The Acquisition of Mandarin Prosody by American Learners of Chinese as a Foreign Language (CFL)

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Graduate School of The Ohio State University

By
Chunsheng Yang, B.A., M.A.

Graduate Program in East Asian Languages and Literatures

The Ohio State University

2011

Dissertation Committee:
Dr. Marjorie K.M. Chan, Advisor
Dr. Mary E. Beckman
Dr. Cynthia Clopper
Dr. Mineharu Nakayama
ABSTRACT

In the acquisition of second language (L2) or foreign language (FL) pronunciation, learners not only learn how to pronounce consonants and vowels (tones as well, in the case of tone languages, such as Mandarin Chinese), they also learn how to produce the vowel reduction, vowel-consonant co-articulation, and prosody. Central to this dissertation is prosody, which refers to the way that an utterance is broken up into smaller units, and the acoustic patterns of each unit at different levels, in terms of fundamental frequency (F0), duration and amplitude. In L2 pronunciation, prosody is as important as -- if not more important than -- consonants and vowels. This dissertation examines the acquisition of Mandarin prosody by American learners of Chinese as a Foreign Language (CFL). Specifically, it examines four aspects of Mandarin prosody: (1) prosodic phrasing (i.e., breaking up of utterances into smaller units); (2) surface F0 and duration patterns of prosodic phrasing in a group of sentence productions elicited from L1 and L2 speakers of Mandarin Chinese; (3) patterns of tones errors in L2 Mandarin productions; and (4) the relationship between tone errors and prosodic phrasing in L2 Mandarin.

The analysis of prosodic phrasing in the corpus shows that prosodic phrasing is closely related to syntactic structure in both L1 and L2 Mandarin productions. Moreover,
results show that the syntactic structure in a prosodic phrase does not influence the
prosodic structure of that constituent in either the learner group or the native group.

Analysis of the duration patterns in the L1 and L2 Mandarin corpus shows that the
most consistent duration pattern that indexes prosodic phrasing is phrase-final
lengthening. In addition, the duration analysis shows that the native group shows phrase-
initial lengthening, the intermediate learner group produces phrase-initial shortening, and
the advanced learner group displays no effect of phrasing on phrase-initial duration. This
pattern of phrase-initial lengthening/shortening indicates a learning effect in that the
advanced learner group patterned more similarly to the native group. We also observed
the transfer of L1 English stress patterns, such as the weak versus strong alternating stress
patterns in the L2 corpus.

With respect to the F0 patterns of prosodic phrasing, it was found that the conflicting
tone sequences (the sequences in which the target at the offset of a preceding tone and the
target at the onset of the following tone are different) posed more difficulty for learners
than the compatible tone sequences (the sequences in which the target at the offset of a
preceding tone and the target at the onset of a following tone are identical). In addition,
the productions by the native speakers involved more target undershoot (namely, the tone
targets are not fully realized) than those by the L2 learners. It was also found that the tone
target undershoot mostly occurred in the first prosodic phrase of an utterance. The
transfer of English intonation patterns was also observed, such as the transfer of a high
phrase accent at the end of a prosodic phrase.
Analysis of tone errors shows that the low and rising tones were the most frequent tone errors produced by the two groups of learners in their L2 Mandarin productions, regardless of the underlying tones. The patterns of tone errors in different tone sequences suggest that the learners not only had difficulty in changing the tone targets quickly in the conflicting tone sequences, they also had difficulty in changing the F0 direction quickly in the compatible tone sequences. It is argued that these tone errors were produced as a consequence of the superimposition of the L1 English utterance-level prosody over tone production by L2 learners.
DEDICATION

This dissertation is dedicated to my parents and my brother.
ACKNOWLEDGEMENTS

This dissertation cannot be completed without the support and assistance of many people. Firstly my sincere gratitude goes to Professor Marjorie K.M. Chan, who has been a great advisor and a role model in both academy and life. I am grateful for all the freedom she has given me in my academic pursuit and all the assistance she has provided in my research and this dissertation. I will always remember those lively and enlightening discussions with her! Xiexie nin, Ma Laoshi!

My great thanks also go to Professor Mary Beckman. It is in Mary’s Phonetics class that I started to become interested in phonetic and prosodic research. I have benefited greatly from Mary’s expertise and knowledge in phonetics and prosody. Again and again she has kept me on track in the process of my dissertation. Thank you, Mary, for always squeezing time from your busy schedule to meet with me and pointing me to new directions whenever I was stuck in my dissertation.

My thanks also go to Professor Cynthia Clopper and Professor Mineharu Nakayama. I was amazed at Cynthia’s knowledge in phonetics and her clarity in presenting even the most abstract concepts. I am also grateful for her help in my candidacy exam and dissertation. I have also benefited from Professor Nakayama’s expertise in second language acquisition. He gave me a lot of suggestions and advice on the experimental
design and the analysis of tone errors in my dissertation. I am also grateful for his great assistance in my job-hunting, even when he was in Japan.

In addition, I would like to thank DEALL for providing me with the graduate associateship for five years and other generous support for attending numerous national and international conferences. My thanks also go to Ms. Debbie Knicely for her availability and readiness for any questions and concerns that I have had during my graduate study. Xiexie ni, Debbie! I also thank Ms. Soyoung Han and Ms. Hui Chu for their assistance.

I would also like to thank Mr. Steven Knicely for his guidance and accommodation in my teaching and his great assistance in my job search. Xiexie nin, Shi Laoshi.

I also want to thank Jing Yan, Hana Kang, Wang Xu, Yi He, Rongbin Zheng, Jia Yang, Bo Zhu, Yifan Pai, Xizhen Qin, Nan Meng, Ying Liu, Haohsiang Liao, Qiong Yang, Man He, Hui Xiong, Song He, Alex Ratté, Seth Wiener, Bryan Wang and many others for their friendship and support over the past six years.

Finally, I would like to thank my parents and my brother for their continuing love and believing in me. Your support is the greatest motivation for the completion of this dissertation.
VITA

2009, M.A. in Chinese Linguistics, The Ohio State University

2005, M.A. in English Languages and Literatures, Ningbo University, China

2000, B.A. in English, The University of Science and Technology of China, China

PUBLICATIONS


http://chinalinks.osu.edu/naccl-20/proceedings/10_yang-c.pdf

viii

http://www.languageatinternet.de/articles/1142/


Major Field: East Asian Languages and Literatures

Specializations: Chinese Linguistics, Phonetics and Phonology, prosody, second language acquisition
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................. ii
DEDICATION ................................................................................................................. v
ACKNOWLEDGEMENTS .............................................................................................. vi
VITA ............................................................................................................................. viii
TABLE OF CONTENTS .............................................................................................. x
LIST OF TABLES ......................................................................................................... xvi
LIST OF FIGURES ....................................................................................................... xviii

CHAPTER 1: INTRODUCTION ......................................................................................1
  1.1. Research questions .......................................................................................... 1
  1.2. Aim and significance of the study ................................................................. 4
  1.3. The organization of the dissertation ............................................................. 6

CHAPTER 2: THEORETICAL BACKGROUND AND PREVIOUS STUDIES ...........8
  2.1. Introduction ..................................................................................................... 8
  2.2. Prosody of Mandarin Chinese ....................................................................... 8
    2.2.1. Lexical prosody of Mandarin Chinese .................................................. 9
    2.2.2. Utterance-level prosody of Mandarin Chinese .................................... 11
    2.2.3. Interaction of tones and intonation ...................................................... 12

x
2.3. Prosodic structure in English and Mandarin Chinese.................................12
  2.3.1. The prosodic word.................................................................15
  2.3.2. The prosodic phrase.............................................................16
  2.3.3. The intonation phrase/breath group/prosodic group.........................18
2.4. Prosodic markings for prosodic phrases in English and Mandarin Chinese.....19
  2.4.1. F0 patterns of prosodic phrases in English and Mandarin Chinese.........20
  2.4.2. Duration patterns of prosodic phrases in English and Mandarin Chinese.....27
2.5. Studies on the acquisition of Mandarin prosody.....................................29
  2.5.1. Tone acquisition.................................................................29
  2.5.2. Acquisition of utterance-level prosody of Mandarin Chinese...............31
2.6. Gaps and problems in previous studies.................................................33
2.7. Summary....................................................................................34
CHAPTER 3: METHODOLOGY.................................................................36
  3.1. Introduction...............................................................................36
  3.2. Stimuli......................................................................................36
  3.3. Subjects....................................................................................43
  3.4. Recording procedure.................................................................43
  3.5. Data transcription.......................................................................45
  3.6. Acoustic measurements...............................................................48
    3.6.1. F0 measurements.................................................................48
    3.6.2. Duration measurements.........................................................49
  3.7. Data analysis.............................................................................49
3.7.1. Speech rate in L1 and L2 Mandarin Chinese........................................49
3.7.2. Overview of F0 patterns in L1 and L2 Mandarin Chinese.....................50
3.7.3. Prosodic phrasing and duration and F0 patterns in L1 and L2 Mandarin
Chinese.................................................................50
3.7.4. Patterns of tone errors in L2 Mandarin Chinese..................................51
3.8. Summary.........................................................................................52

CHAPTER 4: UTTERANCE-LEVEL PROSODY IN L1 AND L2 MANDARIN
Chinese...............................................................................................53

4.1. Introduction.......................................................................................53
4.2. Overview of the speech data in L1 and L2 Mandarin Chinese..................54
  4.2.1. Speech rate in L1 and L2 Mandarin Chinese.....................................54
  4.2.2. Max F0, min F0 and F0 range in L1 and L2 Mandarin Chinese.........56
  4.2.3. Summary.....................................................................................57
4.3. Analysis of prosodic phrasing and surface duration patterns in L1 and L2 Mandarin
Chinese...............................................................................................57
  4.3.1. Break distribution in L1 and L2 Mandarin Chinese.............................58
  4.3.2. Duration patterns of prosodic phrasing in L1 and L2 Mandarin Chinese.....63
  4.3.3. Summary and discussion..................................................................67
4.4. F0 patterns of prosodic phrasing in L1 and L2 Mandarin Chinese.............68
  4.4.1. Pitch reset in L1 and L2 Mandarin Chinese.....................................69
  4.4.2. Time-normalized F0 analysis in L1 and L2 Mandarin Chinese............75
    4.4.2.1. F0 patterns in the T1 sequence....................................................75
4.4.2.1.1. F0 patterns in the T1 sequence: Native group.........75
4.4.2.1.2. F0 patterns in the T1 sequence: Advanced learner group.....78
4.4.2.1.3. F0 patterns in the T1 sequence: Intermediate learner group..........................................................81
4.4.2.1.4. Summary of the F0 patterns in the T1 sequence........83
4.4.2.2. F0 patterns in the T2 sequence.........................................................84
4.4.2.2.1. F0 patterns in the T2 sequence: Native group...............84
4.4.2.2.2. F0 patterns in the T2 sequence: Advanced learner group.....88
4.4.2.2.3. F0 patterns in the T2 sequence: Intermediate learner group..........................................................90
4.4.2.2.4. Summary of the F0 patterns in the T2 sequence.........92
4.4.2.3. F0 patterns in the T4 sequence.........................................................93
4.4.2.3.1. F0 patterns in the T4 sequence: Native group...............94
4.4.2.3.2. F0 patterns in the T4 sequence: Advanced learner group....100
4.4.2.3.3. F0 patterns in the T4 sequence: Intermediate learner group
........................................................................103
4.4.2.3.4. Summary of the F0 patterns in the T4 sequence.........105
4.4.2.4. F0 patterns in the T2T4 alternating sequence................106
4.4.2.5. F0 patterns in the T4T2 alternating sequence...................112
4.4.2.6. Summary of the F0 analysis of prosodic phrasing in L1 and L2
Mandarin Chinese.................................................................116
4.5. Summary...............................................................................118
CHAPTER 5: PATTERNS OF TONE ERRORS IN CONTEXT………………………121

5.1. Introduction…………………………………………………………………121
5.2. Analysis of tone errors in L2 Mandarin Chinese…………………………...121
5.3. Tone errors in different tone sequences……………………………………...124
   5.3.1. Tone errors in the T1 sequence……………………….…………………..124
   5.3.2. Tone errors in the T2 sequence…………………………………..……..133
   5.3.3. Tone errors in the T3 sequence………………………………………..…141
   5.3.4. Tone errors in the T4 sequence………………………………………..…148
   5.3.5. Tone errors in the T2T4 alternating sequence…………………………...153
   5.3.6. Tone errors in the T4T2 alternating sequence…………………………...159
5.4. Discussion of the tone error analysis……………………………………..............162

CHAPTER 6: GENERAL DISCUSSION AND PEDAGOGICAL APPLICATIONS
…………………………………………………………………………...169

6.1. Introduction…………………………………………………………………169
6.2. Summary of the findings in this study…………………………………………169
   6.2.1. Research question 1………………………………………………..……….169
   6.2.2. Research questions 2 and 3……………………………………………..…..170
   6.2.3. Research question 4……………………………………………………..….171
6.3. Theoretical implications…………………………………………...………..…....172
   6.3.1. Are F0 patterns the consequence of duration control in the prosodic phrasing
           of Mandarin Chinese?…………………………………………………………..173
6.3.2. Why are there so many low tone errors? ............................................ 174

6.4. Pedagogical applications of the findings in CFL teaching..........................175
   6.4.1. Audio-visual training on L2 prosody..................................................176
   6.4.2. Training on L2 Mandarin Chinese prosody........................................177
      6.4.2.1. Training on L2 Mandarin tones ..................................................178
      6.4.2.2. Training on L2 utterance-level prosody......................................180
         6.4.2.2.1. Training on Mandarin stress..............................................180
         6.4.2.2.2. Training on tone co-articulation..........................................182
         6.4.2.2.3. Training on quick change of F0 direction and tone
                        targets.............................................................................185

6.5. Summary...............................................................................................186

CHAPTER 7: CONCLUSION.............................................................................187
   7.1. Limitations of the study.................................................................187
   7.2. Directions for further study.............................................................188
   7.3. Concluding remarks.........................................................................189

APPENDIX A: The list of tone sequences used in the study............................190

APPENDIX B: Subject information...............................................................192

APPENDIX C: Numbers of utterances used in the analysis of different tone
                        sequences..................................................................................193

APPENDIX D: Numbers of the occurrence of pitch reset in the T1 and the T4 sequences .................................................................195

REFERENCES..............................................................................................196
LIST OF TABLES

Table 2.1. The four lexical tones in Mandarin Chinese..........................10

Table 3.1. Tone targets in the four lexical tones of Mandarin Chinese..........46

Table 4.1. Summary of the duration and F0 patterns in L1 and L2 Mandarin
Chinese.......................................................................................................119

Table 5.1. Correct utterance percentage (CUP), correct tone percentage (CTP) and
accuracy weighting (AW) in the T1 sequence: SS-24..........................126

Table 5.2. Correct utterance percentage (CUP), correct tone percentage (CTP) and
accuracy weighting (AW) in the T1 sequence: SS-33..........................127

Table 5.3. Correct utterance percentage (CUP), correct tone percentage (CTP) and
accuracy weighting (AW) in the T2 sequence: SS-24.........................134

Table 5.4. Correct utterance percentage (CUP), correct tone percentage (CTP) and
accuracy weighting (AW) in the T2 sequence: SS-33.........................134

Table 5.5. Correct utterance percentage (CUP), correct tone percentage (CTP) and
accuracy weighting (AW) in the T3 sequence: SS-24.........................143

Table 5.6. Correct utterance percentage (CUP), correct tone percentage (CTP) and
accuracy weighting (AW) in the T3 sequence: SS-33.........................143
Table 5.7. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weighting (AW) in the T4 sequence: SS-24 …………………….149

Table 5.8. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weighting (AW) in the T4 sequence: SS-33 …………………….149

Table 5.9. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weighting (AW) in the T2T4 sequence: SS-24 …………………….154

Table 5.10. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weighting (AW) in the T2T4 sequence: SS-33 …………………….154

Table 5.11. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weighting (AW) in the T4T2 sequence: SS-24 …………………….159

Table 5.12. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weighting (AW) in the T4T2 sequence: SS-33 …………………….160

Table 6.1. Examples of the contrast between fully toned stressed syllables and neutrally atonic syllables ……………………………………………………181
LIST OF FIGURES

Figure 2.1. Prosodic hierarchy in English and Mandarin Chinese.........................14
Figure 2.2. An example of prosodic hierarchy in Mandarin Chinese.....................14
Figure 2.3. Schematic F0 contour of five prosodic phrases in a prosodic group in
Mandarin........................................................................................................21
Figure 2.4. F0 contours in a statement “He said you would” spoken by a female
speaker of American English...........................................................................23
Figure 2.5. An interrogative with a sequence of neutral tones spoken by a Beijing
actress in a movie...............................................................................................25
Figure 3.1. Screen capture of one Powerpoint slide used in data collection...........45
Figure 3.2. Transcription illustration of an utterance 陆蔚用慢用药 Lù Wèi yòng màn
yòng yào “Lu Wei uses slow medicine”, produced by a female native
speaker.............................................................................................................47
Figure 4.1. Mean syllable duration of three subject groups.................................55
Figure 4.2. The max F0, min F0 and F0 range across subject groups: Female.........56
Figure 4.3. The max F0, min F0 and F0 range across subject groups: Male ..........57
Figure 4.4. Break distribution in SS-24 and SS-33 across groups.......................59
Figure 4.5. Syllable duration across groups and sentence structures.................64
Figure 4.6. Percentage of pitch reset in SS-24 and SS-33 for the T1 sequence across
   groups..........................................................70
Figure 4.7. Percentage of pitch reset in SS-24 and SS-33 for the T4 sequence across
   groups ........................................................................72
Figure 4.8. F0 contours in the T1 sequence: Female native speakers..................76
Figure 4.9. F0 contours in the T1 sequence: Male native speakers...................77
Figure 4.10. F0 contours in the T1 sequence: Female advanced learners..........78
Figure 4.11. F0 contours in the T1 sequence: Male advanced learners..........79
Figure 4.12. F0 contour in the second prosodic phrase in the T1 sequence: An
   utterance produced by a female advanced learner..........................80
Figure 4.13. F0 contours in the T1 sequence: Female intermediate learners......81
Figure 4.14. F0 contours in the T1 sequence: Male intermediate learners......82
Figure 4.15. F0 contour in the T1 sequence: An utterance produced by a male
   intermediate learner.................................................83
Figure 4.16. F0 contours in the T2 sequence: Female native speakers.............85
Figure 4.17. F0 contours in the T2 sequence: Male native speakers.............86
Figure 4.18. F0 contour in the T2 sequence (SS-24): An utterance produced by a male
   native speaker...........................................................87
Figure 4.19. F0 contour in the T2 sequence (SS-33): An utterance produced by a male
   native speaker...........................................................87
Figure 4.20. F0 contours in the T2 sequence: Female advanced learners........88
Figure 4.21. F0 contours in the T2 sequence: Male advanced leaners............89

xix
Figure 4.22. F0 contours in the T2 sequence: Female intermediate learners........90

Figure 4.23. F0 contour in the T2 sequence (SS-24): An utterance produced by a
female intermediate learner.........................................................91

Figure 4.24. F0 contours in the T2 sequence: Male intermediate learners...........92

Figure 4.25. F0 contours in the T4 sequence: Female native speakers...............94

Figure 4.26. F0 contour in the T4 sequence (SS-24): An utterance produced by a
female native speaker.................................................................95

Figure 4.27. F0 contours in the T4 sequence (SS-24): An utterance produced by a
female native speaker.................................................................96

Figure 4.28. F0 contour in the T2 sequence (SS-33): An utterance produced by a
female native speaker.................................................................97

Figure 4.29. F0 contours in the T4 sequence: Male native speakers...............98

Figure 4.30. F0 contour in the T4 sequence (SS-24): An utterance produced by a male
native speaker.................................................................99

Figure 4.31. F0 contour in the T4 sequence (SS-33): An utterance produced by a male
native speaker.................................................................100

Figure 4.32. F0 contours in the T4 sequence: Female advanced learners..........101

Figure 4.33. F0 contour in the T4 sequence (SS-24): An utterance produced by a female advanced learner.................................101

Figure 4.34. F0 contours in the T4T2 sequence: Male advanced learners........102

Figure 4.35. F0 contour in the T4 sequence (SS-33): An utterance produced by a male advanced learner.........................................................103

xx
Figure 4.36. F0 contours in the T4 sequence: Female intermediate learners……..104
Figure 4.37. F0 contours in the T4 sequence: Male intermediate learners………..105
Figure 4.38. F0 contours in the T2T4 sequence: Female native speakers…………….107
Figure 4.39. F0 contours in the T2T4 sequence: Male native speakers…………….108
Figure 4.40. F0 contour sin the T2T4 sequence: Female advanced learners………..108
Figure 4.41. F0 contours in the T2T4 sequence: Male advanced learners……………109
Figure 4.42. F0 contours in the T2T4 sequence: Female intermediate learners……..109
Figure 4.43. F0 contours in the T2T4 sequence: Male intermediate learners………..110
Figure 4.44. F0 contour in the T2T4 sequence (SS-24): An utterance produced by a male intermediate learner………………………………………………………….111
Figure 4.45. F0 contour in the T2T4 sequence (SS-33): An utterance produced by a male intermediate learner………………………………………………………….111
Figure 4.46. F0 contours in the T4T2 sequence: Female native speakers…………….113
Figure 4.47. F0 contours in the T4T2 sequence: Male native speakers…………….113
Figure 4.48. F0 contours in the T4T2 sequence: Female advanced learners………..114
Figure 4.49. F0 contours in the T4T2 sequence: Male advanced learners…………….114
Figure 4.50. F0 contours in the T4T2 sequence: Female intermediate learners……….115
Figure 4.51. F0 contours in the T4T2 sequence: Male intermediate learners……….115
Figure 5.1. Illustration of the low and rising tone errors in the T1 sequence (SS-24):
An utterance produced by a male intermediate learner…………………..128
Figure 5.2. Another illustration of the low and rising tone errors in the T1 sequence (SS-24): An utterance produced by a male intermediate learner………..129
Figure 5.3. Illustration of the low tone error in the T1 sequence (SS-33): An utterance produced by a male advanced learner..................................................131

Figure 5.4. Illustration of the low tone error in the T2 sequence (SS-24): An utterance produced by a male intermediate learner...........................................135

Figure 5.5. Illustration of the low tone error in the T2 sequence (SS-33): An utterance produced by a female advanced learner..............................................137

Figure 5.6. An example of H target misalignment in the T2 sequence (SS-33): An utterance produced by a female advanced learner.................................138

Figure 5.7. Illustration of the over-application of T3 sandhi: An utterance produced by a male advanced learner.................................................................145

Figure 5.8. An example of H target misalignment in the T3 sequence: An utterance produced by a male advanced learner .....................................................147

Figure 5.9. Illustration of the low tone error in the T3 sequence (SS-24): An utterance produced by a male advanced learner......................................................151

Figure 5.10. Illustration of the rising tones in the T4 sequence (SS-33): An utterance produced by a male intermediate learner..................................................152

Figure 5.11. Illustration of the rising tone errors in the T2T4 sequence (SS-24): An utterance produced by a male advanced learner.............................................155

Figure 5.12. Illustration of the low and rising tone errors in the T2T4 sequence (SS-33): An utterance produced by a male intermediate learner.........................157

Figure 5.13. Illustration of the low tone error in the T2T4 sequence: An utterance produced by a male advanced learner.........................................................158
Figure 5.14. Illustration of the rising tone errors in the T4T2 sequence (SS-24): An utterance produced by a male intermediate learner..........................161

Figure 5.15. A tentative model of learners’ tone error production..........................166

Figure 6.1. Illustrations of the target undershoot in the T2 sequence: An utterance produced by a female native speaker..........................184
CHAPTER 1: INTRODUCTION

1.1. Research questions

Mandarin Chinese is the official language in China, Taiwan, and Singapore. It is also one of the official languages used in the United Nations. Mandarin Chinese has many varieties, such as Beijing Mandarin, Taiwan Mandarin, etc. Even though intelligibility is not an issue at most times, speakers of different varieties of Mandarin Chinese do encounter some communication issues, especially with respect to the lexicon. Standard Mandarin Chinese, so-called Putonghua, is based on the phonology of Beijing Mandarin. However, standard Mandarin Chinese is not exactly the same as Beijing Mandarin. For example, the rhotacization or er-suffixation is more frequently used in Beijing Mandarin than in standard Mandarin Chinese. Thus, strictly speaking, there are no native speakers of standard Mandarin Chinese, because all speakers, including Beijing Mandarin speakers, grew up with the influence of one dialect or another.

Typologically speaking, Mandarin Chinese is a tone language, in which tones are lexically specified, namely the fundamental frequency (F0)\(^1\) or pitch pattern over a syllable can be used to distinguish the meaning of words. In addition, intonation, the F0 pattern over a phrase or an utterance, is also an integral part of Mandarin phonology. In

---

\(^1\) Fundamental frequency (F0) is the rate of the vocal fold vibration of a speaker per second. Pitch is the perceived fundamental frequency. Even though fundamental frequency and pitch are not exactly the same, especially at higher frequency, they are used interchangeably in this dissertation.
this sense, F0 plays a dual role in Mandarin Chinese, namely, representing both lexical tones and intonation.

Due to the rapid economic development in China, the teaching and learning of Mandarin Chinese has gained great momentum in recent years. Chinese courses are offered in more and more countries all over the world. In tandem with the ever increasing popularity of Mandarin Chinese, more and more studies were conducted on the acquisition of Mandarin Chinese in the field of second language acquisition (SLA).

Most studies on the acquisition of Mandarin phonology focus on the acquisition of lexical tones (Bent 2005; Chen 1997; 2000; Miracle 1989; Shen 1989; White 1981, and so on). However, in teaching Chinese as a foreign language (CFL), it is often observed that even though CFL learners can correctly produce tones in isolation, tonal errors often occur when tones are concatenated and produced in utterances. It seems that tone errors produced by second language (L2) learners may not only be due to lexical tones, but may also be due to something beyond lexical tones, involving tone production at the utterance level. That “something” is prosody.

Prosody is a concept couched in the framework of Autosegmental-Metrical (AM) phonology (Pierrehumbert and Beckman 1988; Beckman 1996; Ladd 1996, 2008). Prosody of an utterance refers to the hierarchically organized structure of phonologically defined constituents and heads (Beckman 1996). More specifically, I subscribe to Shattuck-Hufnagel and Turk’s (1996:196) expansion of Beckman’s definition of prosody. In their revised definition, prosody refers to 1) the higher-level structures that best account for the acoustic patterns at lower levels; and 2) the acoustic patterns of F0,
duration, amplitude, spectral tilt, and segmental reduction, and their articulatory correlates, which can be best accounted for by reference to higher-level structures. From these definitions, it can be seen that what I mean by prosody is very general. On the one hand, it concerns how an utterance is chunked into prosodic constituents at lower levels (i.e., prosodic phrasing). On the other hand, it refers to the acoustic patterns of F0, duration, and amplitude, etc. These two aspects of prosody are closely related to each other in that different prosodic phrasing will result in different acoustic patterns with respect to F0, duration and amplitude.

Most previous studies on the acquisition of Mandarin prosody focus on the lexical prosody (Chen 1997; Miracle 1989; Shen 1989), namely they only focus on the acquisition of tones. Very few studies examine utterance-level prosody, which includes intonation. Chen (2000) and Viger (2007) are perhaps the only studies that examine American L2 learners’ acquisition of utterance-level prosody in Mandarin Chinese.

To fill the gaps in the previous research, this dissertation investigates the acquisition of Mandarin prosody by American CFL learners. Specifically, I aim to address the following questions:

1) How do American CFL learners prosodically phrase their utterances (that is, break them up into smaller units), and how does the prosodic phrasing in L2 Mandarin differ from that in L1 Mandarin?

