Korean speakers' top 10 syllables with the highest error rates are also in T2 sequences. It is very similar to the other L1 speakers in that all of them are at word-final position. Some of them are also located at the phrase-final (either AP- or ip-final, depending on the length of focused constituents) positions, such as the items at the end of focused constituents labeled with “f” and those with “n5,” and at utterance-final positions such as the items ending with “n8.”

In summary, the highest error rates mostly take place with target T2 items at prosodic word-final positions, although some of them are also at phrase- or utterance-final positions. It seems the *T2-Final positional effect is also very strong in sentence-level tonal productions. All the focus-marked syllables with the highest error rates are located at the end of focused constituents.

5.2.2. Error rates of syllables inside and outside of focused constituents

Before we proceed to the next section examining the vulnerable positions in neutral constituents, it is necessary to figure out whether being focused has any impact on L2 learners' error rates. By examining the syllables with high error rates, especially those located in longer focused constituents and comparing the average error rates of focused syllables and those of neutral syllables below, it seems being

³⁹ As mentioned in Chapter 3, the syntactic structures of the pre-focused constituents are slightly different across the tone types. For T2, item n1 and n2 form a word *xue xi* “study.” Therefore the item “T2-F4-n2” is at word-final. It is a little different from T1 sequences where item “n2” is at the word-initial position of *tingshuo* “hear.”

focused does not have a big or definite impact on the success rate of tonal productions, especially those in longer focused constituents. It seems that the English speakers' error rate of tones may be affected by the "focus" feature more than Japanese and Korean speakers.

To get a bigger view, the top 50 syllables with highest error rates in each L1 data set were examined. The stimuli contain 228 syllables in total, 60 of them located in focused constituents and 168 of them neutral. So, among the 50 syllables, about 13.6 of them are expected to be within focused constituents according to the proportion of total focused syllables in the stimuli. It turns out that the three L1 groups of speakers have similar numbers of focused syllables with high error rates: English speakers have 14 focus-marked syllables, Japanese speakers have 11 and Korean speakers have 12. None of them are higher or significantly higher than the expected count 13.6, so "focus-marking" effects cannot be claimed. Among the 14 syllables with the highest error rates in the English speakers' data set, 9 of them are located at focused phrase- (and word-) final positions, including two items in monosyllabic focused constituents. Within the 11 syllables with the highest error rates in the Japanese speakers' data set, 7 of them were located at focused phrase- (and word-) final positions, including two items in monosyllabic focused constituents. Within the 12 syllables with highest error rates in the Korean speakers' data set, 6 of them were at focused phrase-final positions, including three of them that have the highest error rate of 95%, and two of the six are monosyllabic focused constituents. It seems, even within the same focused constituents, word- or phrase-final positions are also more vulnerable than other positions.

We also compare the average error rates of all focused syllables, i.e., the syllables coded with "f," and the average error rates of all neutral syllables, i.e., the syllables coded with either "n" or "v" in all tonal productions in the Experiment 2. The Table 5.9 displays the average error rates for each tone of each L1 group. The error rates are labeled with two "*" if they are at least 20% higher than the counterpart error rates for specific target tone. Three stars means the difference between error rates was at least 30%.

Table 5.9. Comparison of average error rates of focused and unfocused syllables in Exp2

speaker Target tone	English speakers		Japanese speakers		Korean speakers	
	Focused	Neutral	Focused	Neutral	Focused	Neutral
T1	33%	33.44%	30.8%	17.78%	23.6%	15.33%
T2	57.2%	59.44%	57.4%	50.11%	53.8%	47.56%
T3	***61.6%	26.67%	**44.2%	24.83%	43.8%	34.83%
T5	43.75%	51%	15.6%	25.25%	29.25%	37%
T4	10%	****62%	21%	25.44%	24.2%	19.44%

The comparative chart of average error rates of focused syllables and unfocused syllables indicates that English speakers are poor in producing, and in fact disfavor, low tones to express focus. English speakers can successfully produce T4 in focused constituents but disfavor making a high falling tone in unfocused constituents. A smaller difference between the error rates of focused and unfocused tones is in Japanese speakers' production of T3. It seems Japanese speakers are better in performing T3 in unfocused constituents than in focused constituents, but it does not mean Japanese disfavor low tone in focus expressions. We will discuss more about Japanese speakers' T3 issue in §5.3.1, where we found they often use FT3 in expressing focus, especially monosyllabic focus. There is no big difference between error rates of focused and unfocused syllables in the Korean speakers' data set.

In summary, we found that whether it is focused or not does not show definite impact on the error or success rates of L2 tonal productions, but English speakers demonstrate more sensibility to the “focus” feature than other speakers. English speakers tend to produce T4 more successfully in focused constituents and poorly produce T3 in expressing focuses. Japanese speakers also demonstrate a little sensibility to focus based on the study of error rates, and we will follow up this issue below. We discuss more about the positional effects in neutral constituents in the next subsection §5.2.3 and focused constituents in §5.3.

5.2.3. The error patterns in neutral constituents

As mentioned above, within the list of the top 50 syllables with the highest error rates for each group of L1 speakers, most of them are syllables located in unfocused (neutral) constituents. In this

section, we look into the details of the positional effects in the unfocused part of the test sentences. Unlike the positions located in focused constituents, all of these neutral positions in specific tone types bear the same morphemes. For example, the position “T1-n1” always bears the morpheme *ta* meaning “he.” The error rates of syllables at the same positions for specific target tones are examined. That is, the discussion is based on the error rates of all types of syllables coded with “n1,” “n2,” “n3,” “n4,” “n5,” “n6,” “n7,” “n8,” or “v” respectively. Due to Pre-Tone 3 sandhi, the target T3 has only six types of syllables, namely “n2,” “n3,” “n5,” “n6,” “n8,” “v,” while the target T5 has those with “n1,” “n4,” “n7,” “v.” The “v” in T3 sequences is in F2 while the “v” in T5 is in F1. The syllables with higher error rates compared to other syllables in the same column (the same tone sequence) receive more attention in the analysis. Those positions with error rates above 50% are bold.

Table 5.10. English speakers’ error rates of syllables in unfocused (neutral) substituents

T1	T2	T3	T3/5	T4
n1 0.57000	n1 0.33000	n2 0.14000	n1 0.39000	n1 0.540
n2 0.10000	n2 0.79000	n3 0.38000	n4 0.36000	n2 0.300
n3 0.32000	n3 0.63000	n5 0.36000	n7 0.29000	n3 0.510
n4 0.25000	n4 0.28000	n6 0.40000	v 1.00000	n4 0.540
n5 0.50000	n5 0.79000	n8 0.07000		n5 0.540
n6 0.43000	n6 0.52000	v 0.25000		n6 0.360
n7 0.13000	n7 0.43000			n7 0.610
n8 0.18000	n8 0.75000			n8 0.170
v 0.52500	v 0.82500			v 0.525

Table 5.11. Japanese speakers’ error rates of syllables in unfocused (neutral) substituents

T1	T2	T3	T3/5	T4
n1 0.22000	n1 0.19000	n2 0.22000	n1 0.48000	n1 0.280
n2 0.18000	n2 0.87000	n3 0.36000	n4 0.10000	n2 0.240
n3 0.09000	n3 0.51000	n5 0.07000	n7 0.13000	n3 0.230
n4 0.21000	n4 0.50000	n6 0.47000	v 0.30000	n4 0.150
n5 0.06000	n5 0.84000	n8 0.07000		n5 0.470
n6 0.34000	n6 0.41000	v 0.30000		n6 0.200
n7 0.15000	n7 0.19000			n7 0.380
n8 0.10000	n8 0.60000			n8 0.190
v 0.25000	v 0.40000			v 0.150

Table 5.12. Korean speakers' error rates of syllables in unfocused (neutral) substituents

T1	T2	T3	T3/5	T4
n1 0.27000	n1 0.26000	n2 0.23000	n1 0.25000	n1 0.390
n2 0.09000	n2 0.84000	n3 0.34000	n4 0.19000	n2 0.170
n3 0.07000	n3 0.52000	n5 0.28000	n7 0.24000	n3 0.190
n4 0.16000	n4 0.15000	n6 0.53000	v 0.80000	n4 0.180
n5 0.15000	n5 0.80000	n8 0.11000		n5 0.180
n6 0.18000	n6 0.26000	v 0.60000		n6 0.140
n7 0.18000	n7 0.13000			n7 0.220
n8 0.13000	n8 0.77000			n8 0.080
v 0.15000	v 0.55000			v 0.200

Across the three L1 sets, the T2 sequences have the biggest number of syllables with above-50% error rates. English speakers have the highest general error rate, so the total number of syllables with high error rates is also the biggest among the three groups. Several of the most vulnerable positions with high error rates are identified by examining the lists across the three L1 groups:

(5) Most vulnerable positions in neutral constituents

1. Word-initial positions for T2 and word-final positions for T4: “n2,” “n5,” and “n8” are also word-final in T2 sequences. “n1,” “n4,” and “n7” are word-initial in T4 sequences.
2. The two positions bearing verbs: the “v” positions before the focused constituents and the verbs at “n6” positions, especially in T3 sequences. They are also located at the ip-initial positions.
3. Major prosodic phrase-final positions: “n3” “n5” “n8” are all located at AP and/or ip-final positions, and “n8” is also at IP and utterance final positions. “n3” also immediately precede the focuses in F3, F4 and F5. We analyze the patterns considering the target tones in the discussion on specific L1 groups below.
4. Sentence-initial (especially for target T1, T4 sequences) and final positions (especially for target T2 sequences).

Next, we discuss the position-related error patterns in the neutral syllable data for each L1. In the English speakers' data set, more syllables in T2 and T4 sequences have high error rates than in T1 and T3 sequences. For target T1 sequences, high error rates take place at sentence-initial positions (“n1”), word-

and ip-final positions (“n5”) and ip-initial positions right before focused constituents (“v”). For target T2 sequences, high error rates are mostly at the word-final positions (“n2,” “n5,” “n8”), ip/IP-final position (“n3,” “n5,” “n8”), or ip-initial position (“v,” “n6”). “n8” is also sentence-final.

For target T3/5 sequences, the high error rates are at either ip-initial positions (“v,” “n6”) or ip-final positions (“n3” in F1). The “v” and sometimes “n3” are right before the focused constituents, and, same as the other “v” cases, the high error rates may be due to the speakers’ intention to make contrast of unfocused and focused. For target T4, high error rates are mostly at word-initial positions (“n1,” “n4,” “n7”), or ip-initial positions (“n1,” “n3,” “v,” “n4”) and word- or ip-final positions (“n5”). It’s very interesting that, compared to the tonal productions in T2 sequences, all word-final syllables in T4 sequences are successfully produced with low error rates except for “n5.” It confirms that the positional effects *Rise-F and *Fall-I we found in Experiment 1 are also very strong in sentence-level tonal productions. In addition to the word-level positional effects, the phrase-level positional effects are also obvious, where the phrase-initial or final positions are more vulnerable than other positions in neutral constituents. The “n5” is at both word-final and ip-final position. The target tones at “n5” are most often substituted by T3, and this substitution pattern seems to be popular in all three L1 groups. This trend indicates that all speakers tend to produce a low tone at major phrase final position.

In the Japanese speaker’s data set, the general error rates are much lower than the English speakers’, but, same as the English speakers, T2 sequences have the highest number of syllables with high error rates. For target T1 sequences, the positions bearing monosyllabic verbs (“v,” “n6”) have higher error rates. The verbs are at ip/AP-initial and right before the focuses. For target T2 sequences, word-final items (“n2,” “n5,” “n8”) have extremely high error rates (over 80%), where “n5” and “n8” are also at ip/AP-final. The items “n3” at ip-final and “v” at ip-initial and before focuses also have relatively high error rates (over 50%). For target T3 and T5 sequences, “n1” and “n6” have relatively high error rates (over 40%) and they are at ip/AP-initial positions. “n1” is actually at sentence-initial. For target T4

Sequences, same as English speakers “n5” located at both word- and ip-final position, has a relatively high error rate, and T3 is the most often substituted tone for target T4.

In the Korean speakers’ data set, for target T1, sentence-initial syllables have the highest error rate among the group. It is surprising because the first syllable in the test sentence is aspirated and supposed to be correctly produced as a high tone if L1 transfer is working. For target T2, word-final positions (“n2,” “n5”) have very high error rates (over 80%), and pre-focused positions (“n3,” “v”) have relatively high error rates (about 55%). For target T3 sequences, Korean speakers have similar error patterns to the English and Japanese speakers that “n6” and “v” have higher error rates than other items, and T3 is most often substituted for target T3 at “n6.” For T4, sentence-initial “n1” has a higher error rate than other syllables.

The two positions bearing verbs “v” and “n6” are both located at phrase-initial positions. Given the same word class, as well as the same status in syntactic and prosodic structures, the two items are expected to have similar behaviors. However, we found many more cases of “v” with high error rates than those of “n6.” It’s very possible that the “v” item’s more often having high error rates is because it is located at a pre-focus position. Speakers often tend to make the pitch shapes of pre-focus syllables and the following focused syllable contrasting.

To summarize, in unfocused constituents, all word-level, phrase-level and sentence-level positional effects are found. The word-level positional effects found in Experiment 2 data confirm the findings we got in Experiment 1. *Rise-F and *Fall-I are strong in sentence-level tonal productions across the three L1 data sets, especially in the English speakers’ data set. Phrase-level effects (*ip/AP-initial or final) are mostly overlapped with the word-level positional effects, except for the case of “n6.” Clearly, the ip/AP- or PhP-final positions are more vulnerable than phrase-initial positions. The fact that all speakers favor a substitute low tone at “n5” indicates a trend of having low pitch at major phrase finals. Sentence-level effects are mostly embodied with *T1 and *T4 at sentence-initial positions, and the *T2 at

sentence-final is overlapped with the word and phrase level *T2 effects. Some positional effects caused by focused constituents but falling in neutral syllables are also found: the pre-focused positions are very vulnerable. As we know, English is a “head prominence” language while Japanese and Korean are “edge prominence” languages. Although this kind of difference is manifested by the different pitch patterns used to express focus, which will be discussed in the next section, the positional effects analyzed above suggest that “edge prominence” languages and the “head prominence” language demonstrate similar effects of “vulnerable phrase edges” in the L2 tonal productions.

5.3. Different types of L1 learners’ expressions of focus

This section focuses more on the hypothesis test of different realizations of focus by three types of L1 speakers. The monosyllabic focus errors and substitution patterns are examined in §5.3.1 to demonstrate the L1-specific focus expressions with the feature of pitch contours. By examining the longer focused constituents, the hypothesis that the realization of longer focused constituents reflects L1 phrasal and disyllabic word level tone patterns is tested and discussed in §5.3.2.

The realization of focus (also called “prosodic stress,” “sentence stress” or “prominence” at the sentence level) is language-dependent. Previous studies show that there are two ways of prominence realization at a postlexical level: “culminatively” by marking the head of a prosodic unit (such as English) and “demarcatively” by marking the edge of a prosodic unit (such as Japanese and Korean) (Hyman 1978; Beckman 1986; Beckman and Edwards 1990; Ladd 1996; Venditti et al. 1996, as cited in Jun 2005). The postlexical pitch accents may be manifested by employing specific pitch accent tones, which are also used for the realization of lexical prosody in English. We expect English speakers to use a lot of high falling tones T4, similar to H*L, in expressing focus in Experiment 2. However, in the Japanese and Korean speakers’ L2 productions, prominence is realized demarcatively, for example, by marking the AP boundaries with L% in the Japanese speakers’ data set and H tone in the Korean speakers’ data set.

In section §5.3.1, we examine the substitute tone patterns for monosyllabic focus to confirm the hypothesis about English speakers and explore how Japanese and Korean speakers realize monosyllabic focus. It is found that T4 and FT3 are the two most popular tones used to express monosyllabic focus across the three sub-data sets. All learners have the highest accuracy with target T4 and the second highest accuracy with T1 (both are high registered tones). However, when errors occur, T4 and FT3 are the two most popular substitute tones used to express monosyllabic focus across the three sub-data sets. English speakers demonstrate an obvious preference for the substitute tone T4 in monosyllabic focus. Japanese speakers usually lengthen low tones and produce a lot of FT3 to substitute target tones in focus. Korean speakers use more T4 than FT3 in substitutions, but do not demonstrate strong preference.

In section §5.3.2, we examine phrasal level tonal production (long focused constituents) and look at a bigger picture of the L2 tonal production patterns. Section §5.3.2.1 tests transfer effects of L1 intonation patterns over long focused constituents, such as whether (1) English speakers use T4 mostly at the rightmost arguments of the focused phrase as in its L1 because native English usually has the strongest prominence at the rightmost word in broad focus; (2) Japanese speakers seldom use rising tone in the second half of the phrase, and the APs are ended with L%, and (3) Korean speakers' APs are ended with H registered tones. In section §5.3.2.2, we describe the concrete tone patterns in L2 tonal productions of sentences. It is hypothesized that the L2 tonal productions are framed by binary foot constraints, and the tone patterns over prosodic words in Experiment 2 may reflect the L1 transfer of the most frequent disyllabic tone patterns found in Experiment 1. A universal trochaic-footed tone pattern is found across the three sub-datasets, which reveals another similarity of these L1 speakers.

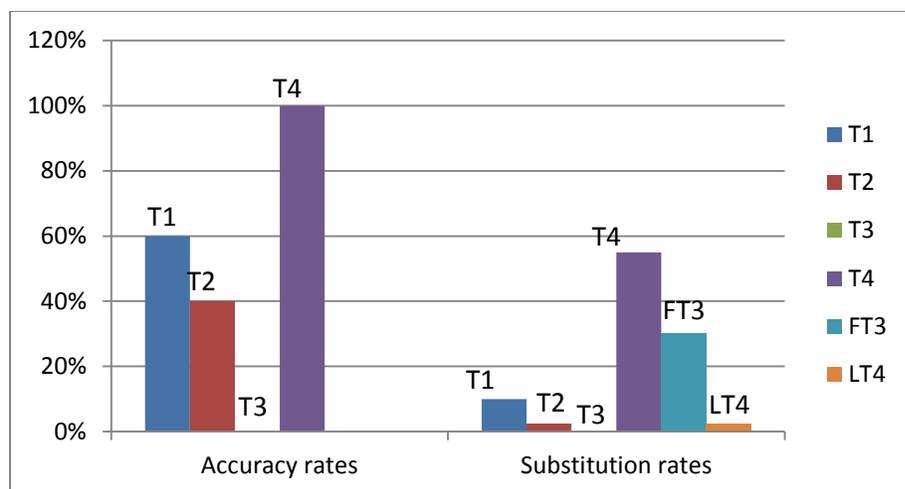
5.3.1. The monosyllabic focused constituents (F1)

This section looks into the pitch patterns of monosyllabic focused constituents to examine how different types of L1 learners express short narrow focus. Native Mandarin speakers usually mark materials as focused by manipulating pitch range (expanding pitch range on focus and narrowing post-

focal pitch range), increasing amplitude and/or duration, but not changing the lexical tone identity or tone types (Flemming 2008). However, we expect that the L2 learners of Mandarin would change the monosyllabic lexical tone identity/type in the process of focus realizations, although they may demonstrate different substitution patterns. Monosyllabic focus is a very good indicator of the different substitution tone patterns, and both error and substitution patterns are examined for each L1. All learners may use a lot of high pitch to mark focus. English speakers probably would tend to express focus by using a specific pitch contour, for example the T4 which sounds like accented H*+L. By contrast, Japanese and Korean speakers may adjust the prosodic phrasing to some degree. In this section, we examine both error and substitution patterns of monosyllabic focus within each sub-dataset.

The error (or accuracy) and substitution patterns of monosyllabic focused constituents in the English speakers' data set are very systematic. The accuracy rates of each tone type are 70% for T1, 45% for T2, 25% for T3 and 90% for T4. The error pattern shows that English speakers are best at doing T4 for monosyllabic focus expression. The substitution pattern confirms English speakers' favor of T4. Within all 40 errors for the target F1-f1 of four tone types, T4 is most often substituted for target tones, and T3 is the least frequently substituted. About 55% of the errors are T4, 30% of errors are FT3, 10% of the errors are T1, and 2.5% are T2 and Low T4 respectively. Both the accuracy rates and substitution rates suggest that English speakers favor T4 in expressing monosyllabic focus. The accurate rates and substitution rates are displayed in Figure 5.4. The accuracy rates are only for the target tones, but substitution rates include the tones in actual L2 productions, such as FT3 and Low T4. Please note the accuracy rates are the percentages of correct/true productions within the productions for specific target tones. However, the substitution rates are out of the total errors, and the total percentages of these substitute tones are 100%. This setting is the same in the following two figures 5.5 and 5.6 for Japanese and Korean speakers.

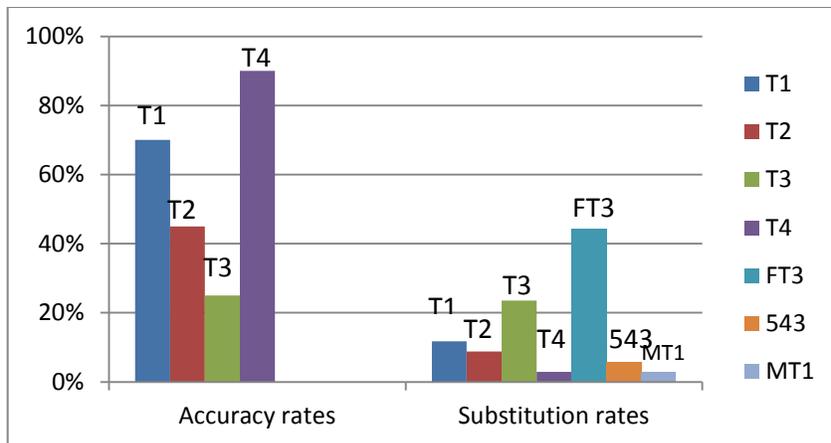
Figure 5.4. Accuracy rates and substitution rates for target F1-f1 in the English speakers' data set



The accuracy rate of T4 is significantly higher than that of the other tones. The occurrence rate of the other tones was compared with T4 using the Poisson test. The results show that the substitution rate of T4 is significantly higher than that of T1, T2, T3, LT4, with P-values all smaller than 0.001. Although the substitution rate of T4 is much higher than FT3, it is not statistically supported ($p=0.1214$). In summary, English speakers' favor of T4 in expressing focus is still very obvious, although the use of FT3 is also noticeable. The T4 is very like H*L, and FT3 is a very long low tone which is like L*H. Remember that T3 is the tone which is most frequently substituted for target tones in general (see §5.1.1), but it is not true in the English speakers' F-f1 case because of the focus-marking.

The error (or accuracy) patterns and substitutions in Japanese speakers' monosyllabic focus productions are a little odd, but we still can identify the patterns. The left part of figure 5.5 displays the accuracy rates of target T1, T2, T3, T4. There are 34 errors in Japanese speakers' F1-f1, and the substitution rates of T1, T2, T3, T4, FT3, MT1, [543] in Japanese speakers' actual tonal productions of monosyllabic focus are displayed in the right part of the figure.

Figure 5.5. Substitute tones for target F1-f1 in the Japanese speakers' data set

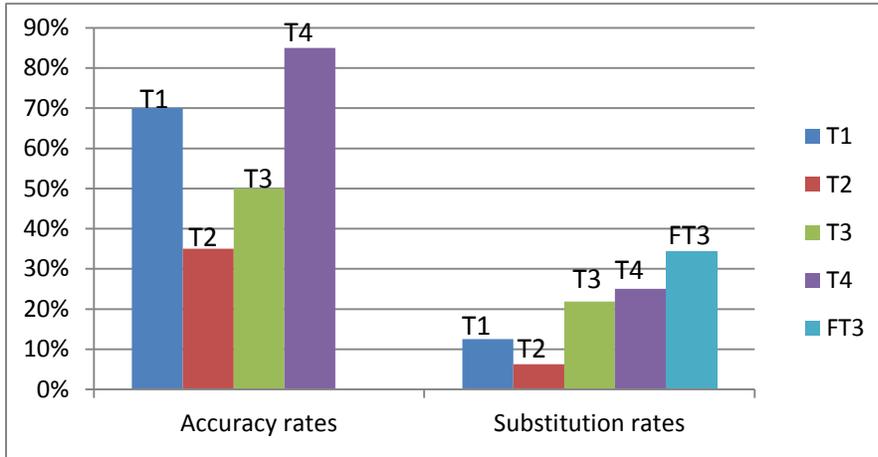


Unlike the English speakers' data where the accuracy rates are correlated with the substitution rates, the Japanese speakers' accuracy pattern conflicts with the substitution pattern. According to the accuracy (or error) rates, Japanese speakers are best at producing T4 to express monosyllabic focus, and poorest at T3. When errors occur, Japanese speakers produce FT3 (15 times) most frequently to substitute the target tone. It is not surprising because FT3 is phonetically long and sounds like a low tone emphasized or stressed by lengthening the rhyme.⁴⁰ However, T3 (8 times) is the second most frequently used substitute tone, and T4 is the least frequently used substitute tone when errors occur, which is just the opposite of the accuracy rate pattern. The occurrence rate of the other tones was compared with FT3 using the Poisson test. It was found that the substitution rate of FT3 is significantly higher than that of T1, T2, T4, [543], MT1 with p-values smaller than 0.05. Although the substitution rate of FT3 is almost double T3's substitute rate, it is not statistically supported ($p=0.21$) due to the small total error numbers. In summary, the error rate patterns show that Japanese speakers are best at producing high registered tones, i.e., T4 and T1, in monosyllabic focus. However, when errors occur, Japanese speakers favor low tones FT3 and T3 to substitute the target tone in monosyllabic focus. Although Japanese is well known as an "edge prominence" language, the speakers also employ some specific pitches, such as FT3 to express focus, especially short narrow focus.

