THE INTERACTION OF LEXICAL AND DISCOURSE-LEVEL CATEGORIES IN SECOND LANGUAGE PHONETICS

A Dissertation submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics

By

Kimberly Teague, M.A.

Washington, DC
July 31, 2011
THE INTERACTION OF LEXICAL AND DISCOURSE-LEVEL CATEGORIES IN SECOND LANGUAGE PHONETICS

Kimberly L. Teague, M.A.

Thesis Advisor: Elizabeth Zsiga, Ph.D.

ABSTRACT

This dissertation examines the interaction between lexical and discourse level categories in second language speech; specifically, the interaction of tone and intonation in the speech of Mandarin learners of English. The topic is investigated by analyzing recordings of 14 monolingual speakers of Mandarin, 16 Mandarin-English bilinguals, and 16 monolingual speakers of English producing word-lists, and statements and questions in both languages. The following questions are addressed:

(1) Does the Speech Learning Model (SLM) extend to tone and intonation?

The SLM predicts merger of L1 and L2 segmental categories, but has never been expanded to the suprasegmental level. The predictions of the SLM at both levels were tested here by examining Voice Onset Time (VOT) of Mandarin /pʰ/ and English /p/, timing of F0 peak in Mandarin H and English H*L, and rate of declination in word-lists. Mandarin-English speakers produced both Mandarin-like and English-like phonetic variants in both languages. Thus, a bi-directional interaction that was similar to segmental interaction was found between Mandarin tone and English intonation. The category merger predicted by the SLM was not found, however, at either the segmental or suprasegmental level.
(2) Which model of tone and intonation, the Autosegmental Metrical (AM) or the Parallel Encoding and Target Approximation (PENTA) model, best accounts for Mandarin-English intonation?

It is argued that the findings of this study are best accounted for by AM because (a) tone and intonation interact bi-directionally at the phonetic level, and (b) Mandarin-English intonation can be accounted for through sequences of level tones.

Two different intonation patterns were found in Mandarin-English statements and questions. In the first pattern, a L boundary tone (L#) marked the right edge of every contrastive focus word in statements and questions; these utterances also had a low post-focus pitch register. In the second pattern, the L# was found in statements, but a H boundary tone occurred in questions; statements had a low post-focus pitch register, while questions had a high one. It is suggested that, since pitch register is predictable from tonal pattern, only tones need be represented to account for Mandarin-English intonation, consistent with an AM account.
ACKNOWLEDGMENTS

It was Lisa Zsiga’s own work on Russian-English phonology that first drug me away from the San Diego beaches to study acoustic phonetics and suprasegmental phonology in Washington, DC. While at first, missing the sunny beaches and flip flops, I wondered whether it would all be worth it- I know now that I couldn’t have found a better advisor than Lisa. Thank you Lisa for the engaging and informative classes, the fun research and teaching assignments, and for reading my many (embarrassing) drafts while I figured out how to write a dissertation. Most of all, thank you for the support and patience you gave me during not one, not two, but three (surprise!) pregnancies. Without your support, none of this would have been possible.

Thank you to Jessie Chiang and Yan Liu for the help finding participants. You two made offered so much help to me and to this study. I also want to thank Cristina Sanz and Dan Loehr for the helpful advice and many meetings along the way. And thank you to Father Sara for feeding me and my family so many Sunday afternoons. I finally took your advice and “wrapped this up!”

A big thank you to Charlotte Webb and Robert Underhill, my first professors of second language acquisition and phonology. I hope you both realize what fond memories I have of graduate school because of you. You are a large part of what made me continue in my studies. Thank you Char for the many dinners, all the laughs, and for lending me your doggie to go running on the beach! Thank you Dr. Underhill for the hours of time you spent answering phonology questions for me… I think I finally got the hang of it! Thank you Deborah Poole for helping me so much along the way. Whenever I’m asked to think about the best class I’ve ever taken, I always think of yours. I admire you greatly for being such a wonderful teacher. Thank
you for the many teaching opportunities you gave me, and for the time you spent directing me and pushing me to become a better teacher myself.

Finally, thank you to my family and friends. Mom and dad, thank you for the encouragement you gave me and for all the babysitting hours you put in so I could write… I promise next time you visit, I’ll actually hang out! Thank you to Aunt Marsha for watching the boys while I prepared for and defended my dissertation…you get extra points for watching Aza after the big fall the night before I defended! Thank you to Aunt MaryAnne who helped introduce me to anyone and everyone in Texas that might be able to help me find participants and actually get done! Thank you Brenda and Jackie for giving my boys so much love… and taking care of them extra hours while I stressed over my dissertation deadlines! Thank you Mike and Tegan Diercks who also helped out with Aza so much!

Last but not least, thank you to my boys: Tom, Aza, and Eli. You reminded me on a daily basis that life existed outside of linguistics and my dissertation! When I look back on my time at Georgetown, I’ll always smile fondly on the hours we poured into building cities and zoos from blocks, playing cars and trains, painting, dancing, climbing, jumping, and laughing.
TABLE OF CONTENTS

Chapter 1: Research Questions ................................................................. 1

1.1. Why Study Mandarin-English pitch? .................................................... 1

1.1.1. The structure of tone and intonation .............................................. 1

1.1.2. Category interaction in bilingual speech ....................................... 2

1.1.3. Representation of register ............................................................. 3

1.1.4. Alternations for mapping tones to prosodic material ...................... 4

1.2. Preview of the study ........................................................................ 5

1.2.1. Structure of the study ................................................................. 5

1.2.2. Results on category interaction in Mandarin-English .................... 6

1.2.3. Results on tonal mapping in Mandarin-English statements and questions ...... 7

1.2.4. Results on pitch register in Mandarin-English statements and questions ...... 8

Chapter 2: Background ........................................................................... 10

2.1. Tone and intonation ....................................................................... 10

2.1.1. Tone and intonation systems ....................................................... 10

2.1.1.1. General overview ................................................................. 10

2.1.1.2. Tone and intonation in Standard Mandarin ............................... 11

2.1.1.3. Intonation in English ............................................................. 15

2.1.1.4. Summary of Mandarin and English tone and intonation .......... 19

2.1.2. Autosegmental-Metrical (AM) model of tone and intonation ............ 20

2.1.2.1. General overview ................................................................. 20

2.1.2.2. Surface contours and autotones ............................................. 22
2.1.2.3. Predicting surface patterns ................................................................. 26

2.1.2.4. Comparison of Mandarin and English from an AM perspective ........... 28

2.1.3. A Parallel Encoding and Target Approximation (PENTA) account of tone and intonation ................................................................. 29

2.1.3.1. General overview ................................................................................. 29

2.1.3.2. Surface contours and local pitch targets ........................................... 31

2.1.3.3. Non-local pitch targets ........................................................................ 32

2.1.3.4. Comparison of Mandarin and English from a PENTA perspective ...... 32

2.1.4. Rationale for research ............................................................................. 34

2.2. Category interaction in second language speech ........................................ 36

2.2.1. Speech Learning Model (SLM) ............................................................... 37

2.2.1.1. General overview .............................................................................. 37

2.2.1.2. Predictions about category interaction .............................................. 37

2.2.1.3. Defining the constructs: Perceptual similarity .................................... 40

2.2.2. Phonetic category interaction of segments ............................................ 42

2.2.2.1. VOT interaction in stop consonants ............................................... 42

2.2.2.2. Formant frequency interaction in vowels .......................................... 44

2.2.3. Phonetic category interaction of pitch contours .................................... 46

2.2.3.1. F0 peak alignment interaction in phonological tones ....................... 46

2.2.3.2. Declination interaction in bilingual pitch contours .......................... 49

2.2.4. Motivation of research questions .......................................................... 50

2.3. Cross-language mapping of tone to prosodic material ............................ 51
2.3.1. Tonal association scale ................................................................. 52
  2.3.1.1. General overview ................................................................. 52
  2.3.1.2. Predictions about tonal mapping .......................................... 54
  2.3.1.3. Considering other constraints with respect to the tonal association scale ................................................................. 55
  2.3.4. Motivation of research questions. ............................................. 57
2.4. Research hypotheses ........................................................................ 57

Chapter 3: Preliminary Perception Study .............................................. 61
3.1. Overview of the perception experiment ......................................... 61
3.2. Participants ...................................................................................... 61
  3.2.1. Participant questionnaire ........................................................... 62
    3.2.1.1. Language ........................................................................... 62
    3.2.1.2. AOA and LOR. ................................................................. 62
  3.3. Perception task 1: Transcription .................................................. 63
    3.3.1. Phonetic stimuli .................................................................... 63
  3.3.2. Materials .................................................................................. 64
  3.3.3. Procedures ................................................................................ 64
  3.3.4. Results of /p/ stimuli ............................................................... 64
  3.3.5. Results of H* over bi-syllabic words ........................................ 65
  3.4. Perception task 2: Similarity judgment .......................................... 66
    3.4.1. Phonetic stimuli .................................................................... 66
    3.4.2. Materials ................................................................................ 67
4.4.6. Mandarin question task worksheet ................................................................. 79
4.5. Recording procedures ........................................................................................ 79
  4.5.1. List task ........................................................................................................ 80
  4.5.2. Statement task ............................................................................................ 80
  4.5.3. Question task ............................................................................................ 81
4.6. Data preparation ................................................................................................ 82
  4.6.1. Transcription ................................................................................................ 83
    4.6.1.1. Prosodic structure .................................................................................. 83
    4.6.1.2. Prominence relations ........................................................................... 83
    4.6.1.3. Tone ........................................................................................................ 84
  4.6.2. Measurement of phonetic data .................................................................... 85
  4.6.3. Data preparation .......................................................................................... 86
    4.6.3.1. VOT values ........................................................................................... 86
    4.6.3.2. Declination ............................................................................................ 87
    4.6.3.3. Peak alignment ...................................................................................... 89
    4.6.3.4. Pitch register ......................................................................................... 90
Chapter 5: Effects of bilingualism on category interaction .................................... 91
  5.1. Category interaction of voiceless bilabial stops /p/ ......................................... 91
  5.2. Rate of declination ........................................................................................ 95
  5.3. F0 peak alignment ......................................................................................... 98
  5.4. Summary and discussion ................................................................................ 102
    5.4.1. SLM hypotheses not supported ............................................................... 102
5.4.2. A parallel in bilingual segmental and supra-segmental production .......... 105
5.4.3. Understanding the data from an AM or PENTA perspective ................. 105
5.4.4. What do these results reveal about tone and intonation structure? .......... 107

Chapter 6: Mandarin-English tone to syllable mapping .................................. 108
6.1. Prosodic structure and syllabic prominence ........................................... 108
6.2. Mapping of tone to syllable ................................................................. 114
   6.2.1. Monolingual English ........................................................................ 114
      6.2.1.1. Statements ............................................................................... 114
      6.2.1.2. Questions ............................................................................... 116
      6.2.1.3. Assessing the tonal association scale ........................................ 121
   6.2.2. Mandarin-English Pattern A ......................................................... 121
      6.2.2.1. Statements ............................................................................... 121
      6.2.2.2. Questions ............................................................................... 126
      6.2.2.3. Assessing the tonal association scale ........................................ 131
   6.2.3. Mandarin-English pattern B ........................................................... 131
      6.2.3.1. Statements ............................................................................... 132
      6.2.3.2. Questions ............................................................................... 133
      6.2.3.3. Assessment of the tonal association scale ................................... 137
6.3. Discussion ............................................................................................... 137
   6.3.1. Review of findings .......................................................................... 137
   6.3.2. The predictions of the tonal association scale ................................. 138
6.3.3. Boundary tones on the edges of words ................................................. 140

6.3.4. Does Mandarin-English intonation resemble Mandarin lexical tone? ....... 146

Chapter 7: Pitch register in Mandarin-English statements and yes/no questions .... 149