2) What are the duration patterns in L2 Mandarin, and how do they differ from those
3) What are the F0 patterns in L2 Mandarin, and how do they differ from those in L1 Mandarin?

4) How does prosodic phrasing influence tone production in L2 Mandarin speech?

1.2. Aim and significance of the study

As a Chinese instructor, I have had students who have a very good command of Chinese grammar, but poor control of lexical tones and utterance-level prosody. In many cases, the poor prosodic production seems to overshadow the good command of grammar, leading to the perception of foreign accent, or at least to the impression of non-standardness.

Many factors have been claimed to contribute to the perception of a foreign accent in L2 speech, with prosodic transfer (i.e., the transfer of intonation, stress, and so on) being one of them (Gut 2003; Nguyen et al. 2008, among others). In examining the contribution of prosody to the perception of a foreign accent, the first step is to examine what are the prosodic differences between L1 and L2 speech. That first step is what this study aims to achieve, in preparation for future investigation into the relationship between prosodic differences and the perception of a foreign accent. However, it is one thing to identify the prosodic differences between L1 and L2 speech, while it is another thing to claim that the identified prosodic differences contribute to the perception of a foreign accent. Evidently,
the scant literature in prosodic acquisition seems to suggest that many researchers have taken for granted that, as long as there are prosodic differences between L1 and L2 speech, these differences will inevitably lead to the perception of foreign accent. However, this is an empirical question which should be examined experimentally.

Another complication in the claim of prosodic transfer is that not all prosodic differences are derived from the transfer of L1 prosody. The term, transfer, comes from contrastive analysis (Weinreich 1953; Lado 1957). According to Lado (1957: 2), L2 learners tend to transfer the forms and meanings in their native language to the second or foreign language. However, studies have shown that L1 transfer can not fully explain the L2 interlanguage, namely, not all L2 deviations are predictable from the contrastive analysis of the two languages involved (Ueyama 2000 and references therein). The modified view on L2 interlanguage tends to treat it as an autonomous linguistic system that changes under the influence of a combination of at least three components: L1, L2, and universals that are shared by all languages (Ueyama 2000:3; Major 2001). Thus, even if L2 prosodic deviations from L1 norms turn out to contribute to the perception of a foreign accent, such prosodic deviations may not merely be the result of L1 prosodic transfer, as developmental universals of language acquisition may lead to prosodic deviations in L2 speech as well. In this sense, this study will not only help show what are the prosodic differences between L1 and L2 Mandarin Chinese, but it will also help identify what are the sources for these prosodic differences, L1 prosodic transfer or developmental universals.
The present study goes beyond low-level tone acquisition and probes the source of prosodic deviations in L2 Mandarin by taking prosodic structure into consideration. There have been some impressionistic observations, but no systematic studies have taken the current approach toward the study of the acquisition of Mandarin prosody. The findings from this study will reveal the prosodic differences between L1 and L2 Mandarin Chinese, which will pave the way for further study on what aspects of prosodic deviations may contribute to the perception of a foreign accent. The findings of this study will also have important pedagogical implications for CFL teaching. Compared to the abundant published materials on training/drilling of the grammar of L2 Mandarin Chinese, prosodic training except for tone training has received very little attention in the field. It is expected that the findings from this study can guide the training on L2 Mandarin prosody. Specifically, based on the patterns of prosodic deviations identified in this study, training materials, including software and programs, can be prepared and used in CFL teaching.

1.3. The organization of the dissertation

Chapter 2 discusses the prosody of Mandarin Chinese in detail. Also presented in Chapter 2 are the prosodic structure of English and Mandarin Chinese, previous studies on the acquisition of Mandarin prosody, and gaps and problems in previous studies. Chapter 3 provides a detailed account of the design of the experiment, including stimuli, subjects, transcription, and data analysis. Chapter 4 analyzes the prosodic phrasing, the surface F0 and duration patterns in L2 Mandarin, as compared to those in L1 Mandarin.
Chapter 5 discusses the tone errors found in L2 Mandarin. Chapter 6 summarizes the findings of the study, provides general discussions of some theoretical issues, and concludes with some pedagogical applications of the findings to CFL teaching. Chapter 7 provides some concluding remarks.
CHAPTER 2: THEORETICAL BACKGROUND AND PREVIOUS STUDIES

2.1 Introduction

This chapter provides some theoretical background and summarizes previous studies on the acquisition of Mandarin prosody. Specifically, section 2.2 describes the prosody of Mandarin Chinese, both lexical prosody and utterance-level prosody. Section 2.3 discusses the prosodic structure in English and Mandarin Chinese. Section 2.4 presents the prosodic markings of prosodic phrases in English and Mandarin Chinese. Section 2.5 summarizes some previous studies on the acquisition of Mandarin prosody. Finally, section 2.6 outlines some gaps and problems in previous studies on the acquisition of Mandarin prosody.

2.2. Prosody of Mandarin Chinese

Mandarin Chinese is a tone language and tones are lexically specified. In addition, there are also acoustic patterns in terms of F0, duration and amplitude at the utterance level. These two different aspects represent prosody at two levels, namely, prosody at the lexical level and prosody at the utterance level. The next few sections will discuss lexical and utterance-level prosody in Mandarin Chinese. Also discussed is the interaction between tones and intonation, which forms part of utterance-level prosody.
2.2.1. Lexical prosody of Mandarin Chinese

There are four lexical tones in Mandarin Chinese. They are often referred to as Tone 1 (high level tone, referred to as T1 hereafter), Tone 2 (mid-rising tone, referred to as T2 hereafter), Tone 3 (low-dipping tone, referred to as T3 hereafter), and Tone 4 (high-falling tone, referred to as T4 hereafter). Chao (1930) designed a five-level numerical scale for representing pitch height, with “1” being the lowest pitch value and “5” the highest pitch value within a speaker’s pitch range. In this scale, T1 through T4 are represented as “55”, “35”, “214”, and “51”, respectively. The low-dipping variant of T3 (“214”) only occurs in isolation or at utterance-final position. At non-utterance-final position, T3 often surfaces as a low tone (“21” or “22”). Moreover, when there are two consecutive T3’s, the first T3 often undergoes tone change, namely tone sandhi, and becomes a rising tone. Thus, the numerical values for T3 in actual speech can be “214”, “21”2 (or “22”), or “35”. In the Pinyin system, the official Romanization system used in the People’s Republic of China since 1958, and widely adopted throughout the world, lexical tones are marked with iconic diacritics above the letters representing vowels, as in ā, á, ǎ, à. Table 2.1 lists the four tones and their names, together with the tone values in numbers and tone diacritics in the examples.

---

2 The low tone (“21”) also frequently surfaces at utterance-final position in Beijing Mandarin.
In addition to the four lexical tones, there is also the neutral tone（轻声 qīngshēng）in Mandarin Chinese. The neutral tone (T0), which occurs on an unstressed (atonic) short syllable at non-initial position in a word or phrase, must be preceded by at least one syllable that carries one of the four lexical tones (Lin 2007: 98). The F0 height and contour of the neutral tone are contingent upon the tones of the preceding syllables. Specifically, 1) if the preceding syllable is T1, T2 or T4, the neutral tone is realized as a low or falling tone; 2) if the preceding syllable carries a low-level T3, the neutral tone is realized as a high or rising tone (Cheng 1973: 56; Yip 2002: 182).

There are two F0 events related to tones in Mandarin Chinese, namely pitch register and pitch contour. Pitch register refers to the relative height of F0, while pitch contour refers to the F0 trajectory over a syllable. Both pitch register and pitch contour are important in tone production and perception (Shen 1989; Yang and Chan 2010, among others).

When tones are concatenated and produced in phrases or utterances, tones may undergo some change, with the phenomenon of tone sandhi being an important case. In addition to the T3 sandhi, there are other tone sandhi phenomena that may occur in

<table>
<thead>
<tr>
<th>Tone</th>
<th>Tone name</th>
<th>Tone values</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>High level tone</td>
<td>55</td>
<td>mā, 妈, ‘mother’</td>
</tr>
<tr>
<td>T2</td>
<td>Mid rising tone</td>
<td>35</td>
<td>má, 麻, ‘hemp’</td>
</tr>
<tr>
<td>T3</td>
<td>Low dipping tone</td>
<td>214</td>
<td>mǎ, 马, ‘horse’</td>
</tr>
<tr>
<td>T4</td>
<td>High falling tone</td>
<td>51</td>
<td>mà, 骂, ‘to scold’</td>
</tr>
</tbody>
</table>

Table 2.1. The four lexical tones in Mandarin Chinese
natural speech. Detailed discussion of tone sandhi phenomena related to other tones will be provided in Chapter 5.

2.2.2. Utterance-level prosody of Mandarin Chinese

Utterance-level prosody refers to prosodic phrasing, and the F0 and duration patterns\(^1\) at the utterance-level. In some cases, prosodic phrasing aligns with syntactic structure. However, prosodic phrasing can override syntactic structure at times, especially when the utterance is very long. Prosodic phrasing is marked by both F0 and duration patterns (Tseng et al. 2005).

Broadly speaking, the overall F0 patterns over an utterance constitute the intonation. Intonation is an important component of prosody. Most previous studies on Mandarin utterance-level prosody were conducted on the intonation of Mandarin Chinese, especially the question intonation (Shen 1985; Lee 2005; Yuan 2004, among others). Question intonation in Mandarin Chinese is marked by both local and global F0 cues. At the local level, there is both utterance-initial F0 rise (Shen 1990) and utterance-final F0 rise (Yuan 2004; Lee 2005, among others). However, there is no definitive conclusion on what is the domain for the utterance-final F0 rise, whether it is the final syllable or the final NP phrase (Lee 2005). At the global level, there is overall F0 rise in questions (Shen 1990; and Lee 2005). With respect to the perception of Mandarin intonation, tones at the end of both statements and questions influence the perception of intonation, and the tone-intonation clash (i.e., cases where T2 occurs at the end of statements and where T4 at the

\(^1\) We do not take amplitude (intensity) into consideration in this study, even though amplitude is worth investigation in further studies. Jin (1996) investigates the sentence stress in Mandarin Chinese by examining F0, duration, and amplitude.
end of questions) poses difficulty for listeners’ perception of intonation (Yuan 2004; Yang and Chan 2010).

2.2.3. Interaction of tones and intonation

Due to the dual role of F0 in representing both tones and intonation, tones interact with intonation in Mandarin Chinese. Chao (1933) has famously characterized the relationship between tones and intonation as “small ripples riding on large waves” and proposed that their relationship is that of an “algebraic sum”. Except for some researchers (e.g., J. Shen 1985, 1992, 1994) who do not agree with Chao, most Chinese linguists subscribe to his characterization. However, they explicitly separate pitch contour from pitch register and argue that the additive effects in Chao’s account pertain to pitch register, but not to pitch contour (Cao 2002, 2004; Liao 1994; Wu 1996). However, it is nearly impossible to separate the “large waves” (or intonation) from the “small ripples” (or tones) in an utterance. Nonetheless, Chao’s account does make it explicit that there are both tones and intonation in Mandarin Chinese.

2.3. Prosodic structure in English and Mandarin Chinese

The prosodic structure of an utterance is the hierarchical organization of prosodic constituents at different levels in the utterance. The prosodic structure is “an abstract entity, which is associated with a separate component of the grammar and must integrate various types of information to determine the appropriate prosodic shape of a spoken utterance” (Shattuck-Hufnagel & Turk 1996: 194). The prosodic structure mediates
between syntax and phonology; prosodic constituents are the domains where phonological/phonetic phenomena occur, such as the tone sandhi domain (Selkirk 2003).

English and Mandarin Chinese share similar prosodic constituents along their prosodic hierarchy, although the same constituents may have different phonetic representations in each language. Both languages have the following prosodic constituents: the syllable, the foot, the prosodic word (PW), the prosodic phrase (PPh)/phonological/intermediate phrase (ip), and the intonation phrase (IP)/breath group (BG)/prosodic group (PG) (Nespor & Vogel 1986; Selkirk 1984; Pierrehumbert & Beckman 1986; Speer et al. 1989; Chu & Qian 2001; Wang 2003; Tseng et al. 2005).

Figure 2.1 illustrates the prosodic hierarchy in English (Nespor & Vogel 1986; Selkirk 1984) and Mandarin Chinese (Chu & Qian 2001; Wang 2003; Tseng et al. 2005). Figure 2.2 gives an example of an actual utterance in Mandarin Chinese and its corresponding prosodic hierarchy.
Figure 2.1. Prosodic hierarchy in English and Mandarin Chinese

<table>
<thead>
<tr>
<th>English</th>
<th>Mandarin Chinese</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;As soon as he got on the bus, he started to read.&quot;</td>
<td>&quot;他一上车就看书来&quot;</td>
</tr>
</tbody>
</table>

Figure 2.2. An example of prosodic hierarchy in Mandarin Chinese
Worth mentioning is that one specific sentence may have different surface prosodic structures, depending on the different prosodic phrasing and prominence-marking.

Of the above-mentioned prosodic constituents, this study focuses on the prosodic phrase. In order to situate the prosodic phrase within the prosodic hierarchy, the prosodic word and the intonation phrase/breath group/prosodic group will be briefly discussed below. It should be noted that in the following sections, reference is made to English when no specific language is mentioned, due to the fact that most of the theories of prosody were originally developed using English. In addition, all the sentences involved are declarative.

2.3.1. The prosodic word

The prosodic word is the minimal unit of articulation in normal communication (Levelt 1989; Wheeldon 2000). A prosodic word often contains a single lexical word together with associated unstressed functional words (i.e., auxiliaries, determiners, conjunctions, or prepositions) (Wheeldon 2000). For instance, if “Give them” becomes “Give’m” in connected English speech (Shibata 2005: 15), “Give’m” is a prosodic word. The prosodic word may be just a lexical word or a syllable in some cases. In prosodic morphology (McCarthy & Prince 1993), a prosodic word is minimally a foot.

In Mandarin Chinese, prosodic words are mostly disyllabic or trisyllabic, even though there are monosyllabic prosodic words as well. For example, 看 kàn, ‘to look’, a morpheme uttered in isolation as an imperative, is a prosodic word in conversation. Prosodic words in Mandarin usually start with a stressed, full-toned syllable and function
as the rhythmic unit. However, prosodic words may or may not coincide with lexical
words. For example, 他每天 tā měitiān ‘he everyday…’, consisting of two lexical words
tā and měitiān, can be uttered as one or two prosodic words in actual speech (Wang 2003).
There are usually perceived minor breaks, but no pause, between prosodic words in
Mandarin (Peng et al. 2005; Tseng & Chou 1999; Tseng et al. 2005). Prosodic words are
also marked by prosodic word-initial syllable shortening, pre-boundary lengthening and
pitch discontinuity of the intonation contour (Yang & Wang 2002). The prosodic word
has been argued to be the domain for tone sandhi in Rugao Mandarin (Huang 1999), and
the domain for obligatory tone deletion in Shanghai Chinese (Selkirk & Shen 1990).

2.3.2. The prosodic phrase

The prosodic phrase is the prosodic constituent immediately higher than the prosodic
word along the prosodic hierarchy. A prosodic phrase is usually composed of two or
three prosodic words, even though sometimes a 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).

Some researchers argue that the boundaries between prosodic phrases are usually
constrained by the syntactic constituent (e.g., a noun phrase with a noun as head or a verb
phrase with a verb as head) (Nespor & Vogel 1986; Selkirk 1986; Shattuck-Hufnagel &
Turk 1996). However, Ladd (2008: 289) argues that the prosodic and syntactic structures
should be defined on their own terms; otherwise the discussion on the prosody-syntax
interaction will be circular. Ladd argues that the definition of the prosodic constituent, with reference to the syntactic constituent, should not be viewed as a definition; instead, it should be interpreted as “hypotheses, predictions about the correspondence between one type of independently definable structure and another” (Ladd 2008: 289).

In the mainstream American English ToBI (Tones and Break Indices) (MAE-ToBI) transcription system (Beckman, Hirschberg, and Shattuck-Hufnagel 2005), the equivalent of the prosodic phrase is the intermediate phrase, which encompasses one or more pitch accents and is marked with a phrase accent. In addition to the tonal marking, the intermediate phrase has other cues associated with the phrasing, such as lengthening of the phrase-final syllable and pitch reset across phrases.

In Mandarin Chinese, the prosodic phrase usually consists of two or three prosodic words that may have a close semantic and/or syntactic relationship, although they may only contain one prosodic word in some cases. The prosodic phrase corresponds to the minor and the major phrase in the Pan-Mandarin ToBI transcription system, depending on whether there is perceived pause after the phrase (Peng et al. 2005). The prosodic phrases in Mandarin are usually marked by a perceived major break, and a minor pause, as well as pre-boundary lengthening, and pitch reset across phrases (Peng et al. 2005; Tseng & Chou 1999; Tseng et al. 2005).

---

4 Pitch accent, phrase accent and boundary tone are explained here. A pitch accent falls on a stressed syllable in an intermediate phrase, marking the prominence of the syllable. The phrase accent, either high (H-) or low (L-), marks the end of an intermediate phrase in English. A boundary tone, high (H%) or low (L%), marks the end of an intonational phrase.

5 Pitch reset refers to the phenomenon that a speaker increases his/her pitch to a higher level in an utterance. Pitch reset is usually an auditory cue to the start of a new prosodic unit.

6 The phrase is a major phrase when there is a perceived pause at the break.
2.3.3. The intonation phrase/breath group/prosodic group

The intonation phrase is the largest prosodic chunk into which utterances are divided. The intonation phrase is the domain of a perceptually coherent intonational contour, or tune (Shattuck-Hufnagel & Turk 1996: 210). It is within this prosodic domain that a contour or tune (defined by a combination of high and low tones) is associated or aligned with the segmental information (Beckman & Pierrehumbert 1986). This constituent is also treated as the intonational unit (IU) or breath group in other frameworks (Chafe 1994; Himmelmann & Ladd 2008). In the American English ToBI transcription system (Beckman, Hirschberg, and Shattuck-Hufnagel 2005), an intonational phrase is comprised of one or more intermediate/prosodic phrases, and ends with a boundary tone, which is high (H) or low (L). The use of boundary tones indicates the relationship between the current utterance and the subsequent 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: 305).

The boundary of the intonational phrase often matches that of the major syntactic boundary. Nonetheless, the intonational phrase boundary of an embedded sentence might be different from its syntactic boundary. Selkirk (1984: 27) mentions that the intonational phrase is not a syntactic unit, because a sequence of components with which an intonational contour is associated may not necessarily be a constituent in the syntactic structure.
The intonational phrase is one of the prosodic constituents analyzed in some studies of Mandarin Chinese (Chu & Qian 2001; Wang 2003, among others). Tseng (2003) argues that phrase-level intonation in Mandarin is not as significant as that in atonal 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. A breath group is what a speaker utters with one full breath. In the Pan-Mandarin ToBI transcription system (Peng et al. 2005), a distinction is made between the breath group and the prosodic group, with the latter marked by a prolonged pause. Due to the lack of categorical difference between breath group and prosodic group, Peng et al. (2005) suggest that these two constituents might be collapsed later. The present study takes the prosodic group as a prosodic constituent above the prosodic phrase in Mandarin Chinese. The prosodic group corresponds to the intonation phrase in English. The prosodic group is marked by pre-boundary lengthening, audible pause and clear pitch reset across boundaries (Peng et al. 2005).

Worth noting is that the breath group and the prosodic group are sometimes considered to be prosodic constituents in discourse prosody (such as in Tseng et al. 2005), which means that both of these constituents can be larger than utterances.

2.4. Prosodic markings for prosodic phrases in English and Mandarin Chinese

The prosodic phrase is a constituent larger than the prosodic word and smaller than the intonational phrase in English and the prosodic group in Mandarin Chinese. The prosodic phrase has been attested to be very important in the prosodic and rhythmic
organization for both English and Mandarin Chinese (English in Kim 2006; Mandarin in Cao 1999, 2002). In the next two subsections, the F0 and duration patterns of prosodic phrases in English and Mandarin Chinese will be discussed in detail.

2.4.1. F0 patterns of prosodic phrases in English and Mandarin Chinese

The prosodic/intermediate phrase in English is marked by one or more pitch accents (prominence on syllables), together with a phrase accent. The phrase accent has a scope over the entire prosodic phrase (Pierrehumbert & Hirschberg 1990: 302). The phrase accents are associated with coherence and cohesion in the discourse. There are two types of phrase accents in English. The high (H) phrase accent indicates that “the current [prosodic] phrase is to be taken as forming part of a larger composite interpretive unit with the following phrase,” whereas the low (L) phrase accent emphasizes the separation of the current phrase from the subsequent phrase. In an English intonational phrase, there might be one or more prosodic phrases. Beckman, Hirschberg, and Shattuck-Hufnagel (2005) point out that there is pitch reset across the boundaries of intermediate phrases. However, Fon (2002) does not observe pitch reset across the boundaries of intermediate phrases in English.

In Mandarin Chinese, the prosodic group (PG)-initial and PG-final prosodic phrases (PPhs) have distinct intonation patterns (Tseng et al. 2005: 289). Specifically, the intonation of the PG-initial PPhs is marked by a F0 reset before declining rapidly, but the decline stops short before reaching the F0 minimum. The PG-final PPh also possesses a F0 reset, but not to the point of the PG-initial reset, and then the contour trails to an
ending with final lengthening. The F0 contours of the PG-medial PPhs are held somewhat flat. Figure 2.3, cited from Tseng et al. (2005: 288), illustrates the F0 patterns of prosodic phrases occurring at different positions of a prosodic group in Mandarin.

![Figure 2.3. Schematic F0 contours of five prosodic phrases in a prosodic group in Mandarin Chinese](image)

Tseng et al. argue that the F0 features of the three PG positions (PG-initial, PG-medial, and PG-final positions) reflect the different functions of intonation. Specifically, the F0 reset and the non-terminal fall in PG-initial prosodic phrase indicate a new beginning to be followed by more speech. The flatter and less distinct F0 pattern in PG-medial prosodic phrases indicates a continuing effect, while another lower reset and the following gradual decline and final lengthening in the PG-final prosodic phrase indicate the approaching of the overall terminal effects. What Tseng et al. define as a prosodic group refers to a group of utterances; nonetheless, the same pattern of F0 may apply to the prosodic phrases within a single utterance.
Pitch reset is a fairly robust cue for the boundary between prosodic phrases in Mandarin Chinese (Fon 2002; Tseng et al. 2005; Yang and Wang 2002). Pitch reset is also used as an important cue for prosodic phrasing in Mandarin Chinese (Peng et al. 2005).

In an English utterance, only some syllables are pitch-accented (stressed and prominent) in addition to the phrase accents and boundary tones. Thus, some syllables in an English utterance may not have any tone targets; their F0 contours are merely the results of the interpolation\(^7\) of the pitch accent, the phrase accent and the boundary tone. Figure 2.4 shows the F0 contours in a statement “He said you would” spoken by a female speaker of American English.\(^8\)

---

\(^7\) F0 interpolation refers to the phonetic process in which the F0 contour over a syllable is derived from the carryover of the F0 contour on the preceding syllable and the anticipation of the F0 contour on the following syllable.

As can be seen in Figure 2.4, there is a high pitch accent (indicated by H*), a low phrase accent (indicated by L-) and a low boundary tone (indicated by L%) in the utterance. In this utterance, the high pitch accent on the syllable “said” and the low phrase accent on the syllable “you” primarily determine the F0 contour of the whole utterance. The F0 contour on the syllable “would” just follows the low phrase accent on “you” and continues to trail off till the end of the utterance (i.e., the effect of the low boundary tone in a declarative sentence).

In Mandarin Chinese, intonation or other pragmatic functions may influence the actual realization of tones, namely the occurrence of tone target undershoot, even to the

---

9 Tone targets refer to the phonological components of tones. T1 has high target only; T2 has low and high targets; T3 has low target or both low and high targets; and T4 has high and low targets. Tone target undershoot refers to the phenomenon that the target in a tone is not realized or only partially realized. For example, a rising tone (T2) may
extent that the lexical tones are entirely overridden. In other words, the F0 contours on those syllables may be just the result of interpolation of the tone targets on the adjacent syllables; thus, the underlying tone targets may become unrecognizable in context. Figure 2.5, cited from Wu (1996), shows the F0 contour of an interrogative utterance with an ending in down-drift, spoken by a Beijing actress in a movie. In Figure 2.5, all tones that become neutral tones are marked by N (neutral). Specifically, the syllables 'little', 'money', 'to come out', 'to come', 'the possessive particle in Chinese', and 'one interrogative sentence-final particle in Chinese', all carry the neural tone. In actual speech, it is not difficult to find such examples, especially in Beijing Mandarin speech.

---

surface as a level tone, due to the low target undershoot.
Figure 2.5. An interrogative with a sequence of neutral tones spoken by a Beijing actress in a movie

It is worth mentioning that the overriding of lexical tones only occurs when the tones on the syllables involved become neutral tones. This phenomenon is particularly salient in Beijing Mandarin, due to its tendency for tone neutralization. Thus, even if American CFL learners transfer the English intonation patterns, L2 Mandarin would be very similar to L1 Mandarin, with the only possibility that CFL learners may produce more unstressed syllables carrying neutral tones, even those which native speakers do not produce in a prosodic constituent.

Having discussed the F0 patterns of prosodic phrases in English and Mandarin Chinese, we can make some predictions about what may occur in L2 Mandarin Chinese.
However, before proceeding to these predictions, keep in mind that, as mentioned in Chapter 1, contrastive analysis can not fully predict what L2 will be like and it is very likely that more F0 patterns may occur in actual L2 data. Three predictions are made here.

**Prediction 1:** Assuming American CFL learners transfer their English patterns of pitch reset (Fon 2002) to their L2 Mandarin, it is expected that pitch reset would not be a salient cue to mark the boundaries between prosodic phrases in L2 Mandarin Chinese;

**Prediction 2:** If American CFL learners unconsciously use English phrase accents across prosodic phrases in their L2 Mandarin, L2 Mandarin speech might carry English-like intonation contours characterized by the H or L phrase accents. That is, if a prosodic phrase is semantically or pragmatically closely related to the subsequent one, the L2 speaker is likely to transfer the English H phrase accent at the end of the preceding prosodic phrase, and will produce a rising intonation. On the other hand, if a prosodic phrase is not closely related to the subsequent one, the L2 speaker is likely to transfer the English L phrase accent and will produce a falling intonation.

**Prediction 3:** American CFL learners may transfer their English intonation patterns to their L2 Mandarin and produce more unstressed syllables carrying neutral tones in a prosodic phrase than native speakers. However, another scenario may occur. CFL learners may not have internalized Mandarin tones, namely their tone system is rather underdeveloped. In this case, CFL learners may attempt to maintain the underlying tone
targets and produce all tones as fully as possible, with the result that the whole utterance may sound very staccato or robot-like. It is expected that if the second scenario were the case, the advanced learners would do better than the intermediate-level learners, assuming that the advanced learners have better prosodic productions than the intermediate learners and, hence, pattern more similarly to the native speakers.

2.4.2. Duration patterns of prosodic phrases in English and Mandarin Chinese

The prosodic phrase in English is marked by initial and final lengthening (Beckman, Edwards and Fletcher 1992; Fon 2002; Keating et al. 2003, among many others). However, lengthening at the phrase-initial and phrase-final positions is motivated by different factors: phrase-final lengthening is the summation of final lengthening and pitch accents, whereas phrase-initial lengthening is primarily due to pitch accent (Fon 2002).\textsuperscript{10} The relationship between pitch accent and lengthening is due to the fact that pitch accent can only be placed on stressed syllables and stressed syllables are inherently long. Moreover, syllable duration, more precisely, vowel duration, decreases as the number of syllables and the segments in the inter-stress interval (ISI) increases (Kim 2006). Even though there might be more than one ISI within a prosodic phrase, Kim’s finding shows that there is shortening effect for the phrase-medial syllables in an English prosodic phrase.