⁴⁰ Kaltenbacher (1997) also found that Japanese learners of German often lengthen the rhythm to express focus in the non-native German productions, and it is claimed to be L1 transfer of focus-expressing strategies. It seems Japanese speakers use similar strategies in producing focus for L2 Chinese and German.

The accuracy rates and substitution rates for F1-f1 in Korean speakers' tonal productions are listed in figure 5.6. For the substitutions, there are 33 errors in total.

Figure 5.6. Substitute tones for target F1-f1 in the Korean speakers' data set



Korean speakers only have conflicting accuracy and substitution rates for T1. As a target tone, the accuracy rate for T1 is as high as 70% and ranked the second. However, the substitution rate of T1 is only 12.5%, and it is ranked the second from the bottom. The accuracy rates and substitution rates of other tones are systematic. It shows that Korean speakers can successfully produce T4 and T1 in most cases of monosyllabic focus. When errors occur, Korean speakers tend to produce FT3 the most frequently, and also T4 less frequently. The occurrence rate of the other tones was compared with FT3 using the Poisson test. The statistical analysis of substitute tones does not support that Korean speakers prefer specific tones to express narrow focus because the substitution rates of all tones are very close. The p-values for equal counts with FT3 are all bigger than 0.05, except for T2 (.0225), which only indicates that Korean speakers disfavor T2 in expressing focus.

Based on all the information about the accuracy rates and substitution rates presented above, it seems that T4 and FT3 are the two most popular tones used to express monosyllabic focus across the three sub-data sets. Both T4 and FT3 indicate some extreme features of lexical tones. T4 is a high registered tone with a sharp falling, the biggest swing in F0; and FT3 is a low tone but with the longest

duration among all Mandarin tone categories. English speakers demonstrate an obvious preference of T4, a high falling tone, in monosyllabic focus, which reflects the L1 transfer of accented pitch tone H*L. Japanese and Korean speakers do not show a strong preference for specific pitch shape in signaling sentence prominence like English speakers. Japanese and Korean are “edge prominence” languages, but they also show a high accuracy rate of T4 in monosyllabic focus and frequent use of FT3 to substitute target tones in focus. Therefore, it seems there is no clear demarcation between the “head-prominence” and “edge-prominence” languages. The data in the present study shows varying degrees of preference of prosodic resources (adjustment of pitch range, shapes, durations, etc.) used in prominence realization by different types of languages.

5.3.2. Long focused constituents (F2-F5) and the disyllabic tone patterns

This section examines the long focused constituents of F2-F5 to examine the L1 transfer effects of phrasal level pitch patterns and disyllabic tone patterns in longer focused constituents.

We check the L1 transfer effects of phrasal pitch patterns mainly in F3, F4 and F5 in section §5.3.2.1.

Section §5.3.2.2 tests the binary foot constraints in long focused constituents and examines the L1 transfer effect of typical disyllabic tone patterns in sentence-level tonal productions. It is argued that the prosodic foot constraints work in the realization of long focused constituents. The results show that the transfer of English phrasal level “strongest prominence placement” is found in the L2 tonal productions. However, the L1 transfer effect of phrasal pitch pattern in the Japanese and Korean speakers’ data set is weak. The typical L1 disyllabic word pitch patterns are found in the English and Japanese speakers’ data sets, but not found in the Korean speakers’ data. All the tonal productions in both experiments show a preference for trochaic foot pattern associated with prosodic words.

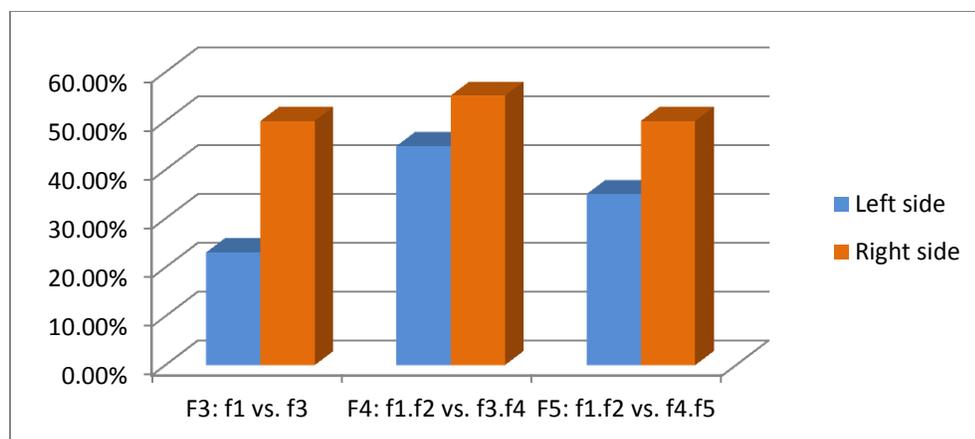
5.3.2.1. L1 transfer of phrasal level pitch patterns

Because English is a “head prominence” language and Japanese and Korean are “edge prominence” languages, we use different ways to examine the L1 transfer effects of phrasal pitch patterns

for different sub-data sets. For the English speakers' L2 data, we examine the location of T4, English speakers' preferred "accent pitch" in L2 productions, in longer focused constituents F3, F4 and F5. If L1 transfer occurs, we would find the strongest prominence at the right part of the long focused constituents. For the Japanese and Korean speakers, we check the edge tones of Accentual Phrases to confirm if the prominence is realized demarcatively. As shown in (4), the long focused constituents of F4-F5 themselves formed Aps, which is different from short focused constituents. So if the L1 phrasal level pitch pattern had been transferred to L2 productions, the Japanese speakers would use no or very few rising tones (T2) at the second half of the long focused constituents and the AP final position would be marked with L%; the Korean speakers would mark the AP-final position with H registered tone, or, use LH pitch patterns (T3-T1/T4 or T2-T3/T4) at the AP-final words signaling the right edge of AP.

English focus is marked by pitch accent, and English usually has the strongest prominence at the rightmost word in the phrase in broad focus. As indicated in section §5.3.1, in L2 tonal productions, English speakers employ T4 to express monosyllabic focus. T4 is also the tone most often substituted in longer focused constituents when errors occur. We expect that T4 appears at the right part of long focused constituents more often. The total number of T4 in all long focused constituents of F3, F4 and F5 were counted. We also counted and compared the frequencies of T4 occurring at the left parts of these constituents, including the positions of "F3-f1," "F4-f1 &f2," "F5-f1&f2," "F3-f3," and at the right part of the long constituents, including positions of "F3-f3," "F4-f3&f4," "F5-f4&f5." Figure 5.7 illustrates the percentages of the T4 occurrences at the left (or first half) and right sides (or second half) of long constituents:

Figure 5.7. The distribution of T4 in long focused constituents (English speakers)



In all F3, F4 and F5 focused constituents, the frequencies of T4 at the right side are all higher than at the left side. It is L1 transfer of the placement of “strongest prominence” at VP-final arguments in broad focus. That is, same as their native English, English speakers also place “T4,” which embodies the greatest stress prominence, mostly at the rightmost word in long focused constituents.

By contrast, the phrasal level L1 transfer effects are weak in Japanese and Korean speakers’ data sets. In Japanese speakers’ F4 and F5 constituents for target T1, T3 and T4 sequences, many T2 are found at the second half of the constituents (starting from the third syllables in F4, F5) as substitute tones. In each erroneous or substitute phrase, T2 appears no more than once. T2 takes place for 5 times in the second half of the 13 erroneous phrases of T1, 6 times in the right part of 14 erroneous phrases of T3, and once in the right part of the 9 erroneous phrases of T4. Compared to the other substitute tones such as T3, T4 and FT3, the occurrence of T2 is relatively low, but the substitute T2 always occurs in the second half of these erroneous phrases, and never at the beginning part of the phrase like the typical Japanese AP pitch pattern. The AP-final tones are checked at the focused constituent-final position of F4 and F5, and only very few low tones are found substituting target tone sequences of T1, T2 and T4. It is found that only 11 out of 40 erroneous F4 sequences and 4 out of 28 erroneous F5 sequences are ended with low tones. If we assume de-phrasing occurs in all sentences and the AP-final is at “n5” positions across all sentences, most of the target tones at “n5” are substituted with low tone T3. However, the “L%” not only occurs in the Japanese speakers’ data set, but in fact is a universal tendency across the three L1 groups as

discussed in §5.2. In the Korean speakers' F4 and F5, there are 50 erroneous phrases for F4 and F5 of T1, T2, T3 and T4. Only one case of T3-T1 (a LH-like pitch shape) was found for a T1-F5 sequence. Even if we assume the de-phrasings happened between the focused constituents and the following nouns, and the right edge of AP moved to "n5" position, the T3 or FT3 is the most frequently substituted tone for the syllable at "n5," and the AP is not ended with a high tone like in a typical Korean AP. The weak L1 transfer effects of Japanese and Korean AP pitch patterns may be due to the different VP word order of the L2 and L1s. Both Japanese and Korean have SOV sentence order, and even in the VP phrases serving as modifiers, the objects precede the verbs. The focused constituents are all VP phrases, and they all have verbs that precede the objects. It seems that in this case of L2 tonal productions, the syntax-based accent placement rules may be interfering with the L1 prosodic transfer (Selkirk 1984, 1985).

In summary, the L1 prosodic transfer effect of the "strongest prominence" locating at the rightmost word of a phrase is found in the English speakers' data set. However, the phrasal level of the L1 prominence realization transfer seems very weak in Japanese and Korean speakers' data sets, and it is proposed that this is due to the interference of the different syntactic structures of the L1s and L2.

5.3.2.2. The transfer of typical L1 disyllabic tone patterns

In addition to the L1 transfer of phrasal level pitch patterns, it is also hypothesized that L2 long tonal productions are constrained by the binary foot constraint, and that the disyllabic tone patterns similar to typical L1 tone patterns on two-syllable words may be transferred to the sentence level of L2 tonal productions. That is, those most frequently produced disyllabic tone patterns found in Experiment 1 can also be found in Experiment 2 and are borne by prosodic words. In this section, we test this hypothesis by examining the most frequently produced tone combinations in F4 and F5 and extend the survey from the prosodic words in focused constituents to those in neutral constituents to confirm this finding. We examine disyllabic word tone patterns to confirm the L1 transfer effects. However, it turns out to be a "universal" trend of the most frequently produced disyllabic tone combinations, because there are a considerable amount of trochaic-footed disyllabic tone combinations across the three sub-data sets.

We first tested if the L2 productions are governed by the binary prosodic foot constraint. Since the Experiment 2 stimuli are identical tone sequences, we thus focus on the erroneous productions, where the target identical tones sequences are modified or changed into non-identical tone sequences. We designed a test of L1 transfer of disyllabic tone patterns based on the role of disyllabic prosodic words in a dissimilation process (T3 sandhi) of native Mandarin. T3 sandhi is a dissimilation process similar to the case of erroneous phrase productions in Experiment 2. In native Mandarin, the domain of T3 sandhi is the prosodic foot (Shih 1986, Lin 2005). So we assume that the dissimilation processes occurring in the Experiment 2 are not simply applied on any adjacent syllables, but the syllables within prosodic words. On the disyllabic prosodic words, we expect that some L1 transfer effects can be found. That is, if the L2 tonal productions in longer focused constituents are constrained by the prosodic structure of the binary foot, the typical L1 disyllabic tone patterns (ie. the most frequently produced tone combinations found in Experiment 1) may emerge in Ft1 as in Ft2 in the L2 productions. However, two adjacent syllables which do not belong to the same prosodic foot may not have such patterns.

F4 constituents provide an ideal ground for the test of prosodic foot constraints, since the syntactic structure and prosodic structure of the prosodic word composed of syllables “f1” and “f2” are not isomorphic. (7) shows this isomorphism. The syntactic structures of focused constituents are the same across tone types, so we take T1-F4 as an example. There are four syllables “f1,” “f2,” “f3,” “f4” in the focused constituents. “f3-f4” is a lexical disyllabic word, so it is also a prosodic word. We call this foot “Ft2.” The morpheme “f1” is an adverb word while “f2” is a monosyllabic verb, so they are not a lexical word syntactically. However, “f1-f2” forms a prosodic word, and we call it “Ft1.” We expect that the L1 transfer may not occur in adjacent syllables which are not a prosodic word. We thus make a dummy prosodic word using “f3” and “f4” in five-syllable focused constituents with F5 as a counterpoint. The “f3-f4” in F5 is not a lexical word, nor a prosodic word.

(6). The syntactic and prosodic structures of F4 and F5

a.		Ft1		Ft2		
FOCUS 4:		f1	f2	f3	f4	
Prosodic feet:		(σ σ)		(σ σ)		
Syntactic structure:		(σ σ		(σ σ)N)vp)		VP
Pinyin:		Duo	He	Ji	Tang	
Gloss:		MORE DRINK		CHICKEN SOUP		
						Dummy prosodic word
b.				f3	f4	← f5
FOCUS 5:	f1	f2				
Prosodic feet:	(σ	σ)	(σ	(σ	σ)foot)	super foot
Syntactic stru:	(σ	σ	(σ	(σ	σ)N)vp)	VP
Pinyin:	Jin Tian	He	Ji	Tang		
Gloss:	TODAY	DRINK	CHICKEN	SOUP		

As shown in (6), “Ft1” is not a lexical disyllabic word but a prosodic word. So, if the L2 tonal productions in longer focused constituents are constrained by the prosodic structure of the binary foot, the substitute tone patterns in “Ft1” would be very close to those in “Ft2” but not the dummy word, and this may reflect the L1 transfer of disyllabic tone patterns. So we examine the most frequently produced substitute tone productions in these three words in the sentence-level tonal productions in Table 5.13. The most frequently produced substitute tone productions include identical tone combinations (ITC), which are mostly correct productions, and non-identical tone combinations (NITC), which are erroneous productions. We are only concerned about the top three erroneous productions (NITC) with the highest frequencies. Therefore, the ITCs are dimmed in grey. The second column contains the top 3 most frequently produced disyllabic tone combinations found in Experiment 1. The following columns display the top 5 most frequently produced tonal productions at “Ft1,” “Ft2” and “dummy word” and followed by the counts of occurrences in Experiment 2. The tone combinations in “Ft1,” “Ft2” and “dummy word” are bold if they are phonetically identical to those tone combinations in the second column.

The examination of most the frequently produced tone sequences is to answer two questions: (1) whether the tone patterns in “Ft1” are more similar to those in “Ft2” or dummy words and (2) whether the most frequently produced tone sequences on prosodic feet in Experiment 2 are the same as the most frequently produced tone sequences found in Experiment 1.

Table 5.13. The most frequently produced tone sequences for “Ft1,” “Ft2” and dummy word

	Most frequent tone sequences in Exp. 1	Ft1		Ft2		Dummy word	
English speakers	T1-T3 (10.54%) T2-T3(10.39%) T3-T4(8.67%)	T1-T1	14	T1-T1	11	T1-T1	19
		T4-T4	14	T5-T3	11	T4-T4	11
		T2-T3	10	T4-T4	6	T2-T2	9
		T5-T3	10	T1-T3	5	T3-T2	7
		T3-T4	7	T1-T4	5	T3-T5	5
Japanese speakers	T1-T3(11.56%) T2-T3(9.84%) T3-T1(9.45%)	T4-T4	17	T4-T4	13	T4-T4	15
		T1-T1	14	T2-T3	10	T1-T1	13
		T5-T3	11	T1-T1	9	T3-T5	9
		T2-T2	8	T5-T3	9	T2-T2	7
		T2-T3	7	T1-T3	4	T2-T5	6
Korean speakers	T3-T4(11.64%) T1-T1(10.47%) T1-T4(9.22%)	T4-T4	16	T2-T3	13	T1-T1	18
		T1-T1	14	T5-T3	13	T4-T4	16
		T2-T2	10	T1-T1	12	T2-T2	12
		T5-T3	7	T4-T4	11	T3-T5	6
		T5-FT3	6	T4-T1	7	FT3-T5	4

We address question one first. It is expected that the tone patterns in “Ft1” are more similar to those in “Ft2” but should be different from the tone patterns in dummy words. This is true for all three groups of speakers. The similarity between “Ft1” and “Ft2” can be viewed in terms of the register pattern. Almost all of the most-frequently-produced NITC in “Ft1” and “Ft2” across all subsets in experiment 2 are trochaic footed (or [+U][-U] registered, such as T1-T3, T2-T3, T5-T3) while those in the dummy words are all iambic footed (or [-U][+U] registered, such as T3-T2, T3-T5, FT3-T5). The strong patterns indicate that the dissimilation processes operating in L2 tonal productions made by all speakers may be constrained by the prosodic word structures. The frequently produced non-identical tone sequences found in all L1 groups suggest a “universally” preferred trochaic footed tone pattern.

The answer to the second question is also very clear. The L1 transfer of disyllabic tone patterns is only found in the English and Japanese speakers’ data. 43 out of 48 cases in the English speakers’ “Ft1/2” and all cases in the Japanese speakers’ “Ft1/2” are identical to the most frequently produced tone combinations in Experiment 1. However, none of the Korean speakers’ tone combinations in “Ft1/2” are the same as those in the second column, which indicates a discrepancy of word-level and sentence-level tests of most frequently produced disyllabic tone combinations.

This L1 transfer effect test above is only conducted within the F4, and we extend the examination of tonal patterns on prosodic words to other focused constituents and non-focused constituents. We also found the preference of trochaic footed tone patterns across the three L1 data sets. The chart 5.14 displays the extended examination of the top five most frequently produced tone sequences on disyllabic prosodic words of all tone types in other focused constituents and neutral constituents. Since the syntactic structure of pre-focus materials in the stimuli are slightly different across the tone types, we only examine the materials starting from position “f1” in F2, F3 and F5. The prosodic words in neutral constituents “n4-n5” and “n7-n8” include all tone productions of “n4,” “n5,” “n7,” and “n8” in the 20 stimulus sentences. Therefore, each focus-marked prosodic word has 80 productions in each L1 data set, while each prosodic word in the neutral constituents has 400 productions made by 20 subjects in total. The numbers following the tone combinations in each cell are the occurrence counts of the productions. The identical tone sequences are correct productions and they are dimmed in gray. We are concerned with the non-identical tone sequences. Most of them are erroneous productions. Those trochaic footed tone combinations or [+U] [-U] registered sequences such as “T1-T3,” “T2/5-T3,” “T4-T3” are bold.

Table 5.14. Prosodic word tone patterns in F2, F3, F5 and neutral constituents

Syllables	Prosodic words in other focused constituents				Prosodic words in neutral constituents	
	F2: f1-f2	F3: f2-f3	F5: f1-f2	F5:f4-f5	n4-n5	n7-n8
English Speakers	T4T4 15	T1T1 12	T1T1 17	T1T4 11	T1T3 57	T1T1 78
	T1T1 10	T1T4 12	T4T4 15	T4T4 9	T5T3 57	T5T3 70
	T2T3 10	T4T4 9	T5T3 8	T1T1 8	T2T3 51	T3T4 39
	T1T4 7	T2T2 5	T1T3 5	T2T3 7	T1T1 43	T2T3 38
	T5T3 6	T5FT3 5	T2T2 5	T3T4 7	T4T4 28	T4T4 36
Japanese speakers	T1T1 13	T4T4 12	T1T1 19	T5T3 14	T5T3 87	T5T3 85
	T5T3 12	T1T1 11	T4T4 17	T4T4 12	T1T1 76	T1T1 83
	T4T4 11	T5T3 11	T5T3 12	T1T1 8	T1T3 62	T4T4 59
	T2T3 8	T1T3 6	T1T3 5	T2T3 6	T4T4 53	T2T3 42
	T2T4 6	T4T3 6	T2T1 5	T3T4 6	T2T3 31	T2T2 31
Korean speakers	T5T3 12	T1T1 15	T1T1 15	T1T1 16	T1T1 73	T1T1 75
	T1T1 11	T2T3 12	T4T4 15	T2T3 12	T4T4 68	T5T3 75
	T4T4 11	T5T3 12	T2T2 10	T5T3 11	T5T3 66	T4T4 73
	T2T3 10	T4T4 12	T5T3 9	T4T4 10	T2T3 42	T2T3 39
	T4T1 5	T4T3 5	T3T4 6	T1T4 5	T1T3 19	T2FT3 26

The chart shows that most of the highest frequency tone sequences on prosodic words in F2, F3, F5 and neutral constituents across the three L1 groups are featured with [+U][-U] register. The tone sequences in neutral constituents demonstrate stronger preference for [+U] [-U] tone patterns than those focus-marked. It may be because the words “n4-n5” and “n7-n8” are located at the major-phrase boundaries, and all speakers demonstrate a stronger tendency to end the major phrase or utterance with a low pitch. Some prosodic words are not located at phrase-final positions, such as the words “F4:f1-f2” (i.e., the “Ft1” word) and “F5:f1-f2,” but still have trochaic footed tone sequences as their most frequent productions. In summary, the L2 productions in Experiment 2 show a universal preference for trochee-like [+U] [-U] pitch patterns on disyllabic prosodic words that are woven into sentences.

Examining the L1 pitch patterns, English lexical level stress patterns and the Japanese lexical pitch accent pattern HL show some connections to the [+U] [-U] pitch pattern. The dominant accent pattern in English is trochaic, and the strong-weak stress pattern may lead to [+U] [-U] pitch pattern. The typical lexical pitch accent pattern of Japanese is HL, and about half of lexical words are accented. This may lead to the [+U] [-U] pitch pattern in L2 productions. However, if we ascribe the trochee-like pitch pattern [+U] [-U] to L1 transfer in the English and Japanese cases, it is hard to make sense of the Korean speakers’ case because a typical Korean disyllabic word is usually L-H, which is in conflict with this prevailing [+U] [-U] pattern. It seems the [+U] [-U] affects Korean speakers’ production also. Remember in Chapter 4 (end of §4.3.3) regarding Korean speakers’ substitution pattern that T3 is most often used to substitute target tones at word-final positions. This may also be due to the popular [+U] [-U] pitch pattern.

Another possible source of the trochee-like pitch pattern is the target language, Mandarin. The possibilities for stress in tone languages are an area of ongoing research, but trochaic stress patterns have been observed in Mandarin Chinese. In this section, we discuss two stress-related issues in native Mandarin which demonstrate some connections to the strong-weak, or the L2 [+U] [-U] pitch pattern. One is weakly stressed syllables in Mandarin, and the other is the prosodic strength pattern in the phonetic realization of tones by Mandarin speakers.

Although Mandarin is a typical tonal language, standard Chinese also possesses both stress and intonation. Weakly stressed syllables (*qīngshēngzì*) in disyllabic words usually involve a trochaic stress pattern in Mandarin which is based on the Beijing dialect⁴¹. For example, the second syllables in the citation forms of *māmā* “mother” [55] [55], *miánhuā* “cotton” [35] [55], *piàoliàng* “pretty” [51][51], *tàiyáng* “sun” [51] [35] may become unstressed and lose their citation tones in spoken Mandarin. The above words thus may surface as *māma*, *miánhua*, *piàoliang*, and *tàiyang*. This process leads to some contrastive stress patterns, namely the word-initial syllable is stressed and the final syllable is neutral within lexical words. Weak stress is an essential feature of Mandarin, and of most northern dialects, although only a small amount of lexical words possess this kind of contrastive stress pattern. The L2 learners are exposed to and acquire this kind of strong-weak stress patterns to some degree.

Research on the phonetic realization of lexical tones in connected speech suggests a strong-weak strength pattern. In Kochanski, et al. (2003), quantitative measurements of prosodic strength are measured in Mandarin by different models.⁴² It is found that all the real segmentations in the study show a clear strong-weak pattern for two-syllable words, three-syllable words (though the patterns are weaker than the two-syllable case) and four-syllable words, which are broken up into two two-syllable words. The stress-like patterns are realized as alternations between syllables where the tones are carefully realized with a relatively large swing in fundamental frequency, and syllables where they are realized "sloppily," typically with a small swing.

The above are some evidence for a strong-weak or trochaic stress pattern in Mandarin, so it is at least possible that L2 learners are picking up that pattern and it is influencing their interlanguage (more than it does in the target L2 grammar). That is, in L2, the “strong-weak” pattern does not affect the lexical tone shapes (otherwise it would change the lexical meaning). However, when L2 learners learned the pattern, it may have affected the tone shape, thereby changing the tone identity in L2 tonal productions.

⁴¹ It is different from Taiwan Mandarin.

⁴² These models use Stem-ML, which is a phenomenological model of the muscle dynamics and planning process that controls the tension of the vocal folds, and therefore the pitch of speech.

The third possible reason for the prevailing [+U] [-U] pitch pattern found in this section is the hypothesis of a universal “trochaic bias.” The role of a trochaic accentual pattern is emphasized in many studies of L1 acquisition, such as the speech of perception and segmentation during the prelinguistic period and early word productions in English. Some studies even posited a universal “trochaic bias” (Vihman et al. 1998). The present SLA study of Mandarin tones also indicates a “trochaic bias” across three types of non-tonal language speakers’ Mandarin tone productions.

5.4. Summary of Chapter 5

Chapter 5 examines the three types of L1 speakers’ tonal productions in the sentence level experiment. In spite of some minor differences, the Experiment 2 results confirm the findings we got from Experiment 1 regarding TMS, OCP and TPC. In addition to the word-level positional effects of *Rise-R and *Fall-I, other vulnerable positions are also identified at the phrase and sentence level, such as the focus-related vulnerable positions, including focus-final positions, and vulnerable positions in neutral (non-focused) constituents, including sentence-initial and -final positions, ip- or AP-final positions, and pre-focus positions. According to the data, the focus-mark does not have a definite impact on the error rates of tonal productions, but English speakers do show more sensitivity to the focus feature in producing T3 and T4. In expressing monosyllabic focus, English speakers use many more T4 than other tones, and T4 is located mostly in the rightmost arguments of longer focused constituents; Japanese speakers use FT3 and T3 more often than other tones; and Korean speakers do not show a strong preference for specific tones. The L1 prosodic transfer effect of the “strongest prominence” located at the rightmost word of a phrase is found in the English speakers’ data set. However, the phrasal level of L1 prominence realization transfer seems very weak in the Japanese and Korean speakers’ data sets, and it seems to interfere with the different syntactic structures of L1s and L2. It is shown that the dissimilation processes in Experiment 2 are constrained by the binary foot constraint. When we examine the prosodic word tonal productions, we see the three L1 groups of speakers all produced a considerable amount of trochaic-footed tone sequences, which indicates a universal “trochaic bias” in L2 tonal productions.