7.1. Pitch register ............................................................................................ 149

7.1.1. Monolingual English ............................................................................. 149

7.1.2. Mandarin-English low pitch register in statements and questions ........... 150

7.1.3. Mandarin-English low pitch register in statements and high pitch register in
questions ........................................................................................................ 152

7.2. Register and tonal pattern ....................................................................... 153

7.3. Summary and discussion ......................................................................... 154

Chapter 8: Conclusion ...................................................................................... 156

8.1. Review of the study .................................................................................. 156

8.1.1. How do speech categories interact in bilingual speech? ......................... 156

8.1.2. Do pitch and segments interact similarly in bilingual speech? ............... 157

8.1.3. Does L1 tone interact with L2 intonation in Mandarin-English speech? .... 157

8.1.4. Does the tonal association scale predict Mandarin-English tonal patterns? . 158

8.1.5. Is register predictable from tonal pattern in Mandarin-English? .............. 159

8.2. Future research on Mandarin in pitch register from a tone-sequence framework 160

Appendix A ....................................................................................................... 162

Appendix B ....................................................................................................... 164

Appendix C ....................................................................................................... 166

Appendix D ....................................................................................................... 167
LIST OF FIGURES

Figure 2.1. Tones in citation form uttered by a male speaker over the syllable “ma.”
Adapted from Xu, Y. (2009)................................................................. 11

Figure 2.2. Beijing speaker’s F0 maximum alignment heard during Mandarin Tone 1
pronounced over the bi-syllabic sequence “mama.”........................................ 13

Figure 2.3. Mandarin question and statement pair with initial focus. The question has a
higher pitch following the focused element than does the statement.. .................. 14

Figure 2.4. Rate of word-list declination in Mandarin heard during a sequence of all high
tones............................................................................................................. 15

Figure 2.5. Rate of word-list declination in Mandarin heard during a sequence of falling
tones............................................................................................................. 15

Figure 2.6. English falling and rising contours on the syllable “ma” changes sentence-
type ............................................................................................................ 15

Figure 2.7. Single rising contour heard in English question with initial focus........... 16

Figure 2.8. Alignment of F0 maxima heard during English falling contour over the
bi-syllabic sequence “mama.”........................................................................ 18

Figure 2.9. Rate of declination in English wordlist; the final word has a steep F0
decline .......................................................................................................... 19

Figure 2.10. A sketch of Penta (Xu, 2009) ....................................................... 30

Figure 2.11. Rising surface contour in English monosyllabic question ................. 35

Figure 2.12. Rising surface contour in English bi-syllabic question with initial focus 35

Figure 2.13. VOT of a voiceless bilabial plosive in /a’pa/ ................................. 38

Figure 2.14. Rising surface contour in English monosyllabic question ............... 35

Figure 2.15. Rising surface contour in English bi-syllabic question with initial focus 35

Figure 4.1. Word-list with one phrase of high tones........................................... 87
Figure 4.2. Word-list with one phrase of falling contours................................. 88

Figure 4.3. Word-list with two phrases of falling contours. Phrase break is indicated through a L% boundary tone symbol............................................................. 88

Figure 4.4. Word-list with one phrase of rising and falling contours ...................... 89

Figure 5.1. Mean VOT of voiceless bilabial stops in Mandarin and English.............. 93

Figure 5.2. Mandarin declination in word-list with Tone 1 heard over the initial syllable of bi-syllabic words ...................................................................................... 96

Figure 5.3. English declination in word-list with H*L contour heard over the initial syllable of bi-syllabic words ...................................................................................... 96

Figure 5.4. Comparison of final position F0 by monolinguals and bilinguals .......... 97

Figure 5.5. Mean F0 peak alignment values by monolinguals and bilinguals.......... 100

Figure 5.6. Bilingual F0 peak alignment values in Mandarin (left) and English (right) ..................................................................................................................... 101

Figure 6.1. Example short question from which contrastive focus could not be identified by Mandarin-English speaker who produced the phrase and then listened to it during a subsequent focus-identification task......................................................... 109

Figure 6.2. Spectrogram of a long utterance divided into two phonological phrases; the phrase boundary is divided by a substantial pause at the end of the focused element “Manny.” .................................................................................................................. 110

Figure 6.3. Monolingual English production of the word “to” in the phrase “Ann will need to win.” The first formant frequency is slightly elevated, and the distance between the first and second formant frequencies are relatively large, as expected in the production of a reduced vowel, or shwa. Moreover, the timing of the vowel is very short .................................................................................................................. 112

Figure 6.4. Bilingual Mandarin-English production of the word “to” in the phrase “Ann will need to win.” The first formant frequency is low, and there is hardly any distance between the first and second formant frequencies, as expected in the production of a fully articulated high, back vowel, /u/ .................................................................................. 113
Figure 6.5. Falling F0 contours exemplary of those produced by native English speakers in short and long statements with initial contrastive focus ........................................... 114

Figure 6.6. F0 peak and valley alignment to the syllable during statements produced by native English speakers ........................................................................................................... 115

Figure 6.7. One level of prominence, sentential prominence, conveyed via tone. The word “marry” with initial lexical stress bears no tone, while the focused element “Manny” involves a H*L contour ........................................................................................................... 116

Figure 6.8. Typical rising F0 contours produced by native English speaker in short and long questions with initial contrastive focus ................................................................. 117

Figure 6.9. F0 minimum alignment to the syllable during focused elements produced by Native English speakers ........................................................................................................... 118

Figure 6.10. Rising contour through the phrase “Manny will marry?” ........................................... 119

Figure 6.11. Monolingual English statements and questions with initial contrastive focus ................................................................................................................................. 120

Figure 6.12. Mandarin-English Pattern A statement F0 contours ........................................... 122

Figure 6.13. Two patterns for F0 peak alignment to the syllable for the H toneme in Mandarin-English: (a) H aligned near the right edge of the vowel; (b) H aligned near the left edge of the vowel and extended to the right edge of the syllable ........................................... 123

Figure 6.14. Mandarin-English late-aligned H tone followed by a short disjuncture... 124

Figure 6.15. H tone inserted on stressed syllables of multisyllabic words of Mandarin-English Pattern A phrases ........................................................................................................... 125

Figure 6.16. No H tone inserted on stressed monosyllabic words of Mandarin-English Pattern A phrases ........................................................................................................... 125

Figure 6.17. L boundary tone inserted on the right edge of focus words in Mandarin-English Pattern A questions ........................................................................................................... 127

Figure 6.18. Mandarin-English Pattern A production of “Ann married?” with final LH% ................................................................................................................................. 127
Figure 6.19. Mandarin-English Pattern A intonation contours in statements and questions ................................................................. 130

Figure 6.20. Mandarin-English Pattern B H tone next to a L boundary tone on the edge of the focus word ........................................ 132

Figure 6.21. Mandarin-English Pattern B production of the phrase “Manny will marry.” No H tone is inserted on the stressed syllable of post focus words ............ 133

Figure 6.22. Mandarin-English Pattern B production of the short question “Ann won?” The post-focus region has a higher pitch range than the focus word ........... 134

Figure 6.23. Mandarin-English Pattern B intonation contours in statements and questions ................................................................. 136

Figure 7.1. A divergence in pitch range emerges between English statements and questions as produced by monolingual speakers of English................................. 150

Figure 7.2. Mandarin-English corresponding statement/question pair with no difference in pitch register ......................................................... 151

Figure 7.3. Mandarin-English corresponding statement/question pair with no difference in pitch register ......................................................... 151

Figure 7.4. Corresponding statement and question pair of the phrase “Ann will need to win with a difference in post-focus pitch register........................................ 152

Figure 7.5. Corresponding statement and question pair of the phrase “Manny will Marry” with a difference in post-focus pitch register........................................ 152
LIST OF TABLES

Table 2.1. Comparison of Mandarin and English pitch systems .................................................. 20

Table 2.2. Comparison of Mandarin and English pitch systems from an AM perspective ................................................................. 29

Table 2.3. Comparison of Mandarin and English pitch systems from a PENTA perspective.. ........................................................................ 33

Table 2.4. OT tableau showing that in Mandarin, σ←T >> Dep [T] ........................................... 56

Table 3.1. Stimulus items produced with a H*L pitch accent ................................................... 63

Table 3.2. Results of /p/ transcriptions .................................................................................. 65

Table 3.3. Results of H*L transcriptions ................................................................................ 65

Table 3.4. Triadic-similarity-judgment trials ............................................................................ 67

Table 3.5. Results of /p/ similarity judgment task ................................................................... 68

Table 3.6. Results of /p/ similarity judgment task ................................................................... 69

Table 4.1. Summary of bilingual Mandarin-English speakers ................................................... 76

Table 4.2. Phonetic measurements........................................................................................... 86

Table 6.1. Monolingual English prosodic structure................................................................. 111

Table 6.2. Mandarin-English prosodic structure..................................................................... 112

Table 6.3. OT tableau showing a marked %T-T* input will emerge as an unmarked T*-T% syllable under the ranking markedness >> faithfulness .................................................. 142

Table 6.4. Four OT tableaux which demonstrate T*-T%, T*, %T-T*, %T*-T% patterns for boundary tones on edges of intonational phrases will emerge when faithfulness>> markedness ................................................................................................. 143

Table 6.5. OT tableaux which demonstrate T*-T% or T* phrases emerge from ranking Dep-[T] >> T% and No%T >> Max[T].............................................................................................................. 144

xix
Table 6.6. *OT tableaux which demonstrate T*-T% or %T-T*-T% phrases emerge from the ranking Dep-IO >> Onset and NoCoda >> Max-IO ........................................... 145

Table 7.1. *Comparison of post-focus pitch register and Mandarin-English tonal pattern* .................................................................................................................................... 153
CHAPTER 1: RESEARCH QUESTIONS

1.1. Why study Mandarin-English pitch?

1.1.1. The structure of tone and intonation

Every spoken language is produced with a melody. When this melody helps to create lexical contrasts in speech, it is termed tone; when used to convey discourse-level contrasts, such as whether an utterance is a statement or question, it is termed intonation. Phonologists generally agree that tone and intonation should be represented as a sequence of high (H) and low (L) autotones positioned on a single layer of linguistic structure (Ladd, 2008; Pierrehumbert, 1980). Particularly for East Asian languages, however, this way of analyzing intonation contours has been problematic. A case in point is the Mandarin Chinese intonation of sentence-type. In Mandarin, statements and questions can be differentiated by their overall pitch range, called register (Xu, 1999). Statements are typically heard at a lower pitch register, while questions at a higher one. Unclear how to represent register through sequences of H and L autotones, researchers’ accounts of Mandarin tone and intonation have either skipped over intonation of sentence-type (Duanmu, 2000), or they have rejected the typical tone-sequence analysis of intonation in lieu of a model which represents intonation contours along different layers of linguistic structure (Bao, 1999; Chen, 2005; Xu, 1999; Yip, 1999; 2002). The different ways of modeling tone and intonation, however, make different predictions about the course of L2 learning. From Mandarin to English, for example, tone-sequence models predict that Mandarin tone and English intonation should interact in bilingual speech since they are along the same structural level. Alternative models that separate tone and intonation into different structural layers, on the other hand, make the opposite prediction. This study addresses the issue by testing
whether tones from Mandarin interact with intonation from English in Mandarin-English bilingual speech. If interaction does indeed occur, this will support a model that represents tone and intonation along the same structural level; however, if interaction does not occur, a model that separates tone and intonation onto different layers of linguistic structure will be supported.