With respect to Mandarin Chinese, Chao (1968) finds that the last syllable in a tri-syllabic phrase has the primary stress, the first syllable the secondary stress and the

\textsuperscript{10} What Fon (2002) finds is that there is initial lengthening only when the initial syllable is pitch-accented. She does not claim that there is always initial lengthening in English.
middle syllable the least stress. Due to the close relationship between duration and stress perception (Duanmu 2000), Chao’s stress patterns of phrases in Mandarin Chinese suggest that the last syllable in a tri-syllabic phrase is the longest, the first syllable is the second longest, and the middle syllable the shortest. Most of the later studies confirm the phrase-final and phrase-medial duration patterns, but show some divergent patterns with respect to the duration patterns at the phrase-initial positions. Tseng et al. (2005) find that in a prosodic phrase there is both phrase-initial shortening and phrase-final lengthening, specifically lengthening of the last two syllables in a phrase. Cao’s findings (1999) partially agree with Tseng et al. (2005): in a PG (prosodic group)-initial PPh, there is slight initial shortening and final lengthening; but in a PG-final PPh, there is initial lengthening and slight final shortening. Cao’s findings suggest the complementary patterns of duration at the initial and the final positions of a prosodic group. Fon (2002) also provides evidence for phrase-final lengthening and shows that final lengthening can reflect the discourse hierarchy in Mandarin Chinese. In contrast, Xu and Wang (2009) find that within a (syntactic) phrase in Mandarin Chinese, there is phrase-initial and phrase-final lengthening, accompanied by phrase-medial shortening. Xu and Wang’s findings suggest that there is no difference in temporal patterns of prosodic phrases between English and Mandarin.

The above discussion shows that in terms of temporal patterns, English is most likely to be different from Mandarin Chinese with respect to phrase-initial lengthening. As for phrase-final lengthening, it is likely to be a language universal phenomenon (Fon 2002). According to Fon (2002), the initial lengthening in English is primarily due to the pitch
accent status of syllables. In most varieties of standard Mandarin Chinese, prosodic phrases always start with full-toned syllables. If L2 Mandarin speakers treat the full-toned syllables as accented, they may lengthen these initial syllables due to their accent status, thus transferring the English phrase-initial lengthening to their L2 Mandarin.

2.5. Studies on the acquisition of Mandarin prosody

2.5.1. Tone acquisition

Tone acquisition has received considerable attention among studies on the acquisition of L2 Mandarin Chinese (Bent 2004; Chen 1997, 2000; Miracle 1989; Shen 1989; Sun 1998; White 1981; Zhang 2007, 2010, among others). One common theme in most of these studies is the acquisition order of tones by L2 learners. Even though findings from different studies vary to some extent, most of the studies report that T2 and T3 are the most difficult for L2 learners, and accordingly, are the last to be acquired (see Sun 1998 and Zhang 2007 for details). However, there are limitations in many of these studies. Three of the limitations are pointed out in Chen (2000: 2): (1) the data used are limited, mostly reading data of a list of words or sentences in isolation, without an elicitation context; (2) in some elicited instances, only syllables at certain positions are examined; and (3) tone errors are only analyzed against the desirable citation forms, but not against the prosodic patterns in native English language (i.e., no attempt is made to determine the potential transfer of English intonation patterns). Chen (2000) takes these problems into consideration in his research design and uses connected natural speech as the data for analysis. Following Sun (1998), Chen (2000) takes a combined phonemic and
phonetic approach (namely, a dual perspective), in his tone error analysis. From this dual perspective, a produced tone was not only matched against the citation tone in Mandarin Chinese, but it was also matched against the learner’s own performance. Chen (2000) argues that in this way it can be seen how often a specific tone is produced correctly as compared to the target tone and how often a tone is misproduced to match a different tone. Sun (1998) and Chen (2000) claim that the dual perspective of tone production analysis can enable the researcher to examine their performance of tone production as well as their tone error patterns. While Chen (2000) improves upon the approach to tone analysis over previous research, his analysis falls prey to another problem. Chen extracted the tones from spontaneous speech and asked human judges to evaluate the accuracy of the produced tones. However, in doing so, even the tones from the speech of native speakers are likely to have a large percentage of “errors,” because tone target undershoot in context may result in tonal productions that differ drastically from the citation tones (Xu 1994; Bent 2005, among others). This is reflected in Chen’s new category of tones, the mid-tone. Even though this new category may reflect the phonetic reality in connected natural speech, it raises the question of the necessity and correctness of proposing such a category. As mentioned above, F0 co-articulation and target undershoot may even change a tone to the extent that it cannot correctly identified.\footnote{However, normal communication may not be impeded because listeners can resort to the context for clarification.} Among these tones, the mid tone (M) variant is most frequently used. However, it should be pointed out that this mid variant is not only present in L2 speech but is also frequently used in the speech of native Mandarin speakers. Thus, it may not be necessary to stipulate such a mid tone category
for Mandarin tones. This dilemma poses serious problems for research investigating tone errors among L2 learners. Since tones may change in context, judgment of tone errors should be conducted within utterances, instead of extracting tones from the context and judging the extracted tones in isolation.

2.5.2. Acquisition of utterance-level prosody of Mandarin Chinese

Compared to the abundant studies on tone acquisition, only two studies (to the best of my knowledge) have been conducted on L2 acquisition of utterance-level prosody in Mandarin Chinese, namely, Chen (2000) and Viger (2007).

Chen (2000) analyzes the tone errors above the word level and proposes three mechanisms to account for the tone errors in his study: 1) transfer of English intonation and sentence stress; 2) avoidance strategy (avoiding some tones), and 3) consistent substitution of a particular tone for another. Chen identifies some cases of interference from English intonation and stress, and some developmental patterns that have nothing to do with the first language. Even though Chen (2000) claims that his analysis of tone errors goes beyond word-level tone errors, his analysis does not differ significantly from previous analyses of tone errors in L2 Mandarin Chinese. This consequence stems from his approach to the interaction of tones and intonation. Chen follows other researchers in prioritizing tones over intonation and in analyzing tone errors. That is to say, tone errors are first identified; then these errors are associated with the potential transfer of English prosody (e.g., intonation and stress, etc). In contrast, White (1981) has already provided a more detailed and systematic analysis of the effects of English intonation on tone errors.
in L2 Mandarin. However, the fact that tone errors tend to occur much more frequently when tones are concatenated in an utterance than in isolation suggests that the analysis of the utterance-level prosody should precede that of word-level tone error analysis. Likewise, if one wants to find the patterns of tree distribution in a forest, he/she may need to see from above the whole forest instead of examining each individual tree first.

Viger (2007) examines the acquisition of Mandarin Chinese utterance-level prosody by American L2 learners, and finds that L2 Mandarin Chinese exhibits a striking absence of the utterance-level prosody that occurs in L1 Mandarin. For example, L1 Mandarin speakers produce a global raise of pitch throughout Mandarin echo and yes/no questions, but that is not exhibited in the speech of L2 learners. Viger also finds the lack of transfer of global English prosodic contours into L2 Mandarin speech. Viger does find some transfer effects on utterance-final syllables; L1 English speakers imported a final-syllable rise on the final syllable in Mandarin echo and yes/no questions. Five possible explanations are offered in Viger (2007) to account for why transfer did not occur more extensively and why in particular it was confined to the utterance-final position. These five explanations identified by Viger (2007: 102-103) are: 1) limited online processing resources, namely, global prosody is sacrificed for the sake of local prosody, 2) insufficient evidence in the L2 input for global prosody, as compared to lexical prosody, 3) lack of explicit instruction, 4) slow speech rate, and 5) lack of long-range intonational planning.

Chen (2000) suggests that the transfer of English intonation and sentence stress is partially responsible for the tone errors in L2 Mandarin speech. However, these findings
are based on observation of his data without taking tone categories into consideration.
Viger (2007) only finds utterance-final prosodic (i.e., intonation) transfer, and does not
find transfer of utterance-level prosody. However, she had examined the utterance-level
prosody without taking into account prosodic constituents below the level of utterance.
As mentioned earlier, there may be prosodic transfer at the prosodic phrase level as well.

2.6. Gaps in previous studies

There are some gaps and problems in the previous research. Five are identified and
discussed here. One, as can be seen from Section 2.4, studies on the acquisition of
utterance-level prosody in Mandarin Chinese are rather limited. However, a better
understanding of the underlying reasons for various kinds of tone errors in L2
spontaneous speech requires an understanding of utterance-level prosody in both L1 and
L2 Mandarin. Two, no previous study has taken the top-down approach (namely,
analyzing utterance-level prosody before that of lexical prosody) in examining the
acquisition of Mandarin prosody. It is expected that the top-down design for examining
the acquisition of prosody will complement the previous, widely-adopted, bottom-up
perspective (namely, analyzing the lexical prosody only or analyzing the lexical prosody
before that of the utterance-level prosody). Three, previous studies have not considered
the interaction of tones and intonation in preparing the stimuli. Four, although it is
claimed that syntax interacts with prosody (Kaisse 1985; Nespor & Vogel 1986; Selkirk
1986; Shattuck-Hufnagel & Turk, 1996), no previous studies have analyzed the effect of
syntactic structure on prosodic phrasing in L2 Mandarin. And lastly, five, as mentioned in
Chen (2000), the data used for tone/prosody analysis are very limited. In most cases, read speech is used. However, read speech is different from connected spontaneous speech.

In designing this study, all the above shortcomings were taken into consideration. However, even though ideally spontaneous speech should be used in examining the acquisition of prosody, this study also used read speech. The use of elicited read speech over that of spontaneous speech was determined by the research objectives of this study. In this study, we aim to examine how American CFL learners acquire Mandarin prosody. In order to separate lexical prosody from utterance-level prosody and to examine their interaction, a set of tone sequences containing all four lexical tones of Mandarin Chinese is needed. Additionally, in order to examine whether different syntactic structures influence the prosodic patterns in L1 and L2 Mandarin Chinese, two different syntactic structures were mapped out for each of the tone sequences. In other words, the stimuli collected are controlled. In spontaneous speech, we cannot control for tones or syntactic structures that the speakers use in their speech and we cannot obtain a full set of contrasting pairs. Thus, spontaneous speech was sacrificed for elicited read speech in this study in order to lay the groundwork for future research using spontaneous speech to test the experiment results.

2.7. Summary

This chapter discussed the prosody of Mandarin Chinese, the prosodic structure and the prosodic markings of prosodic phrases in English and Mandarin Chinese. We also observed some shortcomings and identified some problems in previous research. By
taking into consideration those gaps and problems, the present study aims to contribute to
the understanding of the acquisition of Mandarin prosody and to complement previous
studies. In the next chapter, we will present the experimental design of the study.
CHAPTER 3: METHODOLOGY

3.1. Introduction

This chapter discusses the experimental design of the study. Specifically, we present the stimuli used, the subjects, the recording procedure, and the approach of data transcription, acoustic measurements and data analysis.

3.2. Stimuli

As mentioned in Chapter 2, read speech was used in this study in order to control for the tones and syntactic structures in the data collected. When preparing the stimuli, we used all four lexical tones in Mandarin Chinese. Even though the four lexical tones can be combined in different ways, we only used three types of tone sequences consisting of these tones, namely compatible tone sequences, conflicting tone sequences and other tone sequences. Drawing on Xu’s study (1994), in a compatible tone sequence, the tone target at the offset of the preceding tone and the tone target at the onset of the following tone are identical (i.e., both are H targets or both are L targets), whereas, in a conflicting tone sequence, the tone target at the offset of the preceding tone and the tone target at the onset of the following tone are different (i.e., if one is a H target, the other is a L target; alternatively, if one is a L target, the other is a H target). Two examples of each tone
sequence, together with the schematization of tone targets in each sequence, are given in (1), (2) and (3).\textsuperscript{12}

(1) Compatible tone sequences

T2T4 alternating sequence: e.g., 罗燕谈论名利。Luó Yàn tánlùn míngli

“Luo Yan talks about fame and profit.”

\begin{diagram}
\begin{array}{cccccc}
35 & 51 & 35 & 51 & 35 & 51^\textsuperscript{13}
\end{array}
\end{diagram}

T4T2 alternating sequence: e.g., 孟岩爱读外文。Mèng Yán ài dú wàiwén

“Meng Yan loves reading foreign literature.”

\begin{diagram}
\begin{array}{ccccccc}
51 & 35 & 51 & 35 & 51 & 35 & 35
\end{array}
\end{diagram}

\textsuperscript{12} See Appendix A for a complete list of the tone sequences used in this study. Some tone sequences were constructed based on Liao (1994), Yuan (2004) and Yang and Chan (2010).

\textsuperscript{13} The two-digit number here represents different tones, as discussed in 2.2.1.
(2) Conflicting tone sequences

T2 sequence: e.g., 王明来拿羊毛。Liú Míng lái ná yángmáo

“Liu Ming comes to get wool.”

\[ / / / / / / \]
35 35 35 35 35 35

T4 sequence: e.g., 魏丽要卖腊肉。Wèi Lì yào mài làròu.

“Wei Li will sell bacon.”

\[ \_ \_ \_ \_ \_ \_ \_ \_ \]
51 51 51 51 51 51
(3) Other tone sequences

T1 sequence\textsuperscript{14}: e.g., 殷安轻摸猫咪。Yīn Ān qīng mō māomī.

“Yin An gently pets a kitty.”

\begin{center}
\begin{tabular}{cccccc}
55 & 55 & 55 & 55 & 55 & 55 \\
\end{tabular}
\end{center}

T3 sequence\textsuperscript{15}: e.g., 李伟想买野鸟。Lǐ Wěi xiǎng mǎi yěniǎo.

“Li Wei wants to buy wild birds.”

For each specific tone sequence, two pairs of sentences were used. The two sentences in each pair are identical segmentally and tonally, but differ syntactically. Examples of a pair of sentences for the T3 sequence are given in (4) and (5). Through this dissertation, “SS” refers to “sentence structure”. The two numbers that follow “SS” code the number of the syllables (the first for the subject and the second for the predicate). Hence, “SS-24” refers to the sentence that has a disyllabic subject and a quadrisyllabic predicate, while

\textsuperscript{14} T1 only has a high (H) target. Thus, the T1 sequence is a compatible tone sequence as well. However, there is only one target involved in T1, while there are two targets (H and L) in all the tones in both the compatible tone sequences and conflicting tone sequences. That is why the T1 sequence falls within the category of “other tone sequences” in this study.

\textsuperscript{15} There is no definitive conclusion as to what are the targets in a T3 (“21” or “214”). Thus, no schematic representation of the T3 sequence is provided.
“SS-33” refers to the corresponding sentence in the pair that has a trisyllabic subject and a trisyllabic predicate.


\[
[李敏 ] \text{subject} \ [xiǎng \ mǎi \ mǔ \ mǎ] \text{predicate}
\]

“Li Min wants to buy a mare.”


\[
[李敏 xiǎng] \text{subject} [mǎi \ mǔ \ mǎ] \text{predicate}
\]

“Li Minxiang buys mares.”

In preparing the sentences, efforts were made to make sure that all the segments in the sentences are sonorants. However, when it was impossible to use only sonorants, obstruents were used. Nonetheless, the use of obstruents was kept to the minimal so that the F0 track can be continuous. In two cases when it was impossible to come up with a pair of sentences satisfying the above requirements, one sentence in the pair used another morpheme with similar segments. These two cases are xí (习, ‘to study’) for xué (学, ‘to study’) and yùn (孕, ‘pregnant’) for yòng (用, ‘to use’).
Altogether there are 12 pairs of sentences used in this study. All 24 sentences are embedded in conversational scenarios. One scenario used to elicit the utterance containing the T2 sequence is given in (6).

(6) Mingtiān de gōngzuò ān pái le ma

A: 明天的工作安排了吗?

“Has tomorrow’s work been arranged?”

Ān pái le Wáng Míng lái ná yángmáo

B: 安排了。王明来拿羊毛。

“Yes. Wang Ming will come to get the wool.”

Wáng Míng lái ná yángmáo Hǎo wǒ qù zhùnbèi yìxià

A: 王明来拿羊毛？好，我去准备一下。

“Wang Ming will come to get the wool? Ok. I’ll go prepare for it.”
The 24 scenarios containing the 24 target sentences were randomized, with one practice scenario and one fatigue scenario placed at the beginning and at the end respectively. Each line of script in a scenario was coded as either A or B. In order to facilitate the recording, each subject was only responsible for either line A or line B in the whole process of recording. There are two blocks of recording scenarios and they are all the same except for the line coding. Each block consists of 52 scenarios, with the same 26 scenarios (24 target scenarios plus one practice and one fatigue scenario) repeated twice. The only difference between the first 26 scenarios and the second 26 scenarios lies in the line coding, namely, the same lines in the first 26 scenarios have the opposite line coding as compared to those in the second 26 scenarios. Thus, each and every subject produced two renditions for every single target sentence. In data analysis, only the second rendition was used, except for four cases in which the second rendition had severe disfluency and the first rendition was used instead.

As shown in the example scenario above, the person’s name, 王明 Wáng Míng, is underlined. The reason for doing so is to make sure that the subjects are aware that the underlined part is a proper noun of personal name. Due to the particular design in this experiment, it is not very evident whether the first two or the first three syllables form the name of a person. In addition, to give L2 learners more time to prepare for the recording, the recording materials were sent to them via email one day in advance.

---

16 A practice scenario was placed at the beginning of all the scenarios containing the target sentences, in order to help subjects get ready for the recording. A fatigue scenario was placed at the end of all the scenarios containing the target sentences to make sure that subjects were still attentive during the recording of the penultimate scenario.

17 The line coding refers to the coding of the each line in one scenario, like A or B in (6).
3.3. Subjects

Ten intermediate-level and ten advanced-level American CFL learners were recruited to participate in the recording. These two different levels of learners, intermediate and advanced, were recruited to examine whether there is any learning effect between the two groups of learners, and what are the persisting patterns of prosodic deviations in L2 Mandarin speech. The American CFL learners were recruited from the American CFL learners at the Department of East Asian Languages and Literatures, The Ohio State University. The control group consisted of ten Beijing Mandarin speakers who were recruited from the Chinese community at The Ohio State University. The ten native Mandarin speakers were all born in Beijing and grew up there before coming to the United States to study. All subjects reported no speech or hearing problems. They were paid $10 for their participation in the recording.

3.4. Recording procedure

The recording was conducted in a sound-attenuated studio at The Ohio State University. Prior to recording, the researcher gave instructions to the subjects. Then the subjects signed the consent form. The subjects could ask any question related to the study. They were also told that they could withdraw from the recording any time without any penalty.

In the process of recording, two subjects (all the recording pairs were matched for language status) worked together in the recording studio. The two subjects sat

---

18 See Appendix B for the complete information of the subjects in each group.
comfortably behind their microphones. The two microphones were side by side behind a 13.3” MacBook Pro laptop. The recording materials were presented through PowerPoint slides on the laptop. The investigator sat in front of the recording control panel outside the recording studio. The progress of the slides was controlled by the investigator through a remote control. If a subject misread one syllable or syllables segmentally, or had severe disfluency in his/her recording, he/she would be asked to re-read the whole scenario with his/her partner. The recording input was digitized at 44.1 kHz with a 24-bit resolution.

There was a five-minute break between recordings of the two blocks. The whole recording lasted less than 30 minutes. On each PowerPoint slide, the pinyin Romanization was placed above each character to facilitate learners’ character recognition. However, as can be seen from Figure 3.1, the screen capture of one slide, the font size (font size 20) of the letters of the Romanization is intentionally smaller than that of the characters (font size 32), thus enabling the subjects, especially the learners, to focus on the characters and not the Romanization.
3.5. Data transcription

The speech data were transcribed in accordance with the Pan-Mandarin ToBI transcription system (Peng et al., 2005). In the Pan-Mandarin ToBI transcription system, eight tiers were proposed. However, only four tiers were transcribed in this study. The four tiers are: word (in Chinese character), pinyin Romanization (with citation tones), tone targets (H and L targets on each syllable), and break indices. All the transcription criteria were kept consistent throughout the transcription. For the tone target tier, high (H) targets on T1, low (L) and H targets on T2, L target or L and H targets (if tone sandhi
occurs) on T3, and H and L targets on T4 were transcribed. Table 3.1 lists the tone targets transcribed on each tone.

<table>
<thead>
<tr>
<th>Tone</th>
<th>Tone values</th>
<th>Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>55</td>
<td>H</td>
</tr>
<tr>
<td>T2</td>
<td>35</td>
<td>L and H</td>
</tr>
<tr>
<td>T3</td>
<td>21 or 214</td>
<td>L or L and H</td>
</tr>
<tr>
<td>T4</td>
<td>51</td>
<td>H and L</td>
</tr>
</tbody>
</table>

Table 3.1. Tone targets in the four lexical tones of Mandarin Chinese

For the break indices, five levels of breaks were transcribed. Following Peng et al. (2005), the five levels of breaks (shortened as B)\(^{19}\) are defined as below:

(7) B0: reduced syllable boundary;

   B1: normal syllable boundary: the “default” case;

   B2: minor phrase boundary: no perceived pause;

   B3: major phrase boundary: have audible pause;

   B4: breath group boundary

From the definitions of breaks at different levels, it can be seen that both B2 and B3 correspond to the boundary between prosodic phrases, with the only difference being that there is audible pause for B3. B4 is the boundary between utterances. However, when the

\(^{19}\) Note that the breaks transcribed in this study do not strictly correspond to prosodic hierarchy as illustrated in Figure 2.1 in Chapter 2, due to the focus on the prosodic phrases.
pause between the two prosodic phrases within an utterance is too long, that is, longer than 500 ms, B4 is transcribed instead. B0 corresponds to the boundary between two fused syllables, while B1 is the boundary between two normal syllables. Since B1 is the default boundary between normal syllables, it is not marked in the transcription. Figure 3.2 gives an example of the transcription of one utterance used in this dissertation.

Figure 3.2. Transcription illustration of an utterance 陆蔚用慢用药 Lù Wèi yòng màn yòng yào “Lu Wei uses slow medicine”, produced by a female native speaker

To guarantee the reliability of break transcription, breaks in T2 and T4 sequences were tested for intra- and inter-transcriber reliability. For the intra-transcriber reliability test, the breaks in T2 and T4 sequences were transcribed twice, with a one-and-a-half month’s interval between the two transcriptions. The transcription agreement rate between them is 96.8%. For the cases of discrepancy between the two transcriptions, the
original utterance was listened to again and the final transcription was decided. For the inter-transcriber reliability test, another transcriber who is a Chinese doctoral student of phonetics, was hired to transcribe the same subset of data independently. The transcription agreement rate between her transcription and the author’s transcription is 93.5%. For the cases of discrepancy between the two transcribers, the two transcribers discussed those cases and came up with a final transcription. It is worth mentioning that both the intra- and inter-transcriber reliability tests show that the transcription of break indices is reliable.

3.6. Acoustic measurements

3.6.1. F0 measurements

The F0 values were measured on all tone sequences except for the T3 sequence. Due to the applications of T3 sandhi, the surface tone targets in the T3 sequence were expected to be much more unpredictable and complicated than those in the other tone sequences. The F0 values were measured by running a Praat script (Boersma and Weenink 2009) by Xu (2005-2010). This script measures the ten F0 values on the voiced portion of each labeled syllable by dividing the voiced portion into ten equal intervals. The F0 values obtained this way are time-normalized (namely, the duration of the tone-carrying syllables is treated as the same across all subjects) and accurately represent the F0 contours of the whole utterance. With these F0 values, the F0 contours of SS-24 and SS-33 in each tone sequence were compared and the effect of different syntactic structures on prosodic phrasing was identified.
In addition, the F0 values of the high and low targets on each syllable in all the tone sequences except for the T3 sequence were measured on the target tier of the transcription with Praat.

3.6.2. Duration measurements

In order to examine the duration difference between L1 and L2 Mandarin Chinese, duration values were measured on each syllable on the Romanization tier. One difficulty in measuring duration is how to segment syllables in utterances. The first two formants F1 and F2 on the spectrogram tier were the primary bases for syllable segmentation, though F3 was also referred to at times, especially in the segmentation of the lateral and other liquids. Most importantly, the segmenting criteria were kept consistent throughout the transcription process.

3.7. Data analysis

In the data analysis, speech rate and overall F0 patterns (including the F0 range, the minimal F0, the maximal F0) in each subject group were examined first, followed by the analysis of the prosodic phrasing, F0 and duration patterns at the utterance level, and the analysis of tone errors.

3.7.1. Speech rate in L1 and L2 Mandarin Chinese

The mean syllable duration was used as a measure of speech rate in this study. When the mean syllable duration of each subject was calculated, the utterance-final syllables
were excluded due to the utterance-final lengthening effect (Cao 1999; Fon 2002; Tseng et al. 2005, among others). The mean syllable duration of one each subject was calculated by averaging all the syllable duration values at all positions except for the utterance-final positions.

3.7.2. Overview of F0 patterns in L1 and L2 Mandarin Chinese

To obtain the maximum F0 (max F0), the minimum F0 (min F0) and F0 range in each subject group, only the T2 (namely, the rising tone), the T4 (namely, the falling tone), and the T2T4 and the T4T2 sequences were used, because all syllables except for those involving sandhi or tone errors in these tone sequences have both H and L targets. The F0 values of the H and L targets on each syllable in these four sequences were obtained first. Then, the F0 range was calculated by subtracting the F0 value of the L target from the F0 value of the H target on the same syllable. Finally, the data of the H and L targets and the F0 range of all subjects in each group were compared and contrasted.

Note that all the data from these four tone sequences were used without considering the influence of tone errors.

3.7.3. Prosodic phrasing and the duration and F0 patterns in L1 and L2 Mandarin Chinese

Following the analysis of the overall F0 patterns and speech rate in L1 and L2 Mandarin was the examination of the prosodic phrasing and the duration and F0 patterns in L1 and L2 Mandarin Chinese. To examine the prosodic phrasing in L1 and L2
Mandarin, we calculated the number of different levels of breaks occurring after each syllable position in SS-24 and SS-33 in all tone sequences across subject groups.

As regards the duration patterns for prosodic phrasing in L1 and L2 Mandarin Chinese, we compared and contrasted the differences by plotting the syllable duration of SS-24 and SS-33 in all tone sequences.

With respect to the F0 patterns for prosodic phrasing in L1 and L2 Mandarin Chinese, we first examined the patterns of pitch reset in L1 and L2 Mandarin Chinese. Previous studies have shown that the lower end of the pitch range does not change much for the kinds of pitch manipulation typical of declaratives (Ladd 1988 and Fon 2002). Thus, only high targets in the T1 and the T4 sequences were examined. Specifically, the F0 values on the syl-2 and syl-3 in SS-24 and those on the syl-3 and syl-4 in SS-33 were compared. If the F0 value of the high target on the syl-3 is higher than that on the syl-2 in SS-24 and the F0 value of the high target on the syl-4 is higher than that on the syl-3 in SS-33, it was deemed that pitch reset occurred. Then, the percentages of pitch reset occurring in the T1 and the T4 sequences were compared across subject groups.

Secondly, we compared and contrasted the differences of the time-normalized F0 patterns in SS-24 and SS-33 at the utterance level across groups by plotting the F0 contours of SS-24 and SS-33 in each tone sequence separately.