CHAPTER 6: A DISCUSSION OF THEORETICAL IMPLICATIONS

This chapter discusses theoretical implications of Experiments 1 and 2 on interlanguage grammars (ILG) within the framework of Optimality Theory.

In the previous two chapters, we closely examined L2 tonal productions and tested for evidence of three phonological constraints: the TMS, the OCP, and TPC. It was found that these phonological universals, to varying degrees, constrain the L2 tonal grammars in word-level productions. This result is confirmed in sentence-level productions. We also found L1 transfer effects at the individual tone level, the disyllabic word level, and also at the sentence (focus) level. Additionally, this study uncovered error and substitution tone patterns that appear to be independent of both the L1 (native-language) grammars and the L2 (target-language) grammar. For example, for the English speakers, both the L1 (English) and L2 (Mandarin) allow high level, low level, rising, and falling tones, but learners performed level tones, especially low level tones, better than falling tones. Another example can be seen in the fact that all L2 learners seemed to disfavor rising tones, especially at word-final positions. In both Experiment 1 and Experiment 2, we found a universally-preferred High-Low registered tone sequence in L2 tonal productions. For the Japanese and English speakers, this could be due to L1 transfer. However, for the Korean speakers, this could not be attributed to L1 transfer since Korean always has high tones at word or phrase-final positions. In terms of identical tone sequences, contour tone pairs (rising-rising and falling-falling tone patterns) are not allowed, and are thus equally bad, in all L1s and are equally good in the L2. However, we noticed that in the tonal production of both experiments, there are always many more T1-T1 (level-level sequences) than T4-T4 (falling-falling sequences), and more T4-T4 than T2-T2 (rising-rising sequences) across the three types of learners' data sets.

As discussed in Chapter 2, the acquisition of an L2 can be viewed as the adjustment of the L1 grammar ranking toward the L2 ranking in the spirit of OT. This chapter introduces a conceptual model of

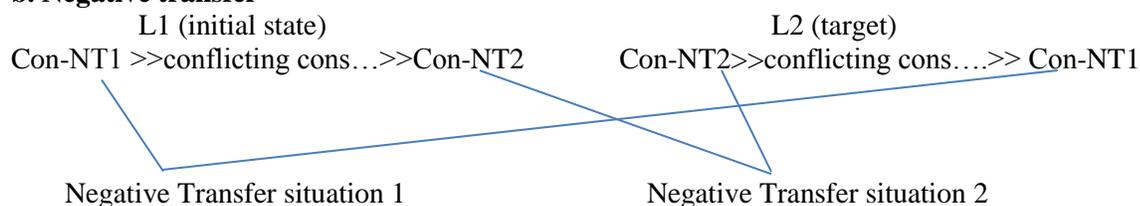
constraint re-ranking for SLA in §6.1, to provide a new way of viewing positive and negative transfer. Since this study mainly focuses on the operation of phonological universals in the SLA of Mandarin tones by English, Japanese, and Korean speakers, we pay particular attention to the re-ranking of Markedness constraints. Two types of Markedness constraints associated with positive transfer (Con-PT1 and Con-PT2) and two types of constraints associated with negative transfer (Con-NT1 and Con-NT2) are identified and effects of the re-ranking of constraints during SLA are exemplified by the tone-related constraints used in the present study.

The discussions in the subsequent sections §6.2 and §6.3 concerning the interlanguage tonal grammars in the present study provide evidence for the movement of Markedness constraints within a conceptual constraint re-ranking model. In section §6.2, part of the ILG properties, or, the choice of specific surface tone forms made by L2 learners, are illustrated through tableaux based on common error patterns found in Chapter 4. Partial interlanguage tonal grammars are worked out through representative tableaux. It is confirmed that phonological universals as well as L1 transfer play a role in shaping interlanguage tonal grammars. Some prominent patterns independent from L1 and L2, such as the patterns of identical tone sequences and the often-seen substitute tone of T3 at word-final positions, are addressed in §6.3 as part of the interlanguage properties. It is found that, even specific constraints which are usually included in some general constraints (i.e., the specific constraint OCP (LH) is usually included in the general constraint OCP (contour)), also experience promotion or demotion separately from the general constraints during the course of SLA. These movements of specific Markedness constraints are evidenced by the error and substitution patterns found in the present study. These cases are argued to present the situation of “the Emergence of the Unmarked” (TETU, McCarthy and Prince, 1994; Broselow et al., 1998), revealing a universal preference for less marked structures in SLA. The chapter ends with a summary of the discussions in §6.4.

6.1. The constraint re-ranking model for second language acquisition (SLA)

6.1.1. The general structure of the conceptual SLA constraint re-ranking model

b. Negative transfer



The re-ranking of Con-NT1 constraints indicates that the L2 learners are trying to acquire a greater range of linguistic structures or a new inventory of linguistic structures. The Con-NT1 constraints are ranked high in L1 which means they place restrictions on the linguistic structures allowed in L1.⁴³ However, during the course of SLA, the demotion of these constraints below the Faithfulness constraints in the L2 means the restrictions of linguistic structures are somewhat weakened or even removed. On the other hand, the re-ranking of Con-NT2 constraints means that learners are trying to delimit the range of possible L2 structures. Con-NT2 is ranked relatively low in L1 but high in L2, meaning learners should place more restrictions on specific linguistic structures, or even avoid producing some L1 structures when they try to speak the L2. These constraint types and the re-ranking process of the constraints are exemplified by the tone-related constraints and corresponding ILG properties in the following subsections §6.1.2 and §6.1.3. Section §6.1.2 introduces constraints related to tone-acquisition. Section §6.1.3 compares the L1, the L2, and ILG grammar rankings.

6.1.2. Constraints used in the analysis of the SLA of Mandarin tones

In OT, the tonal grammars of L1, L2, and the interlanguage can be represented by specific rankings of universal constraints, including Markedness constraints and Faithfulness constraints. Below are the definitions of the constraints used in the analysis of the SLA of Mandarin tones in the present study. Please note that not all constraints used in the discussions in this section will be used in the representative tableaux in §6.2.

⁴³ Either negative Markedness constraints (such as NoContour or *Contour) or positive Markedness constraints (such as COINCIDE specifically restricting some structures to some positions) place restrictions on the occurrence of linguistic structures in the grammar of the language in question.

6.1.2.1. Markedness Constraints

The Tonal Markedness Scale (cf., Hyman & VanBik, 2004) is concerned with the structure of individual tones and penalizes marked forms, such as contour tones, in this study. The TMS constraints include some fixed rankings of constraints (1a), (1b) and (2) attested or discussed in the previous chapters.

(1) Tonal Markedness Scale (TMS)

- a. *T2>>*T4>>*T1: Tone 2 is more disfavored than Tone 4 than Tone 1.
- b. *Contour >>*Level: Contour tones are more disfavored/marked than level tones

As mentioned in the chapter 2, FT3 which contains three tone targets within one TBU is regarded as the tone type with the highest complexity in Mandarin. In terms of level tones, we assume T1, a high level tone, is more marked than T3, a low level tone, following the universal tone complexity scale *High>>*Low (Yip 2002). A complete tonal markedness scale including FT3 and T3 is found in (2).

(2). A complete TMS ranking

*FT3>>*T2>>*T4>>*T1>>*T3: FT3 is more disfavored than T2; T2 is more disfavored than T4; T4 is more disfavored than T1; T1 is more disfavored than T3.

Here we define *FT3 as a general constraint banning more than two component tones associated to a TBU while *T2, *T4, *T1, *T3 are constraints banning rising, falling, high level and low tones.

(3) Other TMS constraints

- a. *FT3: More than two component tones associated to a single TBU are prohibited.
- b. *T2: No rising tones (low and high component tones associated to a TBU).
- c. *T4: No falling tones (high and Low component tones associated to a TBU).
- d. *T1: No high level tones (a high tone associated to a TBU).
- e. *T3: No low level tones (a low tone associated to a TBU).

The Obligatory Contour Principle (OCP) was first proposed by Leben (1973). The OCP (whole tone) and other specific OCP constraints in (5) will be used in this chapter.⁴⁴

⁴⁴ According to Gussenhoven (2004), the least demanding interpretation of the OCP which militates against the occurrence of adjacent, like tones is one whereby like tones cannot occur on adjacent TBUs within a morpheme.

- (4).a. OCP constraints: Adjacent identical autosegments are prohibited. (Leben 1973)
- b. OCP-whole tone (OCP (T)) (Yip, 2002): Identical whole tones cannot occur on adjacent TBUs. A pattern with the same whole tone at the two positions, such as Tone1+Tone1 [+U, hh] [+U, hh], violates this constraint.

(5) Other OCP constraints:

- a. OCP (contour): no two contour tones (such as T2-T2, T4-T4, FT3-FT3) at adjacent TBUs.
- b. OCP (H): no two High tones (such as T1-T1) at adjacent TBUs.
- c. OCP (L): no two Low tones (such as T3-T3) at adjacent TBUs.
- d. OCP (LH): no two rising tones (such as T2-T2) at adjacent TBUs.

As a supplementary finding to Zhang, J (2004)'s claim that "phrase-final syllables are the preferred bearers of contours", this SLA study found that rising tones are disfavored at word-final positions and falling tones are disfavored at word-initial positions in L2 tonal productions. The supplementary finding can be presented as:

(6) Positional constraints of whole tones

- a. *Rise-F: No rising tones (T2 in this study) are allowed at word-final positions. This constraint can also be represented as *T2-F or *[+U, lh]-Final.
- b. *Fall-I: No falling tones (T4 in this study) are allowed at word-initial positions. This constraint can also be represented as *T4-I or *[+U, hl]-Initial.

Because Faithfulness constraints usually do not evaluate prosodic structure, analyses of phonological phrasing or prosodic structures are generally based on rankings of conflicting Markedness constraints (McCarthy 2008; Becker and Potts 2011). The postulation of intonational language grammar in the next section is mainly based on the observation of phrasal level tone patterns. The phrasal tone constraints include COINCIDE (Zoll 2004) ensuring that every phrase have phrasal tones and Alignment constraints requiring tones at specific positions. COINCIDE is a positional Markedness constraint that dictates that linguistic structures should belong to some constituents (Zoll 2004). The present study proposes that phrases should include phrasal tones. The family of Alignment constraints here is an

abbreviation of constraints concerning the association of tone register features and positions. Since the relationship between the phrasal tones and foot structures are complicated, here we only postulate some simplified Alignment constraints and use their effects to explain the L1 transfer. Different from a fixed ranking like (1) above, the family of alignment constraints is flexible and should be able to accommodate specific L1 rankings regarding tone alignment.

(7). Phrasal Tone Constraints⁴⁵

- a. COINCIDE (phrase, tone): For every phrase, there is a phrasal tone.
- b. Align (Pr, L, [+/-U], L) : For every prosodic unit there must be some tone with register feature of [+U] or [-U] such that the left edge of the tone matches the left edge of the prosodic unit.⁴⁶
- c. Align (Pr, R, [+/-U], R) : For every prosodic unit there must be some tone with register feature of [+U] or [-U] such that the right edge of the tone matches the right edge of the prosodic unit.

6.1.2.2. Faithfulness Constraints

Faithfulness Constraints enforce close correspondence between the input and the output.

(8) Faithfulness Constraints

- a. Ident-whole tone (Id-T): let α be a TBU in the input, and β be any TBU corresponding to α in the output; if α has tone T, then β has tone T.
- b. Max-whole tone (Max): Maximize all input whole tones in the output. (Do not delete).

6.1.2.3. The OT treatment of tones and intonations

OT treatments of Mandarin tones and tone sandhi are given in Yip (2002) and other works, such as Lee (2002) and Yin (2003). Intonational structures sometimes display features that are not obviously

⁴⁵ Gussenhoven (2004) uses Alignment constraints to address intonation properties. Basically, Alignment constraints focus on the alignment of tones with specific positions but assume the existence of phrasal tones. That is, if there is no phrasal tone occurred in phrases, Alignment constraints are irrelevant. This study thus include the COINCIDE constraints to ensure the existence of phrasal tones.

⁴⁶ Here I take the general notion of “prosodic unit” as the domain for the sake of simplicity. A full discussion on the relationship among foot structure, stress/pitch accent, and tones requires more references.

found in segmental structures or in lexical tonal structures. Previous works such as Gussenhoven (2004) bring intonational data to bear on OT by making modifications to the concept of some general constraints such as tonal alignment. This section discusses the tonal grammar of L1s and L2 in a conceptual way.

We assume that in L2, Mandarin, Faithfulness constraints dominate most Markedness constraints, such as the TMS, the OCP, and TPC constraints attested in the previous chapters. Very few Markedness constraints, such as the OCP (L), dominate Faithfulness constraints, because two low tones (T3) in a row are not allowed in Mandarin. I summarize the general Mandarin tonal grammar ranking in (9). For the sake of economy, I simplify the group of Faithfulness constraints as FAITH hereafter.

(9) General ranking of Mandarin tonal grammar

OCP (L)>>FAITH >> TMS, OCP, TPC

In the case of intonation languages, the phrasal tone constraints play an important role in making tones appear in surface forms. Since the surface forms of intonation are basically determined by Markedness constraints on phrasal tones, they stand at the top in native intonation language grammar rankings and dominate Faithfulness constraints. At the same time, some Markedness constraints, such as some TMS constraints and OCP (contour), also dominate Faithfulness constraints.⁴⁷ However, other Markedness constraints, such as OCP (H) and OCP (L) are dominated by Faithfulness constraints. This is because the L1s usually don't have two identical contour tones on adjacent TBUs within a word but instead allow two high tones or low tones occurring on adjacent TBUs. I summarize the general ranking of L1 tonal grammars in (10) and simplify the Phrasal tone constraints (such as COINCIDE and Alignment) as PhrTone.

(10) General ranking of L1 intonation grammar

PhrTone >> OCP (contour), TMS, TPC >>FAITH>>OCP (H), OCP (L)

⁴⁷ For example, the ranking *contour in Japanese and Korean dominates FAITH since we assume contour tones are not taken as the basic tone types in Japanese and Korean (see Chapter 2).

In the next section §6.1.3, we explore the operation of the SLA constraint re-ranking model by examining the characteristics of correspondent tone-acquisition-related constraints and comparing the general constraint rankings of ILG with those of L1 and L2.

6.1.3. The promotion and demotion of Markedness constraints

The experiment results found in Chapter 4 and 5 show that English, Japanese, and Korean speakers choose some tone sequences not faithful to the target forms nor can be derived from L1 prosodic structures. This indicates that the learners have developed their ILGs bearing some features that differ from both the native-language and the target-language grammars. This section compares the general ILGs to the L1 or L2 grammars within the idealized SLA constraint re-ranking model. The English, Japanese, and Korean speakers' ILGs share some similarities in the re-ranking behaviors (see the partial grammar rankings of three L1s in § 6.2). Here we take some representative Markedness constraints in the case of Korean speakers as an example to exemplify the characteristics of the four types of constraints. At the same time, we correlate the effects of constraint re-rankings with the actual error and substitution patterns found in Chapter 4 and 5. The data demonstrates that some Markedness constraints are promoted while some are demoted relative to FAITH in the constraint re-ranking process. Some points of the arguments included in this section will be evidenced in §6.2 and §6.3.

6.1.3.1. The postulation of L1, L2, and interlanguage grammar rankings

The rankings of native L1 or L2 grammars here, or the initial state and target state of SLA, are roughly inferred based on the observation of specific prosodic structures, either allowed or not allowed, in the native or target languages. The rough rankings are provided in the next sub-section § 6.1.3.2, item (11). The native L1 intonation grammars are derived mainly from the phrasal phonology. For example, in L1 (Korean as an example) grammar, the basic tone types are H(igh)s and L(ow)s (Jun 1996). We assume no contour tones (two tone targets borne by a single TBU) are allowed. Therefore, *Contour is ranked higher than FAITH while *Level is ranked relatively low. Korean allows H-H and L-L sequences, but

does not allow two contour tones in a row. Therefore the OCP (H) and the OCP (L) are ranked relatively low while the OCP (contour) is ranked high. The Korean phrase level pitch patterns usually end with high tones, so Align (Pr, R, [+U], R) is also ranked high.

In the L2 (Mandarin) grammar, all tones types are allowed except FT3, at least in the test materials of Experiment 1.⁴⁸ Therefore *Level and *Contour are ranked low while *FT3 is ranked high. All identical tone sequences are allowed except for T3-T3, so the OCP (L) is ranked higher than FAITH while other specific OCP constraints are all ranked lower than FAITH. Since Mandarin allows all types of tones at both word-initial and word-final positions, all Alignment constraints regarding the tone register features and positions are ranked low, such as Align (Pr, L, [+/-U], L) and Align (Pr, R, [+/-U], R), which indicates a very rich tone-position combination inventory.

At the very least, ILGs include the rankings for true tonal productions (when the L2 tonal productions are same to the target tone sequences) and for mis-productions (when the L2 tonal productions are different from the target tone sequences). The ILGs of true productions are very similar to or the same as the target language (Mandarin) tonal grammar in which Faithfulness constraints, in general, are promoted to a very high ranking. The rankings in (11b) below roughly describe the mis-productions made by Korean speakers. It generally conforms to the actual partial ILG ranking worked out in section §6.2. In non-native tone productions, word-final T2 is usually not realized in the non-native tone productions for Experiment 1 (see §6.2 for a summary). Experiment 2 also shows a preference of low tone at prosodic word-final positions. Align (Pr, R, [+U], R) is thus ranked relatively low in (11b). The ranking of the OCP (contour) constraint is a little more complicated since variables are found in the Korean speakers' non-native tone productions. Many fewer contour tone pairs were found than expected, and there were fewer contour tone pairs than high level tone pairs in the ILG. Therefore we rank the OCP (contour) relatively high in the ILG ranking at this point. The behavior of OCP (contour) will be

⁴⁸ In Experiment 1, if we follow the traditional view and take FT3 as the default form, the FT3 at word-initial positions are either changed into T5 due to the Pre-T3 Sandhi rule, or, changed into T3 due to Pre-Other Sandhi rule. The word-final FT3 is not realized because of the sandhi rules occur across word/phrase boundaries.

discussed in detail based on the close examination of contour tone pairs in ILG in §6.3. The Korean speakers produced all individual tone types and many T1-T1 sequences (as shown in §4.3.2). Therefore the constraints *contour, *Level, and OCP (H) are ranked relatively low. There are some FT3 in the L2 tonal productions, but not all underlying FT3 were realized in the surface form. As reported in §4.3, very few T3-T3 sequences were found and some FT3 were realized in order to satisfy OCP (L) (see §6.2). Thus we propose *FT3 is dominated by OCP [L], but ranked higher than Id-T (or FAITH).

6.1.3.2. Four types of Markedness constraints and the effects of constraint re-ranking

In the rough grammar rankings below, the four types of constraints (Con-PT1, Con-PT2, Con-NT1, and Con-NT2) are labeled under the corresponding tonal acquisition constraints as well as. The three grammars (L1, ILG, and L2) are composed of the same set of constraints, but these constraints follow different rankings. Please note we assume the general L1 and L2 grammar rankings are those which are summarized in (9) and (10). Since this section focuses more on the re-ranking of specific tone-acquisition related constraints such as the OCP, the TMS, and positional constraints, we do not include many PhrTone constraints in the diagram below.

(11) Example constraints and re-ranking in SLA of Mandarin tones

a. The L1 (Korean) native grammar:

*FT3, OCP(contour), Align (Pr.R.[+U].R), *Contour >> Faith >> *Level, OCP(H), OCP(L),
 Con-PT1 Con-NT1 Con-NT1 Con-NT1 Con-PT2 Con-PT2 Con-NT2

b. ILG: Korean speakers' mis-productions:

OCP(L) >> *FT3... OCP(contour) >> Faith >> *Contour, Align (Pr.R.[+U].R), *Level, OCP(H)
 Con-NT2 Con-PT1 Con-NT1 Con-NT1 Con-NT1 Con-PT2 Con-PT2

c. The L2 (Mandarin) grammar:

OCP(L), *FT3 >> Faith >> *Level, *contour, OCP(H), OCP(contour), Align (Pr.R.[+U].R),
 Con-NT2 Con-PT1 Con-PT2 Con-NT1 Con-PT2 Con-NT1 Con-NT1

In the rankings of (11), all four types of constraints are exemplified by tone-acquisition constraints. The effects of the re-ranking of these constraints are discussed below, addressing the L1 transfer effects, phonological universal effects, and SLA pedagogical problems.

The constraint *FT3 in the Korean speakers' case can be taken as a Con-PT1. This constraint is ranked high in the L1 grammar, the initial state of SLA. We know this because Korean usually does not allow the association of single TBU with three component tones.⁴⁹ *FT3 is also at the top of L2 grammar because the test word in Experiment 1 is followed by a neutral-toned particle, and the word-initial and word-final syllables in the test words not allow FT3. Therefore the *FT3 case is an example of positive transfer. Mandarin learners find it easy to acquire the ranking of *FT3. However, FT3 appears in the non-native productions as a substitute for the target T3 very often, and it is argued in §4.2 that this is caused by the "FT3-First" teaching method. Under the current teaching method, FT3 is taken as the underlying or default form of the Tone 3 by L2 learners. This leads to an unnecessary re-ranking of *FT3. In ILG rankings, *FT3 may be out-ranked by other constraints, such as OCP [L], allowing learners produce some *FT3 in some cases (this point will be further detailed in the next section §6.2.2). The English speakers begin with a similar ranking as the Korean speakers, so *FT3 is ranked relatively low in the ILG. Japanese speakers perform Tone 3 much better than English speakers, and it is thought that *FT3 is ranked very high (as attested in §6.2) in the Japanese speakers' interlanguage grammars.

The constraints *Level and OCP (H) are Con-PT2 since these constraints are ranked very low both in L1 and L2. Both Korean and Mandarin allow individual high and low level tones, and also tone sequences of H-H (or T1-T1). The consistent low rankings of *Level and OCP (H) in the L1, the L2, and also in the ILG are reflected not only in the error rates but also in the substitution patterns. Very low error rates of individual level tones were found in Korean speakers' data set (and also in other speakers' data sets) and a very high proportion of level tones were used as substitute tones (see §4.2). OCP (H) is the

⁴⁹. Jun (1996) also suggests that in Seoul Korean "there seems to be an undershoot of the phonetic value of tones when the tones are sufficiently crowded compared to the number of TBUs".

only OCP constraint ranked low initially in the L1 and stayed low in the ILG. For the Korean speakers, very low error rates were found for the target T1-T1 and a higher proportion of T1-T1 sequences than expected were found in L2 tonal productions (see §4.3). This pattern suggests an L1 transfer of the low ranking of the OCP (H) which leads to a quick acquisition.

The constraints “Align (Pr, R, [+U], R)”, “OCP (contour)” and *Contour are Con-NT1 constraints in for the Korean speakers. This type of constraint is ranked high in the L1, but should be demoted to approximate the L2 grammar. The most typical pitch patterns over disyllabic words and the Accentual Phrases in Korean often end with an H tone, a high registered tone, at word or phrase-final positions. This means Align (Pr, R, [+U], R) is ranked high in the Korean grammar. During the SLA of Mandarin tones, when the input tone sequences have high registered tone at the word-final positions such as T1-T4, T3-T1, T3-T4 (but not those ended with T2 when this Alignment constraint interacts with *T2-R), they are easy for Korean speakers to produce. This can be seen in Table 6.3, which shows that the accuracy rates of target tone sequences ending with T1 and T4, especially T1-T1, T1-T4, T3-T1, and T3-T4 are all above 60%. Meanwhile the accuracy rates of L2 input tone sequences ending with low registered tones are much lower (between 40% and 58%). The target language Mandarin allows any tone combinations, and can end with either high or low registers. Therefore all Alignment constraints are ranked very low in the Mandarin grammar ranking. So for L2 learners, the re-ranking of Con-NT1 is difficult. In this situation, the learners are trying to acquire a greater range of linguistic structures (such as the family of Alignment constraints and OCP constraints in this study) or to acquire a new inventory of linguistic structures. OCP (contour) is ranked higher than OCP (H) in the ranking of L1 Korean grammar, although both of them are ranked at the bottom of L2 ranking. The OCP (contour)’s initial high ranking influence the L2 tonal productions. This is reflected in the significantly higher error rates (see chart 6. 3) and lower substitution rate of T2-T2 and T4-T4 than T1-T1 (see Fig. 4.6).

OCP (L) can be taken as a Con-NT2 constraint. OCP (L) is ranked very low in L1 which means Korean allows tone sequence of L-L which usually takes place near phrase and utterance boundaries. This

constraint is moved up to the top of the ILG grammar approximating the L2 grammar. That is, the learners learned the high ranking of OCP (L) very quickly, although they don't always know what tone sequences they should change into. This may be due to the universal OCP constraint disfavoring tone pairs, or to the quick acquisition of the L2 grammar of Pre-T3 Sandhi. In substitutions, there are very few Low-Low sequences, confirming that Korean learners learned the T3 sandhi rule quickly. This is an NT2 case in which learners are trying to avoid producing Low-Low sequences which are not allowed in the L2. It seems that the re-ranking of NT2 is less difficult for learners than the re-ranking of NT1 based on the data, especially when the L1 transfer effect is compatible to phonological universals.