1.1.2. Category interaction in bilingual speech

A definition of interaction is first needed to test whether interaction occurs between tone in a bilingual’s L1 and intonation in the L2. The Speech Learning Model (SLM), a popular model of second language acquisition of phonology, provides one account of bilingual category interaction; however, the model focuses solely on consonants and vowels, called segments (Flege, 1995). It predicts that bilinguals will merge perceptually similar L1 and L2 segmental categories (vowels and consonants). Notably, this merger will be evident in that phonetic values of bilinguals’ segments will fall in between the phonetic values of monolingual speakers of each language. Will pitch contours interact similarly?

Little research has gone into the exploration of category interaction beyond the segmental level; moreover, the single source which focuses on the topic reports inconclusive results since the participants in the study were actually trilinguals, and since the production of pitch contours was examined without comparing findings to segmental production for the same speakers (Mennen, 2004). What is clear from the study is that the SLM did not make the correct predictions for the participants’ production of intonation; that is to say, the phonetic values (peak alignment) of the participants’ L1/L2 intonation contours did not merge. It is unclear, however, whether this finding occurred because (a) segmental categories and intonation contours interact differently in bilingual speech, (b) the SLM makes the wrong predictions about bilingual
category interaction, or (c) the interaction of bilingual and trilingual speech categories must be represented differently. To investigate these alternatives, this study tests the SLM by comparing monolingual and bilingual Mandarin and English segmental, tone, and intonation categories. Specifically, the following phonetic values are examined in monolingual and bilingual Mandarin, and monolingual and bilingual English: (a) duration of voice onset time in voiceless bilabial stops, (b) rate of word-list declination, and (c) alignment of F0 peaks in falling contours heard over bi-syllabic words with initial stress. If the SLM hypotheses are true for Mandarin-English segments but not for Mandarin-English tone and intonation, we will know that the segments and pitch interact differently in bilingual speech. If, however, the hypotheses are false for both Mandarin-English segments and pitch, we will have evidence that the SLM is not an accurate model of bilingual category interaction. Finally, if the SLM hypotheses are true for segments and pitch, we can conclude that the inconclusive findings from the previous study stemmed from the selection of trilingual participants.

1.1.3. Representation of register

The second part of this dissertation turns its attention toward the phonological representation of pitch register through a tone-sequence model. In Chapter 5, results on category interaction in this study reveal a two-way interaction between tone and intonation. This finding supports a model that represents tone and intonation along a single layer of phonological structure. The problem of how to represent Mandarin register through a sequence of H and L autotones, however, still remains. This study addresses the issue by examining native Mandarin speakers’ intonation of yes/no statements and questions in English. When we examine register in Mandarin, there is a consistency across all speakers’ utterances such that statements are produced
with a lower register and questions with a higher one. When Mandarin speakers begin to speak in English, this pattern between high and low register sometimes occurs in their English as well, however, not always (Teague, 2011). This inconsistency in the use of sentence-type register in the speech of Mandarin bilinguals speaking English provides an opportunity to examine whether register is just an incidental bi-product of a particular tonal sequence. If so, we might expect that the existence of a register difference between statements and questions will be predictable from a particular Mandarin-English tonal pattern. If register cannot be represented as sequences of H and L autotones, then no correlation between tonal sequence and register is expected.

1.1.4. Alternations for mapping tones to prosodic material

The tonal patterns of Mandarin-English statements and questions are first needed to test whether a particular tonal pattern relates to pitch register. The Optimality Theoretic Tonal Association Scale can be used as a starting point for predicting how tonal sequences might be realized differently under different typological systems (Gussenhoven, 2004). The scale provides a universal ranking of constraints which require tone to associate to specific locations within an utterance. At the top of the scale is a constraint which requires all accented syllables to bear tone; directly beneath this is another constraint that requires all stressed syllables to bear tone. Sliding in between the constraints of the Tonal Association Scale are faithfulness constraints which prohibit changes to be made to input tonal sequences. If the faithfulness constraints are low-ranked within the scale, we expect a phonological system in which a tone emerges on every stressed syllable of every utterance; if faithfulness constraints are ranked in the middle of the tonal association scale, we predict a system in which a tone will emerge only on the accented syllable of each utterance. To begin understanding the different Mandarin-English intonation
patterns in English statements and questions, this dissertation tests whether these two alternations are in fact present in Mandarin-English yes/no statements and questions with initial contrastive focus.

1.2. Preview of the study

1.2.1. Structure of the study

The remainder of the study examines Mandarin-English word-lists and yes/no statements and questions guided by the following five research questions:

(1) Does L1 tone interact with L2 intonation in Mandarin-English speech?

(2) Do tone and intonation contours interact similarly to segmental categories in Mandarin-English speech?

(3) Do speech categories merge in bilingual speech?

(4) Do tones associate to syllables in a manner predictable by the tonal association scale?

(5) Is the pitch register of an utterance predictable from its tonal pattern in Mandarin-English speech?

Chapter 2 provides a review of the literature on Mandarin and English tone and intonation, and category interaction during second language acquisition of phonology. It ends with the hypotheses that will be tested in this dissertation. Chapter 3 gives the details of a preliminary study used to confirm that the Mandarin and English sound categories under examination in this study were perceptually similar sounds for Mandarin-English bilinguals. Ten Mandarin-English bilinguals listened to sequences of syllables which differed by only one sound. The center syllable contained an English voiceless bilabial stop, and it was surrounded by syllables which had either a Mandarin voiceless bilabial stop or a Mandarin voiced bilabial stop.
Participants were asked to determine which sound, the first or third, was the same as the center sound. All bilingual speakers paired together Mandarin and English voiceless bilabial stops. Next, participants listened to sequences of bi-syllabic words with initial stress, such as mama. These sequences of sounds differed in terms of the tonal sequence heard over the segmental string. The center word was heard with an English falling contour (H*L), and the two surrounding Mandarin words were heard with Tone 1 (H tone) on the initial syllable, or Tone 1 (H tone) on the initial syllable followed by Tone 3 (L tone) on the final syllable. Bilingual speakers paired together Mandarin Tone 1 (the H tone) and English H*L heard over bi-syllabic sequences with initial stress (mama). These sounds were then used as the basis for designing tasks to examine category interaction in Mandarin-English speech.

Chapter 4 gives details of the methodology used to elicit data in this study. In particular, two main tasks were designed to elicit monolingual Mandarin, monolingual English, and bilingual Mandarin-English word-lists, and corresponding sets of yes/no statements and questions. The word-lists were used to explore category interaction in Mandarin-English speech, while the corresponding sets of yes/no statements and questions were analyzed to explore intonation patterns in Mandarin-English speech.

1.2.2. Results on category interaction in Mandarin-English

Chapter 5 discusses the results on Mandarin-English category interaction. The Mandarin-English speakers in this study produced the voice onset time (VOT) of voiceless bilabial stops, and the peak of Mandarin Tone 1 and English H*L, and the rate of word-list declination in a manner that was typically consistent with monolingual speakers for the language they were speaking. For example, when a bilingual speakers produced voiceless bilabial stops in Mandarin,
the VOT typically matched the VOT of monolingual Mandarin speakers; when bilinguals produced the sound in English, the VOT matched typical values for a monolingual English VOT. What made the data particularly interesting, however, was the large standard deviation found in the VOT values of Mandarin-English bilinguals in comparison to that found in the monolingual VOT values. Closer inspection of the bilingual VOT values revealed a bi-modal distribution in the Mandarin-English VOT values such that most VOT values fell within a range typical for monolingual speakers of the language being spoken, but a handful of tokens were produced with phonetic values that fell within the range of the opposing language. The results of VOT values were paralleled by the data on peak alignment in falling contours, and again in rate of declination. These findings imply that tone and intonation interact in bilingual speech in a similar way as do segments; moreover, this type of category interaction is not predicted by the Speech Learning Model. Rather than modeling similar L1 and L2 categories as undergoing phonetic merger in bilingual speech, it is suggested that these categories should instead be represented as a bi-modal distribution which builds from phonetic values heard during linguistic input. Furthermore, since a bi-directional interaction of Mandarin tone and English intonation was found, this study supports a model which represents tone and intonation along a single layer of linguistic structure.

1.2.3. Results on tonal mapping in Mandarin-English statements and questions

Chapter 6 moves on to discuss the results on Mandarin-English tonal sequences. In particular, three patterns for mapping tones to prosodic material were found. Monolingual English speakers mapped tones to the accented syllable, the syllable following the accented one, and the phrase edge, creating relatively smooth falling (H*L-L%) and rising (L*H-H%) contours
across statements and questions respectively. Mandarin-English speakers, by comparison, mapped tones to the accented syllable, the edge of the focus word, some stressed syllables of multisyllabic words, and the phrase-edge; two different tonal patterns were mapped in this way: Either all focus words ended in a low boundary tone whether in a statement or question, or focus words in statements ended in a low boundary tone while in questions, they ended in a high boundary tone. The results on tonal mapping in Mandarin English were not predicted by the tonal association scale since association between every stressed syllable and tone never occurred; moreover, the patterns found in Mandarin-English were much more complicated than what the tonal association hierarchy could predict. It is suggested that while the scale has provided us with a great starting point for thinking about the structure of intonation contours as required positions to which tones must associate, it is too simplistic of a system to capture the rich and intricate systems that can occur, while still ruling out impossible typologies. Deriving a set of constraints which will appropriately and accurately model the different tonal patterns found in Mandarin-English, and all languages as a whole, is beyond the scope of this dissertation; however, it is suggested that a good starting point may be to mimic the constraints used for syllable structure, ONSET and NoCoda, and to think of typological intonation systems as the result of positive and negative constraints on tonal association. This idea is fleshed out in more detail at the end of Chapter 6.

1.2.4. Results on pitch register in Mandarin-English statements and questions

Chapter 7 moves on to report the findings on pitch register in Mandarin-English statements and questions. A relationship is found between tonal pattern and pitch register. When speakers inserted a L boundary tone at the right edge of the focus word in both statements and
questions, no difference in pitch register was found between corresponding statements and questions. The other patterns, however, emerged with high register for questions and low register in statements. The relationship between tonal pattern and pitch register is used to argue that register is in fact an incidental bi-product of the tonal sequence, and does not, therefore, need to be accounted for on a separate layer of linguistic structure. Once the correct tonal sequence is found, the patterns of pitch register will come for free.

Chapter 8 then concludes and summarizes the study.
CHAPTER 2: BACKGROUND

2.1 Tone and intonation

2.1.1. Tone and intonation systems

2.1.1.1. General overview

In every language, a cadence of rising and falling melodies can be heard over strings of consonants and vowels, called segments. These suprasegmental cadences, when examined closely, involve a rich linguistic structure whereby speakers use their experience of pitch to systematically encode meaning during spoken communication. Sometimes pitch conveys lexical distinctions in language, termed tone. Other times, it signals discourse-level contrasts, such as whether a particular item is the focus of an utterance or not; this is termed intonation. All languages have intonation, but only some have lexical tone. Those that do use lexical tone are termed tonal languages, and are estimated to make up roughly 60-70 percent of languages in the world (Yip, 2002). Languages that do not make use of lexical tone are called non-tonal languages.

While most phonologists agree that tone and intonation systems should be represented as sequences of H and L tones, called autotones, that reside along a single layer of linguistic structure, a handful of linguists, particularly those who focus on tone and intonation in East Asian languages, reject typical tone sequence models for contour interaction models that represent tone and intonation on different layers of linguistic structure. The two types of models account for the facts of tone and intonation systems quite differently, and as a result, they make very different predictions about what should happen during second language acquisition. To exemplify this point, a review of the Mandarin and English tone and intonation systems will be
reviewed. This will be followed by a review of two popular models of tone and intonation, the tone-sequence Autosegmental Model (AM) and the contour interaction Parallel Encoding and Target Approximation Model (PENTA). A discussion concerning the different predictions these models make concerning the nature of Mandarin acquisition of English is then discussed.