3.7.4. Patterns of tone errors in L2 Mandarin Chinese

The tone errors in each tone sequence of L2 Mandarin Chinese were analyzed separately. It is worth mentioning how the correctness of the surface tones was judged in
this study. Unlike previous studies, the correctness of tones in this study was judged within an utterance against the original sentence. In other words, the judge listened to the whole utterance, compared it to the original sentence, and then decided whether a tone is correct or not. In doing so, the effects of both phonetic and phonological sandhi were taken into consideration. Segmental deviation was ignored, since the focus of this study is on prosody. Moreover, this approach of tone judgment is similar to what listeners encounter in actual communication, namely, the judge knew what he/she would hear and based his/her correctness judgment on his prior knowledge of the sentence and his/her native intuition. The author of the dissertation judged all the tone data twice, with a month’s interval between the two judgments. The agreement rate of tone correctness judgment was 98%. For the cases of judgment discrepancy, the tones were judged again and a final judgment was made. Afterwards, another judge, a Ph. D student in Chinese, was hired to judge the correctness of one-sixth of the tone data. The tone judgment agreement rate between the hired judge and the author was 95%. In the cases of discrepancy in judgment, the two judges listened to the original tones in the utterances again and made a final judgment.

3.8. Summary

This chapter provided a detailed account of the methodology of the study. In Chapter 4, we will analyze the prosodic phrasing and the F0 and duration patterns in L1 and L2 Mandarin Chinese. In Chapter 5 we will analyze the tone errors in L2 Mandarin Chinese.
CHAPTER 4: UTTERANCE-LEVEL PROSODY IN L1 AND L2 MANDARIN CHINESE

4.1. Introduction

This chapter examines prosodic phrasing and \textit{F0} and duration patterns at the utterance level, two important aspects of utterance-level prosody. Before analyzing utterance-level prosody, we will provide an overview of the L1 and L2 Mandarin speech data (the min \textit{F0}, the max \textit{F0}, the \textit{F0} range and the speech rate).

To analyze the prosodic phrasing in L1 and L2 Mandarin Chinese, we examined the break distribution in SS-24 and SS-33 of all tone sequences across subject groups. As a reminder, SS-24 refers to the sentence containing a certain tone sequence with a disyllabic subject and a quadrisyllabic predicate, while SS-33 refers to the corresponding sentence of the same tone sequence with a trisyllabic subject and a trisyllabic predicate. With two syntactic structures for each tone sequence, the relationship between prosodic phrasing and syntactic structure can be revealed.

For the acoustic patterns for prosodic phrasing in L1 and L2 Mandarin Chinese, we examined the duration and \textit{F0} patterns in SS-24 and SS-33 within and across subject groups. It was expected that such duration and \textit{F0} analysis at the utterance level could reveal detailed prosodic differences between L1 and L2 Mandarin.
4.2. Overview of the speech data in L1 and L2 Mandarin Chinese

This section provides an overview of the L1 and L2 speech data. Specifically, we examine the speech rate and the overall F0 patterns, including minimum F0 (min F0), maximum F0 (max F0), and F0 range in the speech data of each subject group (native, advanced, and intermediate).

4.2.1. Speech rate in L1 and L2 Mandarin Chinese

In this study, we used the mean syllable duration of all the target sentences produced by all subjects in each subject group as the measure of speech rate. However, we eliminated the utterance-final syllables in order to exclude the effect of the utterance-final lengthening (Cao 1999; Fon 2002; Tseng et al. 2005, among others). From the procedure outlined in section 3.6.2, we obtained the data of the mean syllable duration of all the subjects in each group.

Then, Analyses of variance (ANOVA) were conducted on the data of the mean syllable duration, with subject group as independent variable. The ANOVA results show that there is a group effect on the mean syllable duration ($F_{2,27} = 23.7$, $p < 0.001$). Further t-tests with the Bonferroni correction\(^\text{20}\) were conducted to compare the mean syllable duration of the three subject groups. T-test results show that native speakers have faster speech rate than both the advanced and the intermediate learner groups (advanced: $t_{18} = -4$, $p < 0.001$; intermediate: $t_{18} = -8$, $p < 0.001$), but there is no significant difference in speech rate between the advanced and the intermediate learners.

\(^{20}\) The Bonferroni correction is a statistical adjustment for the multiple comparisons. If $n$ outcomes are tested, the alphas level (i.e., the significance level. 0.05 was used in this study) will be divided by $n$. In this study, the new alpha level is $0.05/3 = 0.017$. 

54
rate between the advanced and intermediate learner groups. The above results can be seen clearly in the three boxplots in Figure 4.1. In each boxplot, the five horizontal lines show the five-number summaries of a group of numerical data. The five-number summaries are the minimal observation, the lower quartile (the 25\textsuperscript{th} percentile which cuts off lowest 25% of the data), the median, the upper quartile (the 75\textsuperscript{th} percentile which cuts off highest 25% of the data), and the maximal observation. In a boxplot, outliers are indicated by discrete dots above the maximal observation or below the minimal observation.

![Mean syllable duration of three groups](image)

Figure 4.1. Mean syllable duration of three subject groups
4.2.2. Max F0, min F0 and F0 range in L1 and L2 Mandarin Chinese

According to the procedure outlined in section 3.7.2, the data of the max F0, min F0 and the F0 range for each subject group were obtained. Figures 4.2 and 4.3 present the max F0, min F0 and the F0 range for the male and female subjects in the three groups respectively.

![Box plots for max F0, min F0, and F0 range across subject groups: Female](image)

a. Max F0  
b. Min F0  
c. F0 range

Figure 4.2: The max F0, min F0 and F0 range across subject groups: Female
From Figure 4.2 pertaining to the female subjects, we can see that the min F0 for the female advanced learners is the lowest among the three female groups, while their max F0 is the highest among the three groups. The cumulative effect is that the female advanced learners have the widest F0 range among the three groups. Also worth noting is that two female learner groups, both the intermediate learners and the advanced learners, have wider F0 ranges than the female native speakers.

From Figure 4.3 it can be seen that for the male subjects, the intermediate learners have the lowest min F0 and max F0, while the native speakers have relatively higher min F0 and max F0 than the advanced learners. With respect to the F0 range, the male subjects in three groups have very close F0 range, even though the two learner groups are on the slightly higher side.
4.2.3. Summary

In this section, we provided an overview of the L1 and L2 Mandarin speech data. The results show that L2 Mandarin Chinese is slower than L1 Mandarin Chinese, and that the F0 range in the two learner groups is (slightly) wider than that in the native group. L2 speech is usually slower than L1 speech (see Gass and Selinker 2008) and the speech data in this study are no exception. However, the (slightly) wider pitch range in the two learner groups than the native group was not expected, because previous studies have shown that L2 usually has narrower pitch range than L1 (Viger 2007), and English speakers generally have narrower pitch range than Mandarin speakers (Chen 1974, in White 1981), too. However, since there are no data of L1 English for either learner group, we will not make any further analysis on the F0 range between L1 and L2. But, it would be worth examining whether the two learner groups in this study have wider pitch range in their L2 Mandarin than their L1 English, considering the fact that the pitch range of their L2 Mandarin is wider than that of the L1 Mandarin by the native speakers.

In the next sections, we will proceed to analyze the prosodic phrasing and the surface duration and F0 patterns in L1 and L2 Mandarin Chinese.

4.3. Analysis of prosodic phrasing and the surface duration patterns in L1 and L2 Mandarin Chinese

In this section, we examine the prosodic phrasing and the duration patterns in L1 and L2 Mandarin Chinese, leaving the surface F0 patterns to be discussed in Section 4.4.
Section 4.3.1 discusses the break distribution in different subject groups and examines how syntax interacts with prosody. Section 4.3.2 examines the durational patterns of prosodic phrasing in L1 and L2 Mandarin. Section 4.3.3 summarizes this section.

4.3.1. Break distribution in L1 and L2 Mandarin Chinese

In this section we analyze the break distribution in SS-24 and SS-33 within and across groups. Five levels of breaks were transcribed across syllable boundaries. Specifically, B0 is the boundary between two fused syllables and B1 is the boundary between two normal syllables. Both B2 and B3 are the boundaries between prosodic phrases. However, when there is an audible pause between the two prosodic phrases, B3 was used; otherwise B2 is used. Furthermore, when the pause between the two prosodic phrases is too long, that is, longer than 500 ms, B4 was transcribed. B4 was also used after utterance-final syllables.

To examine the break distribution in SS-24 and SS-33, we calculated the number of breaks of different levels occurring after each syllable position (namely, the syl-1 through the syl-6 position) in each sentence structure.

Figure 4.4 presents the break distribution in SS-24 and SS-33 across groups. The three figures in the upper panel show the break distribution in SS-24 across groups. The three figures in the lower panel show the break distribution in SS-33 across groups. The x-axis in all six figures represents syllable position (s1 for syl-1 position, s2 for syl-s position, and so on), and the y-axis represents the proportions of each level of breaks.
occurring after each syllable position. The square brackets along the x-axis display the syntactic structure of the syllables.

Figure 4.4. Break distribution in SS-24 and SS-33 across groups

21 I thank Professor Mary Beckman for making Figures 4.4 and 4.5 with R.
From the three figures in the upper panel, it can be seen that after the syl-1, the syl-5 and the syl-6 positions, there is not much difference in breaks across groups, with the only difference being that there are a few cases of use of B0 among the advanced learners and native speakers, indicating that there are some fused syllables in these two groups. Other than that, only B1 was used after the syl-1 and the syl-5 positions and only B4 was used after the syl-6 position. The difference in break distribution across groups mainly occurs after the syl-2, the syl-3 and the syl-4 positions. After the syl-2 position, B3 accounts for 50% of the breaks in the intermediate and advanced learner groups, and B1 and B2 comprise of the rest of the breaks. At the same position, native speakers primarily used B2, even though they also used some cases of B1 and B3. After the syl-3 position, native speakers only used B1, while both learner groups used some B2’s and B3’s. After the syl-4 position, the native speakers primarily used B1, with some cases of B2, while both learner groups used more B2 than the native group.

Syntactically speaking, SS-24 has two syntactic components: the disyllabic subject and the quadrisyllabic predicate. Also, it is worth repeating that all the subjects (the proper nouns of personal names) were underlined in the reading scripts in order to inform the subjects that they are persons’ names. It seems that the underlined names triggered the two groups of learners to make a big break, namely B3. By contrast, the native speakers were not influenced greatly by the underlined names and mainly treated the subject-predicate boundary as a minor prosodic phrase boundary, and mainly used B2. However, it is difficult to interpret the two groups’ use of almost 20% of B2 after the syl-
One possible interpretation is that learners tended to chunk longer phrases into smaller units, making a minor phrase break after the syl-4 position. After the syl-4 position, both learner groups used 20% of B2 and some cases of B3.

In order to see whether the syntactic structure in the predicate of SS-24 played a role, let us analyze the syntactic structure of the predicate in SS-24. Three SS-24’s have the syntactic structure of $[\sigma\sigma]^{22}_{NP}[[\sigma]_V[[\sigma]_V[\sigma\sigma]_{NP}]_{VP} (one T1 sequence, one T2 sequence, and one T3 sequence), two SS-24’s have the syntactic structure of $[\sigma\sigma]_{NP}$ $[[\sigma]_V[\sigma\sigma]_{NP}]_{VP} (one T2T4 sequence and one T4T2 sequences), and one has the syntactic structure of $[\sigma\sigma]_{NP} [[\sigma]_V[[\sigma\sigma]_{Adj}][\sigma]_{NP}]_{VP} (one T4 sequence). A detailed examination of the break distribution in the two learner groups shows that there are almost the same percentage of B2 and B3 after the syl-4 position in $[\sigma\sigma]_{NP}[[\sigma\sigma]_V[\sigma\sigma]_{NP}]_{VP}$ and in $[\sigma\sigma]_{NP}[[\sigma]_V[[\sigma]_V[\sigma\sigma]_{NP}]_{VP}$ and in $[\sigma\sigma]_{NP}[[\sigma]_V[[\sigma\sigma]_{Adj}][\sigma]_{NP}]_{VP}$. Thus, it seems that the different syntactic structure in the predicate has not led to more occurrences of B2 and B3 after syl-4 position in the two learner groups. The reexamination of the original data reveals that the learners’ disfluency and the tendency of using shorter prosodic phrases are two potential factors which led to the break distribution. By contrast, the native speakers did not have the issue of disfluency and mainly used B1 after the syl-4 position.

The three figures in the lower panel in Figure 4.4 show that the break distribution in SS-33 is more consistent across groups than that in SS-24. The biggest difference in the break distribution of SS-33 across the groups lies after the syl-3 position. For the native

---

$^{22}$ $\sigma$ represents a syllable.
speakers, the majority of breaks after the syl-3 position are B2, while for both learner groups, the majority of breaks are B3, with 20% - 30% of B2. Again, this pattern of break distribution after the syl-3 shows that the learners tended to have a greater break at this position than the native speakers. After the syl-4 position, both learner groups have some cases of B3 and a few B2 breaks.

To summarize, our speech data shows that in most cases the prosodic structure coincides with the syntactic structure in all three subject groups. However, learners tended to make a large break between the subject and the predicate, which may be the artifact of the way that the reading materials were presented for the recording, in that the underlined names may have led learners to produce a large break at the boundary between the subject and the predicate. Moreover, we find that the syntactic structure in the predicate does not influence the prosodic structure of that constituent. The more frequent occurrence of B2 and B3 after the syl-4 position in SS-24 in the two learner groups than in SS-33 suggests that the longer prosodic phrase (the predicate in SS-24) posed more difficulty for the learners. Thus, the learners tended to break up the longer prosodic phrases into smaller prosodic units. The difficulty that the longer prosodic phrases posed for the learners may be related to the working memory capacity in second language acquisition (SLA). According to Miyake and Shah (1999: 450), “working memory is those mechanisms and processes that are involved in the control, regulation, and active maintenance of task-relevant information in the service of complex cognition, including novel as well as familiar, skilled tasks”. Due to the limited working memory capacity among L2 learners, they usually needed greater cognitive efforts in producing
longer prosodic phrases. Thus, the use of bigger breaks within prosodic phrases, such as B2 or B3, becomes a useful strategy for L2 learners to produce longer prosodic phrases.

However, due to the short length of the sentences used in this study, the analysis of the prosodic phrasing in this section can only provide limited insight into the complicated interaction of syntax and prosody. An analysis of spontaneous speech is needed in further studies to probe more into the relationship between syntax and prosody.

4.3.2. Duration patterns of prosodic phrasing in L1 and L2 Mandarin Chinese

It is predicted in section 2.4.2 that the most salient difference in duration patterns between L1 and L2 Mandarin may occur at the phrase-initial position. If American CFL learners treat the full-toned syllables as accented, they may lengthen these initial syllables due to their accent status. Therefore, they would transfer their English phrase-initial lengthening to their L2 Mandarin (Fon 2002). However, the phrase- and utterance-final lengthening is more likely to be a universal phenomenon and there may not be significant difference between L1 and L2 Mandarin Chinese.

In analyzing the durational pattern of prosodic phrasing in L1 and L2 Mandarin, we eliminated several utterances from the two learner groups to make sure that SS-24 and SS-33 for the same tone sequence have the same breaks at all positions except for the syl-2 and the syl-3 positions. In doing so, we eliminated the potential confounding effect that the differences in breaks may have on the duration patterns. The mean syllable duration at different syllable positions in all tone sequences for each subject group are plotted in Figure 4.5. In that figure, the x-axis displays the syllable position in the sentence (s1 for
syl-1 position, s2 for syl-2 position, and so on), while the y-axis displays the mean syllable duration of a particular position in that sentence structure. The dotted line with the dark circle represents SS-24, while the solid line with the non-filled-in circle represents SS-33.

Figure 4.5 shows that the duration patterns in SS-24 and SS-33 are very similar across subject groups. Specifically, there is clear phrase-final lengthening at the syl-2

Figure 4.5. Syllable duration across groups and sentence structures

Figure 4.5 shows that the duration patterns in SS-24 and SS-33 are very similar across subject groups. Specifically, there is clear phrase-final lengthening at the syl-2
position in SS-24 and at the syl-3 position in SS-33, together with the utterance-final lengthening.

One of our predictions about duration difference between L1 and L2 Mandarin is that American CFL learners may transfer English phrase-initial lengthening into their L2 Mandarin speech. There are two possible syllable positions, the syl-3 position in SS-24 and the syl-4 position in SS-33, where phrase-initial lengthening can be examined. However, it is difficult to determine whether there is phrase-initial lengthening at syl-3 position in SS-24. In order to see whether there is phrase-initial lengthening at the syl-3 position in SS-24, the duration of the syl-3 in SS-24 should be compared with that of the syl-3 in SS-33. However, the syl-3 in SS-33 is at the phrase-final position and is accompanied by clear phrase-final lengthening. Thus, the syl-4 position is the only position where we can examine the phrase-initial lengthening across subject groups. The syl-4 in SS-33 is at the beginning of the second prosodic phrase, while the syl-4 in SS-24 is in the medial position of the second prosodic phrase. Thus, if syl-4 in SS-33 is longer than that in SS-24, then there is phrase-initial lengthening at the syl-4 position in SS-33. Let us examine the duration at syl-4 position in Figure 4.5. It can be seen that syl-4 in SS-33 is longer than that in SS-24 for the native group, suggesting that there is phrase-initial lengthening in the native group. The syl-4 in SS-33 is almost as long as that in SS-24 for the advanced learner group, indicating that there is no phrase-initial lengthening/shortening for the advanced learner group. And the syl-4 in SS-33 is shorter

---

23 As there is phrase-medial shortening in both English and Mandarin Chinese, the syl-4 in SS-24 may be shortened. Thus, the syl-4 position may not be an ideal position to examine the phrase-initial lengthening, either. However, since the phrase-medial shortening occurs in both languages and the effect of the phrase-medial shortening is usually not so large, it is still valid to examine the phrase-initial lengthening on the syl-4.
than that in SS-24 for the intermediate learner group, suggesting that there is phrase-
initial shortening effect for the intermediate learner group. Thus, the comparison of the
syllable duration at the syl-4 position between SS-24 and SS-33 shows that Prediction 1
in Chapter 2 was not borne out. On the contrary, the native group shows phrase-initial
lengthening, the intermediate learner group shows phrase-initial shortening, and the
advanced learner group shows no effect of phrasing on phrase-initial duration. The
duration patterns for the native group support the findings in Cao (1999) and Xu and
Wang (2009), but not those in Tseng et al. (2005). The duration patterns at the syl-4
position in the two learner groups suggest that the learners did not transfer the English
phrase-initial lengthening to their L2 Mandarin. However, the different duration patterns
at the syl-4 position for the two learner groups does show that there is a learning effect
about the phrase-initial lengthening in that the advanced learner group patterned more
similarly to the native group.

In addition, the two figures on the right of Figure 4.5 show an interesting pattern of
syllable duration in SS-24 for the two learners groups: the six syllables in SS-24 have the
alternating duration pattern of short versus long. It seems that both learner groups further
divided the quadrisyllabic predicate in SS-24 into two disyllabic chunks, regardless of the
original syntactic structure of the predicate. The duration pattern in SS-24 accords with
the more occurrences of B2 after the syl-4 position, as can be seen in Figure 4.4. This
duration pattern may be attributed to the difficulty that learners had in producing longer
prosodic phrases. Another, potentially more likely, interpretation is that learners
transferred the English rhythm patterns into their productions of L2 Mandarin. In English,
weak and strong syllables tend to alternate to form the English rhythm, and weak
syllables tend to be shorter and strong syllables tend to be longer (Beckman 1986). Thus,
the duration pattern in the two learner groups can be attributed to the transfer of English
rhythmic patterns.

4.3.3. Summary and discussion

The analysis of prosodic phrasing shows that prosodic phrasing tends to align with
syntactic structure. However, we did find that learners tended to produce a large break
across the boundary between the subject and the predicate. We attributed this to the
artifact of how the stimuli were presented to the subjects (i.e., the underlying of the
proper names in the subjects). It was also found that both learner groups tended to
produce B2 or B3 after the syl-4 position in SS-24. We attributed this more frequent use
of B2 and B3 to the difficulty that the production of longer prosodic phrases posed for
learners.

The analysis of duration patterns for prosodic phrasing shows that the most salient
duration pattern that indexes prosodic phrasing is the phrase-final lengthening in both L1
and L2 Mandarin Chinese. This result supports the findings in Xu and Wang (2009). In
addition, phrase-initial lengthening was found in the native group, but not in both learner
groups. However, the advanced learner group patterned more similarly to the native
group with respect to phrase-initial lengthening, indicating a learning effect, namely as
learners’ proficiency increases over time, their prosodic production improves. Another
duration pattern is that the two learner groups favored the short versus long alternating
duration pattern in SS-24, and they further divided the quadrisyllabic predicate in SS-24 into two disyllabic phrases. We argued that this duration pattern in SS-24 by the two learner groups is the result of the transfer of English rhythmic pattern in prosodic phrases.

In the next section, we will examine the F0 patterns in L1 and L2 Mandarin Chinese.

4.4. F0 patterns of prosodic phrasing in L1 and L2 Mandarin Chinese

In the analysis of F0 patterns of prosodic phrasing in this section, we made sure that SS-24 and SS-33 for the same tone sequence have the same breaks at all positions except for the syl-2 and the syl-3 positions. In addition to controlling for the breaks in SS-24 and SS-33, we also controlled for tones and used only those utterances that have the correct surface tones (sandhi tones were considered correct). That is to say, only utterances with the correct surface tones and correct breaks were used in the analysis.24 With all the utterances containing wrong breaks and tone errors eliminated, it was predicted that there may not be much difference in the surface F0 patterns between L1 and L2 Mandarin. However, if there are any differences at all, they would constitute real F0 differences between L1 and L2 Mandarin Chinese, and are not due to confounding factors, such as F0 differences derived from tone errors.

According to the procedures outlined in Chapter 3, the surface tones in the utterances were judged for correctness. After the judgment of the correctness of tones, the data were further classified into two groups: those utterances with tone or break errors and those utterances without either tone or break errors. Only the second group, those utterances

---

24 See Appendix C for the number of utterances used in the analyses of each tone sequence.
without any tone or break errors were used in the analysis of this section. Those utterances with tone errors will be discussed separately in Chapter 5.

With respect to the patterns of pitch reset, we analyzed the T1 and the T4 sequences that were produced by different subject groups. We calculated the frequency of pitch reset occurring across the boundary of the subject and the predicate. Specifically, if the H target on the first syllable in the predicate is higher than that on the last syllable in the subject, pitch reset is considered to have occurred.

In addition, the time-normalized F0 analysis was conducted to examine how the tone targets on each syllable were realized and how the tone targets on the adjacent syllables influenced each other. According to the procedures outlined in Chapter 3, the F0 values on each syllable were obtained and then plotted. Due to the usually higher pitch of female speakers as compared to male speakers, the F0 contours by male and female speakers were analyzed separately. The time-normalized F0 analysis did not include the T3 sequence, because the application of the T3 sandhi may lead to great variability in the surface F0 patterns in the T3 sequence. Different from the duration analysis, different tone sequences were analyzed separately in this section, in order to investigate whether different tone sequences would have influenced the realization of tone targets in the same manner in L1 and L2 Mandarin Chinese.

4.4.1. Pitch reset in L1 and L2 Mandarin Chinese

For the analysis of pitch reset, we examined the utterances containing the T1 and the T4 sequences. Figures 4.6 and 4.7 present the percentage of pitch reset occurring at the
syl-3 position in SS-24 and at the syl-4 position in SS-33 in the T1 and the T4 sequences.

In these figures, the x-axis represents the sentence structure, while the y-axis represents the percentage of pitch reset occurring in each subject group.

Figure 4.6. Percentage of pitch reset in SS-24 and SS-33 for the T1 sequence across groups

Figure 4.6 shows that pitch reset tended to occur more frequently in SS-24 than in SS-33 for all subject groups and pitch reset occurred more frequently in the learner groups than in the native group. Specifically, in SS-24, pitch reset occurred more frequently in the intermediate learner group than in the advanced learner group and there was more pitch reset in the advanced learner group than in the native group. With respect

---

25 See Appendix D for the numbers of the occurrence of pitch reset in the T1 and the T4 sequence.
to SS-33, pitch reset occurred most frequently in the advanced learner group. However, the percentages of pitch reset in SS-33 for all three groups are very low (ranging from 25% - 41%).

Figure 4.7 shows that pitch reset occurred much more frequently in the native group than the two learner groups, especially in SS-33. Native speakers used pitch reset most of the times (85%) across the boundary between the subject and the predicate in SS-33. The advanced learner group used pitch reset in only one third of the boundaries (31%) between the subject and the predicate in SS-33, while the intermediate learner group used pitch reset least frequently (21%). With respect to SS-24, the native group used pitch reset in one third of the boundaries between the subject and the predicate (40%). In contrast, the advanced learner group did not use pitch reset at all, while the intermediate learner group used pitch reset only in a small portion of the boundaries between the subject and the predicate (7%).
The occurrences of pitch reset in the T1 and the T4 sequences display different patterns. On the one hand, pitch reset occurred more frequently in SS-24 than in SS-33 and more frequently in the learner groups than in the native group for the T1 sequence. On the other hand, pitch reset occurred more frequently in SS-33 than in SS-24 and more frequently in the native group than in the learner groups for the T4 sequence. Thus, it is worth investigating why the patterns of pitch reset in the T1 and the T4 sequences are so different.

In the T1 sequence, the tones for all the syllables have high targets only. The second prosodic phrase (i.e., the predicate) in SS-24 of the T1 sequence has four syllables. Due to the effect of declination (i.e., the downward trend of F0 in an utterance) on the F0 production, speakers tended to start at higher pitch in the second prosodic phrase in SS-
24 so that they could produce a long interval of high level F0 contour. That is why pitch reset occurred very frequently across the boundary between the subject and the predicate in SS-24 for the three subject groups. In addition, declination tends to influence the L2 learners more than the native speakers, due to the transfer of English declarative intonation. As a result, the learners tended to use pitch reset to offset the effect of declination more frequently than the native speakers in both SS-24 and SS-33. That would explain why pitch reset occurred more frequently in the learner group that in the native group.

In the T4 sequence, all tones have both a high and a low target. The existence of low targets results in downstepping of F0 (i.e., the downward shift of F0 production) (Pierrehumbert 1980). However, low targets may not lead to F0 downstepping in L1 Mandarin, at least not so much as in L2 Mandarin. As a result, the F0 contours in the T4 sequence produced by the learners were affected by both downstepping and declination, but more by downstepping than by declination, while the F0 contours of the T4 sequence produced by the native speakers were only affected by the declining trend of F0 (i.e., declination). The combined effects of downstepping and declination on F0 in the learner groups may have surpassed the effect of pitch reset on F0. Thus, even if pitch reset occurred in the learner groups, it may not have shown up in the surface F0 contours. That would explain why pitch reset occurred more frequently in the native group than in the learner groups for the T4 sequence. Figure 4.7 shows that pitch reset occurred in SS-33 more frequently than in SS-24 for both the learners and native speakers. An examination of the utterances containing the T4 sequence produced by the native
speakers show that some native speakers tended to place narrow focus on the syl-4 in SS-24. Due to the placement of narrow focus on the syl-4, the F0 range over the pre-focus syllable (i.e., the syl-3), was somewhat compressed (Liu and Xu 2005). Thus, pitch reset did not occur in SS-24 as frequently as in SS-33 for the native group. However, a different interpretation is needed for the more frequent occurrence of pitch reset in SS-33 than in SS-24 for the two learner groups because learners did not place narrow focus on the syl-4 in SS-24. In SS-33, the first prosodic phrase has three syllables. The combined effect of downstepping and declination would drop the F0 to a very low level in the learner speech. In order to produce another prosodic phrase with three falling tones in SS-33, learners had to reset their pitch across the boundary between the subject and the predicate; that would result in the more frequent use of pitch reset.