6.1.3.3. A discussion on constraint re-ranking in Second Language Acquisition

Above is a general discussion of the four types of Markedness constraints and the effects of their re-ranking in the SLA of Mandarin tones. As mentioned in Chapter 2 concerning the constraint re-ranking during language acquisition, Tesar and Smolensky (2000) assume that only Markedness constraints are demoted in the course of L1 acquisition while Boersma and Levelt (2000) assume that both Markedness constraint demotion and Faithfulness promotion. In L2 syllable structure studies, Hancin-Bhatt (2008) proposes that in the initial state of L2 acquisition, L1 constraint hierarchy is instantiated onto L2 with a simultaneous demotion of Faithfulness. According to the "Full transfer" hypothesis (Schwartz & Sprouse 1996) we expect a full instantiation of the L1 constraint ranking into the initial state of the ILGs. The study and discussion above demonstrates that some Markedness constraints, such as OCP (contour) and *Contour, are demoted while some Markedness constraints, such as OCP (L) are promoted relative to FAITH during the course of SLA. The promotion of OCP (L) and demotion of *Contour are evidenced by the partial ranking of Korean speakers which will be displayed in next section §6.2. The demotion of OCP (contour) will be attested by the analysis of tone pair error and substitution patterns in §6.3.

Because this study deals with prosodic structures which are usually not evaluated by Faithfulness constraints, the re-ranking of FAITH does not receive much attention in this dissertation. Hancin-Bhatt

(1998, 2008) proposes that FAITH constraints are gradually promoted to a higher rank during SLA because learners are trying to acquire a greater range of linguistic structures in her study, while Hayes (1999) dealing with English speakers' learning Japanese syllable structures suggests that the re-ranking of FAITH indicates an initial high ranking but a gradual demotion. This does not conflict with the proposal of the promotion and demotion of Markedness constraints in present study. The movements of Markedness constraints and Faithfulness constraints are always in a relative relationship. For example, the case that learners are acquiring a greater range of linguistic structures can be manifested by the demotion of Markedness constraints (to a lower position than FAITH) as discussed in this study, or, the promotion of Faithfulness constraints (to a higher position than some specific Markedness constraints) as discussed in Hancin-Bhatt (1998, 2008). The case that learners delimit the range of possible L2 structure can be understood as the promotion of Markedness constraints as discussed in this study, or the demotion of Faithfulness as discovered in Hayes (1999). The proposal in the present study provides a wider view of the constraint re-ranking process.

To summarize, this section proposes a conceptual constraint re-ranking model and identifies four types of Markedness constraints exemplified by tone-related constraints used in this study. By comparing the general L1, L2 and ILG rankings roughly postulated according to the properties of these languages, the research data demonstrates that some Markedness constraints are promoted while some are demoted depending on the specific L1 and L2 grammar rankings. In the next section § 6.2, we look into some common error patterns in the non-native tonal productions sorted from the current research data and work out partial ILG rankings for English, Japanese, and Korean speakers. The actual ILG rankings, although only a very small portion of the complete ILG grammar due to the limited amount of data, verify some points we discussed above.

6.2. The interlanguage tonal grammars

Unlike Chapters 4 and 5 which looked into the concrete error and substitution patterns for specific tests, this section examines the general respondent disyllabic tone types (including true production types and mis-production types as in Experiment 1) to suggest partial ILG rankings for English, Japanese, and Korean learners. Section §6.1.1 displays three summary charts of respondent tone patterns for English, Japanese, and Korean learners. In §6.1.2, partial rankings of ILGs are determined based on the examination of tone combinations whose most frequently produced respondent tone sequences are not identical to the targets. The discussion indicates that not only L1 transfer, but also phonological universals play a role in shaping the ILG tonal grammars. By comparing the partial ILG rankings with L1 and L2 rankings postulated in §6.1.2, we see evidence of Markedness constraints demotion or promotion at this intermediate stage of SLA.

6.2.1. The summary charts of response tone combination patterns

The three summary charts below display the respondent tone sequences made by three groups of learners for the target 16 tone combinations in stimuli. The first row contains the target tone combinations. The second row contains the top three or four most frequently produced respondent tone sequences for specific target tone sequences. The numbers within parentheses are percentages of respondent tone sequence for that target tone sequence. For example, “T1-T1(56%)” under target “T1-T1” in the first column of English speakers’ chart means the target T1-T1 is produced correctly to T1-T1 and it takes 56% of the all respondent tone sequence occurred. The “T1-T3 (10%)” below T1-T1(56%) in the same column means the second most often produced tone sequence for target T1-T1 is T1-T3 and it takes 10% of the all respondent tone sequences. Correct or true productions (respondent tone sequences identical to target tone sequences) are highlighted in bold.

Table 6.1. English speakers’ summary chart of disyllabic respondent tones

Target Tone Combinations	T1-T1	T1-T2	T1-T3	T1-T4	T2-T1	T2-T2	T2-T3	T2-T4
Response Tone sequences	T1-T1(56%) T1-T3(10%) T1-T4(6%) T2-T3(5%)	T1-T3(40%) T1-T2(18%) T3-T1(9%) T1-T1(8%)	T1-T3(51%) T1-FT3(18%) T1-T1(10%) T1-T4(6%)	T1-T4(43%) T1-T1(15%) T1-T3(11%) T3-T4(9%)	T2-T1(25%) T3-T1(19%) T1-T1(13%) T2-T3(10%)	T2-T3(36%) T2-T2(13%) T2-T4(10%)	T2-T3(50%) T2-T4(14%) T2-FT3(9%)	T2-T4(43%) T2-T3(18%) T3-T4(18%)

Target Tone Combinations	T3-T1	T3-T2	T5-T3	T3-T4	T4-T1	T4-T2	T4-T3	T4-T4
Response Tone sequences	T3-T1(50%) FT3-T3(10%) T1-T1(8%)	FT3-T3(24%) T3-T2(18%) T3-T4(10%)	T5-T3(46%) T5-FT3(9%) T3-T1(8%)	T3-T4(54%) T2-T4(9%) T2-T3(8%) FT3-T4(8%)	T4-T1(38%) T1-T1(18%) T3-T1(16%)	T4-T2(21%) T4-T3(16%) T1-T3(14%)	T4-T3(38%) T4-FT3 (11%) T1-T3(9%) T2-T4(9%) T3-FT3(9%)	T4-T4(38%) T3-T4(19%) T1-T4(11%) T2-T4(11%)

Table 6.2. Japanese speakers' summary chart of disyllabic respondent tones

Target Tone Combinations	T1-T1	T1-T2	T1-T3	T1-T4	T2-T1	T2-T2	T2-T3	T2-T4
Response Tone sequences	T1-T1(68%) T3-T1(13%) T1-T4(9%)	T1-T3(48%) T1-T2(20%) T1-T4(9%)	T1-T3(58%) T1-FT3(10%) T1-T1(10%)	T1-T4(68%) T1-T1(10%) T3-T4(10%)	T2-T1(53%) T1-T1(15%) T3-T1(15%)	T2-T3(49%) T1-T3(8%) T2-FT3(8%)	T2-T3(54%) T2-FT3(13%) T3-T1(6%)	T2-T4(61%) T3-T4(20%) T1-T4(6%)

Target Tone Combinations	T3-T1	T3-T2	T5-T3	T3-T4	T4-T1	T4-T2	T4-T3	T4-T4
Respond Tone sequences	T3-T1(75%) T1-T1(13%)	T2-T3 (23%) T3-T2(18%) T3-T1(10%)	T5-T3(81%) T1-T3(6%) T5-T4(4%)	T3-T4(75%) T2-T3(8%) T1-T4(6%)	T4-T1(46%) T1-T1(25%) T2-T1(6%) T3-T1(6%)	T4-T3(43%) T1-T3(29%) T4-T2(11%)	T4-T3(61%) T1-T3(18%) T4-FT3(9%)	T4-T4(53%) T3-T4(9%) T1-T3(6%) T1-T4(6%)

Table 6.3. Korean speakers' summary chart of disyllabic respondent tones

Target Tone Combinations	T1-T1	T1-T2	T1-T3	T1-T4	T2-T1	T2-T2	T2-T3	T2-T4
Respond Tone sequences	T1-T1(64%) T1-T4(15%) T3-T4(6%)	T1-T3(41%) T1-T2(21%) T4-T3(11%)	T1-T3(46%) T1-T1(16%) T1-FT3(13%)	T1-T4(65%) T1-T1(16%) T1-T3(6%)	T2-T1(33%) T3-T4(23%) T3-T1(20%)	T2-T3(48%) T2-T2(15%) T2-T1(8%)	T2-T3(55%) T2-FT3(15%) T3-T4(8%)	T2-T4(51%) T3-T4(26%) T3-T1(8%)

Target Tone Combinations	T3-T1	T3-T2	T5-T3	T3-T4	T4-T1	T4-T2	T4-T3	T4-T4
Respond Tone sequences	T3-T1(60%) T3-T4(14%) T1-T1(13%)	FT3-T3(18%) T3-T2(18%) T3-T1(16%)	T5-T3(81%) T1-T3(6%) T5-T4(4%)	T3-T4(64%) T1-T4(9%) FT3-T4(8%)	T4-T1(43%) T1-T1(15%) T4-T4(14%)	T4-T3(39%) T1-T3(16%) T4-T2(14%)	T4-T3(58%) T1-T3(11%) T1-T4(9%)	T4-T4(49%) T1-T4(13%) T4-T3(9%)

The multiple respondent L2 productions displayed in these three charts reveal variations in strategies for producing surface tone sequences and show that L2 learners may have different rankings of constraints for ILGs. In these charts, most of the target tone combinations are produced correctly. That is, the respondent tone combinations with highest frequencies (percentages) are identical to the target tone sequences. In OT, this means in most cases Faithfulness constraints were ranked high, which requires speakers be faithful to the input and the ILG rankings closely approximate the ranking of target language grammar. However, for some target tone sequences, such as T1-T2, T2-T2, T3-T2, T4-T2 in the charts,

the most frequently produced respondent tone sequences are forms different from the target tones. We call these cases “mismatched pairs” (highlighted with circles) in this study. It seems in these cases, most L2 learners are unable to be faithful to the input tone sequences and instead create ILG rankings different from the target Mandarin ranking. We are more interested in “mismatched pairs” because these ill-forms tell us about which target tone sequences were difficult for the learners and what strategies of substitution these learners employed to tackle difficult target tone sequences. The repair strategies for “mismatch pairs” may involve the factors of phonological universals, L1 transfer, and SLA pedagogical issues as we discussed in previous chapters.

6.2.2. Representative interlanguage tonal grammars

It is important to note that actual interlanguage grammars are very complicated. Each individual likely employs different kinds of grammars in performing the same target tones, so the subject may produce the target tones correctly at one time but mis-produce the same targets at a different time, or, produce different types of errors for the same targets. In other words, the states of ILG developments (such as the intermediate state of SLA) are dynamic, and constraint rankings are not fixed or steady. Rather, constraints may be ranked and re-ranked under certain conditions or even be unranked. This instability is evidenced in learners’ variable outputs (Hancin-Bhatt 2008). In this section, the tableaux only deal with specific inputs and employ a small amount of the related constraints to illustrate some features of the current interlanguage grammars possessed by three L1 groups. We assume the L2 stimuli or target tone combinations as the inputs and the actual T2 productions are the outputs. §6.2.2.1 discusses Japanese speakers’ ILG ranking while §6.2.2.2 concerns English and Korean speakers’ ILG rankings.

As discussed in §2.2 and §4.2 regarding the Tone 3 pedagogical issues, the research subjects received the same classroom instructions and assume that the standard form of Tone 3 is FT3. Due to this problematic assumption, the underlying forms of Tone 3 in the target tone sequences, or the input in the OT grammar, is taken as FT3 in the following analysis. In the evaluation process, T3 and FT3 will be

observed respectively when the constraints evaluate candidate surface forms in determining the optimal output in the tableaux below.

6.2.2.1. Japanese speakers' mismatched pairs

For Japanese speakers, there are four mismatched pairs. This group of speakers most often produced the surface forms (output) of T1-T3, T2-T3, T2-T3 and T4-T4 for the L2 target tone sequences (input or underlying forms) of T1-T2, T2-T2, T3-T2 and T4-T2. It seems only the target tone sequences ending with T2 are included in mismatched pairs and this indicates that undominated *Rise-F⁵⁰ has a strong effect and dominates Faithfulness constraints, such as Id-T, Max. Tableau (12) shows that the faithful candidate T1-T2 violates the higher ranked *Rise-F although it perfectly satisfied all Faithfulness constraints. All the mis-matched pairs are very similar to the example shown here:

(12) Tableau for Japanese speakers' choice of output T1-T3 for input T1-T2

T1-T2	*Rise-F	Id-T
→ T1-T3		*
T1-T2	*!W	L

*Rise-F >> Id-T

Because of *Rise-I, a non-T2 tone should be chosen to substitute the target T2 at word-final position. During this process, it seems the TMS is functioning in speakers' ILGs. The TMS picks the most unmarked T3 as the substitute tone. In tableau (13), we examine all candidates beginning with T1 but varying in the identity of the word-final tone. The faithful candidate is eliminated out because of *Rise-F (or *T2). Other candidates are also failed to be chosen as the final winner because of the TMS. From Tableau (13), we can conclude the ranking of *Rise-F, *FT3, *T2, *T4, *T1 >> *T3

(13). Tableau for Japanese speakers' choice of output T1-T3 for input T1-T2

⁵⁰ The reason why *Rise-F but not Align [-U]-F is used to describe the fact that the final-rises are all substituted by low tones T3 is because none of tone sequences ending with other high registered tones, such as T1, T4 are mismatched pairs.

T1-T2	*Rise-F	*FT3	*T2	*T4	*T1	*T3
→ T1-T3					*	*
T1-T1					**!W	L
T1-T2	*!W		*W		*	L
T1-T4				*!W	*	L
T1-FT3		*!W			*	L

Examining the mis-matched pairs and the most frequently produced tone sequences for other target tone combinations for Japanese speakers, *FT3 seems undominated and no FT3 tones have surfaced for underlying FT3. Based on Tableau (14), we found *FT3 dominates Faithfulness constraints Id-T. The surface form violates Id-T in order to satisfy the higher ranked *FT3.

(14). Tableau for Japanese speakers' choice of T2-T3 for input FT3-T2

FT3-T2	*FT3	Id-T
→T2-T3		**
FT3-T3	*!W	L

*FT3>>Id-T

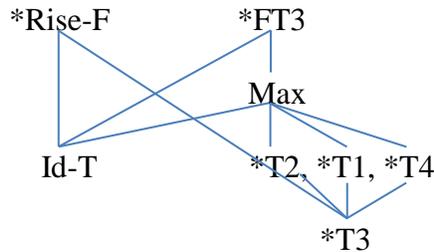
For the target tone sequence of FT3-T2, the output is T2-T3. It seems considering the TMS as the reason for the final pick of T2-T3 as in (13) analysis is insufficient because T2 possesses a much higher tonal complexity than other tones, such as T1, T4 and T3 and there is no reason for T2 to surface as the most optimal form. The tableau below only includes some candidate tone sequences which all end with T3. Comparing the target form and actual productions, it seems the Faithfulness constraint Max is working in speakers' ILGs and dominates *T2. In tableau (15), the winner candidate T2-T3 performs better than other candidates in terms of parsing underlying tones in the surface form. We label the input T2 with an index "A" indicating that the corresponding outputs, such as candidate (a) "T2_A-T3", preserves the input T2_A although it is attached to a different TBU. As a result, candidate (a) only violates Max once because the input FT3 is deleted and a new T3 is inserted. Candidates (b) and (c) are eliminated because the T2_A does not surface in the output. Candidate (e) has the same phonetic tone values to the winner candidate (a). However, it is eliminated because the output T2 is not the same T2 in the input. Candidates (f) is better in being faithful to the input, but it violates the highly ranked *FT3.

(15) Tableau for Japanese speakers' choice of T2-T3 for input FT3-T2

FT3C -T2A	*FT3	Max	Id-T	*T2	*T4	*T1
→ a. T2A -T3		*	**	*		
b. T4B -T3		**!W	**	L	*	
c. T1B -T3		**!W	**	L		*
e. T2B -T3		**!W	**	*		
f. FT3C-T3	*!W	*	*L	L		

To summarize the rankings we worked out above, a Hasse diagram is provided below:

Figure 6. 2.



As the tableaux above illustrate, the Markedness constraints, including the TMS, play an important role in shaping the Japanese speakers' interlanguage grammar. Due to the high ranked constraint of *Rise-F, T2 at word final positions cannot be realized in the L2 tonal productions they are either substituted with a low registered tone, T3, or realized at a word-initial position as in (15). The frequent occurrences of T3 at word-final positions as a substitution may be related to the universal tone Markedness scale in which T3 is taken as the most unmarked tone, or, due to the first language transfer because Japanese speakers' typical pitch accent tone pattern is HL and Align (Pr, R, [-U], R) is ranked high. Even with the strong effects of L1 transfer and L2 grammar, some Markedness constraints reflecting the tonal complexity scale or the TMS cannot be neglected for Japanese speakers. The effects of the ranking *FT3 >> ... >> *T2 ... >> *T3 are clearly represented in the Hasse Diagram above, although some other constraints are ranked between them. Compared with the L1 and L2 grammar rankings in (11a) and (11c), it's clear that some *contour constraints, such as *T2 and *T4, are demoted to a lower ranking than the Faithfulness constraint, Max. This indicates that the demotion of some Markedness constraints is in progress at this intermediate stage of SLA of Mandarin tones.

6.2.2.2. English and Korean speakers' mismatched pairs

English and Korean speakers have very similar substitution patterns for mis-matched pairs in the summary charts 6.1 and 6.3, so the English and Korean speakers' ILG rankings will be discussed together. Please note that this analysis only provides a small part of the English and Korean speakers' ILG ranking and that English and Korean speakers' overall ILGs may be very different from one another. Among the English and Korean speakers' mismatched pairs, the most frequently produced response tone sequences for the target T1-T2, T2-T2 and FT3-T2 are T1-T3, T2-T3 and FT3-T3. In addition, Korean speakers' target T4-T2 is substituted with T4-T3.

For the Korean speakers, *Rise-F seems to be also high ranked and it dominates some Faithfulness constraints, such as Id-T. Tableau (16) shows that for the target T1-T2, the surface form of T1-T3 was chosen because it violates the lower ranked constraints Id-T but satisfies the highly ranked *Rise-F.

(16). Tableau for English and Korean speakers' choice of T1-T3 for input T1-T2

T1-T2	*Rise-F	Id-T
→ T1-T3		*
T1-T2	*!W	L

When the Faithfulness constraint Id-T is violated, the choice of T3 as substitute for the T2 at word-final position is very likely due to TMS. Unlike the Japanese speakers where the choice of T3 is most likely due to the L1 transfer of Japanese pitch accent pattern, L1 transfer cannot also be attributed to Korean speakers because disyllabic words or APs in native Korean usually end with high tones. Tableau (17) displays more information of the choice of T1-T3 as the surface form for the underlying form of T1-T2 and concludes the ranking of *Rise-F, *FT3, *T2, *T4, *T1 >> *T3.

(17). Tableau for Korean speakers' choice of T1-T3 for input T1-T2

T1-T2	*Rise-F	*FT3	*T2	*T4	*T1	*T3
→ T1-T3					*	*
T1-T1					**!W	L
T1-T2	*!W		*W		*	L
T1-T4				*!W	*	L
T1-FT3		*!W			*	L

Unlik the Japanese speakers' data, many FT3 are found in Korean and English speakers' tonal productions. Among the mis-matched pairs, the underlying form FT3+T2 is surfaced as FT3+T3 which means the underlying FT3 (the standard form of Tone 3 as the learners assume) is survived in L2 tonal productions. However it's not clear if the surface form of FT3 survives because of a higher ranked Faithfulness constraint or for some other reason, since the mis-matched data is limited. Looking into the rest of the Korean speakers' data, we see that the underlying FT3 tones have all surfaced as T3. For example, the surface form of input FT3-T4 is T3-T4 which means Id-T is actually dominated by *FT3.

(18). Tableau for Korean and English speakers' choice of T3-T4 for input FT3-T4

FT3-T4	*FT3	Id-T
→T3-T4		*
FT3-T4	*W	L

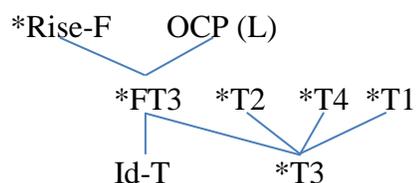
Looking back to the mis-matched pair of input FT3-T2 and the output of FT3-T3, there may be other constraints, for example, OCP [L], standing high in the rank making FT3 win out. I put the two pairs into a combined tableau (19) where the general ranking takes care of the both pairs, with one surface form being faithful to the underlying FT3 and one not.

(19). Tableaux for ranking *Rise-F, OCP (L)>>*FT3>>Id-T

(input) FT3-T2	*Rise-Final	OCP [L]	*FT3	Id-T
→ FT3-T3			*	*
FT3-T2	*!W		*	L
T3-T3		*!W	L	*
(input) FT3-T4				
→ T3-T4				*
FT3-T4			*!W	
T3-T3		*!W		**

The Hasse diagram for Korean speakers' ILG ranking based on the discussion above is:

Figure 6.3.



English speakers' data differs from the Korean speakers' only in the case of T4-T2, where the faithful candidate is selected as the output form. This is probably because there is another constraint dominating *Rise-F in English speakers' grammar, or, the *Rise-F is not violated in this case.

Within the Hasse diagram 6-3, *FT3 is a Con-PT1 constraint and is supposed to still be undominated as in L1 and L2. As discussed in §4.2 and §6.1.3.2, *FT3 may experience unnecessary re-ranking process due to the problematic teaching method of Tone 3 (assuming FT3 as the standard form of Tone 3). *FT3 is dominated by *Rise and OCP (L) as we predicted in §6.1.3.2. The constraints *T2 and *T4 (or, *Contour) are Con-NT1 type constraints since they are ranked high in L1 but low in L2. Although it is hard to see the relationship between *Contour and Faithfulness constraints (such as Id-T), the Hasse diagram of ILG ranking clearly shows that the internal ranking of TMS, *FT3, *T2, *T4, *T1 >>*T3 (or *Contour>>*Level and *High>>*Low) is working in shaping the ILG. OCP (L) is a Con-NT2 typed constraint since it is ranked lower in L1 but higher than FAITH in L2. At this stage of SLA, these learners already acquired the ranking of OCP (L) and it is ranked top as discussed in §6.1.3.2.

6.2.3. Summary of partial Interlanguage rankings

Examining the ILG rankings of Korean and English speakers, some FAITH constraints such as Id-T are violated and ranked low. When FAITH cannot enforce the L2 learners to conform to the target forms, some Markedness constraints, such as the TMS and the OCP (L), emerge and play a decisive role in determining the subject's final choice of surface forms. For example, when FAITH constraints are violated, more than one candidate emerges as the winner, such as the T1-T3 and T1-T4 sequences in tableau (17) and the FT3-T3 and T3-T3 sequences in tableau (19). To decide which candidates are the final outputs, the Markedness constraint TMS reflecting universal preference of simple structured tones and OCP (L) indicating the universal phonological constraints of two identical tone sequences, play decisive roles in picking T1-T3 and FT3-T3 as the final outputs. This evaluation process is similar to the cases we discussed for Japanese speakers' data. It tells us that when learners cannot resort to their L1

grammar for producing some new L2 forms and they have more than one candidate form, learners usually pick those more unmarked forms. The Markedness Differential Hypothesis of Eckman (1977) predicts that the less marked form should be easier for learners than the more marked forms, but it was not clear how the markedness considerations shape an interlanguage grammar. The above analysis within the OT framework displays the operation of Markedness constraints in determining the ILG rankings.

In working out the partial ILG rankings, we also found some evidence of the in-progress promotion or demotion of some tone-related constraints. The Japanese speakers' ILG ranking suggests that *Contour (a CON-NT1 type constraint) is in the process of demotion. *Contour is ranked relatively higher than FAITH in native Japanese grammar but lower than FAITH in L2. At this intermediate stage, the ranking of *FT3...>>FAITH>>*T2, *T4, *T1>>*T3 demonstrates that some sub-constraints of *contour are demoted to lower positions than some FAITH although other constraints interacted with them. The Korean and English speakers' partial ILG ranking suggests that OCP (L) (a Con-NT2 type constraint) is promoted to a higher ranking than FAITH. OCP (L) is ranked low in L1 but high in L2. The ILG ranking indicates that the acquisition of high ranking of OCP (L) is very quick because it is compatible to the phonological universal principle of OCP.

To summarize, for most of the target tone sequences in three summary charts of (6-1,2,3), the three types of learners have similar substitution patterns. These universally preferred structures emerging in the changing state of the grammar being constructed indicate that these L2 speakers have developed their interlanguage grammars that differ from both the L1 and the L2 grammars. That means the three ILGs share some repair strategies and employ some phonological universals in building up ILGs although the learners have different L1 prosodic structures.

In the next section §6.3, we will look into more prominent error and substitution patterns which are independent from both L1 and L2 and represents the situation of the “Emergence of the Unmarked”

(TETU). The analysis of the two TETU cases provides more evidence of Markedness constraints, especially some specific Markedness constraints movements during SLA.

6.3. The “Emergence of the Unmarked” cases in the SLA of tones

The section §6.3 discusses two cases of “The Emergence of the Unmarked”, or cases of L2 learners’ access to UG but not through L1s. As mentioned in the beginning of this chapter, the previous two chapters report many prominent error and substitution patterns which are not obviously motivated by either the L1 nor the L2 grammars. For example, contour tone pairs (rising-rising and falling-falling tone patterns) are not allowed (equally bad) in all L1s and equally good in L2. However, we noticed that in the tonal production of both experiments, there always are many more T1-T1 (level-level sequences) than T4-T4 (falling-falling sequences), and more T4-T4 than T2-T2 (rising-rising sequences) across three types of learners’ data sets. Also, in both Experiment1 and Experiment2, we found a universally preferred High-Low registered tone sequences in the L2 tonal productions. The Japanese and English speakers’ case may be due to L1 transfer, but it is unusual for Korean speakers since native Korean usually has a high tone at word or phrase-final positions. We examine the error and substitution patterns closely in this section and argue that these cases represent the situation of “the Emergence of the Unmarked” (TETU, McCarthy and Prince 1994; Broselow et al. 1998) as part of the ILG properties.