2.1.1.2. Tone and intonation in Standard Mandarin

Standard Mandarin is a tonal language which contrasts four different lexical tones. Traditionally, these have been described as high, rising, low-dipping, and falling. In common speech, they’re referred to as Tone 1, Tone 2, Tone 3, and Tone 4 respectively. The four contrastive lexical tones are transcribed in (1) and are also shown in citation form in Figure 2.1.

(1)   Tone 1    [má]  ‘mother’
      Tone 2    [mã]  ‘horse’
      Tone 3    [mà]  ‘hemp’
      Tone 4    [mâ]  ‘scold’

*Figure 2.1. Tones in citation form uttered by a male speaker over the syllable “ma.” Adapted from Xu, Y. (2009).*
The consonants and vowels of each word displayed in (1) are identical. Speakers recognize the contrast in meaning by listening to the tones which are heard over the segmental string *ma*. When words are strung together to form phrases in Mandarin, the lexical tones create complex melodic contours by where pitch peaks and valleys can be heard on almost every syllable within the utterance.

The timing of these peaks and valleys across a given utterance remains consistent across multiple speakers of the same dialect. In the falling contour heard over the bi-syllabic sequence *mama*, seen in Figure 2.2, for example, speakers of Beijing Mandarin will change from a flat, high pitch to a falling one exactly at the right edge of the first syllable. This precise timing is termed fundamental frequency (F0) peak (or minima) alignment. F0 refers to the number of periods per second that exist in a particular sound wave; physiologically, this measurement correlates to the number of vocal fold vibrations that occur per second.
In addition to creating lexical contrasts through tone, Mandarin speakers also use pitch to convey certain discourse-level distinctions. Mandarin questions, for instance, are typically uttered with at an overall higher post-focus pitch range than are statements as can be seen in Figure 2.3. This type of intonation pattern is commonly termed *register*; questions are uttered in the high register, and statements in the low register. Note that the lexical tones within the utterance remain retrievable, while at the same time, intonation of sentence-type can be heard.
Figure 2.3. Mandarin question and statement pair with initial focus. The question has a higher pitch following the focused element than does the statement.

Just as contrasts between sentence-types are marked through pitch in Mandarin, the beginning and ending of discourse topics are also signaled through a gradual lowering of the F0 across the phrase, known as declination (Strik & Boves, 1995; Umeda, 1982). In Mandarin, declination has been observed in both declarative and interrogative utterances (Shih, 2000; Xu, 1999). In sequences of Tone 1, the high tone, the F0 declines steeply at first, and then tapers off toward the end. In sequences of falling contours, the same general pattern can be seen. Figure 2.4 displays a Mandarin word-list with all high tones; note that, even though all tones are high, a slight decline in F0 is still seen across the utterance. This can be compared to Figure 2.5 which shows a Mandarin word-list with falling contours. In the two figures, the F0 decline follows the same general pattern.
2.1.1.3. Intonation in English

English, by comparison, is a non-tonal language, and so pitch is used only to convey discourse-level distinctions. A rising or falling pitch contour heard over the same sequence of
consonants and vowels will not, therefore, change the lexical referent as is the case for Mandarin. Ma, whether expressed with a rising or falling pitch contour in English, refers to mother; however, in the former pronunciation it is recognized as a question, while in the latter, as a statement as seen in Figure 2.6.

![Figure 2.6](image)

**Figure 2.6.** English falling and rising contours on the syllable “ma” changes sentence-type.

Whereas in Mandarin, a typical phrase is heard with a rich melodic pattern, an English phrase of similar length, and with similar focus, will typically have very few turns in F0 direction. Moreover, unlike Mandarin, English melodies remain consistent even across different lexical items. Thus, the rising pattern seen in Figure 2.6 also occurs in Figure 2.7 even though the words in the utterances differ.
Just as the timing of F0 minima and maxima for a particular melody are consistent across speakers of the same dialect in Mandarin, so too are they for speakers of similar dialect in English (Ladd & Schepman, 2003). An English falling contour, heard over the bi-syllabic sequence seen in Figure 2.8, for example, has a F0 maximum aligned near the middle of the accented syllable. This anchoring of the F0 maximum is expected to be consistent across all speakers. Thus, just as Mandarin speakers display consistent alignment of F0 maxima and minima for a given contour, so do English speakers; however, the timing of F0 maxima and minima in Mandarin and English are not the same, even if the contours in Mandarin and English look similar.
Figure 2.8. Alignment of F0 maximum heard during English falling contour over the bi-syllabic sequence “mama.”

Also comparable to Mandarin, English signals the beginning and ending of discourse topics through declination, albeit the rate of decline differs slightly in English. In Mandarin, the rate of F0 decline begins steep and then tapers off toward the end of the phrase (Figure 2.5), whereas in English the rate of decline starts out steep, decreases toward the middle, and then ends in a final steep decline known as final lowering as seen in Figure 2.9 (Pierrehumbert, 1980).
Figure 2.9. Rate of declination in English wordlist; the final word has a steep F0 decline.

2.1.1.4. Summary of Mandarin and English tone and intonation

Table 2.1 provides a summary of the similarities and differences found between the Mandarin and English pitch systems. Following is a description of how AM and PENTA, two competing models of tone and intonation account for these differences.
Table 2.1. Comparison of Mandarin and English pitch systems

<table>
<thead>
<tr>
<th></th>
<th>Linguistic Function</th>
<th>Melody</th>
<th>F0 Peak Alignment</th>
<th>Register</th>
<th>Declination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandarin</td>
<td>Lexical and discourse-level contrasts</td>
<td>Dense in F0 minima and maxima</td>
<td>Late-aligned peak in bisyllabic sequences with initial focus</td>
<td>Low register for statements and high register for questions</td>
<td>No final lowering</td>
</tr>
<tr>
<td>English</td>
<td>Discourse-level contrasts</td>
<td>Sparse in F0 minima and maxima</td>
<td>Early-aligned peak in bisyllabic sequences with initial focus</td>
<td>No register difference for syntactic information recorded</td>
<td>Final lowering</td>
</tr>
</tbody>
</table>

2.1.2. Autosegmental-Metrical (AM) model of tone and intonation

2.1.2.1. General overview

The most common tone sequence model for representing tone and intonation is the Autosegmental Metrical (AM) model, originated by Pierrehumbert (1980). It integrated autosegmental phonology, used for the representation of lexical tones in tonal languages, and metrical phonology, which treated grouping and prominence (Goldsmith, 1976; Hayes, 1981). In the model, surface pitch contours are decomposed into high (H) and low (L) level tones, called autotones. Each autotone may be part of a simple phonological tone, called a toneme, in which case a single autotone makes up the entire phonological category, or an autotone may be part of a complex toneme, in which case more than one autotone joins to create a single phonological category, such as a rise (LH) or a fall (HL). An AM representation of four Mandarin words whose segmental makeup is the same, but whose tonal contrasts vary, is shown in (2) (Yip, 2002).
(2)

\[
\begin{array}{cccccc}
\text{ma} & \text{ma} & \text{ma} & \text{ma} & \text{ma} & \text{ma} \\
\text{H} & \text{LH} & \text{L} & \text{H} & \text{L} & \text{H}
\end{array}
\]

‘mother’ ‘hemp’ ‘horse’ ‘scold’ ‘mama’

Note that level tones are represented with one association line between the prosodic unit (in this case the syllable) and the level tone, while rises and falls, or complex tones, are indicated through two level tones associated to one prosodic unit. Finally, toneless syllables are shown through their lack of association between a syllable and a tone; toneless syllables have no tonal target; rather, their surface F0 is determined by the surrounding Hs and Ls (Ladd, 2008; Pierrehumbert, 1980).

For the description of intonation, the same system used for the description of lexical tones, is also used for intonation contours; however, in addition to associating to prosodic material such as the syllable, tones can also associate to edges of phrasing. Whereas in tonal languages, almost every word or syllable will bear tone, in non-tonal languages, only prominent syllables tend to bear tone. For example, in English, the stressed syllable of words under focus, called the accented syllable, bears tone, but other syllables which do not have sentential-level stress are typically toneless. When associated to prominent syllables, F0 contours are labeled pitch accents (Pierrehumbert, 1980). They are denoted via a *. Typically, in AM, early-aligned and late-aligned F0 contours are distinguished by placement of the *. L*+H indicates that the L is on the stressed syllable, and the peak of the contour trails behind, whereas L+H* indicates that the L is early, and the peak aligns during the stressed syllable of a prominent word.
When tones associate to edges of phrasing, they are termed phrase accents and boundary tones (Pierrehumbert, 1980). They are denoted as + and % respectively. AM differentiates pitch accents from boundary tones through differences in their timing behavior. While pitch accents align to stressed syllables, phrasal tones line up with the beginning and ends of phrases. An AM representation of the intonation contour in an English utterance with initial contrastive focus is seen in (3).

(3)

Mary wants some.
\[
\begin{array}{c}
\text{H}\star\text{L} \\
\text{L}+\text{L}%
\end{array}
\]

The abstract sequence of tones seen in 2 is then thought to feed into a phonetic level that serves to tie the tones into a smooth contour. So although the abstract phonological representation has clear-cut H and L tones, a listener actually hears the entire utterance as one smooth falling contour across the utterance.

2.1.2.2. Surface contours and autotones

Since surface phrases are heard with fluid and continuous F0 contours, guidelines must be put into place that allow us to determine what the tones of a given phrase are. When a tone associates to a prosodic unit, it is realized as a F0 minimum or maximum on the surface. For example, in Figure 2.2 and 2.8, surface falling contours were seen over the word *mama* in Mandarin and English. Under an AM model, the point where the F0 reaches its height and begins to decline, called the *elbow*, is interpreted as a H. In both languages, this point falls within the boundary of the first syllable, and so a H autotone presumably associates to the initial syllable on the surface.
To explain differences in alignment between two identical phonological tones, AM invokes the phonetics (Ladd, 2008; Pierrehumbert, 1980). Crucially, it is believed that speakers have at least some control over phonetic alignment, and so the same phonological category can be implemented differently cross-linguistically. Relaying this to our Mandarin and English examples (seen in Figure 2.2 and 2.8), we might then account for the patterns by assuming a H autotone associates to the initial syllable in both Mandarin and English; however, in Mandarin H is late-aligned, while in English, it is aligned earlier within the syllable.

Figure 2.8 raises yet another point to be discussed. Typically, the pitch accent of a word under contrastive focus in English, such as the one seen in Figure 2.8, is transcribed as H*L (Pierrehumbert, 1980), a transcription used to denote that the H autotone associates to the stressed syllable, but is followed by a trailing decline in F0. The actual F0 minimum that occurs on the syllable adjacent to the accented one is thus not treated as a direct association between the prosodic unit (the syllable) and a toneme. This treatment of complex contours is different than how complex tonemes are treated in tonal languages, where associations between the prosodic unit and both autotones of a complex toneme are made at the point where a F0 minima or maxima occurs (Goldsmith, 1976). Thus, in non-tonal languages, bi-tonal pitch accents are treated as having one association, while in tonal languages, bi-tonal lexical tones have to have two associations. Since the “rules” laid out for determining how autotones associate to the prosodic material under the AM framework suggest that the location of an autotone is realized at the point of the F0 elbows (Ladd, 2008). In this study, complex tonemes, whether part of a lexical tone system or an intonation system, will be transcribed with an association line drawn between the prosodic unit and the autotone at the point of a F0 minimum or maximum.
Therefore, what is typically transcribed as H*L in AM, will be HL in this study, and temporal alignment will be shown via an association line to the syllable during which the F0 minimum or maximum occurred.