It is predicted in Chapter 2 that the L2 learners may not use pitch reset across the boundaries of the two prosodic phrases. The patterns of pitch reset in the T4 sequence confirmed our prediction. However, the patterns of pitch reset in the T1 sequence did not confirm our prediction. On the contrary, it was found that pitch reset occurred more frequently in the learner groups than in the native group. The findings of the patterns of pitch reset in the three subject groups show that the occurrence of pitch reset in an utterance is dependent upon several factors, including the underlying tones, downstepping, declination, and the weighting of downstepping and declination. Therefore, our prediction 1 was only partially borne out and it was difficult to generalize this pattern of pitch reset to L1 and L2 Mandarin.
4.4.2. Time-normalized F0 analysis in L1 and L2 Mandarin Chinese

4.4.2.1. F0 patterns in the T1 sequence

This section analyzes the F0 patterns in the T1 sequence in the three groups. A pair of sentences containing the T1 sequence used in this study is given in (1) and (2).

(1) 殷安青摸猫咪。Yīn Ānquīng mō māomī. “Yin Anqing pets a kitty.”

(2) 殷安轻摸猫咪。Yīn Ān qīng mō māomī. “Yin An gently pets a kitty.”

As mentioned in Section 3.6.1, Xu’s Praat script (Xu 2005-2010) was used to obtain the ten F0 values on the voiced interval of each syllable. The mean time-normalized F0 values were obtained by averaging all the F0 values at each syllable position in those utterances with correct tones and breaks produced by all subjects in each group. The mean time-normalized F0 values on each syllable for SS-24 and SS-33 in the three groups were plotted, with separate plotting for male and female subjects. In these figures, the x-axis displays the syllable position; the y-axis displays the F0 values (Hz). The legend on the right of each figure shows the different sentence structures.

4.4.2.1.1. F0 patterns in the T1 sequence: Native group

Figure 4.8 shows the F0 contours in the T1 sequence by the female native speakers. It can be seen that there is clear F0 reset at the beginning of the second prosodic phrase (i.e., the predicate) in SS-24. However, there is no clear pitch reset at the beginning of the
second prosodic phrase in SS-33 and the F0 contour in SS-33 is gradually declining, indicating the F0 declination. The bigger pitch reset in SS-24 can be attributed to the length of the second prosodic phrase (i.e., the predicate) in SS-24 in that people tend to raise the F0 height before a long prosodic phrase, which confirms our discussion in 4.4.1. Other than that, there is no significant difference in F0 contours between SS-24 and SS-33.

Figure 4.8. F0 contours in the T1 sequence: Female native speakers

![Figure 4.8. F0 contours in the T1 sequence: Female native speakers](image)

Figure 4.9 shows the F0 contours in the T1 sequence by the male native speakers. One salient difference in F0 contours between SS-24 and SS-33 is the overall higher F0 contour in SS-33 than in SS-24. Other than that, the F0 contours are rather similar.
between SS-24 and SS-33. From both Figures 4.8 and 4.9, it can be seen that the F0 contour over the first syllable in both SS-24 and SS-33 is rising, because the speakers needed to raise F0 from the usual mid level to a higher level. It is also worth mentioning that the F0 contours over most syllables in Figures 4.8 and 4.9 are slightly falling, instead of maintaining the high level contour of T1. Since there is no low target in T1, the falling tendency on those tones can only be attributed to declination in Chinese (Shi 2000).

Figure 4.9: F0 contours in the T1 sequence: Male native speakers

To summarize, the analysis of the F0 contours in the T1 sequence for the native group shows that there is no significant difference in F0 contour between SS-24 and SS-
33, except for pitch reset in SS-24 for the female native group. Thus, the effect of phrasing on the surface F0 patterns in the T1 sequence is very limited.

4.4.2.1.2. F0 patterns in the T1 sequence: Advanced learner group

Figures 4.10 and 4.11 present the F0 contours in the T1 sequence by the female and male advanced learners respectively.

Figure 4.10. F0 contours in the T1 sequence: Female advanced learners
Figures 4.10 and 4.11 show that, similar to the native speakers, the F0 contours on the first syllables in both SS-24 and SS-33 have the rising contour. Moreover, there is pitch reset at the beginning of the second prosodic phrase in SS-24 for both male and female advanced learners. It is also observed that the F0 contours of the second prosodic phrase in both SS-24 and SS-33 except for in SS-33 by the male advanced learners are declining steeply. The steep declining of the F0 contours in the second prosodic phrase suggests that, not only was the F0 production of the advanced learners influenced by the declination, but it was also influenced by the English declarative intonation (i.e., the falling intonation). Otherwise, the F0 contours could not have declined so steeply.
Figure 4.12 presents an utterance containing the T1 sequence produced by a female advanced learner. As shown in the figure, the F0 contour in the second prosodic phrase is falling drastically, which is very similar to the F0 contour of an English declarative sentence, as shown in Figure 2.4 in Chapter 2. Even though the tones produced were not wrong, only the syllable 轻 qīng ‘gently’ was stressed and carried the high tone. However, the other three syllables in the second prosodic phrase were not stressed at all. The F0 contour over the last three syllables in the second prosodic phrase was the result of the interpolation of the high tone on qīng and the low phrase accent on 猫 māo ‘cat’ and the boundary low tone at the end of the phrase. Thus, we see clear transfer of English declarative intonation in this utterance. Inevitably, such as utterance would lead to the perception of foreign accent.
To summarize, it was found that the advanced learner group displayed similar patterns of F0 contours to the native group. However, they did transfer the declarative intonation of English to their utterances and produced steeply declining F0 contour in the second prosodic phrase.

4.4.2.1.3. F0 patterns in the T1 sequence: Intermediate learner group

Figures 4.13 and 4.14 present the F0 contours in the T1 sequence by the female and male intermediate learners respectively.

![Figure 4.13. F0 contours in the T1 sequence: Female intermediate learners](image_url)

Figure 4.13. F0 contours in the T1 sequence: Female intermediate learners
Figure 4.13 shows that the F0 contour of SS-33 is gradually declining. There is pitch reset at the syl-3 and the syl-5 positions in SS-24. It seems that the intermediate learner further divided the second prosodic phrase in SS-24 into two phrases. That is why pitch reset occurred at the syl-5 position as well.

![Figure 4.14. F0 contours in the T1 sequence: Male intermediate learners](image)

Figure 4.14 shows that there is nearly no difference in F0 contour between SS-24 and SS-33 for the male intermediate learners, except for the higher F0 contour in SS-33. However, we did observe individual differences in the surface F0 patterns in the T1 sequence. Figure 4.15 presents an utterance containing the T1 sequence produced by a
male intermediate learner. It can be seen from the figure that the F0 contour in this utterance is very flat, which is very similar to the native pattern.

Figure 4.15: F0 contours in the T1 sequence: An utterance produced by a male intermediate learner

4.4.2.1.4. Summary of the F0 patterns in the T1 sequence

The analysis of F0 patterns in the T1 sequence shows that there is no significant difference in F0 contour between SS-24 and SS-33 in all subject groups, except that the F0 contours are often higher in SS-33 than in SS-24. However, there are some minor differences. For example, the advanced learner group transferred the declarative intonation pattern of English to their L2 utterances and produced declining F0 contours. It
is also worth mentioning there are individual differences in the surface F0 patterns among the learners.

4.4.2.2. F0 patterns in the T2 sequence

In this section, we analyze the F0 patterns in the T2 sequence in the three groups. A pair of sentences containing the T2 sequence is given in (3) and (4).

(3) 刘明来游云南。Liú Míng lái yóu Yúnnán. “Liu Ming comes to travel in Yunnan.”

(4) 刘明来游云南。Liú Míng lái yóu Yúnnán. “Liu Ming comes to travel in Yunnan.”

4.4.2.2.1. F0 patterns in the T2 sequence: Native group

Figure 4.16 displays the F0 contours in the T2 sequence by the female native group. It can be seen from the figure that the F0 contour on the syl-4 in SS-24 is relatively flat, indicating the effect of phrasing on tone target realization. The syl-4 in SS-24 is in the medial position of the second prosodic phrase in SS-24. Due to tone co-articulation (i.e., the effect of adjacent tones upon tone production), the low target on the syl-4 in SS-24 was not fully realized, namely low target undershoot occurred. Even though the F0 contour on the syl-4 in SS-24 is still gradually rising, perceptually it was a high level tone.
As will be discussed in Chapter 5, this is a case of T2 phonetic sandhi. Similarly, there is also low target undershoot on the syl-2 in SS-33.

4.16. F0 contours in the T2 sequence: Female native speakers

Figure 4.17 displays the F0 contours in the T2 sequence by the male native speakers. Similar to the female native speakers, there is low target undershoot on the syl-4 in SS-24 and the syl-2 in SS-33, showing the effect of prosodic phrasing on tone target realization. Figures 4.18 and 4.19 present the F0 contours in two utterances containing the T2 sequence produced by two male native speakers respectively. The low target undershoot in both SS-24 and SS-33 can be clearly seen in these two figures. Figures 4.16 through 4.19 also display another tendency of tone target realization. Both low and high targets
tended to be realized relatively fully when the tones are near the end of an utterance, even though the pitch range becomes narrower towards the utterance-final positions, as compared to the utterance-initial positions.

4.17. F0 contours in the T2 sequence: Male native speakers
Figure 4.18. The F0 contour in the T2 sequence (SS-24): An utterance produced by a male native speaker

Figure 4.19: F0 contour in the T2 sequence (SS-33): An utterance produced by a male native speaker
4.4.2.2. F0 patterns in the T2 sequence: Advanced learner group

Figure 4.20 displays the F0 contours in the T2 sequence by the female advanced learner group. It should be pointed out that only one utterance was used to plot the F0 contour in SS-33 in Figure 4.20. There are two female advanced learners in this study. However, both utterances containing the T2 sequence (SS-33) produced by the other learner and one utterance produced by this learner had tone errors. Thus, their utterances were not used. It can be seen from Figure 4.20 that there is low target undershoot at the syl-2 position in SS-33 and at the syl-4 position in SS-24. The pattern of target undershoot is similar to that of the native speakers.

![Figure 4.20. F0 contours in the T2 sequence: Female advanced learners](image)
Figure 4.21 displays the F0 contours in the T2 sequence by the male advanced learners. It can be seen that there is low tone target undershoot at the syl-2 position in SS-33 and at the syl-4 position in SS-24.

Figure 4.21. F0 contours in the T2 sequence: Male advanced learners

Both Figures 4.20 and 4.21 show that, except for the utterance-final syllable, the pitch range, the pitch range in the second prosodic phrase is narrower than that in the first prosodic phrase in both SS-24 and SS-33. Thus, the advanced learner group patterned with the native group.
4.4.2.2.3. F0 patterns in the T2 sequence: Intermediate learner group

Figure 4.23 display the F0 contours in the T2 sequence by the female intermediate learner group. Different from the F0 patterns in both the native group and the advanced learner group, the F0 patterns are very similar between SS-24 and SS-33 for the female intermediate learners. Even though there is some target undershoot at the syl-4 position in SS-33, the F0 contour is still rising. That is to say, the underlying tone targets were maintained by these learners.

![Figure 4.23. F0 contours in the T2 sequence: Female intermediate learners](image)

Figure 4.23 presents an utterance containing the T2 sequence produced by a female intermediate learner. It can be seen from the figure that all the rising tones maintain the
rising F0 contour, suggesting that her F0 production did not involve target undershoot. In other words, all tones maintained their underlying targets and were fully produced.

Figure 4.23: F0 contour in the T2 sequence (SS-24): An utterance produced by a female intermediate learner

Figure 4.24 displays the F0 contours in the T2 sequence produced by the male intermediate learner group. As shown in the figure, there is clear low target undershoot at the syl-2 position in SS-33. There is slight target undershoot at the syl-4 position in SS-24 and at the syl-5 position in SS-33. However, the rising F0 contours are still maintained.
Figures 4.22 and 4.24 also show that, for intermediate learners as well, with the exception of the utterance-final syllable, the pitch range in the second prosodic phrase is narrower than that in the first prosodic phrase, and that holds true for both SS-24 and SS-33.

4.4.2.2.4. Summary of the F0 patterns in the T2 sequence

The analysis of F0 patterns in the T2 sequence shows that low target undershoot occurred more frequently in the native group and the advanced learner group than in the intermediate learner group. More specifically, the low target undershoot tended to occur more often at the syl-2 position in SS-33 and the syl-4 position in SS-24, i.e., the phrase-medial positions of each prosodic phrase. Another F0 pattern is that, except for on the
final syllable, the F0 range in the second prosodic phrase is narrower than that in the first prosodic phrase. Declination in an utterance is mainly relevant to the high targets and there is declination in both English and Mandarin Chinese (Ladd 1988; Shih 2000). Thus, the narrower pitch range in the second prosodic phrase of the utterances observed above can be attributed to the effect of the declination upon the high targets in an utterance. It was also observed from the F0 contours in the T2 sequence that all three subject groups tended to produce the full rising contours towards the end of the utterances, but with narrower pitch range on the tones toward the end of the utterance. However, the F0 range on the utterance-final syllables in the utterances produced by the two learner groups was not narrowed, but widened. The widened F0 range on the utterance-final syllables in the learner groups suggests that the learners may have placed narrow focus on those syllables. The error in placing narrow focus on the utterance-final syllables may have been due to an artifact of the data-collection method in using read speech.

4.4.2.3. F0 patterns in the T4 sequence

This section analyzes the F0 patterns in the T4 sequence in the three subject groups. A pair of sentences containing the T4 sequence is given in (5) and (6).

(5) 魏丽要卖腊肉。Wèi Lì yào mài làròu. “Wei Li wants to sell bacon.”

(6) 魏立耀卖腊肉。Wèi Lìyào mài làròu. “Wei Liyao sells bacon.”
4.4.2.3.1. F0 patterns in the T4 sequence: Native group

Figure 4.25 displays the F0 contours in the T4 sequence by the female native group. It can be seen from the figure that the F0 contour at the syl-3 position in SS-24 is flat, which indicates the low target undershoot in the T4. Moreover, there is target undershoot at the syl-2 and syl-4 positions in SS-33. The F0 contours over these two positions are relatively flat.

![Figure 4.25. F0 contours in the T4 sequence: Female native speakers](image)

Figure 4.26 displays the F0 contours in an utterance (SS-24) containing the T4 sequence produced by a female native speaker. In this utterance, the syl-1, syl-4 and syl-6
are stressed, while the other three syllables are not. Target undershoot occurred on those unstressed syllables (i.e., the syl-2, syl-3 and syl-5).

![F0 contour in the T4 sequence (SS-24): An utterance produced by a female native speaker](image)

**Figure 4.26.** F0 contour in the T4 sequence (SS-24): An utterance produced by a female native speaker

Figure 4.27 displays the F0 contours of another utterance containing the T4 sequenced (SS-24) produced by a female native speaker. The main focus falls on syl-4 in this utterance. Due to the after-focus compression in pitch range (Liu and Xu 2005), the last two syllables in the utterance lost their tones. In addition, there is some target undershoot on syl-2 as well. The F0 contour over syl-3 became rising, due to its pre-focus position. The same pattern has been found in previous studies (Liao 1994). Not only did this speaker place the focus on syl-4, several other female native speakers did likewise.
The placement of focus on the syl-4 can also account for the F0 contours in SS-24, as shown in Figure 4.25 (i.e., the high F0 peak on the syl-4 position in SS-24).

Figure 4.27. F0 contours in the T4 sequence (SS-24): An utterance produced by a female native speaker

Figure 4.28 presents the F0 contours of an utterance containing the T4 sequence (SS-33) produced by a female native speaker. There is clear low target undershoot at syl-1, syl-2, and syl-4 positions, due to the tone co-articulation in the utterance.
Figure 4.28. F0 contours in the T4 sequence (SS-33): An utterance produced by a female native speaker.

Figure 4.29 displays the F0 contours in the T4 sequence by the male native speaker group. It can be seen that target undershoot occurred at the syl-2 and syl-4 positions in SS-33 and at the syl-3 position in SS-24.
Figure 4.29. F0 contours in the T4 sequence: Male native speakers

Figure 4.30 presents the F0 contour in an utterance containing the T4 sequence produced by a male native speaker. Clearly, there is target undershoot at the syl-3 and syl-4 positions.
Figure 4.30: F0 contours in the T4 sequence (SS-24): An utterance produced by a male native speaker

Figure 4.31 displays the F0 contour in an utterance containing the T4 sequence (SS-33) produced by a male native speaker. Target undershoot occurred at the syl-2 and syl-4 positions. It was also observed that the focus of the utterance fell on syl-3 and syl-5 positions, which explains the expanded F0 range on these two syllables.
Figure 4.31. The F0 contours in the T4 sequence (SS-33): An utterance produced by a male native speaker

4.4.2.3.2. F0 patterns in the T4 sequence: Advanced learner group

Figure 4.32 displays the F0 contours in the T4 sequence produced by the female advanced learners. Different from the F0 contours in the T4 sequence by the native speakers, the F0 contours in the T4 sequence between SS-24 and SS-33 are very similar for the female advanced learners, even though there is slight low target undershoot at the syl-1 position in SS-24. The similarity in F0 patterns between SS-24 and SS-33 suggests that the female advanced learners tended to produce each T4 as fully as possible.
Figure 4.32. F0 contours in the T4 sequence: Female advanced learners

Figure 4.33. F0 contours in the T4 sequence (SS-24): An utterance produced by a female advanced learner
Figure 4.33 presents the F0 contour in an utterance containing the T4 sequence produced by a female advanced learner. It can be seen that all falling tones maintained their falling contour, even though there is slight low target undershoot at the syl-1 position.

Figure 4.34 displays the F0 contours in the T4 sequence by the male advanced learners. Similar to the F0 patterns produced by the female advanced learners, the F0 patterns for the T4 sequence are very similar between SS-24 and SS-33 for the male advanced learners.
Figure 4.35 presents the F0 contours in an utterance containing the T4 sequence (SS-33) produced by a male advanced learner. In this utterance, target undershoot occurred at all syllable positions except for the syl-3 and syl-6 positions. Thus, the pattern of target undershoot in this utterance is very similar to that in the native group. However, most utterances produced by the male advanced learners tended to have similar F0 patterns for SS-24 and SS-33.

4.4.2.3.3. F0 patterns in the T4 sequence: Intermediate learner group

Figures 4.36 and 4.37 display the F0 contours in the T4 sequence by the female intermediate learner group and the male intermediate learner group respectively. In both
figures, the F0 contours are similar between SS-24 and SS-33. Even though slight target undershoot did occur, all T4’s in the intermediate learner group kept the clear falling F0 contours. Again, the F0 patterns suggest that the intermediate learners, similar to the advanced learners, tended to produce tones as fully as possible and did not produce target undershoot in their utterances.

Figure 4.36. F0 contours in the T4 sequence: Female intermediate learners
4.4.2.3.4. Summary of the F0 patterns in the T4 sequence

The analysis of F0 patterns in the T4 sequence shows that low target undershoot occurred more frequently in the native speech than in the speech of both intermediate learners and the advanced learners. The two groups of learners tended to produce each T4 as fully as possible, even though at times target undershoot also occurred. However, we observed individual differences on the target undershoot, namely, some advanced learners patterned very similarly to the native group in the realization of tone targets. It was also found that some native speakers tended to place focus on some key syllables in an utterance (i.e., 慢 màn, ‘slow’). As a result, the post-focus pitch range was compressed. On the other hand, the F0 contour over the pre-focus syllable was often rising, instead of falling, due to the effect of focus on the pre-focus syllable.
4.4.2.4. F0 patterns in the T2T4 alternating sequence

This section analyzes the F0 patterns in the T2T4 sequence in the three groups. A pair of sentences containing the T2T4 sequence is given in (7) and (8).

(7) 南梦来卖燃料。Nán Mèng lái mài ránliào. “Nan Meng comes to sell fuel.”

(8) 南梦来卖燃料。Nán Mènglái mài ránliào. “Nan Menglai sells fuel.”

Different from the F0 patterns in the conflicting tone sequence (i.e., the T2 and the T4 sequences), it is expected that the F0 contours in the compatible tone sequence (i.e., the T2T4 and T4T2 sequences) should be very similar between SS-24 and SS-33. Figures 4.38 through 4.43 display the F0 contours in the T2T4 sequence by different subject groups. It can be seen from these figures that the F0 contours between SS-24 and SS-33 are very similar. The only exception is that the F0 contours are not very smooth for the two learner groups. One potential explanation for the non-smoothness of F0 contours is that some of the utterances produced by each learner group have had been eliminated in the analysis. Due to the fewer utterances used in calculating the mean F0 values in each learner group, the mean F0 values would be more affected by the F0 values in each
individual utterance. Thus, the F0 contours for the learner groups are not as smooth as those for the native group.

Figure 4.38. F0 contours in the T2T4 sequence: Female native speakers
Figure 4.39. F0 contours in the T2T4 sequence: Male native speakers

Figure 4.40. F0 contours in the T2T4 sequence: Female advanced learners
Figure 4.41. F0 contours in the T2T4 sequence: Male advanced learners

Figure 4.42. F0 contours in the T2T4 sequence: Female intermediate learners
Figures 4.44 and 4.45 present the F0 contours of two utterances containing the T2T4 sequence produced by two male intermediate learners respectively. It can be seen from these two figures that, even though the F0 contours over the utterances are not very smooth, they have almost the same F0 contours as those produced by the native group, as shown in Figures 4.38 and 4.39.
Figure 4.44. F0 contours in the T2T4 sequence (SS-24): An utterance produced by a male intermediate learner

Figure 4.45. F0 contours in the T2T4 sequence (SS-33): An utterance produced by a male intermediate learner
To summarize, the analysis of the F0 patterns in the T2T4 sequence shows that the F0 patterns in SS-24 and SS-33 are very similar in all three subject groups. However, the F0 contours produced by the native group are very smooth, while those by the two learner groups are not so smooth.

4.4.2.5. F0 patterns in the T4T2 alternating sequence

This section analyzes the F0 patterns in the T4T2 sequence in the three groups. A pair of sentences containing the T4T2 sequence is given in (9) and (10).

(9) 孟岩爱读外文。Mèng Yán ài dú wàiwén. “Meng Yan loves reading foreign literature.”

(10) 孟言艾读外文。Mèng Yán’ài dú wàiwén. “Meng Yanai studies foreign languages.”

Figures 4.46 through 4.51 display the F0 contours in the T4T2 sequence produced by the three subject groups separately. Similar to the F0 patterns in the T2T4 sequence, the F0 patterns in SS-24 and SS-33 are very similar in all subject groups.
Figure 4.46. F0 contours in the T4T2 sequence: Female native speakers

Figure 4.47. F0 contours in the T4T2 sequence: Male native speakers
Figure 4.48. F0 contours in the T4T2 sequence: Female advanced learners

Figure 4.49. F0 contours in the T4T2 sequence: Male advanced learners
Figure 4.50. F0 contours in the T4T2 sequence: Female intermediate learners

Figure 4.51. F0 contours in the T4T2 sequence: Male intermediate learners
4.4.2.6. Summary of the F0 analysis of prosodic phrasing in L1 and L2 Mandarin Chinese

In this section, we analyzed the F0 patterns of prosodic phrasing in L1 and L2 Mandarin Chinese. With respect to pitch reset, we found different patterns in the T1 and the T4 sequences. In the T1 sequence, pitch reset occurred more frequently in the learner groups than in the native group and more frequently in SS-24 than in SS-33. In the T4 sequence, pitch reset occurred more frequently in the native group than in the learner groups and more frequently in SS-33 than in SS-24. Some interpretations were provided to account for these patterns of pitch reset in these two tone sequences. It was shown that, due to the effects of downstepping and declination, it was difficult to examine pitch reset by merely comparing the high targets on the two syllables across the boundary. The prediction of pitch reset in Chapter 2 was only partially borne out. However, no generalization was made about the overall patterns of pitch reset in L1 and L2 Mandarin Chinese.

The analysis of the time-normalized F0 patterns in different tone sequences shows that there are differences in the tone target realization in different tone sequences. In the T1 sequence and the compatible tone sequences (i.e., the T2T4 and T4T2 sequences), no significant differences were found between SS-24 and SS-33 across groups. However, in the conflicting tone sequence (i.e., the T2 and the T4 sequences), tone target undershoot occurred more frequently in the native speech than in the learner speech. These patterns of tone target realization show that the learners, especially the intermediate learners, have not mastered tone co-articulation; thus they tended to produce each tone as fully as possible. Another general pattern is that target undershoot tended to occur more
frequently in the first prosodic phrase and the beginning part of the second prosodic phrase. Moreover, except for on the utterance-final syllables in the learner groups, the pitch range in the second prosodic phrases is often narrower than that in the first prosodic phrase for all subject groups. This was attributed to the effect of declination on F0 production.

Our analyses also show that the F0 patterns in the T2T4 and the T4T2 alternating sequences are very similar between SS-24 and SS-33 and across groups. In this study, compatible tone sequences and conflicting sequences were used in order to examine whether different tone sequences impose different effects on the target realization across subject groups. Different patterns of tone target realization were found in different tone sequences. In the compatible tone sequences, there were nearly no differences in the tone target realization between SS-24 and SS-33 and across subject groups. Since the tone target at the offset of a preceding tone and the tone target at the onset of the following tone in a compatible tone sequence are identical, there is little possibility of target undershoot. That is why the surface F0 contours between SS-24 and SS-33 are almost the same in the compatible tone sequences in all subject groups. However, the surface F0 contours in the conflicting tone sequences are different between SS-24 and SS-33 and across subject groups. Specifically, the speech of native speakers involves target undershoot more frequently than the speech of learners. The differences in the target realization between native and learner groups can be explained by the tone sequences involved. In the conflicting tone sequence, the target at the offset of a preceding tone and the target at the onset of a following tone are different. Thus, the sudden change of the
tone targets on the adjacent syllables posed more difficulty in F0 production for the learners than for the native speakers. The sudden change of tone targets on the adjacent syllables in the conflicting tone sequences posed some difficulty for the native speakers as well, which led to the frequent target undershoot in the native speech. However, even if the conflicting tone sequences posed great difficulty in F0 production for the learners, they still attempted to maintain the underlying tone targets and produce each tone as fully as possible, with little target undershoot. As will be discussed in the next chapter, learners tended to produce more tone errors in the conflicting tone sequences. Therefore, the conflicting tone sequences were more difficult for the learners to produce than the compatible tone sequences.

To summarize, our F0 analysis shows that the main difference between L1 and L2 Mandarin lies in realization of the tone targets. It is expected that the failure to produce target undershoot at phrase-initial or phrase-medial position in L2 speech may render L2 Mandarin more staccato or robot-like. Further study should be conducted to examine whether the absence of target undershoot may lead to the perception of foreign accent.

4.5. Summary

In this chapter, we examined the prosodic phrasing and the surface F0 and duration patterns at utterance-level in L1 and L2 Mandarin Chinese. Our analyses show that prosodic phrasing mostly aligns with the syntactic structures in both L1 and L2 Mandarin Chinese. With respect to the duration patterns of prosodic phrasing, no significant differences were found between L1 and L2 Mandarin Chinese. The only exceptions are
the phrase-initial lengthening in the native speech and the transfer of English rhythmic pattern in L2 Mandarin Chinese. With respect to the time-normalized F0 patterns, it was found that conflicting tone sequences were more difficult for learners than compatible tone sequences. Specifically, the native productions involved more target undershoot than L2 speech. Individual differences in target undershoot were also observed in the utterances produced by some advanced learners, who also frequently produced target undershoot, similar to native speakers. A summary of the F0 and duration patterns at the utterance level in L1 and L2 Mandarin Chinese is provided in Table 4.1.