Two cases of TETU are discussed in this section. By observing the tone pair patterns in ILG, one TETU case is claimed as emergence of the interacting effects of the TMS and the OCP, or, a strengthened effect of the TMS by a process of Local Conjunction (Smolensky 1993; Ito and Mester 1996; Alderete 1997) in interlanguage grammars. The other case is about the L2 learners’ access to the TMS involving TPC. The analysis of the two cases is based on the concrete error and substitution patterns we found in the previous chapters. The data demonstrates that some subtypes of general Markedness constraints experience promotion or demotion during the course of SLA separately from the general constraints. The

findings also provide evidence for the operation of Markedness promotion or demotion within the constraint re-ranking model proposed in this study.

This section includes three sub-sections. § 6.3.1 provides a brief discussion about Markedness effects and the notion of “The Emergence of the Unmarked” in the studies of SLA. Two TETU cases are explored in §6.3.2 and §6.3.3.

6.3.1. About Markedness and “The Emergence of the Unmarked”

This section discusses the Markedness effects in the study of SLA. Phonological universals are a major factor in L2 phonology and an important corollary of universals is Markedness. Markedness concerns universal preferences in language for certain forms or features and it is widely accepted that unmarked forms (simpler, more basic and more natural forms) have a wider distribution, both within a given language and across languages (Eckman 2008, among others). Markedness effects have long been recognized as playing a role in second language phonology including the work of Eckman (e.g., Eckman 1977, 1991, 2008) and a few of studies in an OT framework (e.g., Broselow, Chen, & Wang 1998; Hancin-Bhatt 1997; Hayes 1999). In his Markedness Differential Hypothesis (MDH), Eckman (1977) codified the observation that L2 learners are more likely to acquire unmarked structures more easily than marked ones. Eckman’s Structural Conformity Hypothesis (SCH) incorporates the notion of typological Markedness and refines his MDH which sounds that are difficult to acquire in the L2 are difficult not simply due to being different from sounds in the L1, but by being different and more marked. Eckman (1991) addresses the important constructions of universals in SCH and even postulates that interlanguages are natural languages and governed by the universals that all natural languages are governed by. Eckman (2008) calls for more SLA phonological studies within OT because “to date, the only phonological theory..... to explicitly and intrinsically incorporate Markedness is Optimality Theory”. Some of the recent studies within OT framework examine the effect of Markedness constraints, as well as transfer, on learners’ acquisition of constraint rankings in the L2. The majority of SLA studies within OT focus on L2 syllable structures (e.g., Hancin-Bhatt and Bhatt 1997, 2000; Broselow et al. 1998; Hayes 1999).

Broselow, Chen and Wang (1998) introduce the notion of “The Emergence of the Unmarked” to the study of second language syllable structures. The term “The Emergence of the Unmarked” (TETU), originally coined by McCarthy and Prince (1994), refers to situations where some marked structure is generally allowed in a language, but banned in particular contexts; the complementary unmarked structure thus “emerges” in those contexts. Broselow et al. (1998) looked at SLA research data involving the simplification of English syllable codas (both voiced and voiceless obstruents are allowed) by native speakers of Mandarin Chinese whose syllable structures are much simpler than those of English (no obstruents are allowed in coda). The study identifies two effects (bisyllabicity and devoicing) that are not obviously motivated by either the native- or the target-language grammars but emerge in the interlanguage grammar and claims this type of Markedness effects that are visible in second-language acquisition represent the situation of TETU. The study revealed variation in strategies for producing surface forms without stop codas, including epenthesis and deletion, with subjects showing a clear preference for bisyllabic over monosyllabic or trisyllabic forms. Broselow et al. (1998) argues that the constraint of WD BIN (constraints requiring that each lexical word minimally contain a foot and that feet are optimally binary) plays an important role in shaping the ILG although WD BIN is ranked below Faithfulness constraint in the L1(Mandarin) grammar and its effects may be masked by higher ranking constraints. It is also found that, although both epenthesis and deletion are employed to transform English codas to fit the syllable structures of Mandarin, some subjects also employ the strategy of devoicing. The corresponding constraint NO VOICED OBS CODA is not visible in both L2 (because English allows both types of voiced and voiceless coda) and L1 (Mandarin forms are subject to the more general constraint NO OBS CODA). Broselow et al.(1998) thus argues that the L2 learners demote NO OBS CODA first and the NO VOICED OBS CODA plays a role in selecting the surface forms of final voiceless obstruents. The TETU in L2 phonology suggests that L2 learners still can access UG even when some aspects of UG are not required in L1.

The situation of TETU in SLA and L2 learners' access to UG has been posited but is still under investigation in the study of L2 phonology. How Markedness constraints move in the course of SLA, especially the interaction of Markedness constraints and its consequent effects in ILG properties, is not very clear. This section extends the L2 phonology study in OT to the topics of lexical tone acquisition and explores more detailed properties of the tonal interlanguage grammars. This study reveals that TETU is not only found in the study of SLA of syllable structures, two cases of TETU which are similar to the TETU situations discussed in Broselow et al. (1998) studying SLA of syllable structures are found in the research data of SLA of lexical tones. One case of TETU is discovered to be associated to some Con-NT1 constraints in §6.3.2 and the other associated to some constraints of Con-PT2 in §6.3.3.

6.3.2. A TETU case of interacting effects of TMS and OCP

As discussed in Chapter 4 and 5, generally speaking, tone pairs (identical tone combination, ITC) are more disfavored by L2 learners than non-identical tone combinations (NITC) and we conclude that the OCP is working in the L2 tonal productions. In the experiment results, these tone pairs (T1-T1, T4-T4 and T2-T2) are found to be “disfavored” by L2 learners to varying degrees. Some systematic error and substitution patterns of identical tone sequences are found in this study and they appear to be independent of L1s and L2 to some degree. The error rates of T1-T1 (high level tone pairs) are always significantly lower than T4-T4 (falling tone pairs), and those of T4-T4 are lower than T2-T2 (rising tone pairs) across the three sub-datasets (see Table 6.1, 6.2 and 6.3). On the other hand, the substitution rates of T1-T1 are higher than T4-T4, and those of T4-T4 are always higher than T2-T2 (see Fig.4.6). They are copied in Table 6.4 and 6.5.

Table 6.4. Accuracy rates of target tone pairs (cited from summary Table 6.1, 6.2 and 6.3)

Target L1	T1-T1	T4-T4	T2-T2
English speakers	56.25%	37.5%	12.5%
Japanese speakers	67.5%	52.5%	6.25%
Korean speakers	63.75%	48.75%	15%

Table 6.5. The proportions of tone pairs in L2 tonal production (cited from Fig.4.6)

Target \ L1	T1-T1	T4-T4	T2-T2
English speakers	8.125%	3.44%	0.78%
Japanese speakers	9.14%	3.75%	0.39%
Korean speakers	10.47%	4.29%	1.02%

The higher accuracy rate and substitution rate of T1-T1 may be due to L1 transfer because the three L1s allow High-High patterns. However in native English, Japanese, and Korean, both two rising tones or two falling tones on adjacent TBUs are unusual. That is, T4-T4 and T2-T2 are equally “bad” in L1s. The target language Mandarin, on the contrary, allows T4-T4 and T2-T2 so that they are equally “good” in Mandarin. Therefore, there is no reason for L2 learners favor T4-T4 over T2-T2 if we only examine the prosodic structures of L1s and L2. Considering the tone types in the tone pair patterns in question, it seems the TMS, working with the OCP, plays a role here. In this section, I argue that this TETU case associated to Con-NT1 constraints, such as OCP (contour) and OCP (LH), reflects a conjoint effect of the OCP and the TMS, or, a strengthened TMS effect by a process of Local Conjunction.

To account for the Markedness effects residing in the OCP, a group of sub-OCP constraints are proposed in the constraint definition section §6.1.2.1. In addition to the general OCP (T), we also proposed OCP (contour), OCP (L), OCP (H) and OCP (LH). All these OCP constraints, except OCP (L) (Con-NT2), are ranked at the bottom of Mandarin tonal grammar ranking because in Mandarin all tone pairs are allowed but T3 pairs. Here we still take Korean speakers’ case as an example. Native Korean allows H-H tone sequences and the OCP (H) constraint is ranked at the bottom of L1 grammar ranking. Because OCP (H) is also at the bottom of the ranking of L2 grammar, OCP (H) is a Con-PT1 and no re-ranking is required. Therefore, it is not surprising that L2 learners perform T1-T1 very well and the proportion of T1-T1 in L2 tonal production is high. So far we have accounted for what drives the highest frequency of T1-T1, and we are in a position to account for the asymmetry between T4-T4 and T2-T2.

Both LH-LH and HL-HL tone strings are new for L2 learners, i.e., Korean speakers in the example. This means the general OCP constraint and OCP (contour) are ranked high at the initial stage of SLA. The sub-OCP constraint, specific OCP (LH), is actually masked by the general OCP (contour) because rising is a type of contour tones and it is inactive in native L1 ranking. The OCP (LH) is also inactive in the L2 grammar because all contour pairs are allowed in Mandarin and the general OCP (contour) is ranked low. The diagram (20) shows the rankings of the OCP constraints in L1 and L2 grammars and OCP (LH) is dimmed in gray:

(20). L1 and L2 grammars

L1(Korean) or the initial state : OCP (Contour), OCP (LH), >> Faithfulness >> OCP (H), OCP(L)

L2 (Mandarin) or the target state: OCP (L) >> Faithfulness >> OCP (contour), OCP (LH), OCP (H)

The diagram suggests that OCP (LH) and OCP (Contour) are both Con-NT1. HL-HL and LH- LH are assumed to be equally difficult for L2 learners to acquire. However, according to the research data, it seems the L2 learners acquire HL-HL sequence earlier than LH-LH sequence since the error rates of T4-T4 is much lower than T2-T2 and the substitution rate of T4-T4 is much higher than T2-T2. Reflected in the SLA re-ranking, it seems the general OCP (contour) is demoted to a low rank earlier than the specific OCP (LH) and at some point of SLA of Mandarin tones, the ILG for some subjects may have the following ranking where OCP (contour) moved to the bottom but left OCP (LH) behind at the top:

(21) ILG: **OCP (LH)**, OCP (L) >> Faithfulness >> OCP (contour), OCP (H)

This case of “general constraint demoted earlier than specific constraint” is very like the “devoicing” case in the L2 acquisition of syllable structures discussed in Broselow et al. (1998). Following this ranking, the HH and HL-HL sequences are allowed but LH-LH is not allowed for some subjects, thus lead to higher frequency of HL-HL than LH-LH in L2 tonal productions. In this case, the demotion of a general Con-NT1 activated a specific Con-NT1 and some relatively unmarked structures are picked as the surface forms. Alternatively, the effects of the specific Con-NT1 (OCP (LH)) are originally masked by other general Con-NT1 (OCP (contour)) in L1 grammar but they become visible

during the course of SLA. This situation represents the situation of TETU. The diagram (22) below sketches the grammar rankings of L1, ILG and L2 and the consequences in the actual productions by some example input tone pairs and output tone sequences (typical dissimilated tone sequences are cited from summary chart 6. 3). In the diagram, we see the same group of input tone pairs but surfaced with different outputs in L1, IG and L2. Only T1-T1 survived in L1 because OCP (contour) is at the top of the ranking and no contour tone pairs are allowed in the output. In an intermediate stage of the SLA of Mandarin tones, T1-T1 and T4-T4 are in the outputs, because OCP (contour) demoted but OCP (LH) is still at the top of the ranking which means LH-LH is not still allowed in the L2 tonal productions. This leads to the asymmetry of rising tone pair and falling tone pair productions we discussed at the beginning of this section. In the L2 grammar and also the grammar for L2 true productions, all contour tone pairs survived. This means learners who become proficient in Mandarin tones eventually will demote OCP (LH) below the Faithfulness constraints, thus allowing the full range of tone pairs in Mandarin. According to the data, very few subjects consistently made all tone pairs and this is very likely because of the instability of the constraint rankings in the developing interlanguage grammar as illustrated below.

(22) The case of “Emergence of the Unmarked” associated to Con-NT1:

Input	L1 grammar: OCP (contour), OCP (LH) >>Faith>>OCP(H)	Output
T1-T1/ T4-T4/ T2-T2/	➡	T1-T1/ T3-T4/ T2-T3/
Input	ILG grammar OCP (LH) >>Faith>>OCP (contour), OCP (H)	Output
T1-T1/ T4-T4/ T2-T2/	➡	T1-T1/ T4-T4/ T2-T3/
Input	L2 (Mandarin) grammar Faith>>OCP (LH), OCP (contour), OCP (H)	Output
T1-T1/ T4-T4/ T2-T2/	➡	T1-T1/ T4-T4/ T2-T2

The diagram (22) clearly displays that not only the general constraint OCP (contour) is demoted but also that the specific OCP (LH) is in the progress of demotion at this intermediate stage of SLA. Because of the separate demotion process of the general constraint OCP (contour) and the specific OCP (LH), an asymmetry of tone pairs in the non-native tonal productions receives a good explanation.

The asymmetry in tone pairs in three groups of learners also indicates a conjoined effect of OCP and TMS. The OCP constraints, stated as “adjacent identical autosegments are prohibited”, actually says nothing about the markedness of the elements involved. On the other hand, the TMS (*Rising>>*Falling>>*Level) is only concerned with individual tone complexity and says nothing about tone pairs. They are two independent constraints, but it seems they jointly shape some properties of the ILG about tone pairs. That is, the constraints of the OCP family in the ILG rankings above are not discrete constraint, but have an inner link with the TMS because we found two violations of LH are more serious than two violations of HL. This kind of situation can be modeled in terms of “Local Conjunction” of a constraint with itself, or, “self-conjunction” as in Smolensky (1993) (Alderete 1997, Ito and Mester 1996). Following Alderete (1997), OCP effects, broadly understood to encompass segmental processes of dissimilation and restrictions on segment co-occurrence, are the result of Markedness constraints which are strengthened by the operation of Local Conjunction. So, in this context of SLA of Mandarin tone pairs, I propose to deal with the TMS in the co-occurrence restrictions by extending the original TMS ranking to the locally-conjoined tone Markedness constraints, as shown in (23).

(23) (*Rising)² >> (*Falling)² >> (*Level)²

This scale states that the local self-conjunction of rising tones is more disfavored than the local self-conjunction of falling tones, and in turn then the local conjunction of level tones. By doing local conjunction of the tone markedness scale, the greater number of T1-T1 sequences found than T4-T4 sequences, and the greater number of T4-T4 sequences found than T2-T2 sequences, can be explained by one general principle: two rising tones in a row are hardest to produce or most marked, two falling tones in a row second hardest, and two level tones in a row are the easiest. This Markedness effect is not active in both the L1s and in the L2, but it emerges in the SLA data as a case of TETU and shapes the interlanguage tonal grammars.

6.3.3. A TETU case of interacting effects of TMS and TPC

In the section of 6.2.2, we discussed several situations about the roles of Markedness constraints, in shaping the partial ILG rankings such as the “TMS” effects. In the partial ILG rankings, the low T3 tone is often substituted for target tones. In native English and Mandarin, all tone types are allowed and in the prosodic grammar rankings, these TMS family constraints (internal rankings) are dominated by other constraints thereby invisible. Japanese and Korean only allows level tones as basic tone types, therefore *Contour >>*Level is already integrated in their native grammar rankings. However, it is not clear about the internal rankings between the two contour tones of T2 and T4, or the two level tones T1 and T3. As we can see in the tableau (13) for Japanese speakers and tableau (17) (duplicated below) for Korean speakers that the domination of T3, the most unmarked tone type, by other tones, such as T1, play a decisive role in picking the surface form ended with T3. The internal ranking of the TMS about a subtype of TMS constraint, *T3, which is inactive in L1 and L2 becomes active in the interlanguage grammars.

(24) Tableau for Korean speakers’ choice of T1-T3 for input T1-T2

T1-T2	*Rise-F	*FT3	*T2	*T4	*T1	*T3
→ T1-T3					*	*
T1-T1					**!W	L
T1-T2	*!W		*W		*	L
T1-T4				*!W	*	L
T1-FT3		*!W			*	L

For those true-productions made by L2 learners, the Faithfulness constraints such as Id-Tones are always at the top of the rankings and the Faithfulness constraints mask these low ranked Markedness constraints. However, in the tonal grammar rankings of the mis-productions which includes those “mismatched pairs” (as in the Hasse diagram of 6-2 and 6-3), we can see the Faithfulness constraints, such as Id-Tones, are violated and demoted in order to satisfy highly ranked *Rise-F. The TMS family of constraints, which are originally dominated by Faithfulness constraints, then become active and shape the ILG. As discussed in §6.2.2, the patterns reveal a preference for less marked structures. These less marked structures, or, the lower ranked constraints may normally have no visible effects in the native speakers’ grammars, but they are still assumed to be present in the grammar. When higher ranking

constraints, usually some Faithfulness constraints, are violated, these Markedness constraints emerge and play a decisive role in picking the more unmarked forms. This kind of TMS effect is manifested in the error and substitution patterns of individual tones in L2 tonal production across three groups of L1 speakers. In both Experiments 1 and 2, all learners have higher error rates and lower substitution rate of rising tone T2 than falling T4 and in turn than level tone T1 since the rising tone is considered most marked according to TMS. It should be noted that although T3 doesn't have the lowest error rate because of the interference of T3 pedagogical problem, the low level tone T3 is likely the least marked tone due to its highest substitution rate.

This kind of TMS effect sometimes is interacted with TPC based on the examination of the disyllabic error patterns in both Experiments 1 and 2. Here again we take the Korean speakers' data as an example discussing the prevailing High-Low registered tone pattern found in this study. The high frequency of High-Low registered tone sequences found in both mis-matched pairs and Experiment 2 (see discussions in §5.3.2.2) is discovered to be due to the high error rate of T2 at word-final positions. This indicates a strong effect of positional effect of TPC. This study of L2 tonal productions found that word-final position is a good bearer for T4 only, but not for T2 since the T2s at word-final positions has significantly higher error rates than at word-initial positions across all three groups of learners. Due to this positional Markedness constraint, the T2 at word-final position becomes extremely vulnerable and it entails substitute tones for this target T2. At this point, it is only the constraint of TPC pertaining to the rising tone which affects L2 tonal productions. The TMS plays a role to pick the most unmarked tone T3 (from other options, such as T4, T1, FT3) for the word-final position and lead to the dominating High-Low registered tone pattern. For English and Japanese speakers, this choice may be due to the prominent L1 prosodic structure (i.e., English's dominant trochaic foot structure and Japanese's typical pitch accent High-Low pattern). However, for Korean speakers whose native language always has High tone at word or phrase-final positions, the low tone at word-final position is clearly not originated from L1. So, in addition to the possible reason of L2 grammar as discussed in §5.3.2.2, it is very likely that the low

substitute tone for the target T2 at word-final position is an interacted effect of universal unmarkedness of low level tone and the vulnerability of rising at word-final positions.

6.4. Summary of Chapter 6

This chapter summarizes some robust error and substitution tone patterns found in the previous two chapters and conducts theoretical analysis accounting for the ILG properties. This section introduces a conceptual constraint re-ranking model for SLA. Four types of Markedness constraints (Con-PT1, Con-PT2, Con-NT1, Con-NT2) are identified and exemplified by tone-related constraints used in the present study. The research data demonstrates that some Markedness constraints (Con-NT1) are demoted while others (Con-NT2) are promoted during the course of SLA. Many prominent error and substitutions, especially those which cannot be obviously motivated by L1 or by L2, are addressed and explained by constraint re-ranking effects. The operation (promotion or demotion) of Markedness constraints is also evidenced by the partial ILG rankings and the cases of TETU based on the observation of the current research data.

Two cases of TETU in SLA, or, L2 learners' access of UG, are discussed in this section. They demonstrate that some unmarked structures, despite their non-obvious existence in either the L1 or the target language, or the both, can surface in interlanguage tonal grammar development. The case of TETU describing the emergence of the Tonal Markedness Scale conjoined with the Obligatory Contour Principle is found to be associated to Con-NT1 constraints. The activation of a specific OCP constraint is discovered to be due to the earlier demotion of a general OCP constraint which is originally ranked high in the initial state ranking. This study proposes that the interacting effect of the TMS and the OCP can be taken as a strengthened tonal markedness effect by local self-conjunction. The other case of TETU pertaining to the emergence of an internal ranking of the TMS (*contour, *High>>*Low) is discovered to be interacted TPC constraint (*Rise-F) and the interacting effects may lead to the prevailing High-Low registered tone sequences found in results of Experiment 2.

CHAPTER 7: CONCLUSIONS AND IMPLICATIONS

This study has shown that, while a learner's L1 affects the second language acquisition of Mandarin lexical tones, universal phonological principles also constrain non-native tonal productions. This study has also suggested that the current pedagogical method related to the teaching of Tone 3 interferes with the acquisition of Tone 3 and Tone 3 Sandhi. The objective of this chapter is to summarize research findings and describe the pedagogical implications. The first section §7.1 outlines experiment results as they relate to: universal phonological constraints (in §7.1.1), the effects of L1 transfer (in §7.1.2), and the conceptual constraint re-ranking model for SLA (in §7.1.3). Section §7.2 describes pedagogical implications of this study. §7.3 offers brief concluding remarks for this study.

7.1. Summary of research findings

This cross-linguistic study within the theoretical framework of Optimality Theory explores how phonological universals, first language transfer, and L2 pedagogy method jointly shape English, Japanese, and Korean speakers' tonal interlanguages. Hypotheses regarding three phonological universals (i.e., the Tonal Markedness Scale or TMS, the Obligatory Contour Principle or OCP, and Tone-Position Constraints or TPC) and first language transfer of typical pitch patterns are tested in two phonological experiments, with Experiment 1 focusing on word-level non-native tonal productions and Experiment 2 focusing on sentence-level tonal productions.

7.1.1. Three universal phonological constraints in the SLA of Mandarin tones

This study has found that the TMS, the OCP, and TPC are relevant in English, Japanese, and Korean speakers' interlanguages, although to different degrees. This result is obtained from Experiment 1 and confirmed in Experiment 2. Although neither native L1s nor the target L2 show strong effects of

these phonological universals, effects of the TMS, the OCP, and TPC can be seen in interlanguage grammars.

Mandarin Chinese has four basic tone types: T1, T2, T3, and T4. All of these tone types are freely combined in disyllabic words, with the exception of the sequence T3-T3. The three L1s analyzed in this study vary in how they make use of pitch: Japanese and Korean allow high and low level tones as basic tone types; English has a very loose association of tones and TBU, and allows both contour and level tones. Despite these differences, strong effects of the Tonal Markedness Scale *Rise>>*Fall>>*Level (*T2>>*T4>>*T1 for this study) are detected in all three groups of learners in both error rates and substitution rates, at both word-level and sentence-level experiments. The unmarkedness of T3 [21] is supported by the experimental results as T3 is found to be used most often as a substitute tone across all sub-datasets.

Two tests were conducted to detect whether there were any OCP effects in the L2 tonal productions. Test 1 concerning error rates demonstrates possible weak OCP effects, but this is not supported statistically. Test 2 compared the proportions of identical tone combinations, and shows more obvious OCP effects. That is, although the target language allows tone pairs such as T1-T1, T2-T2, and T4-T4, and L1s usually do not allow two identical contour tones on adjacent TBUs, L2 learners disfavor tone pairs and produce much fewer contour tone pairs than expected. In addition, many more T1-T1 sequences were found than T4-T4, and more T4-T4 sequences than T2-T2 in the three interlanguages. The higher occurrence of T1-T1 than the occurrence of contour tone pairs (T2-T2 and T4-T4) is attributed to an interacting effect of L1 transfer and the OCP, in this case with the L1 transfer effect being stronger than the OCP effect. The asymmetry of T4-T4 and T2-T2 productions provides evidence for the emergence of interacting effects of the TMS and the OCP. The error and substitution patterns are argued to be a result of a demotion of a general OCP (contour) constraint occurring earlier than a demotion of a specific OCP (LH) constraint during the course of the acquisition of Mandarin tones. This represents the situation of “the Emergence of the Unmarked” (TETU) (McCarthy and Prince, 1994; Broselow et al.,

1998). This effect can also be modeled as a strengthened effect of the TMS by a process of Local Conjunction (Smolensky 1993; Ito and Mester 1996; Alderete 1997), wherein the Tonal Markedness Scale in the co-occurrence restrictions is dealt with by extending the original TMS ranking to $(*\text{Rising})^2 \gg (*\text{Falling})^2 \gg (*\text{Level})^2$.

Effects of Tone-Position Constraints, such as $*\text{Fall-I} \gg * \text{Fall-F}$ and $*\text{Rise-F} \gg * \text{Fall-F}$ are also found in the L2 tonal productions. For both the word-level and the sentence-level experiments, all types of learners in this study performed the rising tone (T2) most poorly at word-final positions while performing the high falling tone (T4) most poorly at word-initial positions. These findings have been supported by statistical tests. In the sentence-level experiment, it was found that non-native tonal productions are governed by binary foot constraints, and that most disyllabic prosodic words in sentences bear trochee-like tone pattern, i.e., [+U][-U] register tone sequences (i.e., high-low register tone sequences, such as T1-T3, T2-T3, T4-T3), across the three L1 groups. Several possible reasons are discussed to account for the prevailing [+U] [-U] pitch pattern. One possibility is that the strong-weak (trochaic) stress pattern found in Mandarin may influence the interlanguage. Another possibility is that this is an effect of the interaction between the TMS and TPC. That is, a large number of rising tones at word-final positions are disfavored in the interlanguage because of the highly-ranked $*\text{Rise-F}$, and the rising tones (T2) are substituted with the least marked tone, T3, in accordance with the TMS.