Our discussion of how autotones are recognized on the surface would not be complete without mentioning the possibility for an autotone to have an effect without being associated to a syllable at all. Various phonologists have argued that downsteps and upsteps result from floating tones (Clements, 1979; Hyman & Shuh, 1974; Williams, 1976). This happens when an autotone remains within the tonal sequence, but does not associate to a prosodic unit. This phenomenon is especially common in African tonal languages, such as Tiv (Clements, 1979). In Tiv, for example, words with surface H tones are of two types: one class of nouns has simple H tones, while the other has an unassociated L preceding the H tone as seen in (4).

(4)

\[
\begin{array}{cc}
\text{kwa} & \text{kwa} \\
| & | \\
H & L \uparrow H \\
\end{array}
\]

‘ring of huts’ ‘leaf’

When following other H tones, the distinction between these words is apparent in the surface contour; all H tones that follow the unassociated L become downstepped as seen in (5).
Example (5) provides a contrast of two phrases. Column (a) provides the tonal representation of the words in the phrase. Note that all of the words are made of H tones except for the unassociated L in the word kwa, meaning leaf. Column (b) provides the relative level of F0 at which the tones are realized. Note that after the unassociated L in kwa, meaning leaf, all H tones are produced at a lower F0; when the L tone is not present, as in the first example, H tones remain at a consistent level. The unassociated L is thus realized as downstep during subsequent H tones.

Intonation specialists who work on Mandarin or English systems of pitch have not, at present, identified evidence to suggest that there are unassociated autotones which cause upstep of downstep of the pitch range to occur in either of these languages. Note, however, that the intonation of sentence-type in Mandarin is similar to the upstep and downstep data of Tiv in that the divergence in pitch register between Mandarin statements and questions begins following the focus word; the difference in pitch range does not extend across the entirety of the utterance in Mandarin. One possibility may be that upstep or downstep occurs following the focus item in Mandarin; however, such an analysis would require evidence for unassociated tones, and this
sort of analysis goes beyond the discussion presented in this dissertation. What this dissertation
does offer though, is a look at tonal patterns in Mandarin-English, which occur in the absence of
lexical tone. The relation of these tonal patterns to statement and question pitch register are then
explored to find out whether a particular tonal pattern might be related to pitch register.

2.1.2.3. Predicting surface patterns

More recently, the AM framework for representing tone and intonation has begun to be
embedded into Optimality Theory (OT) (McCarthy & Prince, 1993, 1995). To determine how a
particular phonological sequence will surface in OT, possible input-output pairings are compared
against a set of universal constraints. There are two types of constraints. Faithfulness constraints
require conformity between input and output forms. Some common faithfulness constraints
prevent deletion, insertion, changes in tonal value, or changes in tonal association from
occurring. These constraints are called MAX-IO [T], DEP-IO [T], ID-IO [T], and *LINK respectively. A
description of each is given in 6-9.

(6) MAX-INPUTOUTPUT [T] (MAX-IO): Each input tone has an output
correspondent (McCarthy & Prince, 1995; Yip, 2002)

(7) DEP-INPUTOUTPUT [T] (DEP-IO): Each output tone has an input
correspondent (McCarthy & Prince, 1995; Yip, 2002).

(8) IDENTITY-INPUTOUTPUT [T] (ID-IO): Input and output correspondents are identical
relative to tonal specifications (McCarthy & Prince, 1995; Yip, 2002).

(9) *LINK: Associations between tone and prosodic unit must be the same in the input
and output (Moren, 2004).
Markedness constraints are the second type of constraint. They are what drive changes to occur between input and output forms. These constraints prohibit certain phonological structures from surfacing. They must have perceptual, physiological, or typological grounding. Some examples of markedness constraints on toneme sequences are NoContour and ToneToStress. They are described in 10 and 11.

(10) **NoContour**: Only one tone can be within the domain of a syllable (Yip, 2002).

(11) **ToneToStress**: Tones must associate to stressed syllables (DeLacy, 2002).

Differences between how tones map to the segmental string are captured through ranking markedness and faithfulness constraints in different orders cross-linguistically. Constraints which are placed at the bottom of the constraint hierarchy can be violated in order to satisfy higher-ranked constraints. Thus, all languages are thought to have the same set of constraints (grounded in perception, articulation, and acoustics), but differences between the languages are captured through differences in how each constraint is ranked in a given language.

To start, an input is placed into the system. Next, a set of all possible outputs for each given input is generated, and these input-output pairs are evaluated using the constraint hierarchy. Those candidates that violate higher-ranked constraints are eliminated. The best possible candidate from the set of all possible candidates is the one with the least offensive violations; in other words, the candidate whose first violation is the lowest-ranked in the hierarchy is the winner, or the output.

Through considering all possible rankings of the constraints, researchers predict phonological systems that should emerge cross-linguistically, while ruling out impossible grammars. For example, using the constraints MaxIO [T] and NoContour (presented in 6 and
10 respectively), OT predicts two possible patterns for tonal contours. For languages where $\text{MaxIO}^{[T]}$ outranks $\text{NoContour}$, two types of contours will occur on syllables: (a) simple tonemes, and (b) complex tonemes. For languages where $\text{NoContour}$ outranks $\text{MaxIO}^{[T]}$, however, all syllables in the language are predicted to have simple tonemes. Crucially, no ranking yields a grammar that prohibits simple tonemes on all of its syllables, and so a language with such a typology must not exist if the OT constraints are correct. Incidentally, many African languages, such as Tembo, have only syllables with simple tonemes, and both English and Mandarin make use of syllables with simple and complex tonemes. Moreover, no language, to my knowledge, has yet been documented that allows only complex tonemes to occur on syllables, and so all possible rankings of the constraints yield true grammatical typologies, while correctly ruling out impossible grammars.

2.1.2.4. Comparison of Mandarin and English from an AM perspective

Table 2.2, is a partial replication of Table 2.1 which laid out a number of similarities and differences found between Mandarin and English systems of pitch. Added to this table is a description of how these similarities and differences are accounted for under an AM account.
Table 2.2. Comparison of Mandarin and English pitch systems from an AM perspective

<table>
<thead>
<tr>
<th></th>
<th>Linguistic Function</th>
<th>Melody</th>
<th>F0 Peak Alignment</th>
<th>Register</th>
<th>Declination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandarin</td>
<td>Lexical and discourse-level contrasts</td>
<td>Dense in F0 minima and maxima</td>
<td>Late-aligned peak for high tone in bi-syllabic sequences</td>
<td>Low register for statements and high register for questions</td>
<td>No final lowering</td>
</tr>
<tr>
<td>English</td>
<td>Discourse-level contrasts</td>
<td>Sparse in F0 minima and maxima</td>
<td>Early-aligned peak for high tone in bi-syllabic sequences</td>
<td>No register difference for syntactic information</td>
<td>Final lowering</td>
</tr>
<tr>
<td>AM Account</td>
<td>Lexical and discourse contrasts are represented as sequences of H and L autotones on the same layer of linguistic structure in both Mandarin and English.</td>
<td>Mandarin has a denser tonal tier than does English.</td>
<td>Mandarin and English have phonologically similar contours with different specifications for phonetic alignment; Mandarin has late-alignment of H, while English alignment of H is earlier.</td>
<td>At present, it is unclear as to how Mandarin register should be represented using sequences of H and L autotones.</td>
<td>Identical phonological contours in Mandarin and English word lists involve different phonetic specifications for the rate of declination; English declination is steeper toward the end of phrases in comparison to Mandarin.</td>
</tr>
</tbody>
</table>

2.1.3. A Parallel Encoding and Target Approximation (PENTA) account of tone and intonation

2.1.3.1. General overview

Under a contour-interaction model of tone and intonation, surface pitch contours are represented as several layers of linguistic structure which add together. While many versions of contour-interaction models have been proposed, probably the most popular model is the Parallel Encoding and Target Approximation Model (PENTA), proposed by Xu (1999).
PENTA is a functional approach to modeling tone and intonation. Each linguistic function is conveyed through a different level of structure constructed from pitch targets and approximation parameters. There are five possible pitch targets: high (H), rising (R), low (L), falling (F), and mid (M), called *local pitch targets*. The local pitch targets are similar to autotones in the AM model in that they are thought to be abstract tones; however, rather than being composed of sequences of level tones, local pitch targets can be dynamic. That is, while in AM, an accelerating F0 observed across a syllable is thought to surface from a sequence of two autotones (L and H), in PENTA, the F0 acceleration is the target itself. The approximation parameters, on the other hand, include local target pitch range, duration, or strength; these are termed *non-local pitch targets*.

Every syllable in each language has local and non-local pitch targets. The pitch targets add together during each syllable to make up the surface contour. Figure 2.10 presents an overview of PENTA taken from Xu (2009).

![Figure 2.10. A sketch of PENTA (Xu, 2009).](image)
2.1.3.2. Surface contours and local pitch targets

Under PENTA, the melodies heard during speech are made up of local pitch targets. In Mandarin, changes to the melody result in new lexical contrasts, and thus local pitch targets are linked to the lexical communicative function. Stressed syllables are implemented with one of four local pitch targets: H, R, L, or F, while unstressed syllables are realized with a M pitch target. When syllables are produced in succession, the varying pitch targets are realized with many turns in F0 direction as a speaker moves from one pitch target to the next. English, like Mandarin, is uttered with a local pitch target on every single syllable; however, in English, local pitch targets do not connote a lexical communicative function; rather they portray accent; the precise communicative function behind accents is still being researched (Xu & Xu, 2005). Unaccented syllables in English typically are implemented with a M pitch target. Moreover, most syllables are unaccented, and so far fewer changes in F0 direction are heard during English utterances as compared to Mandarin utterances. Accented syllables and unaccented sentence-final syllables can have a H, L, or F pitch target; a R pitch target in English has not yet been observed.

Just as the F0 maxima and minima are thought to be incidental from the change in pitch target during each syllable, the alignment of the peaks and valleys with respect to the syllable are also assumed to be incidental. Speakers are not expected to have sufficient control over their rapidly-moving articulators so as to time a peak or valley as pleased. Alignment is instead the consequence of the sequence of pitch targets. The differential alignment values for the falling contours seen in Figures 2.6 and 2.7 for Mandarin and English are therefore accounted for in that
Mandarin has a H pitch target followed by a M pitch target; English presumably has a H pitch target, followed by a F. Since the pitch targets are different, so is the F0 peak alignment.

2.1.3.3. Non-local pitch targets

PENTA accounts for intonation through non-local pitch targets. In Mandarin, for example, the higher overall register in questions is predicted by the addition of a non-local pitch target that raises the F0 of each syllable to a higher frequency. Declination, seen both in Mandarin and English, on the other hand, is triggered partially by a non-local pitch target that raises the overall pitch during the focused item, and partially by a universal tendency to downstep pitch when producing sequences of H and L tones in succession. The fact that English is produced with an extra bit of final lowering than Mandarin is not really discussed in research on PENTA, but it might be accounted for through description of local pitch targets since English utterances are said to have a falling local pitch target in final position, whereas Mandarin may have H, L, or R as well (Xu & Xu, 2005). This fall during phrase final position may give the appearance of extra lowering in phrase-final position.