<table>
<thead>
<tr>
<th>Group</th>
<th>Initial Length*</th>
<th>Initial shorten*</th>
<th>Final Length*</th>
<th>Compatible sequence</th>
<th>Conflicting sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native group</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Same between SS-24 and SS-33</td>
<td>Frequent target undershoot</td>
</tr>
<tr>
<td>Advanc. group</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Same between SS-24 and SS-33;</td>
<td>1. Little target undershoot; 2. Most tones maintain underlying targets;** 3. More tone errors.</td>
</tr>
<tr>
<td>Intern. group</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Same between SS-24 and SS-33;</td>
<td>1. Little target undershoot; 2. Most tones maintain underlying targets; 3. More tone errors.</td>
</tr>
</tbody>
</table>

(*Length for lengthening; shorten for shortening.)  
(**Some advanced learners produced frequent target undershoot as well, similar to native speakers.)  
Table 4.1. Summary of the duration and F0 patterns in L1 and L2 Mandarin Chinese

This chapter investigated utterance-level prosody in L1 and L2 Mandarin Chinese by examining only those utterances with correct tones. In the next chapter, we will examine
those utterances containing tone errors and attempt to identify the patterns of tone errors in L2 Mandarin Chinese.
CHAPTER 5: PATTERNS OF TONE ERRORS IN CONTEXT

5.1. Introduction

In Chapter 4 we removed the utterances containing tone errors from the F0 analysis at the utterance level. This chapter will focus on those utterances containing tone errors and attempt to identify the patterns of tone errors in L2 Mandarin Chinese. Also examined in this chapter is the relationship between prosodic phrasing and tone errors. As previous studies have shown, different tones are acquired in different order by L2 learners and different tones may involve different error patterns in L2 speech. Thus, different tone sequences will be examined separately in this chapter.

5.2. Analysis of tone errors in L2 Mandarin Chinese

In connected speech, including read speech, Chinese shows great variability in the surface realizations of tones, due to the application of phonetic and phonological sandhi rules. Take the T3 sandhi for example. When T3 occurs before another T3 and these two T3’s are within the same prosodic domain, the first T3 will become a T2, namely a rising tone. In addition to T3 sandhi, there are many other phonetic sandhi applications which may change the underlying tones. In the following sections, the sandhi processes which may involve different tones will be discussed in detail. However, even though the
phonetic and phonological sandhi application may change the surface realizations of underlying tones, the sandhi tones are not errors, because they are predictable by applying the sandhi rules. In contrast, tone errors are not predictable from the application of any sandhi rule. Sometimes the tone errors in L2 speech may be random, especially among beginning learners of Mandarin, and no distributive patterns of tone errors can be discerned. In this study, the two groups of learners have studied Chinese for at least two years. Thus, their tone system may have started to display some patterns. However, there is no denying that, even though learners have some patterned errors in their L2 Mandarin, it does not mean that their tones cannot be further improved. With appropriate training and through their own efforts, learners’ tone production can be improved. The purpose of this chapter is to identify some of the patterns of tone errors in L2 Mandarin Chinese.

As mentioned in Chapter 3, the judgment of tone errors in this study was done in the context of utterances. Specifically, the judge listened to the utterance and compared it to the original sentence before making judgment on the correctness of tones in that utterance. Then, the utterances containing at least one tone error identified this way were singled out for analysis.

To analyze the tone productions, we calculated the correct utterance percentage (CUP, the percentage of the correct utterances produced by each subject group) and the correct tone percentage (CTP, the percentage of the correct tones produced by each subject group) for each tone sequence. A correct tone includes a sandhi tone as well. In order to examine the relationship between syllable position and tone errors, another parameter, accuracy weighting (AW), was calculated for all the utterances containing at
least one tone error. When AW was calculated, a surface correct tone was assigned the value of ‘1’, while a surface wrong tone was assigned the value of ‘0’. Then the mean AW values for all the utterances with tones errors in each subject group were obtained by averaging the values at each syllable position for all subjects in that group. The purpose of calculating AW is to see whether the surface tone errors have some position effects, such as phrase/utterance boundary effects.

Careful examination of the utterances containing at least one tone error shows that there are a fixed number of learners in both learner groups who were responsible for most of the errors in that group. Thus, each learner group was further categorized into two subgroups according to the frequency of tone errors in all tone sequences. That is to say, the intermediate learner group was further divided into the upper-intermediate group and the lower-intermediate group, and the advanced learner group into the upper-advanced group and the lower-advanced group. When subjects were recruited in this study, their class level, such as students from the third-year Chinese class or from the M.A. or Ph. D program in Chinese, was the sole criterion for selection. However, learners’ class level may not necessarily reflect their competence and performance in their linguistic production, especially prosodic production. It is not unusual to find some advanced L2 Mandarin learners who have very poor command of tones and prosody, even though their overall proficiency in L2 Mandarin Chinese may be very high.

---

26 See Appendix B for further categorization of each learner in the two learner groups.
In the following sections the surface tones realized in each tone sequence will be discussed respectively. Following that is the discussion and summary of the overall pattern of the tone errors in L2 Mandarin Chinese.

5.3. Tone errors in different tone sequences

5.3.1. Tone errors in the T1 sequence

T1 is a high-level tone (“55”) in Mandarin Chinese. There are no phonological or phonetic sandhi processes which involve T1. However, T1 may become a neutral tone when it occurs at the end of an utterance. Moreover, declination may affect the overall F0 contour of an utterance with only T1’s and leads to the gradually declining F0 contour over the utterance (Shih 2000). However, declination does not change the tone identity of T1. Therefore, in a sentence with only T1’s, the surface tones should still be T1 except for that on the last syllable, which might become a neutral tone or toneless. However, the tones on the last syllables of the sentences used in this study usually do not surface as neutral tones in standard Mandarin Chinese. Thus, if the tone on the last syllable of the sentence is toneless, it is considered a wrong tone.

Before discussing the tone errors in the T1 sequence, I will illustrate how accuracy weight (AW) was calculated by taking, as an example, the tone errors in the T1 sequence (SS-33) that were produced by the advanced learner group. The four surface tone sequences containing wrong tones are cited below (square brackets indicate the subject and the predicate in the tone sequence):
(1)  [H-H-H] [L-H-H]
     [H-H-H] [L-H-H]
     [H-H-H] [LH-H-H]
     [H-H-H] [H-H-L]

Note that T1 is a high-level tone. Thus, all non-High tones in (1) are tone errors. According to the definition of accuracy weighting (AW), all non-High wrong tones are assigned the value of ‘0’, while all surface high correct tones are assigned the value of ‘1’. Hence, (1) yields (2):

(2)  [1-1-1] [0-1-1]
     [1-1-1] [0-1-1]
     [1-1-1] [0-1-1]
     [1-1-1] [1-1-0]

Then, the average AW values for the group were obtained by averaging the values across each syllable position, as shown in (3):
The square brackets in (3) and in the tables below were used to mark the subject and the predicate in the target tone sequence, such as \([1-1-1]\) subject \([0.25-1-0.75]\) predicate.

Tables 5.1 and 5.2 list the correct utterance percentage (CUP), the correct tone percentage (CTP), and the accuracy weight (AW) values for the T1 sequence in the three groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>100%</td>
<td>100%</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced</td>
<td>70%</td>
<td>93%</td>
<td>[1-1] [0.5-0.33-0.83-0.83]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>70%</td>
<td>82%</td>
<td>[0.67-0.33] [0.5-0.17-0.17-0.5]</td>
</tr>
</tbody>
</table>

Table 5.1. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weight (AW) in the T1 sequence: SS-24
From Tables 5.1 and 5.2, it can be seen that both learner groups produced some utterances with tone errors, while native speakers did not. The CUP and CTP values show that the advanced learners did overall better than the intermediate learners. With respect to SS-24, the AW values show that the advanced learners have the most tone errors at the syl-4 position, followed by the syl-3 position, while the intermediate learners have the most tone errors at the syl-4 and the syl-5 positions, followed by the syl-2 position. A careful examination of the surface tones shows that the low ("21" or "22") and rising ("35") tones are the most frequent surface tones involved in those tone errors. To have a better idea of the error patterns, one utterance containing both the low tone and the rising tone errors produced by a male intermediate learner is presented in Figure 5.1. The utterance in Figure 5.1 is given in (4).

27 Note that the AW values were obtained by examining only the utterances with tone errors. Thus, there is no direct relationship between AW values and CTP values.
Surname safe should repair fly machines

“Wu An should repair the plane. “

Figure 5.1. Illustration of the low and rising tone errors in the T1 sequence (SS-24): An utterance produced by a male intermediate learner

The three circled areas in Figure 5.1 show the three errors in this utterance. The first circled T1 was produced as a rising tone and the last two highlighted T1’s were produced both as low tones. However, due to the production of creaky voice in this learner, a portion of the F0 on the two surface low tones was mis-tracked. Figure 5.1 also shows that there is an abrupt change of F0 height in the first two correct syllables and that the F0
contour on the last correct syllable is uneven. These F0 patterns suggest that this learner had difficulty in maintaining the high level F0 contour, even on a single syllable or two consecutive syllables. Thus, even if the learner produced the correct tones in the first two syllables, the F0 pattern are still different from those produced by native speakers.

Figure 5.2 gives another example of the low and rising tone errors in the T1 sequence produced by a different subject. The utterance in Figure 5.2 comes from the same target sentence as in Figure 5.1.

Figure 5.2. Another illustration of the low and rising tone errors in the T1 sequence (SS-24): An utterance produced by a male intermediate learner

In this utterance, all surface tones are wrong. The original high tone on the first syllable was produced as a rising tone, and the high tone on the second syllable became a
low tone. The high tones on the third and fourth syllables were produced as a rising tone and a low tone respectively. The high tones on the last two syllables both surfaced as rising tones. Perceptually, the six syllables form three feet, with the metrical structure in (5).²⁸

(5)    x
        x   x   x   x
        x   x   x   x   x   x
Wū ān yīng xiū fēi jī

Interestingly, even though this learner was speaking Mandarin, he was using English stress patterns, namely SW SW SW.²⁹ Not surprisingly, this utterance sounds like speech with a strong foreign accent.

The frequent tone errors occurring at the syl-4 and the syl-5 positions in SS-24 suggest that learners, especially the intermediate learners, had difficulty in maintaining a long-interval of high flat F0 contour. Thus, the intermediate learners alternated T1 with the low and rising tones in between.

With respect to SS-33, the AW values show that the advanced learners have the most tone errors at the syl-4 position, while the intermediate learners have the most tone errors at the syl-4 and the syl-5 position. Tone errors occurring at the syl-4 position mainly

²⁸ The “x” indicates prominence. The more x’s on a syllable, the more prominent that syllable is.
²⁹ S represents “strong”; W represents “weak”.

130
involve T1 surfacing as low tones or rising tones for both learner groups. Figure 5.3 is an example of a low tone error produced by an advanced learner, with the utterance presented in (6).

(6) Yīn Ān qīng mō māo mī.
殷 安 青 摸 猫 咪。
Surname    safe   green fondle cat cat sound

“Yin Anqing pets a kitty.”

Figure 5.3. Illustration of the low tone error in the T1 sequence (SS-33): An utterance produced by a male advanced learner
The circled area in Figure 5.3 shows that this learner produced T1 as a low tone in this utterance.30 Again, the long interval of high level F0 contour posed difficulty for the learner; in what should be a straight sequence of high level tones in this SS-33 structure, the subject produced a low tone on syl-4 followed by high level tones.

The AW values in Tables 5.1 and 5.2 show that tone errors are more likely to occur in the second prosodic phrase, especially at the beginning of the second prosodic phrase, such as at the syl-4 position in SS-33, than in the first prosodic phrase. I argue that these tone errors are due to the transfer of English intonation. Having produced the high level tones in the first prosodic phrase, it was very difficult for some learners to continue to maintain the high level F0 contour in the second prosodic phrase. Thus, learners tended to transfer the English intonation at the beginning of the second prosodic phrase. Due to the existence of rising and falling intonation in English, it is easy for the L2 learners to produce rising and falling F0 contours. However, the syllables at the beginning of the second prosodic phrase are semantically/syntactically closely related to the following syllables in the same phrase. Thus, the falling intonation would disrupt this close semantic/syntactic relationship between the phrase-initial syllable and the following syllables. Therefore, learners transferred the English rising intonation and used the rising F0 contours on the syllables at the beginning of the second prosodic phrases to index the close semantic/syntactic relationship.

30 Note that the beginning portion of F0 on this syllable was mis-tracked by Praat due to creaky voice.
It is worth mentioning that nearly all the tone errors in the T1 sequence and other tone sequences to be discussed below were made by the lower-intermediate and lower-advanced learners.

5.3.2. Tone errors in the T2 sequence

T2 in Mandarin Chinese is a mid-rising tone (“35”). The T2 sandhi pattern that is most widely discussed in the literature is that T2 becomes T1 when T2 occurs between a T1 or T2 and a non-neutral tone (Chao 1968; Hyman 1975; Lin 2007). Previous research has shown that this pattern of T2 sandhi most frequently occurs in prosodically weak positions, such as word-medial positions in casual fast speech (Xu 2007: 15). However, due to the different speech rate by different speakers, whether this sandhi rule applies depends on the actual speech production. This is why T2 sandhi was treated as a phonetic tone co-articulation rule in some literature (Chao 1968; Lin 2007:201). It is predicted that the surface tones in the T2 sequence would exhibit greater variability than those in the T1 sequence.

Tables 5.3 and 5.4 show the correct utterance percentage (CUP), the correct tone percentage (CTP), and the accuracy weight (AW) values for the T2 sequence in each group.
<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>100%</td>
<td>100%</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced</td>
<td>70%</td>
<td>93%</td>
<td>[1-0.67] [0.83-0.33-0.83-1]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>65%</td>
<td>90%</td>
<td>[0.71-1] [1-0.29-0.71-0.57]</td>
</tr>
</tbody>
</table>

Table 5.3. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weight (AW) in the T2 sequence: SS-24

<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>100%</td>
<td>100%</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced</td>
<td>45%</td>
<td>84%</td>
<td>[0.91-0.64-0.82] [0.36-0.73-0.82]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>50%</td>
<td>89%</td>
<td>[0.9-0.9-0.7] [0.4-0.9-0.9]</td>
</tr>
</tbody>
</table>

Table 5.4. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weight (AW) in the T2 sequence: SS-33

As predicted, there is much more variability in the surface tones of the T2 sequence in both the native group and the two learner groups. Examination of the original data shows that in SS-24, both the native group and the advanced learner group have only one utterance that does not involve any phonetic sandhi, while the intermediate learner group has six utterances without any phonetic sandhi. Thus, the advanced learners’ pattern is more similar to the native speakers in terms of co-articulation.

From the CUP and CTP values in Table 5.3, it can be seen that in SS-24 the advanced learners did better overall than the intermediate learners, while native speakers did the best among the three groups. The AW values show that in both learner groups, the most tone errors occur at the syl-4 position. The common pattern of tone errors in the two learner groups is that T2 surfaced as a low tone.
Figure 5.4 gives an example of such a tone error. The utterance in Figure 5.4 is given in (7).

(7) Wáng Míng lái ná yáng máo.

Surname bright come take sheep hair

“Wang Ming comes to get the wool.”

Figure 5.4. Illustration of the low tone error in the T2 sequence (SS-24): An utterance produced by a male intermediate learner
The circled area in Figure 5.4 shows that the learner produced the rising tone as a low tone, which perceptually sounds like T3 (i.e., the low variant) in Mandarin Chinese.

With respect to SS-33, the native group has only one surface tone sequence that does not involve any phonetic T2 sandhi (but with final syllable neutralization), namely, all surface tones maintained the underlying low and high targets. In contrast, the advanced learner group has three surface tone sequences that do not involve any T2 sandhi, and for the intermediate learners, four surface tone sequences (one sequence with final syllable neutralization).

Table 5.4 shows that both the CUP and CTP values for the advanced learner group in SS-33 are slightly lower than those in the intermediate learner groups. Even though statistically insignificant, these numbers show that the advanced learners did not do better than the intermediate learners in this tone sequence. In contrast, in SS-24, both CUP and CTP values for the advanced learner group are higher than those for the intermediate learner group.

Most of the tone errors in SS-33 occurred at the beginning of the second prosodic phrase, namely at the syl-4 position. The error pattern is the same as that in SS-24, namely, T2 surfaced as a low tone. Figure 5.5 gives an example of such tone errors in SS-33. The utterance in Figure 5.5 is given in (8).
Liu Minglai travels in Yunnan.

Surname bright waste field travel cloud south

“Liu Minglai travels in Yunnan.”

Figure 5.5. Illustration of the low tone error in the T2 sequence (SS-33): An utterance produced by a female advanced learner

The circled area in Figure 5.5 shows that this advanced learner produced the rising tone as a low tone.

Another typical tone error pattern in the T2 sequence for the advanced learner group, including both SS-24 and SS-33, is the misalignment of the H target of the preceding tone with the following tone. Figure 5.6 gives an example of such misalignment of H target in
the T2 sequence produced by a female advanced learner. The utterance in Figure 5.6 is given in (9).

(9) Wáng Míng lái ná yáng máo.

“Wang Minglai gets the wool.”

Figure 5.6. An example of H target misalignment in the T2 sequence (SS-33): An utterance produced by a female advanced learner
As can be seen in Figure 5.6, the H target of the first syllable on 王 Wáng is mis-aligned with the second syllable, rendering the second syllable a falling tone (HL). The slow speech rate of this learner may be one of the factors that lead to the H target misalignment, in that the learner might have no time to reach the low target on the second syllable “ming” occurring in a sequence of rising tones. However, even in slow speech, the native speakers will not produce such H target misalignment. Thus, the H target alignment shows that the learner has not fully acquired the tone phonology of Mandarin Chinese. Wu (1984) and Liao (1994) notice that T2 may become a T4 when it is weakened and occurs between two strong or stressed T2’s. However, in the above case, the misaligned T2 occurred on a strong syllable; thus, it was treated as a tone error. It seems that this advanced learner raised her F0 to a level so high on the first syllable that it was very difficult for her to make a quick transition in dropping her pitch to the low target on the second syllable. Thus, instead of dropping to the low target and raising the F0 to the high target on the second syllable, and dropping to the low target on the third syllable again, she just dropped the F0 to the low target of the third syllable. The result is that the F0 contour on the second syllable was merely the interpolation between the high target on the first syllable and the low target on the third syllable. The tone error here suggests that this learner had difficulty in quickly changing the tone targets.

The AW values in Table 5.3 show that more tone errors occur at the second syllable position of the second prosodic phrase, namely at the syl-4 position in SS-24. In SS-24, the few tone errors at the beginning of the second prosodic phrase suggest that T2 is easier when it is at the phrase-initial position. The more tone errors at the syl-4 position
in SS-24 shows that learners tended to make more errors when there is a sequence of rising tones. That is, a sequence of rising tones poses more difficulty for learners’ F0 control, as they have difficulty negotiating quick changes to reach the tone targets.

The AW values in Table 5.4 show that more tone errors occur at the initial position of the second prosodic phrase in SS-33. It seems that T2 is difficult for the learners when it is at the beginning of a prosodic phrase. However, the tone errors in SS-24 show that T2 is easier when it is at the beginning of a prosodic phrase. How can such divergent tone error patterns occur? An examination of the tone error patterns shows that the most frequent error pattern in SS-33 involves T2 surfacing as a low tone. When we were analyzing the tone errors in the T1 sequence, we suggested that the learners tended to produce rising and low tones to replace T1 in their speech. It seems that similar tone errors occurred in the utterances containing the T2 sequence. In SS-24, the few errors at syl-3 position (the beginning of the second prosodic phrase) may be accidental, because rising tones are the frequent error tones that learners used at this position. However, such tone errors were concealed in the utterances containing the T2 sequence, because the surface rising tones may have be produced as errors but were deemed correct, based on the judgment of the surface tone values. In SS-33, there are more tone errors (low tones) at the syl-4 position (the beginning of the second prosodic phrase). The question is why low tones were produced instead of rising tones at the syl-4 position in SS-33. This may be related to the length of the two prosodic phrases in SS-33. Both prosodic phrases in SS-33 contain three syllables and more tone target changes would be involved if all three syllables were T2. Thus, the difficulty of producing a sequence of rising tones led the
learners to produce a tone error, namely the low tone, at the syl-4 position in SS-33. In this sense, I argue here that the low tone error serves as a buffer zone for the subsequent rising F0 contours. Cheng (1968) argues that English learners tended to associate T3 (the low variant) with the unstressed syllables in English in that English unstressed syllables usually have lower F0. This association between the low tone (i.e., T3) and the unstressed syllables led some learners to produce the low tone errors when they transferred their English stress patterns (i.e., the alternation of the stressed syllables versus the unstressed syllables) to their production of L2 Mandarin. However, it is worth mentioning that the low tone errors produced by learners are often stressed, which are more similar to the low pitch accents in English. The reason for the frequent occurrence of stressed, low tone errors is that learners paid conscious efforts to producing each tone. Therefore, the production of the low tone errors is the joint consequence of the transfer of English stress patterns and the learners’ conscious efforts to produce each tone.

The analysis of tones errors in the T1 and the T2 sequences suggests that the low and rising tones are the default tones for some lower-level learners. We also note that there are few tone errors at the utterance-initial position in both SS-24 and SS-33, even though it is also the phrase-initial position. In Section 5.3, some interpretations will be provided to account for this phenomenon.

5.3.3. Tone errors in the T3 sequence

T3 in Mandarin Chinese is a low tone (“21” or “22”) in non-boundary positions and a low-dipping one (“214”) in boundary positions. The third tone sandhi is widely studied in
literature (Chao 1968; Chen 2000; Shih 1986, 1997; Zhang 1988, among many others). The T3 sandhi rule is as follows: A T3 becomes a T2 when it precedes another T3. In the literature, there is already abundant discussion with respect to the domain for the application of T3 sandhi (Chen 2000; Cheng 1973; Duanmu 2000; Zhang 1988, among others). Whether the T3 sandhi will occur in a sequence of T3’s, and if so, where it will occur, depends on a number of factors; these include prosodic structure, speech rate, and so on (See Lin 2007: chapter 9). For example, in very fast speech, all the syllables except for the last syllable in a sequence of T3’s may become T2. One example is given in (10), where all but the last syllable in the sentence undergoes the T3 sandhi rule (T3-T3-T3-T3 → T2-T2-T2-T3).

(10) Lǎo Lǐ zhǎo bǐ. → Láo Lí zháo bǐ.

“Old Li looks for pens.”

To add to the complexity of T3 sandhi, T2 sandhi may occur after the T3 sandhi rule applies to the first two syllables in a sequence of three T3’s, namely, T3-T3-T3 → T2-T2-T3 → T2-T1-T3. However, as mentioned in section 5.3.2, the application of T2 sandhi is more of a phonetic co-articulation process; whether this rule applies in actual
speech depends on speech rate and the individual speaker. Therefore, it is predicted that
the surface tone sequences for the T3 sequence may have even more variability than
those for the T2 sequence.

Tables 5.5 and 5.6 show the correct utterance percentage (CUP), the correct tone
percentage (CTP), and the accuracy weight (AW) values for the T3 sequence in each
group.

<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>95%</td>
<td>99%</td>
<td>[1-0] [1-1-1-1]</td>
</tr>
<tr>
<td>Advanced</td>
<td>60%</td>
<td>90%</td>
<td>[0.88-0.5] [1-0.63-1-0.75]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>50%</td>
<td>89%</td>
<td>[1-0.2] [0.9-1-0.9-0.8]</td>
</tr>
</tbody>
</table>

Table 5.5. Correct utterance percentage (CUP), correct tone percentage (CTP) and
accuracy weight (AW) in the T3 sequence: SS-24

<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>100%</td>
<td>100%</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced</td>
<td>60%</td>
<td>85%</td>
<td>[1-0.83-0.3] [0.83-0.92-0.58]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>30%</td>
<td>79%</td>
<td>[0.79-0.93-0.07] [0.86-0.86-0.71]</td>
</tr>
</tbody>
</table>

Table 5.6. Correct utterance percentage (CUP), correct tone percentage (CTP) and
accuracy weight (AW) in the T3 sequence: SS-33

The surface tones in SS-24 for the native group are quite consistent. Except for one
utterance with a mistake (a slip of the tongue), all the other nineteen utterances have the
same surface tone sequence, namely LH-L-LH-L-LH-L. Thus, nearly all native speakers
treated disyllables, or the disyllabic foot, as the T3 sandhi domain. This sandhi pattern
provides further evidence for the disyllabification tendency in Mandarin Chinese. In contrast, the native speakers had much more variability in the surface tone sequences in SS-33 for the T3 sequence; this is largely due to complications of potentially different T3 sandhi domains. The high AW values at the syl-3 position for the two learner groups in Table 5.5 and the relatively high AW values at the syl-4 position for the two learner groups in Table 5.6 seem to confirm the default status of rising tones at the utterance-medial but phrase-initial position for the learners. Due to the application of the T3 sandhi, T3 at the syl-3 position in SS-24 and at the syl-4 position in SS-33 becomes a rising tone. Since the rising tone is the default tone at these positions for some learners, it is expected that there would have been fewer tone errors at these positions.

In both SS-24 and SS-33, the native speaker group had nearly no utterances containing errors, while both advanced and intermediate learner groups had some utterances with tone errors, especially in the intermediate learner group. In the T1 and T2 sequences, most of the utterances with tone errors were produced by the lower-level subgroup in both learner groups. With respect to the T3 sequence, the lower-level subgroup of each learner group still produced more utterances with errors. However, the upper-level learner groups produced some utterances containing errors, too.

The AW values in both SS-24 and SS-33 for the T3 sequence show that many tone errors occurred at the end of the first prosodic phrase, namely at the syl-2 position in SS-24 and at the syl-3 position in SS-33. One example of such tone error is given in Figure 5.7. The utterance in Figure 5.7 is shown in (10).
(10) Li Min xǐng mǔ mǔ mǔ

李敏 想 买 母 马。

"Li Min wants to buy a mare."

Figure 5.7. Illustration of the over-application of the T3 sandhi: An utterance produced by a male advanced learner

The circled syllable in Figure 5.7 was produced as a rising tone. Perceptually, there is a clear break between the second and the third syllables in the utterance. Theoretically speaking, no T3 sandhi rule should apply, because the second and the third syllables do
not belong to the same prosodic domain. In spite of that, this learner still applied the T3 sandhi and produced a rising tone at the syl-2 position. The over-application of the T3 sandhi rule occurred in both the intermediate and the advanced learner groups, even when T3 was followed by an even bigger break, such as B3 (i.e., the break with an audible pause between two prosodic phrases). This over-application (or inappropriate application) of T3 sandhi indicates the over-generalization of L2 phonological rules in second language acquisition (SLA).

Similar to the misalignment that occurred in the T2 sequences, there are also some cases of misalignment of the tone target, especially in SS-33, for the advanced learners. One example of the misalignment in the T3 sequence is given in Figure 5.8. The utterance in 5.8 is given in (11).

(11) Li  Mǐn xiǎng mǎi mǔ mǎ.
李敏响买母马。
Surname smart sound buy female horse
“Li Minxiang buys mares.”
Figure 5.8. An example of H target misalignment in the T3 sequence: An utterance produced by a male advanced learner.