The two cases of interacting effects of the TMS with the OCP and the TMS with TPC represent the situation of “the Emergence of the Unmarked” to some degree. The effects are manifested in error and substitution patterns, neither of which is obviously motivated by the L1 grammars or the L2 grammar. While constraints such as $*\text{T2} \gg * \text{T4} \gg * \text{T1}$, $*\text{Rise-F}$, and $(*\text{Rising})^2 \gg (*\text{Falling})^2 \gg (*\text{Level})^2$ are inactive in native language grammars, they become active in interlanguage grammars during the course of SLA of Mandarin tones. In tableaux, we can often see more than one candidate emerge when the L1-like grammar is used to account for new L2 input. In these cases, usually the more unmarked one is picked as the surface form in the interlanguage grammar.

7.1.2. L1 transfer effects

L2 learners often substitute other tones for target tones in production. These substitution patterns provide evidence for L1 transfer.

At the individual tone level, this study found that English speakers use contour tones as substitutes for other tones much more frequently than Japanese and Korean speakers. This difference is attributed to the differing nature of the association of tones and TBUs in the three L1s.

The L1 transfer of focus-expression strategies is closely examined in both Experiment 1 (contrastive focus borne by disyllabic word) and Experiment 2 (narrow focus constituents of varying size). At the disyllabic word level, possible L1 pitch patterns over the test disyllabic words were predicted based on the prosodic structures of each of the three L1s. Strong effects were found of the transfer of typical L1 pitch patterns on contrastive focus borne by disyllabic words. The counterpart Mandarin tone sequences mimicking typical L1 pitch patterns had higher frequencies than expected in the non-native tone productions obtained from the word-level experiment. The tone sequences mimicking L1 pitch patterns with high occurrence rates include English speakers' T1-T3, T2-T3, T3-T1, T2-T4, T3-T4; Japanese speakers' T1-T3, T2-T1, T2-T4, T3-T1, T3-T4 and T4-T3 and Korean speakers' T1-T1, T1-T4, T3-T1, T3-T4.

In addition to some similarities shared by three groups of learners in focus expressions, L1 focus realization strategies are found to be transferred to L2 tonal productions in the sentence-level experiment, Experiment 2. The lexical tones of focused constituents were difficult for L2 learners to perform. However, whether the syllable is focused or not is not completely predictive of the error or success rates of L2 tone productions, although English speakers demonstrated more sensibility to the "focused" features than other speakers. Some positions outside of the focused constituent also had extremely high error rates (for example, pre-focus positions and initial or final positions of prosodic units).

For monosyllabic focus, all learners showed the greatest accuracy with the target tone T4, and the second greatest accuracy with T1 (both are high register tones). However, when errors did occur, T4 and FT3 were the two most popular substitute tones used to express mono-syllabic focus. English (as a “head-prominence” language) speakers showed an obvious preference of substitute tone of T4, a high falling tone, in monosyllabic focus, which reflects the L1 transfer of accented pitch tone H*L. Japanese speakers usually lengthened low tones and often used FT3 to substitute target tones in focus constituents. Korean speakers used more T4 than FT3 in substitutions, but did not demonstrate a strong preference of one over the other.

In longer focused constituents, the syllables located at the end of focused constituents are usually with the highest error rates. English speakers showed the L1 prosodic transfer effect of the ‘strongest prominence’ at the rightmost word of the phrase. However, this did not seem to also be the case for Japanese and Korean speakers. It is thought that this difference is due to the different syntactic structures of L1 and L2.

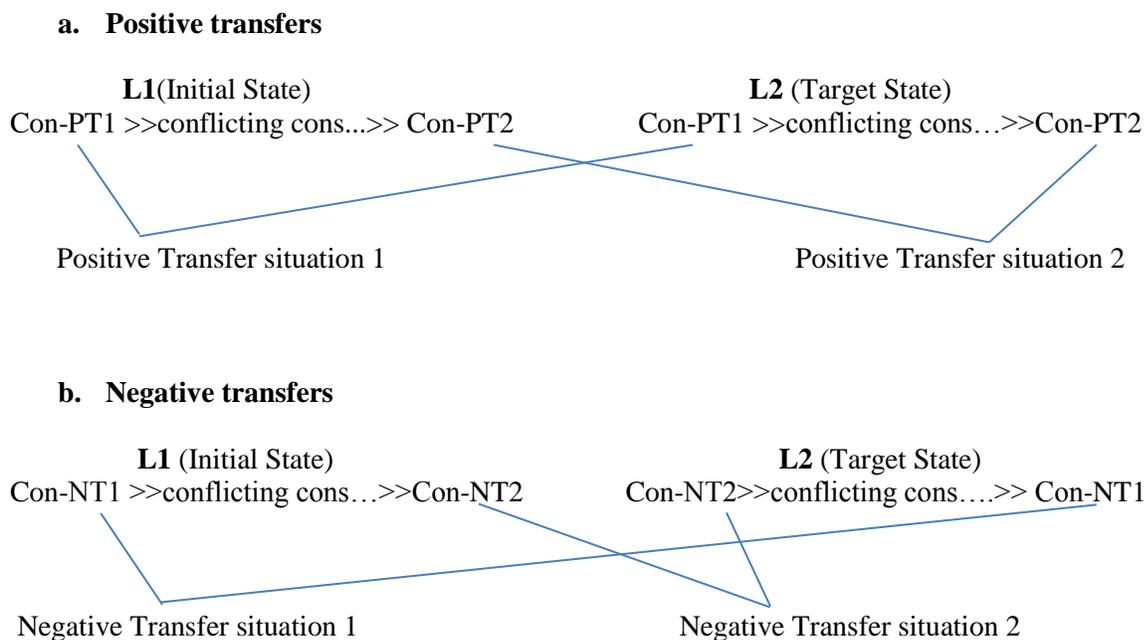
The interactions between universal phonological constraints and the transfer of L1 prosodic structures have also been found. Together these two factors determine which tones will be most frequently produced by L2 learners. The English speakers’ top three most-frequently produced tone sequences were T1-T3, T2-T3, and T3-T4; the Japanese speakers’ most-frequently produced tone sequences were T1-T3, T2-T3, and T3-T1; and the Korean speakers’ were T3-T4, T1-T1, and T1-T4.

7.1.3. The constraint re-ranking model

Based on the analysis of error and substitution patterns, both phonological universals such as the TMS and the OCP, as well as L1 transfer of typical pitch patterns work in interlanguage grammars. The operations of these factors are examined within a conceptual model of Markedness constraint re-ranking for SLA proposed in this study. Within the Optimality Theory framework, Second Language Acquisition is regarded as a process of re-ranking relevant constraints in the L1 grammar to approximate those in the

L2 grammar. It is found that some Markedness constraints are promoted while some are demoted during the course of acquisition of Mandarin tones. The constraint re-ranking model provides a new way of viewing positive and negative transfer.

Figure 7.1. Constraint re-ranking model for SLA



According to Hancin-Bhatt (2008), L2 structures that are similar to or are the same as their counterparts in the L1 can have a generally facilitative effect on learning, whereas L2 structures that are not present in the L1 grammar are a substantial challenge to language learners. Assuming a Full Transfer/ Full Access model (Schwartz & Sprouse 1996), this study proposes that L1 grammar rankings similar to or the same as the L2 structures are transferred to the interlanguage grammar easily and quickly and correspond to what is known as Positive Transfer (PT). If the L1 grammar rankings are very different from the L2's rankings or are not visible (due to very low rankings) in L2 grammar thus requiring much re-ranking, Negative Transfers (NT) occurs. The demotion of Con-NT1 indicates that L2 learners are trying to acquire a greater range of linguistic structures or are trying to acquire a new inventory of

linguistic structures. However the re-ranking of Con-NT2 means L2 learners are trying to delimit the range of possible L2 structures.

Four types of re-ranking situations and corresponding constraint types (Con-PT1/2 and Con-NT1/2) are identified and exemplified by the tone-related constraints used in this study. Constraint re-ranking manifests itself in error and substitution patterns found in the experiment results. In addition, the constraint re-ranking (promotion and demotion of Markedness constraints in relevant to Faithfulness constraints) proposal is supported by the partial interlanguage grammar rankings inferred from the current research data, as shown in Figure 7-2 and 7-3 below.

Figure 7. 2. Japanese speakers' partial interlanguage tonal grammar (i.e., Fig. 6.2)

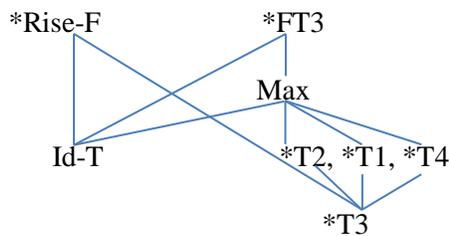
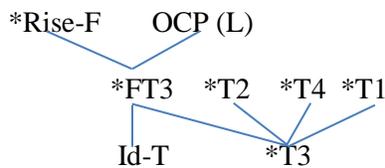


Figure 7. 3. English and Korean speakers' partial interlanguage tonal grammars (i.e., Fig. 6.3)



The partial rankings are worked out based on common error and substitution patterns found in the data sets of the present study. Several Con-NT1 type constraints, such as *contour (that is, *T2, *T4), are found in the process of demotion according to the Japanese speakers' partial interlanguage grammar ranking. The OCP (L), a type of Con-NT2, is promoted to the top of the partial interlanguage grammar rankings in English and Korean speakers' grammars. In the present study, it is found that the acquisition of the ranking of Con-NT2 occurs more quickly than that of Con-NT1. The “Emergence of the Unmarked” (TETU) cases, especially the case reflecting OCP & TMS effects, suggest that some subtypes of

constraints also experience constraint re-ranking separate from the general constraint in the case of Mandarin tone acquisition. For example, as mentioned in §7.1.1, the OCP (LH), which is a subtype of OCP (contour) and a specific Con-NT1 constraint, is demoted earlier than OCP (contour), which is a general Con-NT1 constraint. This leads to a pattern of higher substitution of T1-T1 than T4-T4 and T2-T2 in interlanguages. More studies are suggested to test for the evidence of constraint re-rankings during the course of second language acquisition.

7.2. Pedagogical implications and future directions

This study also attempts to shed some light on tone pedagogy to better facilitate learners of different types of non-tonal languages. Various studies have discussed concrete strategies for tone training including those involving production and perception, such as Bluhme, H., and Burr, R. (1971), Zhao (1988), Wang (1995), Zhao and Cheng (1997), Wang et al (1999), So (2006), Chen and Massaro (2008), Yang (2011), and Holgate (2013). This section does not intend to provide detailed classroom strategies, but instead aims to discuss some theoretical and pedagogical implications of the phonological experiment results in the present study, especially those which have not been detailed discussed in previous studies. Section §7.2.1 outlines the issues regarding the default form of Tone 3 and different teaching methods that have developed from different assumptions. It is argued in this section that FT3 [214] is more like an intonation form of Tone 3 and not the default form, as it has been assumed by traditional teaching methods. The alternative assumption taking T3 [21] as the default form of lexical Tone 3 will dramatically simplify Tone 3 Sandhi rules, thereby relieving L2 learners of unnecessary memorization. Future studies are called upon to revisit the “T3 [21]-First” method. Section §7.2.2 summarizes the most difficult tone types and tone sequences for the three types of learners based on experiment results.

7.2.1. The pedagogical problem of Tone 3

Two main observations can be made regarding subjects’ performance of Tone 3. First, T3 error and substitution patterns are in conflict with one another. T3 [21] is used as a substitute for other tones

more often than any other tone is used as a substitute. Therefore T3 is likely phonetically easy for L2 learners. However, the error rate of target T3 is very high, and it is frequently substituted by FT3 [214] at both word-initial positions (where the Pre-Other-Tone Sandhi should occur to it) and at word-final positions (where the Pre-T3 Sandhi should occur). This was found to be true for English, Japanese, and Korean speakers. Second, in general L2 learners performed the Pre-T3 Sandhi better than the Pre-Other-Tone Sandhi.⁵¹ Native Mandarin speakers on the other hand show the opposite pattern. According to Zhang and Lai (2010), native Mandarin speakers process the Pre-Other-Tone Sandhi, which has a clear phonetic motivation, with a greater accuracy than the Pre-T3 Sandhi, whose phonetic motivation is less clear. It is argued in this study that the conflicting error and substitution patterns of Tone 3 and the poor performance of Pre-Other-Tone Sandhi by L2 learners may be caused by the traditional “FT3-First” teaching method.

Traditionally, FT3 [214] is taken as the default or standard form of Tone 3.⁵² Under this assumption, [35] and [21] are two allotones resulting from the Tone 3 Sandhi rules, with [35] occurring before T3 (Pre-T3 Sandhi) and [21] occurring before other tones (Pre-Other-Tone Sandhi). The “FT3-First” teaching method developed from this assumption. Learners of the “FT3-First” teaching method take FT3 as the default form of Tone 3 and may overuse FT3 in L2 tonal productions.⁵³ This “FT3-First” teaching method leads to complications since two Tone 3 sandhi rules must be memorized and applied in the correct situations. Learners must change the default form [214] into either [35] or [21] at almost every occurrence of Tone 3.

The new findings from the second language acquisition data make us reconsider the default form of Tone 3. People likely take FT3 as the default form because Tone 3 is usually produced as [214] in

⁵¹ It is very clear in Experiment 1 but not very obvious in Experiment 2.

⁵² “Default form” here refers to the form to be used most of the time, when nothing special is going on.

⁵³ For example, the most popular non-native tone patterns for the word *měi guó* 美国 “USA” which should undergo the Pre-Other-Tone Sandhi process (and therefore should surface as [21] [35]) are [214] [21] or [214] [35], or, sometimes surface as [35] [21] in fast speech.

isolation (as well as at the end of an utterance). However, without substantial evidence for the underlying form of FT3, it is not clear that FT3 is the default form⁵⁴. It is widely acknowledged that T3 [21] has a much wider distribution than FT3 [214] in Mandarin, and even at utterance-final positions not all FT3 are realized (Shi and Li, 1997).⁵⁵ Based on the distribution of the allotones of Tone 3 and the evidence obtained from the present study, we propose that FT3 [214] is more like an “intonation form” of Tone 3. FT3 always occurs in isolation or at the end of utterances where phrase or sentence level intonation structures are often placed to the linguistic texts. It is suggested here that the phonetic form of FT3 [214] is the lexical tone T3 [21] with added intonation structure. Future studies are needed to support or reject this proposal.

Some research studies on Mandarin tones assume that T3 [21] is the default form of Tone 3 (such as Yip, 2002; Todo, 1980), but few studies, whether on native Mandarin adults, Mandarin infants, or second language learners, have been done to find evidence for this claim. The findings in Zhang and Lai (2010) that native Mandarin speakers have higher accuracy with the Pre-Other-Tone Sandhi rule is compatible with the proposal that T3 as default form. If T3, rather than FT3, is assumed to be the default form, native Mandarin speakers only have to process one sandhi rule (Pre-T3 Sandhi). If this is the case, there is no “Pre-Other-Tone Sandhi” rule in native Mandarin speakers’ tonal grammar, and no extra processing is required by speakers.

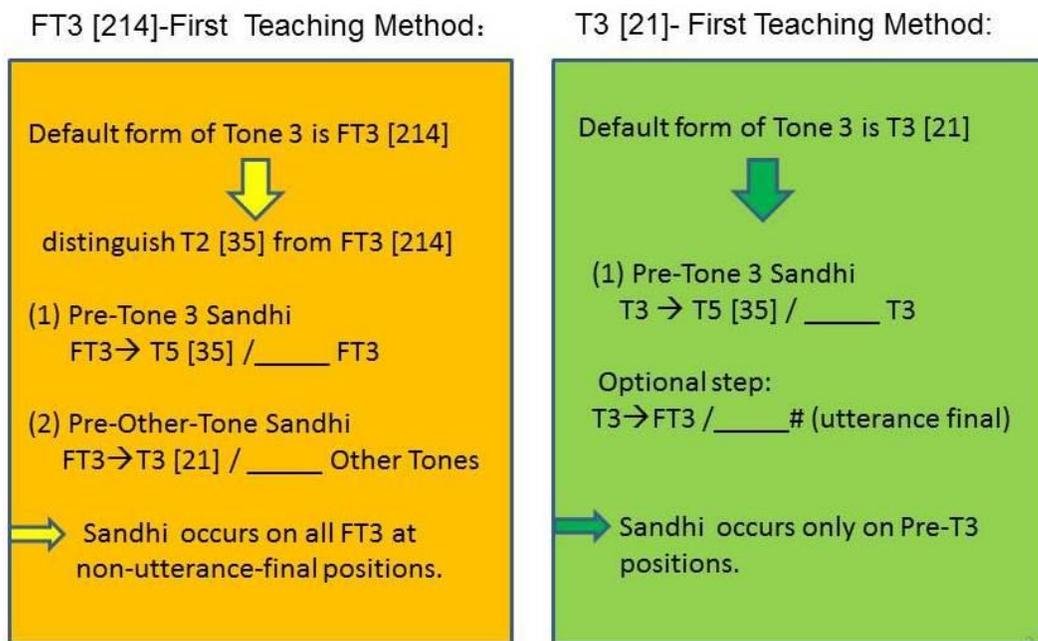
The findings of Tone 3 error and substitution patterns in the present study provide some evidence for the unmarkedness of T3 [21]. Because of the new findings, it is suggested to take T3 [21] as the

⁵⁴ Following the idea of “richness of the base” in OT, if FT3 [214] and T3 [21] are in complementary distribution, then constraints need to be ranked so that either one of them can be in the input and still produce the correct form in the output in all environments. That is, neither form really has the status of “underlying form” separately from the other one. However, it is argued here that it is useful for learners to think of T3 [21] as the default form - the form to use most of the time, when nothing special is going on.

⁵⁵ Shi and Li (1997) argue against the “T3 [21]-first” method advocated here by listing the use of FT3 [214] in native Mandarin speech. However the native Mandarin Chinese data suggests that not all syllables in isolation or at utterance-final positions actually bear FT3.

default form. The issue of the default form of Tone 3 is especially important for the studies of second language acquisition of Mandarin Chinese because different assumptions lead to very different teaching methods and may greatly affect the classroom teaching and learning of Tone 3. Under the assumption of T3 [21] as the default form, the corresponding “T3-First” method leads to only one sandhi process, the Pre-Tone Sandhi rule. The diagram below compares the teaching procedures of “FT3-First” method and “T3-First” method. As can be seen, the “T3-first” method clearly simplifies the process.

Figure 7. 4. Two teaching methods of Tone 3:



The “FT3- First” method must go through at least three steps and learners must memorize and apply two sandhi processes for almost all occurrences of Tone 3. By contrast, the “T3-First” method only requires the memorization and application of the Pre-T3 Sandhi rule, which was found to be applied with great accuracy in this study. Learners would not need to learn the Pre-Other-Tone Sandhi rule, which was found to be difficult for the L2 learners in this study. However, it can only be suggested that further research be done on the effectiveness of the “T3-First” method. Any claim to promote the “T3-First”

method cannot be made here since no empirical studies comparing the actual effectiveness of these two teaching methods have been conducted.

The “T3-First” teaching method has been briefly mentioned in previous studies (Chao, Y., 1948; Lin, T., 1979; Zhao, J-M., 1988; Yue-Hashimoto, A.O., 1986; Wang, 1995; Zhang, 2010). However, most of the previous studies are only concerned with the linear teaching order of T3 [21] and FT3 [214] at the beginning stage, and still take FT3 [214] as the default or base form of Tone 3 in teaching Tone 3 sandhi rules. Even the idea of “teaching T3 [21] first” which has been argued to demonstrate advantages in avoiding the difficult identification task of T2 [35] and FT3 [214], is not accepted by the majority of the researchers and educators in the field of Teaching Chinese as a Foreign Language (CFL). Shi and Li (1997) argue against the idea, claiming that it is unwise to teach T3 [21] first because “Full Tone 3 is the foundation of Tone 3 and is the base form for all allotones and sandhi rules. It will be difficult for learners to study other forms of Tone 3 if students cannot first perform the standard Tone 3, full Tone 3, well” (p.126)⁵⁶. However, very few research studies have been conducted to determine the merit or validity of the “T3-First” method. Since the different assumptions of the identity of the underlying form lead to very different teaching methods, it is hoped that this research will inspire more phonetic, phonological, and psycholinguistic studies to test for evidence of the status of T3 [21] in both Mandarin phonology and language acquisition. More pedagogy-oriented studies are also suggested to revisit the “T3-First” method and compare the actual effects of the two teaching methods.

The author of this dissertation compared these two methods in a small-scale pedagogical study. Preliminary results show that teaching T3 [21] first is not as simple as expected, especially in the beginning, but that it demonstrates some “delayed effects” or “staying power” in a long term study of Mandarin tones. An advantage of T3-First method at the very beginning stage is learners are not necessary to deal with the notorious identification problem of T2 versus FT3, since T2 and FT3 are

⁵⁶ The original statement is “如果我们对语言做一些实际考察之后，就会发现主张只教半三声或先教半三声的想法未免有偏颇之处。我们认为：第三声的本调（我们称之为全三声）是变化的基础，不掌握好本调就很难掌握好它的变调。”

acoustically similar to one another. However, a problem with the T3-First method at the very beginning stage is the distinction of T3 and T4. As we know, T3 is usually not produced in isolation. Teaching or practicing individual T3 [21] without any reference tones next to it, or speaking T3 in isolation, is difficult in practice. It is found in the preliminary study that if learners speak T3 in isolation, some of them may confuse T3 and T4 at the beginning stage. This is because T3 [21] and T4 [51] share some acoustic similarities, as they both can be counted as falling tones. Fortunately, the distinction of T3 and T4 seems much easier to deal with than that of T2 and FT3. In practice, putting T3 in the context of disyllabic or tri-syllabic words in teaching materials is strongly recommended for the T3-First teaching. Another difficulty found in the preliminary study is the confusing orthography of traditional Tone 3 marks in which Tone 3 is signaled with a dipping mark. Despite these difficulties, the T3-First method showed advantages in lexical tone acquisition in this small-scale study after the first several weeks, especially in the performance of connected speech. The greater accuracy of T3 in the connected speech productions found in the preliminary study may be due to the low tonal complexity of T3 [21], as well as the much simpler Sandhi process. Diversified teaching methods are strongly encouraged in the field of Teaching Chinese as a Foreign Language and more empirical studies are suggested to observe and evaluate the two different teaching methods and their effects.

7.2.2. Other pedagogical implications regarding difficult tones in context

This section summarizes which positions were found to be most difficult for each tone, as well as the most difficult tone sequences based on the experiment results. Some specific difficult tones for learners with different L1 background will be addressed first, followed by a summary of the most common difficult tones or tone sequences for all three L1 groups.

This study found many L1 transfer effects of prosodic structures. For example, English speakers found it more difficult to maintain level tones than Japanese and Korean speakers. Since English intonation usually has alternating highs and lows, English speaking learners seem to have opted for

changing tone targets freely. On the other hand, Japanese and Korean speakers have more experience using high and low level tones and seem to have paid more attention to the register feature of tones than the contour features. In their L2 tonal productions, many tone sequences composed of high and low level tones are found, but there are fewer contour tones than there are in the English speakers' data. In addition, Korean speakers found it very difficult to discriminate T1 and T4. It is concluded from this that the training of adhering specific tones to TBUs should be emphasized more for English speaking learners. On the other hand, for Japanese and Korean speakers, more training of switch tone targets is needed.

The three types of learners share some similarities in expressing sentence-level narrow focus, but the learners demonstrate different preferences in employing substitution tones. English speakers usually used T4, a high falling tone, to substitute mono-syllabic focus and often put T4 on the rightmost word in longer focused constituents. Japanese speakers usually lengthened a low tone to express mono-syllabic focus. Korean speakers often used both T4 and T3, not showing a strong preference for either. Sentence-level focused expressions are an important topic in the training of tone performance in connected speech. Based on the training of basic tone types to ensure that learners can make tones correctly and firmly in unfocused constituents, learning how to manipulate pitch range (expanding pitch range for narrow-focused syllables and narrowing pitch range for post-focus constituents) without changing tone identity is an important skill. Expansion of pitch range entails making high tone targets higher and low targets lower, while preserving the original contour shape of tones in question. Increasing the duration and putting more intensity to the focused syllables also helps the focused constituents stand out in tone flows.

Below are summaries of difficult tones at specific positions for all L2 learners. Disyllabic words are believed to be the most frequent type of Mandarin words, and there are a large number of identical tone combinations (such as T2+T2, T4+T4) in modern Chinese. This study finds that these identical tone combinations are especially difficult for L2 learners across all three types of non-tonal language speakers. Both perception and production training is needed for the acquisition of contour tone pairs. Producing two identical contours in a row is a completely new task for these non-tonal language speakers since very few

languages possess this kind of prosodic structure. Being exposed to audio or native speakers' productions of contour pairs is believed to be helpful in familiarizing learners with these new linguistic structures. Some research has even claimed that perception training is more helpful than production training in helping learners to improve L2 Mandarin tones (So, 2006). Audio-visual feedback may also facilitate the acquisition of tone pairs (Chan, 1995; So, 2003, 2006).

With respect to position, it is suggested that intensive training and focus should be placed into T2 at word-final positions, T2 preceding high tones (i.e. a T2-T1 sequence), and T4 at word-initial positions, since L2 learners in this study made many errors of these tones at these positions. At the sentence level, it seems that the realization of tones at prosodic unit boundaries is more difficult than it is at other positions for all types of learners. The four types of most vulnerable positions (the positions with the highest error rates of tones for all L2 learners) are as follows:

- (1) Prosodic word-initial positions for T2 and word-final positions for T4
- (2) The syllables at the beginning of a super foot or at pre-focus positions.
- (3) Major prosodic phrase-final positions.
- (4) Sentence-initial positions (especially for target T1 and T4) and sentence-final positions (especially for target T2).