2.1.3.4. Comparison of Mandarin and English from a PENTA perspective

Table 2.3, is a partial replication from Table 2.1 which laid out a number of similarities and differences found between Mandarin and English systems of pitch. Added to this table is a description of how these similarities and differences are accounted for under PENTA.
Table 2.4. Comparison of Mandarin and English pitch systems from a PENTA perspective

<table>
<thead>
<tr>
<th>Mandarin</th>
<th>Lexical and discourse-level contrasts</th>
<th>Dense in F0 minima and maxima</th>
<th>Late-aligned peak for falling contour in bi-syllabic sequences with initial H tone</th>
<th>Low register for statements and high register for questions</th>
<th>No final lowering</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Discourse-level contrasts</td>
<td>Sparse in F0 minima and maxima; moreover, melodies stay consistent across utterances with different words and lengths.</td>
<td>Early-aligned peak for falling contour in bi-syllabic sequences</td>
<td>No register difference for syntactic information</td>
<td>Final lowering</td>
</tr>
<tr>
<td>PENTA Account</td>
<td>Tone and intonation are built through adding local and non-local pitch targets; they reside in different layers of linguistic structure</td>
<td>English uses longer spans of M pitch targets in utterances than does Mandarin; this results in the generally simpler English melodies.</td>
<td>The difference in F0 maxima alignment is incidental from differences in local pitch targets in Mandarin and English; Mandarin uses a H followed by a M, while English uses a H followed by a F.</td>
<td>Mandarin implements a non-local register pitch target for sentence-type but English does not.</td>
<td>Mandarin and English have a non-local pitch target that raises the F0 when a new topic is introduced, resulting in declination. The extra final lowering heard in English is due to the final F local pitch target at the end of English statements.</td>
</tr>
</tbody>
</table>
2.1.4. Rationale for research

Tone-sequence models, such as AM, and contour-interaction models, such as PENTA, provide competing accounts for the linguistic representation and structure of pitch systems. While tone-sequence models are particularly strong at accounting for the varied tone-to-segment mappings of similar intonation contours in non-tonal languages such as English, accounting for linguistic contrasts conveyed through pitch range, such as Mandarin intonation of questions and statements, has proven to be difficult. On the other hand, contour-interaction models provide a clear explanation for the adjustments to pitch range made during discourse, such as the Mandarin intonation of questions and statements; yet, they fail to capture an important generalization about non-tonal intonation melodies; that is, they remain consistent across different word and syllable combinations. For example, in English questions, there is a rising surface contour that stretches across the utterance; the surface contour is very similar across different words and phrase-lengths as shown in Figures 2.11 and 2.12. Whereas AM captures the similarity between these phrases through their identical pitch accent and boundary tone, under PENTA, each syllable must be assigned a local pitch target. The F0 contour of the accented syllable in a monosyllabic question, however, is not necessarily the same as the accented syllable in a bi-syllabic question as can be seen by comparing Figure 2.11 and Figure 2.12. What remains unclear in PENTA is why questions have a rising local pitch target on the initial accented syllable in some questions, but a low one in others.
Experts in tone and intonation are thus divided in the choice of model they use. Those who focus on non-tonal languages, such as English, carry out research under a tone-sequence model such as the AM tone-sequence model; those who focus on East Asian tonal languages, such as Mandarin, prefer a contour-interaction model such as PENTA. Tone sequence models and contour interaction models, however, make different predictions about the course of L2 learning. Since tone and intonation are represented along the same structural layer in tone sequence models, L1 lexical tones are expected to interfere with acquisition of L2 intonation. By comparison, tone and intonation are not represented along similar structural layers in contour
interaction models, and so a Mandarin lexical tone is not expected to interfere with the acquisition of an English intonation contour in Mandarin-English speech, even if the resulting surface contours have acoustic similarity. By testing whether tone and intonation interfere in L2 speech, we can therefore learn more about the phonological structure of tone and intonation. The first goal of this dissertation is therefore to confirm whether Mandarin tone interferes with English intonation in Mandarin-English bilingual speech. If so, the interference of tone and intonation will provide strong support for modeling tone and intonation along the same layer of linguistic structure. If not, a model that treats tone and intonation on separate layers of linguistic structure may be more appropriate.

2.2. Category interaction in second language speech

To test whether tone interferes with intonation in Mandarin-English speech, we must first provide a definition of interference. What does it mean for first language (L1) tone to interfere with second language (L2) intonation? To answer this question, we must turn to models of L2 category production. When doing so, however, a major complication arises; models of L2 speech acquisition have focused primarily on the way L2 speakers learn to perceive and produce consonants and vowels, called segments. Section 2.2 explores how one of these models, the Speech Learning Model (SLM), might be used as a stepping-stone for examining the interaction of tone and intonation categories, or *tonemes*. This section therefore provides a general overview of the SLM and its predictions for segmental category interaction. It then assesses the strengths and weaknesses of the model against empirical studies that have been used in its support. Next, findings from studies on L2 phonological pitch production are compared to studies on L2 segmental production in order to gauge how closely pitch and segmental acquisition may
resemble one another. The section then ends by motivating the research questions used to examine the interaction of tone and intonation in this study.

2.2.1. *Speech Learning Model (SLM)*

2.2.1.1. General overview

Through use of the Speech Learning Model (SLM), researchers try to account for changes that occur in individuals’ ability to perceive and produce L2 consonant and vowel sounds, called segments, across a lifespan. Advocates of the model propose that the learning mechanisms infants use when establishing the sound system of the L1 remain intact across the life span and remain accessible for use during L2 acquisition. While this claim leaves open the possibility for an adult to learn L2 contrasts with the same accuracy as a child learns his or her L1, it is not believed that such a situation is likely to occur. This is because other factors that are confounded with age will prevent the intact speech learning mechanisms from developing long-term memory representations for speech categories that are identical to those developed for a single language by monolinguals. The primary reason why most L2 learners continue to differ from monolingual native speakers is that they used a shared phonological space for the perception and production of L1 and L2 segments, and so the two systems mutually influence one another. One concern for researchers who use the SLM is to describe how, and to what extent, the phonetic system developed for the L1 interacts with the L2 phonetic system.

2.2.1.2. Predictions about category interaction

The SLM targets a very specific group of bilingual speakers. The term *bilingual*, as used in the SLM, refers to those who first establish a native tongue and then learn a second language later in life. This type of bilingual has been termed *sequential bilingual* in other literature on
bilingualism (Bhatia & Ritchie, 1999). Participants whose speech is studied under the SLM are immigrants living in an area where the L2 is dominant; consequently, participants have the opportunity to communicate with native speakers on a regular basis, thereby receiving exposure to phonetic stimuli as produced by native speakers of the L2.

Through exposure to native L2 phonetic stimuli, segmental categories in the L1 and L2 are expected to interfere in the following ways: sometimes, acoustically similar L1 and L2 sounds will be conceptualized as instances of the same category. As bilinguals gain more and more exposure to the L2, they hear more and more instances of the category produced with L2 phonetic norms; consequently, over time, they’ll begin to shift the way they produce the merged L1-L2 category toward the L2 phonetic values. The result will be a bilingual new category whose phonetic values are intermediate to monolingual values of both languages.

An example of this hypothesis can be made through examining voiceless aspirated stop consonants in Mandarin and English. Stop consonants involve a period of low amplitude, followed by a release or burst as seen in Figure 2.13.

![Figure 2.13. VOT of a voiceless bilabial plosive in /a'pa/](image)
Multiple tokens of a stop consonant can be compared phonetically by measuring the time that occurs between the start of the release to the onset of the nearest periodic wave as seen in Figure 2.13. (Note that the periodic wave correlates physiologically to vocal fold vibration needed to produce a voiced segment). This particular acoustic measurement, used to compare stop consonants, is called the voice onset time (VOT). Mandarin voiceless aspirated stops have typically longer VOT than English voiceless aspirated stops: the average VOT for aspirated /p/ in Mandarin is 90 ms and for aspirated /p/ in English is 75 ms (Ladefoged, 2006); nonetheless, many Mandarin and English speakers perceive Mandarin and English /p/ as being the same sound (a point made clear in Chapter 3 of this study). One would therefore say that Mandarin /p/ and English /p/ are identical at the phonological level, because they are perceived as being instances of the same category, yet differ at the phonetic level since their acoustic measurements differ significantly.

The SLM predicts that perceptually similar sounds will be produced with phonetic values that fall in between the typical values of monolingual speakers from each language. Mandarin-English bilinguals are thus predicted to have VOT values between 90 ms and 75 ms (averaging around 83 ms perhaps).

The type of phonetic interaction just described has been termed category assimilation, and is expected to occur when an L1 and L2 category are perceptually similar (Flege, 1995). Other times, bilinguals will create a new category for an L2 sound. If existing L1 categories are close in phonological space, the L1 and L2 sounds will dissimilate in order to preserve linguistic contrast. This type of phonetic interaction has been termed category dissimilation, and is expected to occur when acoustically similar L1 and L2 categories are perceptually distinct.
According to the SLM, the extent of interaction between two phonetic systems is determined by amount of exposure to native stimuli; the SLM uses time as an index for the amount of L2 exposure. It is thought that in order for changes to occur in the joint L1/L2 phonetic subsystems, a bilingual must hear sufficient L2 phonetic stimuli to shift acoustic values of a given category toward the L2 norm. Since it is assumed that the L1 and L2 phonetic subsystems are linked, greater shifts toward L2 phonetic norms are expected to coincide with greater changes to the L1 phonetic system (Flege & Port, 1981; MacKay, Flege, Piske, & Schirru, 2001). Two variables in particular, length of residency in an L2-dominant country (LOR) and frequency of continued L1-use are generally controlled for and used to determine whether sufficient L2 exposure has taken place to cause shifts in the L1/L2 phonetic subsystems. If so, the end result should be a bilingual phonological system that differs from either monolingual version.

The assumption that a bilingual’s speech differs from that of a monolingual’s is not unique to the SLM. In fact, this idea has been proposed by others, and for all levels of language (Cook, 2002, 2003; Grosjean, 1989, 1992). Grosjean (1989), for example, argues that a bilingual speaker is not the sum of two monolinguals in one, and so the success of a bilingual speaker should not be judged against a monolingual norm, rather whether or not communication is able to take place.

2.2.1.3. Defining the constructs: Perceptual similarity

Predicting instances of category assimilation or dissimilation is dependent upon one’s ability to assess the perceptual similarity of cross-linguistic sound categories. A number of techniques have been used to accomplish this in the past. These include the use of articulatory or
acoustic description as a measurement for perceptual similarity listener judgments of change or no-change stimuli (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Bent, 2006; Gandour & Harshman, 1978; Grabe, Rosner, Garcia-Albea, & Zhou, 2003; Iverson & Kuhl, 1996; Iverson et al., 2003), listener transcriptions using a native spelling system (Best et al., 2001) and judgments of non-native sounds fit into native categories (Guion et al., 2000; Iverson et al., 2003). Each technique will be considered in turn.

One technique used to determine perceptual similarity is to assume that if two speech sounds are acoustically or articulatorily similar, they are also perceptually similar. While this may prove to be the simplest technique in terms of researcher preparation, one should be cautious of its use. Suppose that particular L1 and L2 speech sounds are acoustically similar, and so they are considered to be perceptually similar by the researcher. Using the SLM, one would therefore predict that the phonetic implementation of the acoustically similar sounds should merge in bilingual speech. If, however, bilinguals dissimilate the two sounds, it will not be clear whether (a) the SLM failed to account for bilingual segmental production, or (b) the acoustically similar sounds were in fact perceptually distinct for the listener. This is problematic for use of the model and should therefore be approached with caution.

The second method which has been used to determine perceptual similarity is to place speech sounds into pairs or triads and have listeners decide whether the stimuli are similar or different. This technique has been used in many studies on segmental perception, and provides a clear way of determining perceptual similarity. For studies comparing L1 and L2 pitch movements, however, this technique could cause some confusion. Pitch contours, and the segments on which they are realized, cannot be separated. Therefore, if a study tries to determine
perceptual similarity between L1 and L2 pitch contours by creating change and no-change trials of syllables bearing the L1 and L2 contours, and if participants claim that the syllables are diverse, it will not be clear whether the stimuli were considered different because the pitch contours were heard as being different, or whether the differences were heard at the segmental level.