In the circled syllable in Figure 5.8, the H target on 敏 mǐn ‘smart’ is misaligned with the third syllable 响 xiǎng ‘sound’, rendering it a falling tone (HL). Worth noting is that the misalignment in the T2 and the T3 sequences seems to occur more frequently among the advanced learners than among the intermediate learners. However, native speakers did not produce such H target misalignment even in very slow speech. Similar to the H target misalignment in the T2 sequence, the misalignment in the utterance containing the T3 sequence shows that this learner had difficulty in quickly changing the F0 target to low on the third syllable. Thus, the F0 contour over the third syllable was just the interpolation between the high target on the second syllable and the low target on the third syllable.
In addition, the AW values in Tables 5.6 and 5.7 suggest that, except for the over-application of the T3 sandhi rule at the end of the first prosodic phrase, there are not many tone errors in the T3 sequence for both learner groups. In the discussion of tone errors occurring in T1 and T2 sequences, it was proposed that the low and rising tones are the default tones for some learners, namely, they tended to produce these tones to replace other underlying tones, especially in the second prosodic phrase of an utterance. With the exception of the over-application of T3 sandhi, the less frequent occurrence of tone errors in the T3 sequence further confirms the default status of the low and rising tones. The application of the T3 sandhi changes a T3 to a T2, a rising tone. Thus, in a T3 sequence that contains no tone errors, all syllables carry either T2 (rising tone) or T3 (low tone). Due to the default status of the low and the rising tones for some learners, these sequences were relatively easier for them to produce, including for the lower-level learners. The result is that there are fewer tone errors in the utterances containing the T3 sequence. However, there might be another possible explanation. It may be that some learners intended to produce a low or rising tone, due to their production habits. However, the surfacing low and rising tones happened to be what should have been produced. As a result, the tone errors were concealed in the surface low and rising tones.

5.3.4. Tone errors in T4 sequence

T4 in Mandarin Chinese is a high-falling tone (“51”). One typical sandhi involving T4 is that a T4 becomes a half T4 or a high level tone when it occurs before another T4 or a tone beginning with a high target (Chao 1968). The application of the T4 sandhi is
optional. Whether the sandhi will occur depends on the individual speakers, speech rate, and the context.

Tables 5.7 and 5.8 show the correct utterance percentage (CUP), the correct tone percentage (CTP), and the accuracy weight (AW) values for each T4 sequence in the three subject groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>100%</td>
<td>100%</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced</td>
<td>65%</td>
<td>93%</td>
<td>[0.85-1] [0.71-0.71-0.43-1]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>60%</td>
<td>87%</td>
<td>[0.63-1] [0.38-0.5-0.38-0.88]</td>
</tr>
</tbody>
</table>

Table 5.7. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weight (AW) in the T4 sequence: SS-24

<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>90%</td>
<td>98%</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced</td>
<td>75%</td>
<td>96%</td>
<td>[1-0.8-1] [0.6-0.6-1]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>60%</td>
<td>83%</td>
<td>[0.88-0.63-0.75] [0.25-0.25-0.63]</td>
</tr>
</tbody>
</table>

Table 5.8. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weight (AW) in the T4 sequence: SS-33

Due to the application of the T4 sandhi, there is great variability in the surface tones for the T4 sequence in all three groups. Tables 5.7 and 5.8 show that the two learner groups did not do as well as the native group, and that the advanced learner group did better than the intermediate learner group. In SS-33, there are two utterances with tone errors produced by two native speakers. Similar to the case that a T2 may become a T4
when it is weakened and occurs between two strong T2’s (Wu 1984; Liao 1994), a T4 may become a T2 when it is weakened and occurring between two strong T2’s. However, the surface rising tone occurs after the prosodic phrase boundary at the syl-3 position, which should have prevented the occurrence of such phonetic sandhi. Therefore, these two cases are considered to be slips of the tongue in connected speech.

The AW values in Table 5.7 show that in SS-24, most tone errors occur at the syl-3 and the syl-5 positions for the intermediate learners. For the advanced learner group, the most tone errors occurred at the syl-5 position. The error pattern in both learner groups is that T4 surfaced as a low tone or a rising tone. The AW values in Table 5.8 show that in SS-33, most tone errors occurred at the syl-4 and the syl-5 positions for the two learner groups. Again, the surface tone errors are rising and low tones.

Figure 5.9 gives an example of the low tone error in SS-24. The utterance in Figure 5.9 is given in (12).

(12) Wèi     Lì    yào   mài   là      ròu.
Surname pretty  will   sell   air-dried  meat
   魏  丽  要  卖  腊  肉.
   “Wei Li will sell bacon.”
Figure 5.9. Illustration of the low tone error in the T3 sequence (SS-24): An utterance produced by a male advanced learner.

The tone on the circled syllable in Figure 5.9 was produced as a low tone in this utterance. Also notice that the F0 tendency before the syllable 腊 là ‘air-dried’. The F0 on the two syllables 要 yào ‘will’ and 卖 mài ‘to sell’ continuously falls to a very low point in the learner’s pitch range. The F0 contours over two consecutive T4’s produced by the native speakers usually do not drop continuously as the learner did. The consequence of the continuous F0 fall is that it was difficult for the F0 to reach the high target of the next T4. Thus, a low tone was produced by this learner as a buffer zone so that he could have enough time to raise his pitch to produce the high target on the subsequent T4 syllable.

Figure 5.10 gives an example of the rising tone error in the T4 sequence. The utterance in Figure 5.10 is given in (13).
(13) Lù Wèi yòng mài yùn yào.

Surname guard use sell pregnant medicine

“Lu Weiyong sells contraceptive medicine.”

Figure 5.10. Illustration of the rising tone error in the T4 sequence (SS-33): An utterance produced by a male intermediate learner

T4 on the circled syllable 卖 mài in Figure 5.10 was produced as a rising tone by this learner. Also note the continuous F0 fall on the first three syllables in this utterance. Even though there is a large break between the first prosodic phrase and the second prosodic phrase, it seems that the continuous F0 fall in the first prosodic phrase also imposed some effect on the F0 production on the syllable 卖 mài ‘to sell’ in the second prosodic phrase.
The tone errors in the T4 sequence for the two learner groups have two patterns. Firstly, more tone errors tend to occur after the break boundary, namely the utterance-medial but phrase-initial positions. Secondly, the low and rising tones are the most frequent tone errors in both learner groups. Thus, the tone errors in the T4 sequence further confirm our proposal that the low and rising tones are the default tones for some learners. These tone errors suggest that the learners had difficulty in making quick changes of tone targets. Additionally, the tone errors in the T3 and other tone sequences discussed above have shown that tone errors also display some positional effects, namely, tone errors tend to occur at the utterance-medial, phrase-initial positions (syl-3 in SS-24 and syl-4 in SS-33).

5.3.5. Tone errors in the T2T4 alternating sequence

In the T2T4 alternating sequence, no possible sandhi, whether phonological or phonetic, may occur. Thus, it is predictable that there would be much less variability in the surface tones in this sequence compared to other tone sequences.

Tables 5.9 and 5.10 show the correct utterance percentage (CUP), the correct tone percentage (CTP), and the accuracy weight (AW) values in each group for the T2T4 alternating sequence.
In Table 5.9, even though the two learner groups have the same CUP values in SS-24, the intermediate learner group has lower CTP values than the advanced learner group. For the advanced learner group, the lowest AW value occurs at the syl-2 position, while for the intermediate learner group, the syl-5 and the syl-6 positions have the most tone errors. In SS-24 for the T2T4 alternating sequence, the syl-3, the first syllable in the second prosodic phrase, carries the rising tone. According to the proposal of the default status of low and rising tones, T2 on the syl-3 should be easier for both learner groups. As shown in Table 5.9, the relatively high AW values (0.71 for the advanced learner group and 0.86 for the intermediate learner group) at this position as shown in Table 5.9 support our proposal. The frequent error pattern in the advanced learner group is that the second syllable surfaces as a rising tone. Figure 5.11 gives an example of the rising tone error in the T2T4 sequence. The utterance in Figure 5.11 is given in (14).
Figure 5.11. Illustration of the rising tone error in the T2T4 sequence (SS-24): An utterance produced by a male advanced learner

T4 on the circled syllable in Figure 5.11 was produced as a rising tone. This error is hard to interpret. Because the first two syllables form a rising-falling sequence, the tones at these two syllables should have been easier for the learners, since the target at the offset of the rising tone and the target at the onset of the falling tone are identical. One
interpretation for the tone errors at this position may be that the learner attempted to produce a high phrase accent and a high boundary tone at the end of the first prosodic phrase (i.e., the rising intonation), due to the incompleteness of this phrase relative to the whole utterance. If so, the tone error in Figure 5.11 may be an example of transfer of English intonation pattern into L2 Mandarin.

For the intermediate learner group, the most frequent tone error occurring at the syl-5 position in SS-24 is the replacement of T4 by a rising or a low tone. For the intermediate learner group, some tone errors also occurred at the syl-1 and the syl-2 positions. Again, the frequent pattern is T4 being replaced by a low or rising tone.

In Table 5.10, pertaining to SS-33, both learner groups produced some utterances with tone errors. The intermediate group produced more utterances with tone errors than the advanced learner group. The AW values show that the most tone errors occurred at the syl-4 and the syl-5 positions for the intermediate learners. Again, the rising tones and low tones are the tone errors. The syl-4 has a falling tone, while the syl-5 has a rising tone. The frequent tone errors occurring at the syl-4 position confirm our proposal that the low and rising tones tend to be used by the learners after the break between the subject and the predicate. For the advanced learner group, most tone errors occurred at the syl-5 position. The error pattern is that the original T2 surfaced as a low tone, which served as a buffer zone for the production of the following syllable. Again, the tone errors in the T2T4 sequences confirm that the low and low-rising tones are the default tones for some lower-level learners.
Figure 5.12 gives an example of the low tone error in the T2T4 sequence in the intermediate learner group. The utterance in Figure 5.11 is given in (15).

(15) Nán Mèng lái mài rán liào.

“Nan Menglai sells fuel.”
The tones errors are circled in Figure 5.12. Note that due to the creakiness, the F0 on the syllable 燃 rán was mis-tracked. The correct F0 contour over this syllable should have been very low. In Figure 5.12, the tone on 卖 mài was produced as a rising tone and the rising tone on 燃 rán was produced as a low tone.

Figure 5.13 gives another example of the low tone error in the T2T4 sequence. The utterance is the same as in (15), but this utterance was produced by a male advanced learner.

Figure 5.13. Illustration of the low tone error in the T2T4 sequence (SS-33): An utterance produced by a male advanced learner
The T2 on 燃 rán ‘to burn’ in both Figures 5.12 and 5.13 was produced as a low tone. However, our previous analysis has shown that the rising tone is also the default tone for some learners. The question is why the T2 was produced as a low tone here. I argue that the frequent F0 rise and fall or the quick changes of the F0 direction in the T2T4 alternating sequence posed difficulty for some learners. In spite of the identical terminal targets on the adjacent tones, the F0 direction needs to be changed quickly. Even though not as difficult as the quick change of tone targets in the conflicting tones sequences, the quick change of the F0 direction in the compatible tone sequences was difficult for learners. Therefore, the low tone was produced by the learner as a buffer zone so that he could gain time to raise the F0 to the high target on the following syllable.

5.3.6. Tone errors in the T4T2 alternating sequence

Similar to the T2T4 alternating sequences, it is predicted that there would not be too much surface tone variability in the realization of the T4T2 alternating sequence.

Tables 5.11 and 5.12 show the correct utterance percentage (CUP), the correct tone percentage (CTP), and the accuracy weight (AW) values in each group for the T4T2 alternating tone sequence.

<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>100%</td>
<td>100%</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced</td>
<td>75%</td>
<td>90%</td>
<td>[0.2-0.6]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>75%</td>
<td>87%</td>
<td>[0.4-1.0]</td>
</tr>
</tbody>
</table>

Table 5.11. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weight (AW) in the T4T2 alternating sequence: SS-24
Table 5.12. Correct utterance percentage (CUP), correct tone percentage (CTP) and accuracy weight (AW) in the T4T2 alternating sequence: SS-33

<table>
<thead>
<tr>
<th>Group</th>
<th>CUP</th>
<th>CTP</th>
<th>AW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>100%</td>
<td>100%</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced</td>
<td>65%</td>
<td>88%</td>
<td>[0.57-0.71-0.86]</td>
</tr>
<tr>
<td>Intermediate</td>
<td>40%</td>
<td>82%</td>
<td>[0.67-0.67-0.67]</td>
</tr>
</tbody>
</table>

In Table 5.11 pertaining to SS-24, even though the CUP in two learner groups are the same, the CTP in the intermediate learner group is lower than that in the advanced learner group. For the intermediate learners, the AW values show that more tone errors occurred at the syl-1, syl-3 and syl-5 positions than at other positions. The error pattern at these three positions is that T4 surfaced as low or rising tones. The low AW values at these positions indicate that some learners had difficulty in F0 production in the T4T2 (i.e., falling-rising) alternating tone sequence. Similar to the F0 production in the T2T4 tone sequence, the difficulty that the intermediate learners encountered was attributed to the quick changes of the F0 direction, namely from rising to falling or from falling to rising. Figure 5.14 gives an example of the low and rising tone errors in the intermediate group. The utterance in Figure 5.14 is given in (16).
(16) Mèng Yán ài dú wài wén.

孟岩爱读外文。

Surname rock love read foreign language

“Meng Yan loves reading foreign literature.”

![Pitch-Time Graph](image)

Figure 5.14. Illustration of the rising tone errors in the T4T2 sequence (SS-24): An utterance produced by a male intermediate learner

The T4’s (falling tones) on the first two circled syllables in Figure 5.14 were both produced as rising tones in this utterance. As a result, all the tones in this utterance except on 外 wài are rising. In addition, the T4 on the third circled syllable 外 wài ‘external’ was produced as a low-dipping tone. The examination of the F0 contour over the third syllable shows that the beginning portion of the F0 track is very low and the ending
portion of the F0 on this syllable is rising. Thus, the overall surface tone surfaced as a low-dipping tone.

In addition, the AW value at the syl-3 position for the intermediate learner group as in Table 5.11 is 0. The syl-3 in SS-24 is at the beginning of the second prosodic phrase. Thus, the tone errors in the intermediate learner group again show that learners tended to produce more tone errors at the utterance-medial but phrase-initial positions. For the advanced learners, the AW values in Table 5.11 show that more tone errors occurred at the syl-1 and the syl-3 positions. At both positions, T4 frequently surfaced as rising or low tones, which is rather similar to the pattern in the intermediate learner group.

In Table 5.12 pertaining to SS-33, the lower AW values (0.14) show that there are more errors after the break for the advanced learner group. The frequent error pattern is that the rising tone surfaced as either a low tone or a falling tone. In the intermediate learner group, there are more tone errors at the syl-4 and the syl-5 position than at other positions. At the syl-4 position, the rising tone frequently surfaced as a low tone. At the syl-5 position, the falling tone frequently surfaced as a low or a rising tone.

5.4. Discussion of the tone error analysis

Our analyses of the tone errors in different tone sequences show that the low and rising tones are the most frequent tone errors produced by the two groups of learners. This is especially true for the lower-level learners in both groups, and that is regardless of the underlying tones. In the T1 sequence, the low and rising tone errors were produced because some lower-level learners in both learner groups did not have a good command
of the production of a long interval of high level F0 contour. That is, undoubtedly, due to the absence of such F0 patterns in English intonation. Thus, these learners produced the low and rising tones to alternate the tones in their F0 production. In the T2 sequence, the low tones were produced because the learners failed in their attempt to produce the rising F0 contours and only produced the low tone target. In so doing, they alleviated the difficulty of dropping the F0 to the low target before producing the next T2. In the T3 sequence, the most salient pattern of tone errors is the over-application of the T3 sandhi at the end of the first prosodic phrase. This over-application of T3 sandhi rule suggests that the learners had acquired the T3 sandhi, but had over-generalized their application. In the T4 sequence, the low and rising tones are the most frequent tone errors. These tone errors were produced because the learners had difficulty in the quick change of the tone targets. In the T2T4 and the T4T2 alternating sequences, there are two frequent tone error patterns: (1) the falling tone (T4) was produced as a low or rising tone; and (2) the rising tone (T2) was produced as a low tone. Unlike the T2 or the T4 sequence, the terminal tone targets on the adjacent syllables in the T2T4 and the T4T2 sequences are identical. It was expected that the tones in these compatible tone sequences should not have posed much difficulty for the learners. However, for the lower-level learners in both learner groups, the low and rising tone errors were frequently produced. In the T2T4 and the T4T2 alternating sequences, the replacement of T2 with low tones by learners suggest that even the compatible tone sequences posed difficulty for them, because in these alternating tone sequences, the learners needed to produce the rising and falling F0 contours alternatively and quickly. Such frequent F0 fall and rise were difficult for
learners to produce; thus, they stopped short of reaching the high target in T2 and produced a low tone. For the replacement of T4 with the low tone, the low tone was produced because the learners had difficulty producing frequent F0 rise and fall. Thus, the low tone was produced to serve as a buffer zone for them, in preparing to produce the following T2. The replacement of T4 with the rising tone seems to be difficult to interpret because the consequence of such tone errors is that the learners needed to produce a sequence of rising tones. We attributed this tone error pattern to potential transfer of the high phrase accent and boundary tone in L1 English (i.e., the English rising intonation).

Another general pattern of tone errors in different tone sequences is that the rising tones were more frequently produced at the utterance-medial, phrase-initial positions, while the low tones were more frequently produced at the utterance-medial, and phrase-medial positions. Both tone errors occurred in other positions as well. Our interpretation for such position-related tone errors is that rising tones were produced due to the transfer of the English rising intonation, while the low tones at the phrase-medial positions served as a buffer zone for the learners to proceed with the production of the following tones. The low tone error is also potentially the consequence of the transfer of English stress pattern in that learners tended to associate the low tone in L2 Mandarin with the unstressed syllables in English (Cheng 1968). However, the syllables carrying the low tones were still stressed in the L2 Mandarin speech, due to learners’ conscious efforts to produce each individual tone.

In summary, our analysis of tone errors shows that learners not only had difficulty producing quick changes of tone targets in the conflicting tone sequences, but they also
had difficulty in the quick changes of F0 direction in the compatible tone sequences. However, the conflicting tone sequences are more difficult for the learners than the compatible tone sequences. All of these show that the F0 production has not been internalized by the learners.

One of the research questions in this dissertation is to examine the relationship between prosodic phrasing and the tone errors in L2 Mandarin productions. Two concepts from psychology -- controlled processing and automatic processing -- are relevant here. These two concepts in learning are proposed by psychologists such as Shiffrin and Schneider (1977). Applied to second language learning, controlled processing involves the temporary activation of a selection of information nodes in the memory in a new configuration. Such processing requires a lot of attentional control and is constrained by the limitations of short-term memory. Through repeated activation, the controlled processing may become automatic. Then, the automatic processing can be made available very rapidly whenever the situation requires it, with minimal attentional control on the part of the subject (Mitchell and Myles 1998: 86). We argued that the tone errors, especially the frequent tones errors (i.e., the low tones and the rising tones), were produced as a consequence of the superimposition of the English utterance-level prosody over tone production among L2 learners. It was also argued that L2 learners had difficulty in the quick changes of both tone targets and F0 direction. However, the difficulty in the quick changes of tone targets and F0 direction stems from the superimposition of English prosody. In an English utterance, the tone targets and the F0 direction do not change so quickly and so frequently. Thus, L2 learners need to retrieve
the “controlled” knowledge of tone production, because automatic knowledge of tone production is not yet available to them. While the retrieval of the “controlled” knowledge of tones guarantees that some tones are produced correctly, it depletes the production resources of other tones at the same time. Thus, for the tones whose production resources were depleted, L1 English prosody inevitably surfaced. The result is the production of tone errors. More specifically, low tones were produced as a buffer zone. Rising tones were produced because they were relatively easier to produce, due to the similarity to English rising intonation. To capture the patterns of the tone errors in L2 Mandarin Chinese, a tentative model of learners’ tone error production is proposed in Figure 5.15.

![A tentative model of learners’ tone errors](image)

Figure 5.15. A tentative model of learners’ tone error production

166
As mentioned in the literature review, previous studies have shown that T1 and T4 are usually the easiest tones to be acquired by L2 learners, whereas T2 and T3 are often the most difficult tones to be acquired by L2 learners. However, the analysis of tone errors in different tone sequences show that the lower-level learners tended to produce the low and rising tones in their L2 speech to replace other underlying tones. Thus, it seems that the difficulty that these learners had in tone production not only lies in T2 and T3 per se, but also lies in other tones. On the one hand, the learners had difficulty producing the rising and the low tones. On the other hand, they produced the low and rising tones when they intended to produce other tones. We argue that these tone errors in L2 Mandarin are the product of the superimposition of the phrase- or utterance-level prosody upon the lexical prosody. The learners may have attempted to produce the correct tones. However, the utterance-level prosody, due to the transfer of English prosody, and the incomplete internalization of Mandarin prosody made them unable to produce tones in a consistent manner.

In addition, we found that there were fewer tone errors at the utterance-initial positions, regardless of the tones involved. This pattern of tone errors can also be explained by the controlled processing and automatic processing. L2 learners’ production of tones has not been internalized as part of their L2 phonology; in other words, their tone production has not become automatic. Thus, at the beginning of an utterance, learners resorted to the controlled processing to produce tones, giving them the required resources to produce the correct tones, as well as the phrase prosody. However, in the other
positions, especially towards the end of the utterance, learners would have depleted their production resources (i.e., their knowledge of tones) due to have already used controlled processing in the beginning of the utterance. Thus, the transfer of English intonation and stress patterns would tend to occur, which caused the learners to produce more tone errors at those positions.

It should be pointed out that, with the exception of errors in the T3 sequence in both learner groups, most of the tone errors were made by the lower-level learners in each learner group. This observation highlights the importance of making more accurate categorization of the learner groups in examining L2 acquisition.

To conclude, the examination of tone errors has also provided some insight into the training of tone production in CFL teaching. Our analysis of the patterns of tone errors suggests that tone errors were made not only because some tones were more difficult to produce, but also because the prosodic phrasing in an utterance influences the production of tones in context. Thus, the learners should be trained on tones in isolation as well as tones in context. More importantly, the training of the utterance-level prosody should be integrated into CFL teaching.
CHAPTER 6: GENERAL DISCUSSION & PEDAGOGICAL APPLICATIONS

6.1. Introduction

In this chapter, we summarize the findings in this study by returning to the research questions raised in Chapter 1. We will also discuss the theoretical implications of these findings, and conclude with exploring some pedagogical applications of the findings in CFL teaching and learning.

6.2. Summary of the findings in this study

A summary of the findings in this study is presented in this section with reference to the research questions raised in Chapter 1.

6.2.1. Research question 1: How do American CFL learners prosodically phrase their utterances, and how does the prosodic phrasing in L2 Mandarin differ from that in L1 Mandarin?

Our analysis on prosodic phrasing shows that the prosodic phrasing in both L1 and L2 Mandarin is closely related to syntactic structures. However, L2 learners tended to make a larger break between the subject and the predicate, which may be the artifact of the way of presenting reading materials in making the recordings, namely, all proper
nouns of personal names were underlined. Moreover, we found that the verb phrase (VP) internal structure in the predicate does not influence the prosodic structure of that constituent in either learner or native group. Among the two learner groups, the more frequent occurrences of B2 and B3 at syl-4 position in SS-24 than in SS-33 suggest that the longer prosodic phrase (the predicate in SS-24) poses more difficulty for learners. The result is that learners tended to break up a longer phrase into small units.

6.2.2. Research questions 2 and 3: What are the F0 and duration patterns in L2 Mandarin, and how do they differ from those in L1 Mandarin?

Our analysis of the duration patterns of prosodic phrasing in L1 and L2 Mandarin shows that the most consistent duration pattern indexing prosodic phrasing is the phrase-final lengthening in both L1 and L2 Mandarin Chinese. The overall duration patterns in L1 and L2 Mandarin Chinese are very similar. We also found phrase-initial lengthening in the native group. Overall, the advanced learner group was more similar to the native group. In addition, we observed some evidence of the transfer of English stress patterns, namely, in SS-24 both learner groups tended to form WS WS WS stress patterns.

The analysis of pitch reset shows that the native speakers used pitch reset more frequently than the L2 learners in the utterances containing the T4 sequence. However, the L2 learners used pitch reset more frequently than the native speakers in the utterances containing the T1 sequence. We argued that the occurrence of pitch reset in an utterance depends on a number of factors, such as underlying tones, downstepping, declination, and
the weighting of downstepping and declination. As a result, no generalization of the overall patterns of pitch reset was made concerning L1 and L2 Mandarin.

With respect to the F0 patterns of prosodic phrasing, our analysis shows that the conflicting tone sequences (i.e., the T2 and T4 sequences) are more difficult for learners to produce than the compatible tone sequences (i.e., the T2T4 alternating sequence and the T4T2 alternating sequence). Specifically, native speech involved more F0 co-articulation in the form of target undershoot than L2 speech, even though L2 speech produced by the advanced learners involved more target undershoot than that by the intermediate learners. It was also found that target undershoot mostly occurred in the first prosodic phrase. The lack of target undershoot among the learners suggest that the learners have not fully acquired utterance-level prosody of Mandarin Chinese and tended to produce each tone as fully as possible. It is suggested that such lack of target undershoot may lead to the perception of a foreign accent, even though further perception study is needed to confirm it.

6.2.3. Research question 4: How does prosodic phrasing influence tone production in L2 Mandarin Chinese

Our analyses on tone errors show that the low and rising tones were the most frequent tone errors used by the two groups of learners in their L2 Mandarin Chinese, regardless of the underlying tones. Specifically, the rising tone was used more frequently at utterance-medial, phrase-initial positions, while the low tones were more used at
utterance-medial and phrase-medial positions. However, both tone errors occurred in other positions as well.

Our analysis of tone errors shows that learners not only had difficulty in quickly changing the tone targets in the conflicting tone sequences, but they also had difficulty in quickly changing the F0 direction in the compatible tone sequences. However, the conflicting tone sequences posed more difficulty for the learners than the compatible tone sequences. These findings show that F0 control has not been internalized by some of the learners.

The frequent low and rising tone errors show that learners not only had difficulty in T2 and T3 per se and they also had difficulty in producing other tones in that they produced low (T3) and rising tones (T2) when they intended to produce other tones. It is argued that the superimposition of utterance-level prosody from English can account for these tone errors to a large extent.

6.3. Theoretical implications

In this study, we examined the prosodic phrasing and the surface F0 and duration patterns of the prosodic phrasing in L1 and L2 Mandarin Chinese, together with the relationship between tone errors and the prosodic phrasing in L2 speech. Our findings show that it is important to examine the acquisition of lexical tones together with the acquisition of utterance-level prosody. Without considering utterance-level prosody, it is difficult to tease apart the effect of the difficulty in producing some tones from the effect of utterance-level prosody superimposed upon tone production.
In the reminder of this section, some theoretical issues related to this study will be discussed.

6.3.1. Are F0 patterns the consequence of duration control in the prosodic phrasing of Mandarin Chinese?

Xu and Wang (2009) argue that syllable grouping in Mandarin Chinese is most consistently reflected in the patterns of syllable duration, while the F0 patterns are the consequence of duration control in the syllable grouping of Mandarin Chinese. Our finding that the duration patterns for the prosodic phrasing in Mandarin are more consistent than the F0 patterns has provided further evidence for Xu and Wang’s argument. What is more, we have shown that the patterns of target realization may vary with the tones involved. However, in this study and in Xu and Wang (2009), only short read sentences were used. Thus, the lack of consistent F0 patterns only applies to the sentence or utterance level in Mandarin Chinese. Tseng et al. (2005) find that Mandarin Chinese also has consistent F0 patterns at the discourse level, similar to English. In this study and Xu and Wang (2009), the surface F0 patterns varied with the tones involved because the surface F0 patterns tend to be affected by lexical tones more in short utterances than in discourse. Chao’s characterization (1933 and 1968) of the relationship between tones and intonation can help explain the different F0 patterns at the utterance level and at the discourse level. At the utterance level, the small ripples may not have any patterns. However, at the discourse level, the large waves can form some consistent patterns. Thus, even though the primary function of F0 is to represent lexical tones in
Mandarin Chinese and native speakers of Mandarin do not have so much “freedom” for manipulating F0 for intonational functions in Mandarin as in English, the F0 patterns in longer discourse also can consistently index the prosodic phrasing in Mandarin Chinese.