In the realization of long focus constituents, syllables at phrase boundaries often have higher error rates. The experiment results show that different types of languages, both the 'edge prominent' languages (Japanese and Korean) and the 'head prominent' language (English), demonstrate similar effects of "vulnerable phrase edges" in the L2 tonal productions. Furthermore, it is found in this study that, even within the same focused constituents, word- or phrase-final positions are usually more vulnerable than other positions. Tone training at the sentence level thus should take the above positional information into consideration. To help learners attain more native-like Mandarin tonal performance, instructors not only need to work on the teaching of lexical tones *per se*, but should also pay special attention to the tonal environment. The word-level training of T2 and T4 regarding difficult positions are relatively simple. The

training of tones at the sentence level most likely will require more effort. In addition to various tone training methods proposed in Zhao and Cheng (1997) and other works (Chan, 1995; So, 2006; Chen and Massaro, 2008; Yang, 2011; Holgate, 2013), the Echo Method (Chung, 2013) may also be helpful in the training of sentence-level tone realizations. The Echo Method is a very simple practice method designed to train English language learners in Taiwan and requires no special equipment. The method consists of several basic steps such as listening to a good model speaking short phrases (followed by short sentences, then long sentences),⁵⁷ repeating or “echoing” the utterances silently to oneself, and then saying the utterance out loud. The first two steps will help develop an internal model that learners can use as reference for their pronunciation. This method is argued to fit nicely into current models of working memory in psychology (Contreras 2013). In addition to various concrete training methods, high-variability of training materials are also strongly encouraged to be used in tone training at various prosodic unit boundaries at the sentence level. The training materials may include audio files (for perception) and speaking tasks (for production) designed to have specific tones at different positions (especially at rightmost phrase boundaries) in different prosodic constituents, with varying lengths of utterances, and in both read speech and spontaneous speech contexts.

7.3. Concluding remarks

This dissertation project explores non-tonal language speakers' second language acquisition of Mandarin Chinese tones within the framework of Optimality Theory. A word-level and a sentence-level phonological experiment are conducted to test for evidence of three sources for L2 tonal error: namely, universal phonological constraints (the Tonal Markedness Scale, the Obligatory Contour Principle, and Tone-Position Constraints), the transfer of L1 pitch patterns, and the L2 pedagogical problem of Tone 3. The data shows that these three factors jointly shape the properties of interlanguage grammars.

⁵⁷ This additional modification was made by the author of this dissertation.

This study has limitations. For example, the English-speaking learners and the Japanese and Korean-speaking learners were recruited from different universities, and the time of Mandarin learning experience ranges widely, from 6 months to 18 months. Additionally, only reading tasks were employed to collect the research data. Also, the synchronic analysis of interlanguage tonal grammar shows limited evidence of the promotion or demotion of Markedness constraints relative to Faithfulness constraints. It is believed that a longitudinal study of a similar format to the present study will demonstrate more substantial evidence of the dynamics of interlanguage grammars. Furthermore, this initial research putting L2 tonal error analysis in the framework of Optimality Theory (OT) leaves much to be desired in terms of the incorporation of OT constraints in accounting for the interaction of lexical tones and intonations. It is hoped that this study will inspire more phonetic and phonological studies along these lines.

The value of OT to the field of second language acquisition lies in its account of an interlanguage grammar, restructuring, and the variability of L2 productions observed in the course of second language acquisition. With its rich design, OT, as a generalized model of the grammar, provides us with a set of constructs to model the unconscious knowledge underlying L2 productions with a set of stratified constraints, to explore developing interlanguage grammars and how they are affected by the input. The majority of this kind of studies focuses on L2 syllables. The present study extends the L2 phonology study in OT to the topics of lexical tone acquisition and explores more detailed properties of non-native tonal grammars. The current research data shows “the Emergence of the Unmarked” (TETU) effects outside syllable structure and provides evidence of L2 learners’ access to phonological universals. The appearance of structures in interlanguages that differ from both the L1 and L2 provides potential evidence for universal constraints, as well as for unmarked constraint rankings. Studying L2 sound patterns within the OT framework is still undeveloped and requires more studies to help us explore hypotheses concerning various levels of grammar and fully understand the properties of interlanguages.

The data from this study also shows that L2 pedagogy is relevant in shaping interlanguage grammars. Tone 3 and Tone 3 Sandhi have been a recurrent theme in the study of Chinese phonology. The discussion based on the L2 acquisition data suggests that we reconsider the identity of the default form of Tone 3. FT3 [214] and T3 [21] are more or less in complementary distribution with one another. Based on the L2 research data, it seems at least for L2 learners, learning T3 [21] as the default form of Tone 3 may dramatically simplify the sandhi processes and facilitate acquisition. In our view, FT3 [214] is an intonation form of Tone 3. It is hoped that the discussions developed here motivate further work with the Tone 3 issue, which will not only be important for the field of general Chinese linguistics but also for the field of L2 acquisition of Mandarin Chinese,.

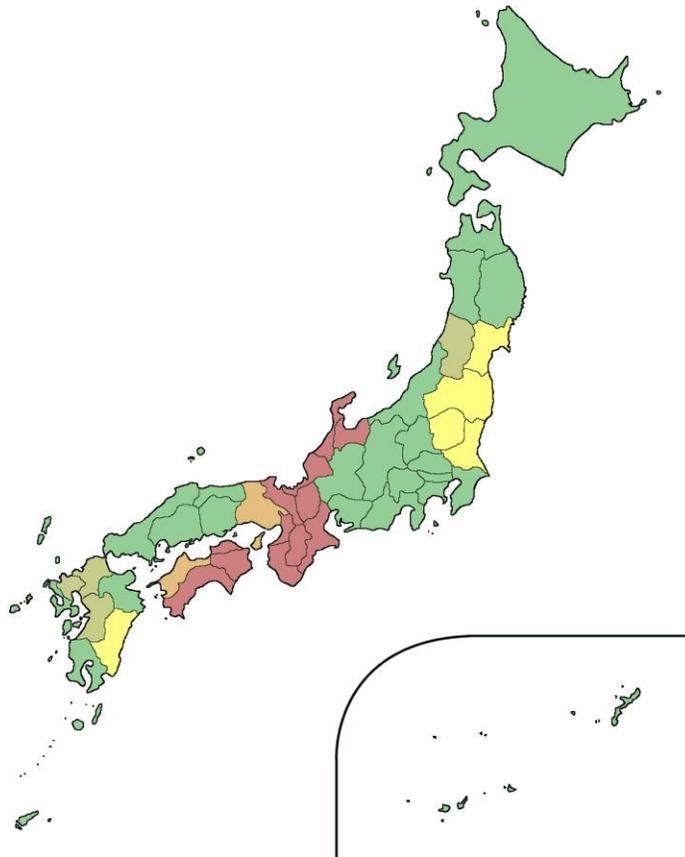
APPENDIX 1

Tokyo-type Accent Map and Recruitment Criteria Regarding Language Background for Japanese Speaking Learners

Speakers from Tokyo and Hokkaido areas are eligible and preferred; other green areas are acceptable. However, speakers from red and yellow areas are not eligible.

日本語が母語で日常生活で以下の県の方言を話す者（緑色で示されている地域）：**東京周辺地域、北海道**、群馬県、埼玉県、神奈川県、千葉県、静岡県、山梨県、長野県、岐阜県、愛知県、青森県、秋田県、岩手県、山形県、新潟県、山口県、島根県、広島県、岡山県、鳥取県、福岡県、大分県

日本語が母語であっても日常生活で以下の県の方言を話す者：（兵庫県、大阪府、和歌山県、奈良県、京都府、滋賀県、三重県、福井県、石川県、富山県、宮城県、福島県、栃木県、茨城県、佐賀県、長崎県、熊本県、宮崎県、沖縄県）。



APPENDIX 2
Sample Language Background Survey (for Japanese Speakers)

Language Background Questionnaire

Dear Participants: please fill out this form in English or Chinese before you make the recordings. You can use the Japanese version form to help you understand the questions.

Name _____ Age _____ Gender : Male Female

Your contact information: telephone # _____ Email: _____

Native/First language(s) _____ Dialect(s) _____

If Japanese, which province/dialect region are you from? _____ (_____ 県)

Do your parents speak dialects? If yes, what dialects: Mother _____ Father _____

Foreign languages and levels: (elementary, intermediate, advanced, native-like)

Foreign language1 (second language) _____ Level _____

Foreign language2 (third language) _____ Level _____

Chinese language study experience:

1. How long have you studied Chinese: _____ year(s) _____ months (average _____ hours/per week)

2. What's the current Chinese class you are in: _____ class at _____ University

3. Have you lived or studied Chinese in a Chinese-speaking country? No. Yes.

If yes, where and how long? _____

4. How do you rate your Chinese skills in....? (None/ Low/ Intermediate/ Advanced/ Native Like)

speaking _____ listening _____ writing _____ reading _____

Thank you very much for your participation!

Investigator's notes: _____

APPENDIX 3

Pre-Test Materials

Pre-test: Please read these characters aloud.

- | | | | | | | | |
|-------------|-------------|------------|-------------|-------------|------------|---------|---------|
| 1. mā | 2. má | 3. mǎ | 4. mà | 5. má | 6. mà | 7. mā | 8. Mǎ |
| 9. wán | 10. wǎn | 11. wàn | 12. wān | 13. wàn | 14. wán | 15. wǎn | 16. wān |
| 17. 书(shu) | 18. 桌(zhuo) | 19. 飞(fei) | 20. 机(ji) | 21. 开(kai) | 22. 学(xue) | | |
| 23. 花(hua) | 24. 茶(cha) | 25. 出(chu) | 26. 口(kou) | 27. 西(xi) | 28. 北(bei) | | |
| 29. 会(hui) | 30. 架(jia) | 31. 期(qi) | 32. 图(tu) | 33. 留(liu) | 34. 足(zu) | | |
| 35. 球(qiu) | 36. 牛(niu) | 37. 奶(nai) | 38. 白(bai) | 39. 水(shui) | 40. 菜(cai) | | |
| 41. 校(xiao) | 42. 手(shou) | 43. 机(ji) | 44. 老(lao) | 45. 家(jia) | 46. 打(da) | | |
| 47. 小(xiao) | 48. 语(yu) | 49. 法(fa) | 50. 果(guo) | 51. 五(wu) | 52. 月(yue) | | |
| 53. 舞(wu) | 54. 汽(qi) | 55. 车(che) | 56. 大(da) | 57. 复(fu) | 58. 习(xi) | | |
| 59. 跳(tiao) | 60. 下(xia) | 61. 雨(yu) | 62. 睡(shui) | 63. 觉(jiao) | 64. 内(nei) | | |

APPENDIX 4

Sample Experiment 1 Reading Materials

EXPERIMENT 1-1 (first page, for English speakers)

During the recording, it's fine if you make pronunciation mistakes. However, you should read the sentences at NORMAL SPEED or a little faster than normal speed. If you have pauses in reading a sentence in the middle of recording, you can read that sentence AGAIN at normal speed, and then continue to read other sentences.

- (1)、我觉得开学的东西很好。Wo jue de kaixue de dongxi hen hao.
I feel that the things used for starting the semester are very good.
- (2)、我觉得西北的东西很不错。Wo jue de xibei de dongxi hen bucuo.
I feel that the things from northwest are very good.
- (3)、我觉得牛奶的东西很好。Wo jue de niunai de dongxi hen hao.
I feel that the things in the milk are very good.
- (4)、我觉得图书的东西很好。Wo jue de tushu de dongxi hen hao.
I feel that the things in the books are very good.
- (5)、我觉得足球的东西很不错。Wo jue de zuqiu de dongxi hen bucuo.
I feel that the things like footballs are very good.
- (6)、我觉得花茶的东西很好。Wo jue de huacha de dongxi hen hao.
I feel that the things for (drinking) Jasmine tea are very good.
- (7)、我觉得开会的东西很好。Wo jue de kaihui de dongxi hen hao.
I feel that the things for the meeting are very good.
- (8)、我觉得白菜的东西很不错。Wo jue de baicai de dongxi hen bucuo.
I feel that the things like Chinese cabbage are very good.
- (9)、我觉得白水的东西很好。Wo jue de baishui de dongxi hen hao.
I feel that the things in the water are very good.
- (10)、我觉得飞机的东西很不错。Wo jue de feiji de dongxi hen bucuo .
I feel that the things like airplanes are very good.
- (11)、我觉得老家的东西很好。Wo jue de laojia de dongxi hen hao.
I feel that the things in my hometown are very good.
- (12)、我觉得出口的东西很好。Wo jue de chukou de dongxi hen hao.
I feel that the things to export are very good.

APPENDIX 5
Sample Experiment 2 Test Materials
(First page materials for English Speakers)

- (1). A: 学习前玩什么的同学能留学? xuéxí qián wán **shénme** de tóng xué néng liúxué?
The classmates that can go study abroad, what do they play before study?
B: 学习前玩球的同学能留学。 xuéxí qián wán **Qiú** de tóng xué néng liúxué.
The classmates who play **BALLS** before study can go study abroad.
- (2). A: 要我给买什么的小姐写语法? yào wǒ gěi mǎi **shénme** de xiǎojiě xiě yǔfǎ?
Which ladies you want me to write the grammar for? What do they buy?
B: 请你给买水果的小姐写语法。 qǐng nǐ gěi mǎi **ShuǐGuǒ** de xiǎojiě xiě yǔfǎ.
Please write the grammars for the ladies who buy **FRUITS**.
- (3). A: 吃什么的医生搬书桌? Chī **shénme** de yīshēng bān shūzhuō?
The doctors that will move desks, what do they eat?
B: 他听说吃鸡的医生搬书桌。 tā tīngshuō chī **Jī** de yīshēng bān shūzhuō.
He heard that the doctors who eat **CHICKEN** will move desks.
- (4). A: 学习前做什么的同学能留学? xuéxí qián **zuò shén me** de tóng xué néng liúxué?
The classmates that can go study abroad, what do they do before study?
B: 学习前玩足球的同学能留学。 xuéxí qián **Wán Zú Qiú** de tóng xué néng liúxué.
The classmates who **PLAY FOOTBALL** before study can go study abroad.
- (5). A: 要我给怎么样的小姐写语法? yào wǒ gěi **zěnmeyàng** de xiǎojiě xiě yǔfǎ?
What kinds of ladies you want me to write the grammars for?
B: 请你给五点买水果的小姐写语法。 qǐng nǐ gěi **Wǔ Diǎn Mǎi Shuǐ Guǒ** de xiǎojiě xiě yǔfǎ.
Please write the grammars for the ladies who **BUY FRUITS AT FIVE O’CLOCK**.
- (6). A: 你计划让看什么的弟妹去睡觉? nǐ jì huà ràng kàn **shén me** de dì mèi qù shuì jiào?
Which brothers and sisters you plan to have (them) go to sleep? What do they watch?
B: 计划让看电视的弟妹去睡觉。 jì huà ràng kàn **Diàn Shì** de dì mèi qù shuì jiào.
(I) plan to have the brothers and sisters who watch **TV** go to sleep.
- (7). A: 怎么样的医生搬书桌? **zěnmeyàng** de yīshēng bān shūzhuō?
What kinds of doctors will move desks?
B: 他听说多喝鸡汤的医生搬书桌。 tā tīngshuō **Duō Hē Jī Tāng** de yīshēng bān shūzhuō.
He heard that the doctors who **DRINK MORE CHICKEN SOUP** will move desks.

REFERENCES

- Alderete, J. (1997). Dissimilation as Local Conjunction. *Proceedings of the North East Linguistics Society* 27: 17-32. Amherst, MA:GLSA.
- Adjemian, C. (1976). On the nature of interlanguage systems. *Language Learning*, 26,(2), 297-320.
- Archangeli, D. (1999) Introducing Optimality Theory. *Annual Review of Anthopology* 28:531-552.
- Arvaniti, A. (2011). The Representation of Intonation. *The Blackwell Companion to Phonology*. Volume II, 757-780
- Bao, Zhiming (1990). On the nature of tone. PhD dissertation, MIT.
- Bao, Zhiming (1999). Tonal Contour and Register Harmony in ChaoZhou. *Linguistic Inquiry* 30.3, 485-493
- Becker, M. and Potts, K. (2011) The Emergence of The Unmarked. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice (eds.) *The Blackwell Companion to Phonology*, pp. 1363–1379
- Beckman, J. (1997). Positional faithfulness. PhD dissertation, University of Massachusetts.Amherst.
- Beckman, M. (1986). *Stress and Non-Stress Accent*. Dordrecht: Foris Publications.
- Beckman, M. and Pierrehumbert, J. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook* 3: 255-309.
- Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (ed.) *Prosodic Typology - The Phonology of Intonation and Phrasing*.
- Bluhme, H and R. Burr (1971) An audio-visual display of pitch for teaching Chinese tone, *Studies in Linguistics* 22:51-57.
- Boersma, P. & C. Levelt. (2000) Gradual constraint-ranking learning algorithm predicts acquisition order. *Proceedings of Child Language Research Forum* 30, Stanford, California, pp. 229-237.
- Bolinger, D. (1951). Intonation: levels versus configurations. *Word* 7. 199-210.
- Bolinger, D. (1964) Round the edge of language: Intonation. *Harvard Educational Review* 34. 282-293. Reprinted in Bolinger (1972), 19-29
- Broselow, E., R. Hurtig, and C. Ringen (1987) "The perception of second language prosody," in G. Ioup and S. Weinberger, eds., *Interlanguage Phonology*: 350-362. Cambridge, MA.
- Büring, D. (2006). Been there, marked that—A tentative theory of second occurrence focus. Manuscript, UCLA.

- Broselow, E., R. Hurtig, and C. Ringen (1987) "The perception of second language prosody," in G. Ioup and S. Weinberger, eds., *Interlanguage Phonology*: 350-362. Cambridge.
- Broselow, E. Chen, S & Wang, C (1998). "The Emergence of the Unmarked in Second Language Phonology." *Studies in Second Language Acquisition*, 20.
- Boersma, P., Weenink, D., 2011. Praat: doing phonetics by computer [Computer program]. Version 5.2.17, retrieved November 2011 from <http://www.praat.org>
- Blicher, D.L., Diehl, R. and Cohen, L. B.(1990). Effects of syllable duration on the perception of the Mandarin Tone 2/ Tone 3 distinction: evidence of auditory enhancement. *Journal of Phonetics* 18, 37-49.
- Bluhme, H., and Burr, R. (1971). An audio-visual display of pitch for teaching Chinese tones, *Linguistics* 22, 51–57.
- Chan, M. K.M. (1995) Students' tone production and audio-visual feedback. Presented at the 1995 Annual Meeting of the Chinese Language Teachers Association. Anaheim, California, 18-20 November, 1995.
- Chao, Y-R. (1930). "A system of tone-letters". *Le maitre phonetique* 45.24-27. (Reprinted in English orthography in *Fangyan* 1980, 2: 81-83.)
- Chao, Y-R. (1948). *Mandarin Primer: an Intensive Course in Spoken Chinese*. Harvard University Press.
- Chao, Y-R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chen, Ao. (2013) *Universal Biases in the Perception of Mandarin tones, from infancy to adulthood*. Netherlands Graduate School of Linguistics/Landelijke-LOT
- Chen, G-T. (1974). "Pitch Range of English and Chinese Speakers." *Journal of Chinese Linguistics*. 2, 2: 159-171.
- Chen, M. (1979). Metrical structure: evidence from Chinese poetry. *Linguistic Inquiry* 10: p.371-420.
- Chen, M. (2000) *Tone sandhi: patterns across Chinese dialects*. (Cambridge Studies in Linguistics 92.) Cambridge: Cambridge University Press.
- Chen, Q. H. (2000). Analysis of Mandarin Tonal Errors in Connected Speech by English-speaking American Adult Learners. Brigham Young University: Dissertation manuscript.
- Chen, Q. (1997). Toward a sequential approach for tonal error analysis. *Journal of the Chinese Language Teachers Association* 32: 21-39.
- Cheng, C.C. (1973). A quantitative study of tone in Chinese. *Journal of Chinese Linguistics*, 1, 93-110.
- Chen, T.H. and Massaro, D.W. (2008) Seeing pitch: Visual information for lexical tones of Mandarin-Chinese. *Journal of the Acoustical Society of America*. 2008. April; 123(4): 2356-2366.
- Chen, Y-Y. (2004). Focus and Intonational Phrase Boundary in Standard Chinese. ISCA Archive. *International Symposium on Chinese Spoken Language Processing (ISCSLP 2004)*.

<http://www/isca-speech.org/archive>

- Chiang, T. (1979). Some interferences of English intonations with Chinese tones. *International Review of Applied Linguistics*, 17 (3), 245-250
- Chung, K.S. (2013). "Karen Chung talks about the Echo Method", podcast in Talking Taiwan website <http://www.talkingtaiwan.com/about/> (accessed on July 4, 2013)
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1975). *Reflections on Language*, New York: Pantheon.
- Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Holland: Foris Publications. Reprint. 7th Edition. Berlin and New York: Mouton de Gruyter, 1993.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (1999). "On the Nature, Use and Acquisition of Language." In T. Bhatia and W. Ritchie (eds.), *Handbook of Child Language Acquisition* (pp.33-54). San Diego: Academic Press.
- Chomsky, N. and H. Morris (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Corder, S. P. (1967). "The significance of learners' errors". *International Review of Applied Linguistics* 5: 161-9.
- Chuang, C. K., Hiki, S., Sone, T., and Nimura, T. (1972). The acoustical features and perceptual cues of the four tones of Standard Colloquial Chinese. *Proceedings of the Seventh International Congress on Acoustics* (Adadémial Kiado, Budapest), 297-300.
- Contreras, D. (2013) The Echo Method and the Teaching of the Four Mandarin Chinese Tones. Senior Honors Thesis. University of North Carolina at Chapel Hill.
- De Lacy, Paul (2002). The interaction of tone and stress in Optimality Theory. *Phonology*. 2002:1-32.
- Delattre, P.C. (1965). *Comparing the Phonetic Features of English, French, German and Spanish*. Heidelberg: Julius Groos Verlag.
- Dreher, J. and Lee, P.C. (1966). Instrumental investigation of single and paired Mandarin tonemes. *Research Communication* 13, Douglas Advanced Research Laboratories.
- Duanmu, San (1990). *A formal study of syllable, tone, stress and domain in Chinese languages*. PhD dissertation, MIT.
- Duanmu, S. (1999). Metrical structure and tone: Evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8.1: 1-38.
- Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning* 27: 315-330.
- Eckman, F. (1991). The Structural Conformity Hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition* 13:23-41
- Eckman, F. R. (2004). "From Phonemic Differences to Constraint Rankings: Research on Second

- Language Phonology.” *Studies in Second Language Acquisition* 26: 513-549.
- Eckman, F. (2008) Typological Markedness and Second Language Phonology, in *Phonology and Second Language Acquisition*. (ed.) Jette G. Hansen Edwards, Mary L. Zampini. (2008) John Benjamins.
- Elliot, C.E. (1991). The relationship between the perception and production of Mandarin tones: An exploratory study. *University of Hawai‘ I Working Papers in ESL*, 10 (2), 177-204
- Eric Holgate (2013) The Effect of Tone Production in Lexical Tone Discrimination Training. Thesis manuscript. University of Florida. www.honors.ufl.edu/apps/Thesis.aspx/Download/2193
- Epstein, S., Flynn, S. & Martohardjono, G. (1996). “Second Language Acquisition: Theoretical and Experimental Issues in Contemporary Research.” *Behavioral and Brain Sciences*, 19: 677-758.
- Feng, S-L. (1998) Prosodic structure and compound words in classical Chinese. In *New approaches to Chinese word formation: morphology, phonology and the lexicon in modern and ancient Chinese*, ed. Packard, 197-260. Berlin: Mouton de Gruyter.
- Feng, S-L. (2001) Minimal Word in Mandarin Chinese. ROA: 507-0302. Roa.rutgers.edu.
- Fery, Caroline and Vieri Samek-Lodovici. (2006). Focus projection and prosodic prominence in nested foci. *Language* 82 (1): 131-150
- Flemming, E. (2008) The role of pitch range in focus marking. Slides from a talk given at the Workshop on Information Structure and Prosody, Studiecentrum Soeterbeeck (2008). <http://web.mit.edu/flemming/www/paper/FlemmingSoeterbeeck.pdf>
- Francis, A., and Ciocca, V. (2003) “Stimulus presentation order and the perception of lexical tones in Cantonese”. *Journal of Acoustical Society of America* 114(3) 1611-1621
- Gandour, J. T. and Harshman, R. A. (1978). “Crosslanguage differences in tone perception: a multidimensional scaling investigation”. *Language and Speech* 21: 1-33.
- Gnanadesikan, A. (1995). *Markedness and Faithfulness Constraints in Child Phonology*. Ms, University of Massachusetts, Amherst. Rutgers Optimality Archive 67. <http://ruccs.rutgers.edu/roa.html>
- Goldsmith, J.A. (1976). “An overview of autosegmental phonology.” *Linguistic Analysis*. 2. 23-68
- Grabe, E. et al. (2003). Quantitative modeling of intonational variation. *Proceedings of Speech Analysis and Recognition in Technology, Linguistics and Medicine 2003*. www.phon.ox.ac.uk/oxygen/publications.php.
- Guo, J-F. (1993). *Hanyu shengdiao yudiao chanyao yu tansuo* [Elucidation and exploration of tone and intonation in Chinese]. Beijing: Beijing Language Institute.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation* (Research Surveys in Linguistics). Cambridge: Cambridge University Press, 2004
- Hancin-Bhatt, B. (2008). “Second Language Phonology in Optimality Theory.” In J. Edwards & J. Zampini (eds.), *Phonology and Second Language Acquisition*. John Benjamins

- Hancin-Bhatt, B., & Bhatt, R. (1997). Optimal L2 syllables: Interaction of transfer and developmental effects. *Studies in Second Language Acquisition* 19: 331-378.
- Hancin-Bhatt, B. (2000). Optimality in second language phonology: Codas in Thai ESL. *Second Language Research* 16 (3): 201-232.
- Hao, Y-C. (2012) Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics* 40 (2012) p. 269-279.
- Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones*, Cambridge University Press, Cambridge.
- Hayes, B. (1989b). The prosodic hierarchy in meter. In Paul Kiparsky and Gilbert Youmans, eds., *Rhythm and Meter*. Orlando: Academic Press, 201-260.
- Hayes, R. (1999). Reranking stages in OT analysis of the acquisition of Japanese as a second language. *Carolina Working Papers in Linguistics* 1.
- Hirschberg, J. et al. (1987a). Intonation and the Intentional Structure of Discourse. Proceedings of IJCAI-87, Milan, August.
- Hirschberg, J. and J. Pierrehumbert. (1987) The Intonational Structuring of Discourse. *Proceedings of ACL-86*, New York, pp. 136–144.
- Hirst, D. & A. Di Cristo. (1998). A survey of intonation systems. In Daniel Hirst & Albert Di Cristo (eds.) *Intonation systems: A survey of twenty languages*, 1-44. Cambridge: Cambridge University Press.
- Holgate, E. (2013). The Effect of Tone Production in Lexical Tone Discrimination Training. Manuscript. University of Florida.
- HSK Department, Chinese government (1992) The handbook of *Outline of the Graded Vocabulary for HSK*, Beijing Language Institute Press <<汉语水平词汇与汉字等级大纲>>, 国家对外汉语教学领导小组办公室汉语水平考试部, 北京语言学院出版社 1992 年
- Hulst, Harry Van der, Bernadet Hendriks and Jeroen van de Weijer (1999). A survey of word prosodic systems of European languages. In: Harry van der Hulst (ed.), *Word prosodic systems in the languages of Europe*, Walter de Gruyter, chap. 7, pp. 425–476.
- Hu, Z.L.(2008) Zuowei waiyu de hanyu jiaoxue [Teaching Chinese as a Foreign Language] in *Foreign Language Education in China*, Vol.1, No.2, May 2008
- Hyman, L.M. (ed.) (1973) Consonant types and tone. *South California Occasional Papers in Linguistics*.1.
- Hyman, L.M. (2011) Tone: Is it Different? in *The Handbook of Phonological Theory*, second edition, (ed.) Goldsmith, J, Riggle, J and Yu, A. Blackwell. Page 197-239.
- Hyman, L. M. & VanBik, K. (2004). “Directional Rule Application and Output Problems in Hakha Lai Tone”. *Language and Linguistics* 5.4:821-861, 2004.
- Itô, J. and Mester, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Dickey, & S.