The final two methods that have been used to test for perceptual similarity, involve having participants listen to non-native contrasts and then compare them to native ones. This can be done through having participants transcribe the non-native sounds using native transcription systems, or through directing participants to choose from a list of sounds, which native sound is most like the non-native one. One problem with this technique is that it forces listeners to fit non-native sounds into their native system; it is difficult to discern, using such a task, whether the speakers would have naturally associated the L1 and L2 sounds together or not. Another problem may be that of transcription systems. Many languages simply do not transcribe pitch in their writing system; when this is the case, this method simply cannot be used.

2.2.2. Phonetic category interaction of segments

2.2.2.1. VOT interaction in stop consonants

One of the early studies that recognized a bi-directional influence of phonetic values in L1 and L2 segmental categories examined monolingual and bilingual speech in French and English (Flege, 1987). Participants included native French speakers who also spoken English, native English speakers who also spoke French, and monolingual speakers of French and English. Voice onset time (VOT) was measured from seven French and seven English words beginning with /t/. 42
The VOT values for each speaker were extracted, and an average VOT value was calculated per speaker. Mean phonetic values were then compared between monolingual and bilingual speakers using an Analysis of Variance (ANOVA) design.

The VOT values of monolingual French and English /t/ differed significantly in that the average VOT value is shorter in French than in English. This difference in language-specific phonetic values was then compared to bilingual speakers in order to assess whether phonetic values would merge in bilingual speech. It was found that American women living in France produced a mean VOT value that was intermediate to monolingual speakers in both languages. Likewise, French women living in the United States produced a mean VOT value that was intermediate to monolingual speakers in both languages. The findings were argued in confirmation of a merger hypothesis by where bilingual speakers are expected to restructure phonetic space after learning a second language. In particular, perceptually similar L1/L2 segmental categories undergo a process of category assimilation by where the phonetic values become intermediate to L1 and L2 values.

A particular strength of this study is that it reveals a clear bi-directional change at the phonetic level in bilingual speech. When a second language is learned, the phonetic categories in the L1 and the L2 undergo distributional changes that cause bilingual phonetic values to differ from that of monolinguals in both languages. The precise changes in bilingual distributions, however, are unclear from the study. Recall that speakers produced 7 tokens of each target sound. Each participant’s phonetic tokens were then aggregated, or averaged. The phonetic data was then compared using a series of ANOVA designs where the dependent variables were average phonetic value.
Since the phonetic values examined were aggregated, there are two possible explanations for the finding that bilinguals have an intermediate value to monolinguals: (a) every time a bilingual speaker produced a phonetic token, it was intermediate to the two monolingual groups, or (b) sometimes the bilingual speaker produced phonetic values that were within the norms of monolinguals from their L1 and other times the bilingual produced values within the norms of monolinguals from their L2. In the latter case, the intermediate values discovered may have only been a statistical artifact arising from the aggregated dependent variables. These two scenarios cannot be differentiated in a design that uses aggregated dependent variables.

It is important to note that, at the time this study was done, aggregating variables was probably the only statistical methodology available. Today, however, more robust techniques can be used to help us identify which scenario (a) or (b) is more likely. In particular, we can include raw phonetic values in our comparisons to help distinguish between the two scenarios described in the previous paragraph. If bilinguals produce all phonetic tokens at a point that is intermediate to monolingual speakers, we should find a uni-modal distribution for bilingual phonetic values centered intermediate to the phonetic values of either monolingual version. On the other hand, if bilinguals produce phonetic tokens sometimes within norms of monolinguals of the L1 and other times within monolingual norms of the L2, then a bi-modal distribution for bilingual phonetic values should emerge.

2.2.2.2. Formant frequency interaction in vowels

Not only do the phonetic values of stop consonants appear to undergo change as a result of learning an L2, but phonetic cues used to contrast vowels also do so. One study, for example, examined whether Italian-English bilinguals produce /ei/ differently from monolingual English
speakers as a result of two mechanisms: phonetic category assimilation or phonetic category
dissimilation. Italian /e/ and English /ei/ were first compared between monolingual speakers. It
was found that English /ei/ uses substantially more movement of the tongue than does Italian;
moreover, Italian /e/ differed from English /ei/ in that the onset of the vowel in Italian had a
significantly higher and more fronted quality.

The production of /ei/, as produced by four groups of Italian-English bilinguals, was
compared to monolingual English /ei/ tokens. It was found that bilinguals who arrived in Canada
at a later age often undershot the tongue movement needed to produce English /ei/ within
monolingual English norms. Those who arrived at an earlier age, however, often overshot the
movement needed to produce monolingual English norms. Results were used to argue that the
late bilinguals' vowels had undergone a process of category assimilation by where Italian and
English vowel categories were merged, while late bilinguals had established a new category for
the English vowel /ei/ and the category had then undergone a process of dissimilation.

One drawback to this study is that Italian-English bilinguals produced only tokens of
English /ei/ and so the bi-directional influence caused by learning English can only be assumed.
We do not know, therefore, whether a process of dissimilation or assimilation truly occurred
between early-bilingual Italian and English vowels. This sort of confirmation would require was
to have information about the bilingual L1 values as well.

Still, information regarding the distributional properties of bilingual categories can be
gleamed from this study. Bilingual speakers only produced two tokens of the vowel /ei/ per
speaker; moreover, the tokens were elicited in different tasks, and so were reported separately
rather than being aggregated. This allows us to look more closely at the distributional properties
of the bilingual L2 category. Of particular interest in light of the discussion in Section 2.2.2.1, are the phonetic values reported for the late bilinguals. Twenty-six tokens were reported to have fallen below the means of monolingual English speakers. By subtracting this number from the possible tokens present, we can reconstruct that another 26 tokens were within the means of monolingual English speakers. Thus, it is possible that a bi-modal distribution was present in the late bilingual category distribution of English /ei/.

Since data did not include tokens of bilingual Italian /e/, nor was it confirmed whether the 26 undershoot phonetic values of the bilingual speakers’ was within norms of monolingual Italians, we cannot confirm whether bilingual categories truly involve a bi-modal distribution such that some phonetic values are within monolingual norms of the L1 and other within monolingual norms of the L2; however, the study reported here makes this a suspect possibility. For clarifying the facts in future studies, it may be best to (a) use a statistical analysis that does not involve aggregated variables, and (b) elicit bilingual phonetic data from the L1 and L2.

2.2.3. Phonetic category interaction of pitch contours

2.2.3.1. F0 peak alignment interaction in phonological tones

A study on bilingual pitch contours reported that, like phonemes, tonemes may also be subject to bi-directional influence at the phonetic level. Mennen (2004) examined rising pitch contours (L*H) as produced by native Dutch speakers of Greek with two goals in sight: (a) to determine whether Dutch speakers of Greek could produce cross-linguistic differences in the timing of L*H in both of their languages, and (b) to find out whether the SLM could be extended to also account for L2 acquisition of pitch categories.
Dutch and Greek each have a rising pitch accent that appears in pre-focus, or pre-nuclear, position. At the phonological level, these pitch accents are identical in Dutch and Greek; however, at the phonetic level, they differ slightly in alignment to the segmental material. In Greek, L*H reaches its peak following the accented syllable (Arvaniti, Ladd, & Mennen, 1998), while in Dutch, L*H peaks within the accented syllable (Mennen, 1999). Additionally, peak alignment for Dutch rising tones is sensitive to vowel weight, whereas it is not in Greek; the peak aligns later when the accented syllable contains a long vowel than when it has a short vowel (Mennen, 1999).

The timing of LH* peaks as produced by 5 near-native Dutch-Greek bilinguals, 5 Dutch monolinguals, and 5 Greek monolinguals were compared. Four out of five of the bilingual participants produced the Greek rise with an earlier peak than monolingual Greek speakers, and they produced a Dutch rise without regard to vowel weight; instead, all rises peaked at a position intermediate to monolingual Dutch short and long vowel peaks. The results of Mennen’s (1999, 2004) experiments suggest that (a) although difficult, bilingual speakers can produce cross-linguistic differences in both of their languages (as evidenced by her fifth participant), and (b) pitch categories, like that of segments, are highly prone to bi-directional interference, although the way in which they affect one another may be different than for segments.

Mennen’s (1999, 2004) study is important in that it suggests that intonation contours, like segments, are subject to bi-directional influence in the speech of bilinguals. It thus provides us with a way to measure whether interference between L1 tone and L2 intonation presents itself in Mandarin-English speech. The way in which bi-directional influence should be modeled, however, remains unclear. The SLM predicts that L1/L2 perceptually similar categories should
merge; this means that the Dutch LH* should have had a later peak in bilingual speech if intonation categories are influenced in a way that is predicted by the SLM (Flege, 1995). The bilingual speakers in Mennen’s study, however, were not immigrants in Greece, although they were highly proficient in the Greek language (Mennen, 1999). All bilingual participants were university professors of the Greek language living abroad. The SLM specifically targets immigrants who are living in an area where the L2 is the dominant language, and who receive continuous exposure to native L2 phonetic stimuli. Since Mennen’s (1999) participants were not immigrants in an area where the L2 was dominant, they may not have received enough native phonetic stimuli for their pitch contours to merge. Moreover, it is not clear how exposure to the English language may have affected the timing of rising pitch accents.

To help clarify studies on L2 acquisition of pitch in the future, it may be necessary to compare segmental and pitch production for bilingual participants. Mennen’s study, for example, did not provide information regarding whether the multilingual participants’ implemented segments in a manner consistent with the SLM. It is therefore impossible to know whether her results for pitch production differed from SLM predictions because (a) L2 pitch acquisition is different than L2 segmental acquisition, or (b) the SLM predictions for bilingual phonetic values are false. The picture is even more puzzling in light of the fact that earlier studies on segmental category interaction used a statistical methodology where phonetic variables were aggregated. If segmental and pitch production for the same speakers were to be analyzed in conjunction during future studies, it would be clearer as to how pitch production fits into models of L2 acquisition of phonology.
2.2.3.2. Declination interaction in bilingual pitch contours

There is some evidence suggesting that L2 declination rates are also influenced by native declination patterns (Backman, 1978; Willems, 1982). Willems (1982), for example, examined the declination rates in English utterances as produced by 10 Dutch-English speakers. The Dutch-English bilinguals lived in a Dutch-speaking area. These declination rates were compared to that of 10 native speakers of English. All the Dutch-English speakers used a smaller range of declination than native English speakers; in particular, pitch was reset at a higher pitch in native English speech than in Dutch-English. Willems (1982) notes that such a declination pattern is typical of native Dutch speakers speaking Dutch, and so he surmises that the smaller rate of declination may stem from Dutch influence. Backman (1978) studied declination rates of Spanish-English bilinguals, reporting a very similar finding. That is, the bilinguals in Backman’s study also used a smaller rate of declination than did the monolingual English speakers, and just as Dutch typically has a smaller rate of declination than English, so does Spanish.

Together, the studies suggest that L2 declination may be influenced by native declination patterns, although more research is needed if the SLM is to be adapted to account for declination patterns in an L2. First, it is not truly clear from the studies whether a smaller rate of declination is typical of all L2 speech, or whether the less steep decline in the bilinguals’ speech was influenced from the L1. Second, neither study surveyed bi-directional influences of L1 and L2 declination rates; while Mennen’s study on phonetic alignment would seem to suggest that, like phonetic alignment values, declination values should too undergo bi-directional changes, this must be confirmed. Finally, the participants in Willem’s study were not immigrants in an English speaking country; it is not clear whether declination rate would change more significantly if a
speaker were to receive more input from native-speakers of the target language. A study on bilingual declination that investigates the bi-directional influence of declination in the speech of bilingual immigrants is therefore needed if declination is to be accounted for by the SLM.