6.3.2. Why are there so many low tone errors?

Our analysis of tone errors in L2 Mandarin Chinese suggests that the low tones, together with the rising tones, were the most frequent tone errors in L2 Mandarin. Thus, it is worth examining why these two tone errors, especially the low tone errors, are more likely to occur in L2 Mandarin than other tone errors.

Previous studies have shown that in both L1 and L2 Mandarin Chinese the low (T3) and rising tones (T2) are more difficult to acquire than the high-level (T1) and falling tones (T4) (Li and Thompson 1977; Chen 1997, 2000; Sun 1998; Zhang 2007, among others). Zhang (2010) argues that low tone (termed “half-T3” by Zhang) is more “unmarked” than T2 because it is often used to replace other tones at the word-initial positions. We have argued that low and rising tones are the default tones in some learners’ phonology of Mandarin. The question is why the low tone is “unmarked” or default among some learners. Yeh (2001) reports on an emerging linguistic phenomenon in Taiwan Mandarin. In the nominal reduplications in Taiwan Mandarin, the second syllable is usually neutralized, such as in 妈妈 māma, “mother”, or 弟弟 dìdì “younger brother”. However, this has been replaced by another pattern, namely, the T3-T2 (the low-rising) pattern, regardless of the underlying tones of the reduplicated syllable. Thus, 妈妈 māma becomes māmá, and 弟弟 dìdì becomes dīdí. Yeh (2010) argues that the low
tone at the beginning position of a nominal reduplication in this emerging pattern is phonetically natural in that T3 only has a low target and is short in duration. Thus, fewer articulatory efforts are needed in producing this tone. This emerging pattern is believed to start among the young people and then is spreading to all speech communities in Taiwan. Nonetheless, this pattern is still mostly used in an intimate relationship, such as lovers, parents and children, good friends, and siblings. The tone pattern in the nominal reduplications of Taiwan Mandarin and the frequent use of low tone errors in L2 Mandarin seem to suggest that the low tone (T3) is preferred in both L1 and L2 acquisition. The underlying reason for these two phenomena is the fewer articulatory efforts involved in the production of the low tone (T3). This similarity suggests that in terms of the acquisition of T3, L2 is similar to L1.

6.4. Pedagogical applications of the findings in CFL teaching

Not only can studies in second language acquisition (SLA) provide insight into linguistic theories, they also have significant pedagogical implications and can guide the practice of L2 teaching. In this section, an attempt is made to apply the findings in this study to CFL teaching. Section 6.4.1 gives an overview of the audio-visual training on L2 prosody. Section 6.4.2 addresses the issues of how to train L2 learners’ Mandarin prosody. Section 6.4.3 concludes this section.
6.4.1. Audio-visual training on L2 prosody

In training learners on L2 pronunciation, especially L2 prosody, the audio-visual training, in which both the audio and visual input and feedback are provided to learners, has proved to be very effective and helpful (Anderson-Hsieh 1992; Chan 1995; Chun 2002; de Bot 1983; So 2003, 2006, among others). Previous studies have shown that audio-visual training can provide learners with dual channel input to assist them with learning sound segments, prosody and other aspects of pronunciation. Audio-visual training can enhance learners’ awareness of phonological/phonetic acoustic characteristics of L2 prosody, which, in turn, assists the learners in establishing the sound categories or prosodic patterns.

Audio-visual training can be done in two different ways, namely, perception training and production training. It was found that perception training is more helpful than production training in helping learners to improve L2 Mandarin tones (So 2006). Even though production is closely related to perception, they do not have a straightforward relationship (Goto 1971; Sheldon and Strange 1982). Thus, I argue that perception training and production training are equally important and that they complement and promote each other. Perception training helps modify L2 learners’ perception of the non-native sounds and prosody by shifting their attention to phonetic information that is previously ignored. Production training trains L2 learners to employ the same acoustic parameters as native speakers, at least acoustic cues equivalent to those in L1 speech.
Thus, effective training should consist of both perception training and production training.\textsuperscript{31}

Most previous studies of audio-visual training on L2 pronunciation were conducted on naïve subjects, namely those who have no L2 background at all (Hardison 1998; So 2006). Even though the training on naïve L2 subjects can shed light on the effectiveness of the specific training, it requires further research to test the actual effectiveness of the training on the L2 learners.

6.4.2. Training on L2 Mandarin Chinese prosody

Our analysis in Chapter 5 has shown that in both the intermediate learner group and the advanced learner group, there are some learners who were responsible for most of the tone errors in that particular group. Thus, it is necessary to train those learners on tone production.\textsuperscript{32} However, it is not enough to train these learners on the production of tones in isolation alone, because even if learners know how to perceive and produce individual tones, when tones are concatenated and produced in context, the tone co-articulation will still pose new difficulty for them. Thus, learners should also be trained on phrase- and utterance-level prosody.

\textsuperscript{31} Chan (1995) is one of the earliest attempts at production and perception training on Mandarin tones. Unfortunately, few similar studies have been conducted ever since.

\textsuperscript{32} Unless learners are proved to be tone deaf or tone mute.
6.4.2.1. Training on L2 Mandarin tones

Tone acquisition is one of the most important aspects in L2 Mandarin acquisition. However, the common practice in an L2 Mandarin class is that Chinese teachers drill on learners’ tone production in monosyllables and at most in disyllables. The training on tones in mono- and disyllabic contexts provides learners with very limited exposure to the variability of tones in context. Even though it is essential to guide learners in acquiring the four lexical tones, the great variability in both F0 contour and F0 register in actual speech, together with the tone errors found in all the tone sequences in this study, suggests that it is equally important to train learners to produce and perceive tones in context.

So (2006: 33) formulates the criteria for perception training, namely: (1) to use high-variability training; (2) to use natural stimuli to approximate natural speech; (3) to use identification tasks rather than discrimination tasks; (4) to include both generalization and retention tests to examine the effectiveness of the training. These criteria are useful guidelines for designing any perception (and production) training of L2 prosody.

With respect to tone training, the high-variability refers to both the high-variability in contexts and the high-variability in speakers. The high-variability in contexts means that learners should be provided with tones produced in different contexts, spanning different numbers of syllables (monosyllabic, disyllabic and multisyllabic) and including both read speech and spontaneous speech. The high-variability in speakers means exposure to tones that are produced by different speakers, including native and non-native, male and female speakers.
Specifically, with respect to the training of the four tones, learners should be provided with both audio (sound files of tones) and visual (the F0 track of tones displayed in Praat or other speech analysis programs) input, together with textual instruction, such as the T3 sandhi and the difference in F0 contour and register for different tones. Note that this has been done in most of the CFL teaching. However, most of the input was given auditorily. Thus, both visual input and the textual instruction should accompany the auditory input. Following the audio-visual input and textual instruction of tones are the perception and production training. Both discrimination (the discrimination task requires the subject to judge whether two tones are the same or not, such T2 versus T3 and T1 versus T4) and identification (the identification task requires the subject to identify which tone category he/she has just heard) tests should be employed. In tone production, not only should the syllables in the tone perception training be used, novel syllables, even non-words, which conform to the phonotactic constraints of Mandarin Chinese though, should also be used to help learners internalize the Mandarin tone system.

The perception training on the high-variability of contexts and speakers follows the same procedure as in tone training above. The difference is that the perception training on the high-variability of contexts employs tones from different contexts, while the tone perception training on the high-variability of speaker uses tones produced by different speakers. It is expected that such training will enable learners to detect both the essential and the peripheral cues for Mandarin tones so that they can tease apart these two types of cues in actual communication.
6.4.2.2. Training on L2 utterance-level prosody

Our analysis of duration patterns in L1 and L2 Mandarin shows that some learners, especially the lower-level intermediate learners, tended to transfer the English stress pattern and form strong versus weak alternating stress patterns in their L2 Mandarin. In addition, our analysis of F0 patterns shows that L2 learners did not produce appropriate target undershoot. In the analysis of tone errors in L2 Mandarin, we found that learners had difficulty in the quick changes of F0 direction in the compatible tone sequences and in the quick changes of tone targets in the conflicting tone sequences. Moreover, the tone errors are position-related. However, the training on tones proposed in section 6.4.2.1 cannot help learners improve their L2 prosody in these aspects. Thus, learners should be trained on utterance-level prosody as well.

Specifically, the training on the utterance-level prosody should address the following aspects: (1) to acquaint learners with the Mandarin stress system; (2) to train learners to produce the appropriate target undershoot, such as the target undershoot in the middle syllable of a trisyllabic prosodic phrase; (3) to train learners on F0 production, namely, to quickly change the F0 direction and to quickly change the tone targets. To help learners with these three aspects of the utterance-level prosody, both prosodic phrases and whole utterances should be employed in training.

6.4.2.2.1. Training on Mandarin stress

To train L2 learners on Mandarin stress, they should be aware that the stress in Mandarin Chinese is different from that in English. In Mandarin Chinese, stress does not
distinguish words in most cases, except for distinguishing the contrast between fully
toned stressed syllables and neutrally toned syllables. Two pairs of examples of the
contrast between fully toned stressed syllables and neutrally atonic syllables are given in
Table 6.1, cited from Chen (2000: 286).

<table>
<thead>
<tr>
<th>Pinyin &amp; characters</th>
<th>Tone numbers</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dong xi</td>
<td>55. 55</td>
<td>i. “east and west”</td>
</tr>
<tr>
<td>东西</td>
<td>55. 0</td>
<td>ii. “things, stuff”</td>
</tr>
<tr>
<td>Di dao</td>
<td>51.51</td>
<td>i. “tunnel”</td>
</tr>
<tr>
<td>地道</td>
<td>51.0</td>
<td>ii. “genuine”</td>
</tr>
</tbody>
</table>

Table 6.1. Examples of the contrast between fully toned stressed syllable and neutrally
atonic syllables

In Table 6.1, the second syllable in each pair can be either a fully toned syllable or an
unstressed tonic syllable. “0” in the table indicates that that syllable is toneless or is said
to be in the neutral tone. In English, stress, in addition to other functions, serves to
differential lexical items, such as SUBject versus subJECT (capitalization indicates stress)
or WHITE House versus white HOUSE. In Mandarin Chinese, tone is the condition for
stress in that there are tonic unstressed syllables, but not atonic stressed syllables
(Kratochvil 1968). In an English utterance, the status of stress of a word is determined by
the lexicon and the prominence of syllables involved. Thus, there are stressed and
unstressed syllables in an English utterance and some stressed syllables are accented. In a
Mandarin utterance, the stress status of a syllable/morpheme is determined by
prominence-assigning in the utterance, together with the lexicon at times when the lexical
unstressed syllables are involved. Generally speaking, there are more unstressed syllables in an English utterance than there are in a Mandarin utterance of the same length (i.e., with the same number of syllables). With respect to the acoustic correlates of stress, Mandarin Chinese and English are rather similar, namely, stress is indexed by F0, duration and intensity. F0 is the primary cue to stress in English (Fry 1958; Beckman 1986), whereas duration is the primary cue to stress in Mandarin Chinese. However, since tone is the condition for stress in Mandarin Chinese, F0 is also very important in representing stress in Mandarin.

Different from tone training, it is difficult to train learners on stress alone, because it is always intertwined with the other aspects of utterance-level prosody. Thus, the training of learners on Mandarin stress should be done together with the training of the other aspects of utterance-level prosody.

6.4.2.2.2. Training on tone co-articulation

Our findings about F0 patterns of prosodic phrasing in Chapter 4 show that the biggest F0 difference between L1 and L2 Mandarin Chinese is target undershoot in that L2 learners did not produce the appropriate target undershoot at phrase-initial and phrase-medial positions. Our analysis of tone errors in Chapter 5 shows that tone errors tended to occur more frequently in the second prosodic phrase. Thus, the training on tone co-articulation should take the above two aspects into consideration and use utterances consisting of at least two prosodic phrases.
Even though the lack of target undershoot may not lead to unintelligibility, it does create the impression of unnaturalness, which, in turn, may lead to the perception of a foreign accent. However, further study is needed to confirm this. To train CFL learners on target undershoot, both perception and production trainings should be conducted. Perception training should be carried out to examine whether CFL learners can detect such target undershoot in L1 speech. For that purpose, audio-visual training is very helpful. Figure 6.1 gives an example of such audio-visual training. The utterance in Figure 6.1 is given in (1).

(1) 王 Ming 来 ná 羊 máo.
王明来拿羊毛。
Surname bright come get sheep wool
“Wang Ming will come to get the wool.”
Figure 6.1. Illustration of the target undershoot in the T2 sequence: An utterance produced by a female native speaker.

The rising tone on the circled syllable in Figure 6.1 was produced as a level tone, due to tone co-articulation. Moreover, the tone on the syllable 拿 ná was also reduced, even though the surface F0 does display a slightly rising contour. By listening to the sound file and observing the F0 contour in Figure 6.1, learners would have gained vivid impression of such target undershoot. With CFL learners’ heightened awareness of target undershoot, production training should be followed by asking learners to imitate the utterance containing the target undershoot. By using whole utterances as training materials, not only would learners acquire the target undershoot, they could also learn how to avoid the transfer of English prosody.

However, it may be argued that training on such target undershoot may not be necessary, because target undershoot is mostly the result of tone co-articulation between
adjacent syllables. However, without appropriate training, learners may not be able to produce such target undershoot and, even if they do, they may be influenced by their native prosodic patterns. Thus, training of target undershoot is not only necessary but also important.

6.4.2.2.3. Training on the quick changes of F0 direction and tone targets

Our analysis of tone errors shows that learners had difficulty in quickly changing the F0 direction in the compatible tone sequence and in quickly changing the tone targets in the conflicting tone sequence. To train learners on these two aspects, native speech materials collected in this study (i.e., the T2 sequence, the T4 sequence, the T2T4 alternating sequence, and the T4T2 alternating sequence) should be used. Only production training is relevant here.

In production training on the quick changes of F0 direction and tone targets, phrases should be used first, such as the di- or trisyllabic subjects and the tri-and quadrisyllabic predicates, in order to reduce the level of difficulty in the beginning. Next, whole utterances containing the compatible and conflicting tone sequences should be used. For some learners, such tone sequences may sound like tongue twisters, and pose great difficulty for them. However, with considerable practice, it is expected that learners can become more proficient in changing F0 direction and tone targets.
6.5. Summary

In this chapter, we summarized the findings of this study and discussed some theoretical implications of the findings. Based on the findings in Chapters 4 and 5, we also provided some suggestions on how to train L2 Mandarin learners on their prosodic production.
CHAPTER 7: CONCLUSION

7.1. Limitations of the study

One limitation of this study is the use of read speech, instead of spontaneous speech. However, in order to address the research questions we raised, we needed to control for the tones and the syntactic structure of the speech data collected. Thus, read speech was used.

Another limitation in this study is the small number of subjects. Only ten subjects were recruited for each group. Moreover, the categorization of each learner group into two subgroups further reduced the number of subjects in each group. In most North American universities, it is currently still difficult to find many advanced learners of Mandarin Chinese except for Chinese heritage learners. Actually, the further categorization of each learner group suggests that learners could have been recruited by other criteria instead of by the class they are enrolled in. In addition, the distribution of male and female subjects in each group is not balanced, especially in the advanced learner group (2 female advanced learners versus 8 male advanced learners). Further study may consider expanding the learner pool by recruiting subjects from different universities.
One more limitation is the short length of the sentences used. In order to prepare two sentences that are identical tonally and segmentally, but differ syntactically, the length of the sentences was set at six syllables in this study.

7.2. Directions for further study

In this study, read speech was collected because only by using read sentences would we be able to control for tones and syntactic structure. In that way, it is possible to examine the influences of tones on prosodic phrasing and the relationship between prosodic phrasing and syntactic structure. However, such design could conceal some phenomena that occur in spontaneous speech. Thus, further study should examine the prosodic phrasing in read passages or spontaneous speech.

Our analysis of utterance-level prosody shows that the biggest F0 difference between L1 and L2 Mandarin Chinese lies in the presence or lack of target undershoot. Specifically, native Mandarin speakers produced more target undershoot than did L2 Mandarin learners. It is worth examining whether native speakers can perceive this acoustic difference in L1 and L2 Mandarin speech. If they do, a subsequent question would be how native speakers might view the lack of such target undershoot in L2 Mandarin Chinese. Does the lack of target undershoot carry any social connotations, such as stigma?
7.3. Concluding remarks

In this study, we examined the acquisition of Mandarin prosody by American CFL learners. We studied prosodic phrasing and F0 and duration patterns in L1 and L2 Mandarin Chinese. We also analyzed the patterns of tone errors that occurred in L2 Mandarin. Based on the findings in this study, we provided some pedagogical suggestions as to how to train L2 learners on their prosodic production. Similar studies, especially studies on the acquisition of utterance-level prosody, are still scarce in the literature. It is hoped that this study can lead to more studies of this type, which will contribute to the field of theoretical linguistics and second language acquisition. At the same time, it is hoped that this study and the suggestions for training the utterance-level prosody, will enable CFL teachers and researchers to widen their research horizon, and to propose more innovative approaches to investigating the acquisition of Mandarin Chinese prosody.
APPENDIX A: A COMPLETE LIST OF ALL TONE SEQUENCES USED IN THIS STUDY

1. 邬安英修飞机。Wū Ānyīng xiū fēijī.
   “Wu Anying repairs planes.”

2. 邬安应修飞机。Wū Ān yīng xiū fēijī.
   “Wu An should repair planes.”

3. 殷安青摸猫咪。Yīn Ānqīng mō māomī.
   “Yin Anqing pets a kitty.”

4. 殷安轻摸猫咪。Yīn Ān qīngmō māomī.
   “Yin An gently pets a kitty.”

5. 王明来拿羊毛。Liú Míng lái ná yángmáo.
   “Wang Ming comes to get wool.”

6. 王明涞拿羊毛。Liú Mínglái ná yángmáo.
   “Wang Minglai gets wool.”

7. 刘明来游云南。Liú Míng lái yóu Yúnnán.
   “Liu Ming comes to travel in Yunnan.”

8. 刘明来游云南。Liú Míng lái yóu Yúnnán.
   “Liu Ming comes to travel in Yunnan.”

9. 李敏响买母马。Lí Mǐnxǐāng mǎi mǔmǎ.
    “Li Minxiang buys a mare.”

10. 李敏想买母马。Lí Mǐn xiǎng mǎi mǔmǎ.
    “Li Min wants to buy a mare.”
11. 李伟想买野鸟。Lǐ Wěi xiǎng mǎi yěniǎo.
   “Li Wei wants to buy a wild bird.”

12. 李伟响买野鸟。Lǐ Wēixiǎng mǎi yěniǎo.
    “Li Weixiang buys a wild bird.”

13. 陆蔚用慢用药。Lù Wèi yòng màn yòng yào.
    “Lu Wei uses the slow medicine.”

14. 陆卫用卖孕药。Lù Wèiyòng mài yùnyào.
    “Lu Weiyong uses the contraceptive medicine.”

15. 魏丽要卖腊肉。Wèi Lì yào mài làróu.
    “Wei Li wants to sell bacon.”

16. 魏立耀卖腊肉。Wèi Lìyào mài làróu.
    “Wei Liyao sells bacon.”

17. 南梦来卖燃料。Nán Mèng lái mài ránliào.
    “Nan Meng comes to sell fuels.”

18. 南梦莱卖燃料。Nán Mènglái mài ránliào.
    “Nan Menglai sells fuel.”

19. 罗燕谈论名利。Luó Yàn tánlùn míngli.
    “Luo Yan talks about fame and profit.”

20. 罗彦坛论名利。Luó Yàntán lùn míngli.
    “Luo Yantan talks about fame and profit.”

21. 陆岩练习育苗。Lù Yán liàn xí yùmiáo.
    “Lu Yan practises growing seeds.”

22. 陆言练学育苗。Lù Yánliàn xué yùmiáo.
    “Lu Yanlian practices growing seeds.”

23. 孟岩爱读外文。Mèng Yán ài dú wàiwén.
    “Meng Yan likes studying foreign languages.”

24. 孟言艾读外文。Mèng Yán’ài dú wàiwén
    “Meng Yanai studies foreign languages.”
### APPENDIX B: SUBJECT INFORMATION

<table>
<thead>
<tr>
<th>Group</th>
<th>No.</th>
<th>Gender</th>
<th>age</th>
<th>Dur. of lang. Study (yrs)</th>
<th>Dur. of study abroad (yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>1</td>
<td>F</td>
<td>23</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>2</td>
<td>F</td>
<td>24</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>3</td>
<td>F</td>
<td>19</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>4</td>
<td>M</td>
<td>21</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>9</td>
<td>F</td>
<td>23</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>10</td>
<td>M</td>
<td>22</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>12</td>
<td>M</td>
<td>31</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>13</td>
<td>F</td>
<td>30</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>14</td>
<td>F</td>
<td>26</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Native</td>
<td>15</td>
<td>F</td>
<td>35</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Advanced-upper</td>
<td>2</td>
<td>M</td>
<td>23</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Advanced-lower</td>
<td>3</td>
<td>M</td>
<td>24</td>
<td>4</td>
<td>0.5</td>
</tr>
<tr>
<td>Advanced-upper</td>
<td>4</td>
<td>F</td>
<td>26</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>Advanced-lower</td>
<td>5</td>
<td>M</td>
<td>28</td>
<td>5</td>
<td>2.5</td>
</tr>
<tr>
<td>Advanced-upper</td>
<td>6</td>
<td>M</td>
<td>27</td>
<td>6</td>
<td>2.5</td>
</tr>
<tr>
<td>Advanced-lower</td>
<td>7</td>
<td>M</td>
<td>24</td>
<td>5</td>
<td>1.5</td>
</tr>
<tr>
<td>Advanced-lower</td>
<td>9</td>
<td>M</td>
<td>23</td>
<td>5</td>
<td>1.5</td>
</tr>
<tr>
<td>Advanced-lower</td>
<td>10</td>
<td>M</td>
<td>22</td>
<td>4</td>
<td>0.75</td>
</tr>
<tr>
<td>Advanced-upper</td>
<td>11</td>
<td>F</td>
<td>29</td>
<td>13</td>
<td>2.5</td>
</tr>
<tr>
<td>Advanced-lower</td>
<td>12</td>
<td>M</td>
<td>25</td>
<td>3</td>
<td>0.25</td>
</tr>
<tr>
<td>Interm-upper</td>
<td>2</td>
<td>F</td>
<td>20</td>
<td>2</td>
<td>0.17</td>
</tr>
<tr>
<td>Interm-lower</td>
<td>3</td>
<td>F</td>
<td>20</td>
<td>2.5</td>
<td>0.17</td>
</tr>
<tr>
<td>Interm-lower</td>
<td>4</td>
<td>M</td>
<td>23</td>
<td>3</td>
<td>0.25</td>
</tr>
<tr>
<td>Interm-lower</td>
<td>5</td>
<td>M</td>
<td>22</td>
<td>1.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Interm-upper</td>
<td>6</td>
<td>F</td>
<td>20</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Interm-lower</td>
<td>7</td>
<td>M</td>
<td>21</td>
<td>3</td>
<td>0.5</td>
</tr>
<tr>
<td>Interm-upper</td>
<td>8</td>
<td>M</td>
<td>20</td>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td>Interm-upper</td>
<td>9</td>
<td>F</td>
<td>26</td>
<td>2.5</td>
<td>0.17</td>
</tr>
<tr>
<td>Interm-lower</td>
<td>10</td>
<td>M</td>
<td>33</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Interm-upper</td>
<td>11</td>
<td>M</td>
<td>29</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>
APPENDIX C: NUMBERS OF UTTERANCES USED IN THE ANALYSIS OF DIFFERENT TONE SEQUENCES

1. The T1 sequence

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>SS-24</th>
<th>SS-33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>Female</td>
<td>8 (8) *</td>
<td>8 (8)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6 (12)</td>
<td>4 (12)</td>
</tr>
<tr>
<td>Advanced</td>
<td>Female</td>
<td>4 (4)</td>
<td>4 (4)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>9 (16)</td>
<td>13 (16)</td>
</tr>
<tr>
<td>Native</td>
<td>Female</td>
<td>14 (14)</td>
<td>14 (14)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6 (6)</td>
<td>6 (6)</td>
</tr>
</tbody>
</table>

*The first number here refers to the number of the utterances used in the analysis, while the second number in the brackets refers to the total number of utterances in that category (ibid).

2. The T2 sequence

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>SS-24</th>
<th>SS-33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>Female</td>
<td>8 (8)</td>
<td>8 (8)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>8 (12)</td>
<td>4 (12)</td>
</tr>
<tr>
<td>Advanced</td>
<td>Female</td>
<td>4 (4)</td>
<td>2 (12)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>10 (16)</td>
<td>11 (16)</td>
</tr>
<tr>
<td>Native</td>
<td>Female</td>
<td>14 (14)</td>
<td>14 (14)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6 (6)</td>
<td>6 (6)</td>
</tr>
</tbody>
</table>
3. The T4 sequence

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>SS-24</th>
<th>SS-33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>Female</td>
<td>8 (8)</td>
<td>8 (8)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6 (12)</td>
<td>6 (12)</td>
</tr>
<tr>
<td>Advanced</td>
<td>Female</td>
<td>3 (4)</td>
<td>4 (4)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>10 (16)</td>
<td>12 (16)</td>
</tr>
<tr>
<td>Native</td>
<td>Female</td>
<td>14 (14)</td>
<td>14 (14)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6 (6)</td>
<td>6 (6)</td>
</tr>
</tbody>
</table>

4. The T2T4 sequence

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>SS-24</th>
<th>SS-33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>Female</td>
<td>8 (8)</td>
<td>6 (8)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6 (12)</td>
<td>5 (12)</td>
</tr>
<tr>
<td>Advanced</td>
<td>Female</td>
<td>4 (4)</td>
<td>4 (4)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>9 (16)</td>
<td>12 (16)</td>
</tr>
<tr>
<td>Native</td>
<td>Female</td>
<td>14 (14)</td>
<td>14 (14)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6 (6)</td>
<td>6 (6)</td>
</tr>
</tbody>
</table>

5. The T4T2 sequence

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>SS-24</th>
<th>SS-33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>Female</td>
<td>8 (8)</td>
<td>6 (8)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>9 (12)</td>
<td>9 (12)</td>
</tr>
<tr>
<td>Advanced</td>
<td>Female</td>
<td>4 (4)</td>
<td>4 (4)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>12 (16)</td>
<td>10 (16)</td>
</tr>
<tr>
<td>Native</td>
<td>Female</td>
<td>14 (14)</td>
<td>14 (14)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6 (6)</td>
<td>6 (6)</td>
</tr>
</tbody>
</table>
APPENDIX D: NUMBERS OF THE OCCURRENCE OF PITCH RESET IN THE T1 AND THE T4 SEQUENCES

1. The T1 sequence

<table>
<thead>
<tr>
<th>Group</th>
<th>SS-24</th>
<th>SS-33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>15 (16) *</td>
<td>4 (13)</td>
</tr>
<tr>
<td>Advanced</td>
<td>11 (14)</td>
<td>7 (17)</td>
</tr>
<tr>
<td>Native</td>
<td>12 (20)</td>
<td>5 (20)</td>
</tr>
</tbody>
</table>

*The first number here refers to the number of the occurrence of pitch reset, while the second number in the brackets refers to the total number of utterances in that category (ibid).

2. The T4 sequence

<table>
<thead>
<tr>
<th>Group</th>
<th>SS-24</th>
<th>SS-33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>1 (14)</td>
<td>3 (14)</td>
</tr>
<tr>
<td>Advanced</td>
<td>0 (13)</td>
<td>5 (16)</td>
</tr>
<tr>
<td>Native</td>
<td>8 (20)</td>
<td>17 (20)</td>
</tr>
</tbody>
</table>
REFERENCES


203