- Urbanczyk (Eds.), *Papers in Optimality Theory*, University of Massachusetts Occasional Papers, 18, 249-384
- Itô, J. and Mester, A. (1996) Rendaku I: Constraint Conjunction and the OCP. Kobe Phonology Forum 1996. roa.rutgers.edu/files/144-0996/144-0996-ITO-0-0.PDF.gz
- Itô, J. and Mester, A. (2007) "Prosodic adjunction in Japanese compounds". *MIT Working Papers in Linguistics 55: Formal Approaches to Japanese Linguistics 4*. Cambridge, Massachusetts, pp. 97-111.
- Jones, D. (1972). *An outline of English phonetics*. 9th ed. Cambridge: Heffer & Sons.
- Jongman, A., Wang, Y., Moore, C., and Sereno, J. (2006) Perception and Production of Mandarin Chinese Tones. *Handbook of Chinese Psycholinguistics*. E.Bates, L.H.Tan, and O.J. L. Tzeng (eds.) Cambridge University Press
- Jun, S-A. (1990) The Prosodic Structure of Korean - in terms of voicing, in E-J. Baek (ed.) *Proceedings of the Seventh International Conference on Korean Linguistics* . Vol. 7: 87-104, Univ. of Toronto Press.
- Jun, S-A. (1993). The Phonetics and Phonology of Korean Prosody . Ph.D. dissertation. The Ohio State University , Columbus , Ohio .
- Jun, S-A. (1996). *The phonetics and phonology of Korean prosody*. Garland Publishing, Inc. New York & London, 1996
- Jun, S-A & H-J. Lee (1998) Phonetic and Phonological Markers of Contrastive Focus in Korean. in *Proceedings of the 5th International Conference on Spoken Language Processing* , Vol. 4, p. 1295-1298, Sydney, Australia.
- Jun, S-A. (2003) Prosodic Phrasing and Attachment Preferences, *Journal of Psycholinguistic Research* 32(2) pp. 219-249.
- Jun, S-A. (2005). Korean Intonational Phonology and Prosodic Transcription. in Jun, S-A. (eds.) *Prosodic Typology*. pp.201-230. Oxford, UK.
- Jun, S-A. (2011) Prosodic Markings of Complex NP Focus, Syntax, and the Pre-Post-Focus String. In the *Proceedings of the 28th West Coast Conference on Formal Linguistics (WCCFL)* (eds. by Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz), pp. 214-230, Somerville, MA.
- Kanerva, J. (1990) *Focus and Phrasing in Chichewa Phonology*. New York: Garland.
- Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
- Kaltenbacher E. (1997) German speech rhythm in L2 acquisition. Leather J, James A. New Sounds 97: *Proceedings of the Third International Symposium on the Acquisition of Second Language Speech*. Klagenfurt, Austria: University of Klagenfurt, 158-166.
- Kim, J.J. et al. (1997). An analysis of some prosodic aspects of Korean utterances using K-ToBI labeling system, in the *Proceedings of ICSP '97* (Seoul, Korea), 87-92

- Kindaichi, H. (1947). Tookyoogo ni okeru “hana” to “hana” no kubetsu—tookyoo akusento shinnidankan kyoochooron [The distinction between “hana ‘flower’” and “hana ‘nose’” in the Tokyo dialect—the emphasis of the new two-level theory of the Tokyo accent].” *Kokugo*, Summer.
- Kiriloff, C. (1969). ‘On the auditory discrimination of tones in Mandarin, *Phonetica* **20**, 63–67.
- Klatt, D. H. (1975). Vowel lengthening is syntactically determined in connected discourse. *JPh* 3:129–40.
- Ko, E-S. (1999) Interaction of tone and stress in Chonnam and Seoul Korean. *Proceedings of the 18th West Coast Conference on Formal Linguistics*, Cascadilla Press, pp. 246-259.
- Ko, E-S. (2010) “Stress and long vowels in Korean: Chicken or egg, first?” *Japanese Korean Linguistics* 17, CSLI, pp. 377-390.
- Ko, E-S. (2013) “A Metrical Theory of Korean Word Prosody”, *The Linguistic Review*. Volume 30, Issue 1, Pages 79–115.
- Kochanski, G., Shih, C-L and Jing H-Y. (2003) Quantitative measurement and prediction of prosodic strength in Mandarin. *Speech Communication*. V. 41, No. 4, pp. 625-645.
- Kratochvil, P. (1968). *The Chinese Language Today*. London: Hutchinson University Library.
- Kratochvil, P. (1971). An experiment in the perception of Peking dialect tones. A Symposium on Chinese Grammar: Scandinavian Institute of Asian Studies Monograph Series 6, Scandinavian Institute of Asian Studies, (Lund, Sweden), 7-31.
- Kuang, J.J.& Wang, H.J.(2006) T3 sandhi at the boundaries of different prosodic hierarchies, (in Chinese), *Journal of Chinese phonetics*, Volume 1, 125-131.
- Kubozono, H. (1999). Mora and syllable. In *The Handbook of Japanese Linguistics*, ed. N. Tsujimura, 31–61. Oxford: Blackwell
- Kubozono, H. (2011). Japanese Pitch Accent. *The Blackwell Companion to Phonology*. 2879-2907
- Ladd, D. R. (1996). *Intonational phonology*. (Cambridge Studies in Linguistics 79.) Cambridge: Cambridge University Press.
- Ladd, D.R. (1988). Declination “reset” and the hierarchical organization of utterances. *The Journal of the Acoustical Society of America*. 84,2:530-544.
- Ladd, D. R. (2008). *Intonational Phonology*. 2nd ed. Cambridge: Cambridge University Press.
- Lai, C, Sui, Y and Yuan, J (2010). A Corpus Study of the Prosody of Polysyllabic Words in Mandarin Chinese. *Proceedings of Speech Prosody 2010*, Chicago, USA, May 2010.
<http://www/ling.upenn.edu/~laic/laic10poly.pdf> (accessed on March 16, 2013)
- Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather and A. James (eds.): *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second Language Speech*. Amsterdam: University of Amsterdam, pp.305-341.

- Leben, W. (1973). *Suprasegmental Phonology*. MIT: PhD dissertation.
- Lee, K. (2002) Chinese tone sandhi and prosody. A revised version of master thesis, University of Illinois at Urbana-Champaign. ROA file # 549-1002.
<http://roa.rutgers.edu/files/549-1002/549-1002-LEE-0-0.PDF>
- Lee, Vakoch and Wurm (1996) Tone Perception in Cantonese and Mandarin: a cross linguistic comparison. *Journal of Pshycholinguistic Research* Vol.25. No.5, 1996
- Lee, Ok Joo. (2005). The Prosody of Questions in Beijing Mandarin. Ph.D dissertation. The Ohio State University.
- Li, C. and Thompson, S (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language* 4: 185-99
- Lieberman, M. (1975). The intonational system of English. Ph.D. dissertation, MIT.
- Lieberman, M. and A. Prince. (1977). *On stress and linguistic rhythm*. *LI* 8: 249-336
- Lim, B-J. (2001). The production and perception of word-level prosody in Korean. *IULC Working Papers*. Volume 01 (2001). (ed.) Clopper & de Jong, Indiana University.
- Lin, W. (1985). Teaching Mandarin Tones to Adult ENG speakers: Analysis of Difficulties and Suggested Remedies. *RELC Journal*, 16, 2: 31-47.
- Lin, Hua & Qian Wang (2007). Mandarin rhythm: An acoustic study. 普通话节律的声学研究。
Journal of Chinese Linguistics and Computing, 17 (3): 127-140.
- Lin, H-S. (2005) Prosodic Correspondence in Tone Sandhi. *USTWPL* 1: 229-265, 2005
- Lombardi, L. (2003) Second language data and constraints on manner: Explaining substitutions for the English interdentals. *Second Language Research* 19 (3): 225-250
- Lu, J-M. (1992). On the perception and production of Mandarin tones by adult English-speaking learners. *Dissertation Abstracts International* 53(7), 2350A (University Microfilms No. 9233909)
- Lü, S. X. (1963). Xiandai Hanyu dan shuang yinjie wenti chu tan [A preliminary study of the problem of mono- and disyllabic expressions in modern Chinese]. *Zhongguo Yuwen* 1963.1:11-23.
 (Reprinted in 1999: 415-444).
- Major, R.C. (2001). Foreign accent: The ontogeny and phylogeny of second language phonology. Mahwah, NJ: Lawrence Erlbaum Associates.
- Major, R.C. (2008) Transfer in second language phonology: A review. In *Phonology and second language acquisition* (ed.) Jette G. Hansen Edwards, Mary L. Zampini. John Benjamins.
- Miracle, C (1989). Tone Production of American Students of Chinese: A Preliminary Acoustic Study. *Journal of the Chinese Language Teachers' Association*, 24, 3.
- McCarthy, J. (1986) OCP Effects: Gemination and Antigemination, *Linguistic Inquiry* 17: 207-264.

- McCarthy, J. & A. Prince (1993). *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Technical Report #3, Rutgers University Center for Cognitive Science.
- McCarthy, J. & A. Prince (1994) The Emergence of the Unmarked: Optimality in Prosodic Morphology. *Papers from the Annual Meeting of the North East Linuistic Society* 24. 333-379
- McCarthy, J. & A. Prince (1995). "Faithfulness and Reduplicative Identity." In J. Beckman, L. Dickey, and S. Urbanczyk (eds.). *Papers in Optimality Theory*. University of Massachusetts Press.
- McCarthy, J. (2008) *Doing Optimality Theory: applying theory to data*. Blackwell Publishing
- McGinnis, S. (1996). Tonal distinction errors by beginning Chinese language students: A comparative study of American English and Japanese native speakers. In S. McGinnis (Ed.), *Chinese pedagogy: An emerging field* (pp. 81-91). Columbus, OH: Foreign Language Publications.
- Moore, C. and Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of Acoustic Society of America*. 102 (3): 1864-1877
- Nagahara, H. (1994). *Phonological Phrasing in Japanese*. Ph.D. dissertation. UCLA.
- Nemser, W. (1971). "Approximative systems of foreign language learners". *IRAL*, 9, (2), 115-124.
- Nespor, M. & I. Vogel. 1986. *Prosodic Phonology*. Foris Publication Holland: Netherlands.
- Norman, Jerry (1988) *Chinese*. Cambridge University Press: Cambridge, UK.
- Nordenhake, M., and Svantesson, J.-O. (1983). Duration of Standard Chinese word tones in different sentence environments. *Working Papers* 25, (Lund, Sweden), 105-111.
- Ohala, J.J., & Ewan, W.G. (1973) Speed of pitch change. *Journal of the Acoustical Society of America*, 53, 345.
- Ohala, J. J. (1978). Production of Tone. *Tone: A Linguistic Survey*. Ed. By Victoria A. Fromkin. 3-39. New York: Academic Press.
- Ourso, M.. 1989. *Lama Phonology and Morphology*. Dissertation. Champaign: University of Illinois, Urbana Champaign,
- Pierrehumbert, J.B. (1980). *The phonology and phonetics of English intonation*. Ph.D. dissertation, MIT.
- Pierrehumbert, J. B. and Beckman, M. (1988) *Japanese Tone Structure*, Linguistic Inquiry Monograph 15, MIT Press, Cambridge.
- Pierrehumbert, J. and Hirschberg, J. (1990) The Meaning of Intonation in the Interpretation of Discourse, in P. Cohen, J. Morgan, and M. Pollack, (eds). *Intentions in Communication*, MIT Press: Cambridge, MA, 271-311.
- Pike, K. (1948). *Tone Languages*. Ann Arbor: University of Michigan Press.
- Prince, A. & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science.

- Prince, A. (2002) Arguing optimality. In: Angela Carpenter, Andries Coetzee, and Paul de Lacy (eds.) *University of Massachusetts Occasional Papers in Linguistics 26: Papers in Optimality Theory II*. Amherst, MA:GLSA Publications, pp. 269-304. [Available on Rutgers Optimality Archive, ROA-562.]
- Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1994). *The Language Instinct*. London: Penguin.
- Rooth, M. (1996b). On the interface principles for intonational focus. In *Proceedings of SALT VI*, T. Galloway and J. Spence (eds.) Ithaca, NY: CLC, 202-226
- Selkirk, E. (1978). On prosodic structure and its relation to syntactic structure. In T. Fretheim, (ed.), *Nordic Prosody II*. Trondheim: TAPIR, 111-140.
- Selkirk, E. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge: MIT Press.
- Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology* 3: 371-405.
- Selkirk, E. (2002). Contrastive FOCUS vs. presentational focus: Prosodic Evidence from Right Node Raising in English. In *Speech Prosody. 2002: Proceedings of the First International Speech Prosody Conference*, B. Bel and I. Marlin (eds.). Aix-en-Provence: Laboratoire Parol et Language, 643-646
- Selkirk, E. (2006). Bengali intonation revisited: An optimality theoretic analysis in which FOCUS stress prominence drives FOCUS phrasing. In *Topic and Focus: Cross-linguistic Perspectives on meaning and Intonation*, C. Lee, M. Gordon and D. Buring (eds.). Dordrecht: Springer, 215-244.
- Selkirk, E. (2009) On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure, Gengo Kenkyu (*Journal of the Linguistics Society of Japan*), 136, 1-39. Special Issue on Linguistic Interfaces, ed. By Haruo Kubozono)
- Selkirk, E. (2011) The Syntax-Phonology Interface, in J. Goldsmith, J. Riggle, and A. Yu, eds., *The Handbook of Phonological Theory*, 2nd edition. Oxford: Blackwell, 2011.
- Schwartz, B., & Sprouse, R. 1996. L2 cognitive states and the full transfer/full access model. *Second Language Research*. 12:34-66
- Selinker, L. (1972). "Interlanguage". *IRAL*, 10, (3), 209-231.
- Shen, J. 1985. Beijing yuyin shiyan lu[Pitch range of tone and intonation in Beijing dialect]. In. *Working papers in experimental phonetics*, ed. Lin,T. and Wang,L.,73-130. Beijing University Press.
- Shen, X. (1989). "Toward a Register Approach in Teaching Mandarin Tones." *Journal of Chinese Language Teachers Association*, 24: 27-47.
- Shen, X., and Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech* 34, 145-156.

- Shen, X., Lin, M., and Yan, J. (1993). F0 turning point as an F0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *Journal of the Acoustical Society of America* 93, 2241-2243.
- Shi, P-W. and Li, M.(1997) Sansheng wenti yanjiu [A study on Tone 3] In *Yuyin yanjiu yu duiwai hanyu jiaoxue [Phonetic study and Chinese language teaching]* (ed.) by Zhao, J-M and Meng, Z-M. Beijing Language and Culture University.
- Shih, C-L. (1986). *The prosodic domain of tone sandhi in Chinese*. Doctoral dissertation, University of California, San Diego.
- Shih, C-L. (1997) Mandarin third tone sandhi and prosodic structure. *Studies in Chinese Phonology*. Ed. Jialing Wang and Norval Smith. Walter de Gruyter, 1997. 81-123.
- So, C. K. L (2003). Training non-native listeners to acquire Mandarin tones with visual and auditory feedback. Paper presented at the WorldCALL conference 2003, Banff, Alberta, Canada.
- So, C. K. Lai. (2006). *Effects of L1 Prosodic Background and AV training on Learning Mandarin Tones by Speakers of Cantonese, Japanese, and English*. Ph.D. dissertation. Simon Fraser University.
- So, C. K., and Best, C.T. (2010). Cross-language perception of nonnative tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech* 53, p. 273-293.
- Stagray, J. R. and Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics* 21: 143-63.
- Steriade, D. (1993). *Neutralization and the expression of contrast*. Paper presented at NELS 24. University of Massachusetts, Amherst.
- Steriade, D. (1995). Underspecification and markedness. In John Goldsmith (ed.), *The handbook of phonological theory*. Oxford: Blackwell 114-74.
- Sugiyama, Y. (2006). Japanese Pitch Accent: examination of final-accented and unaccented minimal pairs. *Toronto Working Papers in Linguistics*. 26:73-88.
- Sun, S. (1998). *The Development of a Lexical Tone Phonology in American Adult Learners of Standard Mandarin Chinese*. Honolulu: University of Hawai'i Press.
- Sundberg, J. (1979). Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* 7:71-79.
- Tesar, B and Paul Smolensky (2000) *Learnability in Optimality*. MIT Press. 2000
- Todo, Akiyasu (藤堂明保) (1980) *Chinese Phonology* 《中国语音韵论》 光生馆.
- Tseng, C-Y. and Chou, F-C. et al. (1999) A prosodic labeling system for Mandarin speech database. In *Proceedings of the XIV International Congress of Phonetic Science*
- Tseng, C-Y. (2003). Mandarin Speech prosody: Issues, pitfalls and directions. In proceedings of *EuroSpeech 2003*, Geneva.
- Tseng, C-Y. and Pin, S-H. (2005). Fluent speech prosody: Framework and modeling. *Speech*

Communication, 46: 284-309.

- Trafer, G. & H. Smith. (1951) *An outline of English structure*. Norman, OK: Battenburg Press.
- Truckenbrodt, H. (1995) *Phonological phrases: their relation to syntax focus and prominence*. Ph.D. dissertation, MIT.
- Ueyama, M. (1998). *Speech Rate Effects on Phrasing in English and Japanese*. Manuscript. UCLA.
- Venditti, J. (2005). The J_ToBI Model of Japanese Intonation. In Jun, S-A. (eds.) *Prosodic Typology*. Oxford, UK.
- Vihman et al. (1998). Is There a “Trochaic Bias” in Early Word Learning? Evidence from Infant Production in English and French. *Child Development*. August 1998. Volume 69. Number 4, Page 935-949.
- Villafañá, Christina. (2000) *Emergence of the Unmarked in Interlanguage Coda Production*, *Working paper in Linguistics*, Volume 7, George Mason University, http://www.gmu.edu/org/lingclub/WP/texts/7_Villafana.pdf (accessed on May 25, 2013)
- Wang, J. (2003). *Rhythmic Grouping, Tone Sandhi and Stress in Beijing Mandarin*. Ph.D dissertation. Macquarie University.
- Wang, X.C. (2006). Perception of L2 tones: L1 lexical tone experience may not help in *Proceedings of Speech Prosody*, Dresden, Germany.
- Wang, Y.J. (1995). *Ye tan meiguo ren xuexi hanyu shengdiao [On American learners’ tone acquisition]*, *Language Teaching and Research* 2: 126-140
- Wang, Y.J. (1997). *Yangping de xietong fayin yu waiguoren xuexi yangping [The coarticulation of Tone 2 and the acquisition of T2 by foreigners]*, *Language Teaching and Research* Vol 4: 94-104
- Wang, Y., Spence, M., Jongman, A., and Sereno, J. (1999) Training American listeners to perceive Mandarin tones. *1999 Acoustical Society of America*. 106, P.3649–3658
- Wang, Y., Jongman, A, and Sereno, J.A. (2006). L2 acquisition and processing of Mandarin Chinese tones. in P. Li, E. Bates, L.H. Tan, and O. Tseng (eds.), *The Handbook of East Asian Psycholinguistics*, Cambridge University Press, pp.250-256.
- Warner, N., and A. Takayuki. 2001. The role of the mora in the timing of spontaneous Japanese Speech. *Journal of the Acoustical Society of America* 109:1144-1156.
- Wayland, R. and S. Guion. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning* 54:681-712.
- Wheeldon, L. (2000) Generating prosodic structure. In Wheeldon, L. (ed.) *Aspects of Language Production*. Psychology Press. Pages 249-274.
- White, C. (1981). Tonal Pronunciation Errors and Interference from English Intonation. *Journal of the Chinese Language Teachers’ Association*, 16, 2: 27-56.

- White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- Wightman, C. W., et al. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *JASA* 91: 1707–17.
- Wusler, L. (1994). The perception of second language prosody—Investigating the conditions on transfer: A revised study. Unpublished paper, Department of ESL, University of Hawai‘i at Manoa
- Xu, Y. (1997). Contextual Tonal Variations in Mandarin. *Journal of Phonetics*. (1997) 25, 61-83.
- Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46. 220-251.
- Yang, C-S. (2011) The Acquisition of Mandarin Prosody by American Learners of Chinese as a Foreign Language (CFL). Dissertation. Ohio State University.
- Yin, H. (2003) The blocking of tone sandhi in Mandarin Chinese. *Proceedings of the 2003 annual conference of the Canadian Linguistic Association*. P. 296-307.
- Yip, M. (1980) *The Tonal Phonology of Chinese*. MIT: PhD dissertation (published 1990). New York: Garland Publishing.
- Yip, M. (1989). Contour tones. *Phonology*. 6: 149–74.
- Yip, M. (1995). Tone in East Asian languages. In John. A. Goldsmith (ed.), *Handbook of theoretical phonology*. Oxford, UK: Blackwell, 476–94.
- Yip, M. (2001). Tonal features, tonal inventories, and phonetic targets. *UCL Working Papers in Linguistics* 13, 161-188
- Yip, M. (2001). Tonal features, tonal inventories and phonetic targets, in *UCL Working Papers in Linguistics* 13, C. Iten and A. Neeleman, editors, Department of Phonetics and Linguistics, University College London. P. 303-329.
- Yip, M. (2002). *Tone*. Cambridge University Press. Cambridge, UK.
- Yuan, J-H. (2004) *Intonation in Mandarin Chinese: Acoustic, Perception, and Computational Modeling*. Ph.D dissertation. Cornell University.
- Zhang, H. (2007) A phonological study of second language acquisition of Mandarin Chinese tones. Master’s thesis, University of North Carolina at Chapel Hill.
- Zhang, H. (2010). “Phonological Universals and Tone Acquisition.” *Journal of Chinese Language Teachers Association*. 45, 1: 39-65.
- Zhang, J. (1998). Checked tones in Chinese dialects – evidence for phonetically-driven phonology. *CLS* 34: 439–53.
- Zhang, J. (2002). *The effects of duration and sonority on contour tone distribution--A typological survey and formal analysis*. Routledge, New York.
- Zhang, J. (2004a). The role of contrast-specific and language-specific phonetics in contour tone

- distribution. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.), *Phonetically based phonology*. Cambridge University Press, Cambridge, UK. 157-190.)
- Zhang, J (2004b): Contour Tone Licensing and Contour Tone Representation. *Language and Linguistics* 5.4:925-968, 2004
- Zhang, J. and Lai, Y-W. (2010) Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology* 27.1: 153-201.
- Zhao, J-M.(1988). Cong yixie shengdiao yuyan de shengdiao shuodao hanyu shengdiao [Talking about tone in Chinese from the perspective of some tone languages]. *Selected papers from the second international symposium on teaching Chinese as a foreign language* (pp.171-181). Beijing: Beijing Language Institute.
- Zhao, J-M. and Cheng, M-Z. (1997). Jichu hanyu yuyin jiaoxue de ruogan wenti [Several issues of teaching Mandarin sound system]. In *Yuyin yanjiu yu duiwai hanyu jiaoxue [Phonetic study and Chinese language teaching]* (ed.) by Zhao, J-M and Meng, Z-M. Beijing Language and Culture University.