2.2.4. Motivation of research questions

The first goal of this dissertation is to confirm whether Mandarin tone interferes with English intonation in Mandarin-English bilingual speech. Toward this goal, the SLM hypotheses are extended to tone and intonation; thus, this study tests whether bilingual phonetic values of phonologically identical tones, used for lexical purposes in Mandarin but discourse-level purposes in English, are intermediate to monolingual values in both languages. Recall, however, that the SLM was not written to account for pitch. We must therefore use this extension of the model with caution. It is possible that the hypotheses will come out negative. If they do, what will we have learned? One possibility is that tone and intonation don’t interact. Another possibility is that pitch and segments interact in different ways. A final explanation might be that the SLM makes the wrong predictions about phonetic category interaction.

To know for sure whether tone and intonation interact in bilingual speech, we must therefore also figure out how phonological pitch contours that are used for the same linguistic function in the L1 and L2 interact in bilingual speech. A starting place will be to test whether phonological pitch contours interact in the same way as do segmental categories. This study will tests this possibility by examining the duration of VOT in voiceless aspirated bilabial stops (/p/), the timing of the F0 peak in falling contours over bi-syllabic words with initial stress (mama), and the rate of declination in word lists. Phonetic values will be examined in Mandarin and English as produced by monolingual speakers of Mandarin and English and bilingual Mandarin-
English speakers. If the SLM predictions are false for all categories tested, we will know that the SLM doesn’t make the right predictions. If the SLM predictions are true for segmental categories but not for categories of pitch, we will know that segments and pitch contours interact differently in bilingual speech; finally, if the SLM hypotheses are true of the /p/ VOT values and the word list declination rates, we will know that pitch contours and segmental contours behave similarly as long as the linguistic function for those categories is the same in both languages.

2.3. Cross-language mapping of tone to prosodic material

Confirming whether to model tone and intonation along the same level of linguistic structure is only half of the battle. Once we establish the need for a tone-sequence model, we still have left to uncover the phonological representation that leads to differences in sentence-type register. That is to say, what pattern of H and L autotones might cause a register difference between questions and statements to emerge in Mandarin intonation? A look into Mandarin-English intonation might provide new clues about register and its phonological representation. Teague (resubmitted) examined statement and question intonation as produced by Mandarin-English learners to find out whether the use of register for differentiating sentence-types would transfer into the learners’ L2. This study found that some, but not all, Mandarin-English speakers did in fact produce statements and questions with a register difference in English. If pitch register is in fact the outcome of mapping H and L level tones to the prosodic material in a particular pattern, then Mandarin-English may provide a perfect context for understanding what this pattern is. Those speakers whose statement and question pairs do involve register differences should use a different pattern of mapping H and L tones to the prosodic material than do speakers whose speech does not involve register differences between sentence-types; moreover, in English
speech there is an absence of lexical tones, and so the tonal tier should be less dense, making the patterns of L and H autotones easier to detect during phonological analysis. The second aim of this study is therefore to explore whether a certain typological pattern for associating tones to the prosodic material corresponds to a difference in register between questions and statements in Mandarin-English. This study uses the tonal-association scale, proposed by Gussenhoven (2004), as the starting place for examining tone-to-text typologies in Mandarin-English. Section 2.3 reviews the tonal-association scale. It then considers previous studies on Mandarin-English intonation, examining whether the predictions of the scale hold up in light of empirical data. The section then ends by motivating the research questions used to examine the phonological representation of register in this study.

2.3.1. Tonal association scale

2.3.1.1. General overview

Within OT, tone and intonation specialists strive to predict all possible patterns that languages may use to map phonological tones to prosodic material through ranking a set of universal constraints in all possible orders. Since many studies have confirmed a strong tendency for tones to associate to prominent syllables and edges of phrasing, our discussion will begin with constraints that center on this point. First, we will consider the relation between tone and prominence. While the strong relationship between tone and stress has been previously noted, the precise definition of stress remains somewhat elusive (DeLacy 2002). In some languages, such as English, tones seem to be attracted to sentential-level stress; in these languages, tones typically occur on or near the stressed syllable of focused items, referred to as the *accented syllable*. In other languages, there is a relationship between lexical stress and tone. In Mandarin,
for instance, researchers have often noted that stressed syllables have underlying tone, while inherently unstressed syllables do not (Xu 1999). Gussenhoven’s (2004) tonal association scale provides one way to capture how tones associate to different levels of prominence in different languages. Constraints within the tonal association scale are part of a fixed hierarchy; they are introduced in 12-15.

(12) $\sigma^* \leftarrow T$ (Gussenhoven, 2004; Pierrehumbert, 2000)

Accented syllables must bear tone.

(13) $\sigma' \leftarrow T$ (Gussenhoven, 2004)

Lexically stressed syllables must bear tone.

(14) $\sigma \leftarrow T$ (Anttila & Bodomo, 2000; Yip, 2002)

Syllables must bear tone.

(15) $\text{NoAssoc}$ (Gussenhoven, 2004)

TBUs are not associated with tones.

The constraints in 12-15 fall into the fixed hierarchy $\sigma^* \leftarrow T >> \sigma' \leftarrow T >> \sigma \leftarrow T$; note that satisfaction of a higher ranked constraint automatically satisfies the lower ranked constraints within the hierarchy. That is, if a tone associates to the syllable in a monosyllabic word that is stressed, then both $\sigma' \leftarrow T$ and $\sigma \leftarrow T$ are satisfied by that mapping, since it is true that all stressed syllables are syllables. Gussenhoven (2004) suggests that interaction of constraints 12 – 14 with a markedness constraint against tones, $\text{NoAssoc}$, (15 above), will determine how tones map to prosodic structure in different languages. For example, if NoAssoc is ranked in between (12) and (13), then only accented syllables will bear tone. If NoAssoc is ranked between (13) and
(14), then accented and stressed syllables, but not unstressed ones, will bear tone. If NoAssoc is at the bottom of the ranking, all syllables will bear tone.

Gussenhoven (2004) also notes that additional constraints that stipulate how tones map to edges of phrasing will also needed to account for F0 maxima and minima that occur at the beginning and ends of phonological and intermediary phrases. Borrowing from the work of Pierrehumbert (2000), Gussenhoven suggests the constraints described in 16 and 17.

(16) \( ji \leftarrow T \) (Pierrehumbert, 2000)
A tone must occur at the right edge of the phonological phrase.

(17) \( i[ \leftarrow T \) (Pierrehumbert, 2000)
A tone must occur at the left edge of the phonological phrase.

When \( ji \leftarrow T \) is ranked above the tonal association scale, syllables that fall at the end of a phrase must have a tone within its domain no matter what level of prominence the syllable has. Thus, syllables that occur at edges of phrasing are exceptions to the tonal prominence rules stipulated by the tonal association scale.

2.3.1.2. Predictions about tonal mapping

The tonal-association hierarchy presented in Gussenhoven (2004) makes a set of predictions regarding typological patterns of tone-to-text mapping that should (and should not) be present cross-linguistically. Starting by ranking NoAssoc at the top of the hierarchy, and then moving downward, four typologies are predicted to occur: (a) tones cannot associate to any syllable (unless that syllable occurs at the edge of a phonological phrase); (b) tones cannot associate to unaccented syllables unless those syllables occur at the edge of a phonological
phrase; (c) tones cannot associate to unstressed syllables unless those syllables occur at the edge of a phonological phrase, or (d) tones can associate to any syllable.

2.3.1.3. Considering other constraints with respect to the tonal association scale

The use of a markedness constraint (NOASSOC) alone to capture how tones will align to prosodic units paints too simplistic a picture. Faithfulness constraints, such as Max [T] or Dep [T] must also be ranked with respect to the system in order to account for less typical inputs. For example, what might happen when a Mandarin speaker encounters an English word where tones do not occur on every stressed syllable? While a toneless stressed syllable is an unlikely input for most Mandarin speakers, the use of OT requires us to still consider what would happen under such a situation. The faithfulness constraints, DEP [T], described in 18, must therefore also be ranked within the tonal-association-hierarchy to help compare input-output pairings where the input does not have tone.

(18) DEP [T]: Each output tone has an input correspondent (McCarthy & Prince, 1995)

Since, in Mandarin, stressed syllables always bear tone, we assume that Ω' ← T outranks DEP [T]. Inputs without tones on stressed syllables thus surface with tone as shown in Table 2.4.
Table 2.4. OT tableau showing that in Mandarin, $\sigma' \leftarrow T >> \text{Dep} \ [T]$.

<table>
<thead>
<tr>
<th>/$\sigma'$/</th>
<th>$\sigma' \leftarrow T$</th>
<th>Dep [T]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma'$</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>$\rightarrow \sigma'$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\leftarrow T$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Use of OT assumes that every possible ranking of the scale must result in a possible linguistic system. Within the hierarchy, there are four possible rankings of $\text{Dep} \ [T]$; each ranking yields a different grammar as seen in (19).

(19)

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^* \leftarrow T$</td>
<td>$\sigma^* \leftarrow T$</td>
<td>$\sigma^* \leftarrow T$</td>
<td>$\text{DEP} \ [T]$</td>
</tr>
<tr>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
</tr>
<tr>
<td>$\rightarrow \sigma'$</td>
<td>$\rightarrow \sigma'$</td>
<td>$\rightarrow \sigma'$</td>
<td>$\rightarrow \sigma'$</td>
</tr>
<tr>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
</tr>
<tr>
<td>$\rightarrow \text{DEP} \ [T]$</td>
<td>$\rightarrow \text{DEP} \ [T]$</td>
<td>$\rightarrow \text{DEP} \ [T]$</td>
<td>$\rightarrow \text{DEP} \ [T]$</td>
</tr>
<tr>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
<td>$\leftarrow T$</td>
</tr>
</tbody>
</table>

Of interest to this study, Mandarin follows pattern (b) where stressed syllables emerge with tones (Xu, 1999). English follows pattern (c) where accented syllables emerge with tones (Pierrehumbert, 1980). What is crucial about this system is that tonal insertions are expected to
be all or none. For instance, systems that insert tones on some lexically stressed syllables, but not all of them (with exception to accented ones), are expected.

2.3.4. Motivation of research questions

In order to explore whether a register difference between statements and questions emerges from a particular pattern of associating L and H tones to the segmental material, the following steps can be taken:

Step 1: Starting from a set of typological patterns predicted to emerge in languages, this study will test whether ME speakers produce typological patterns that fall outside the typologies predicted by the interaction of the tonal association scale and faithfulness constraints. This will be used to determine a set of tonal patterns found in Mandarin-English intonation of statements and questions.

Step 2: Once a set of tonal patterns for Mandarin-English statements and questions is in place, the representation of register can be investigated by determining whether a certain tonal pattern corresponds to a difference in pitch register between questions and statements. Such exploration may lead us toward a sequential representation of pitch register.

2.4. Research hypotheses

To summarize the problem, there is currently a divide in how experts in tone and intonation systems choose to model systems of pitch; some use a tone-sequence model while others conduct research under a contour-interaction model. While contour-interaction models, such as PENTA, capture the facts of Mandarin tone and intonation well, they are problematic for non-tonal languages, such as English; on the other hand, tone-sequence models, such as AM, do a better job capturing the facts of non-tonal intonation, such as English, although accounting for
intonation in tonal languages such as Mandarin, has caused great difficulties. Thus, depending upon language specialty, linguists use different models for the representation of tone and intonation. The two types of models, however, make different predictions about the course of L2 learning; namely, tone sequence models, such as the AM framework adopted into OT, predict that native tone should interfere with acquisition of second language intonation during the course of L2 learning. Contour-interaction models, such as PENTA, predict no interference should occur. In order to gain clearer insights toward the phonological structure of tone and intonation, the first part of this study begins by investigating interference between similar speech categories in bilingual speech, and testing whether Mandarin tone and English intonation interfere at the phonetic level during Mandarin-English bilingual production of word-lists. Extending the SLM hypotheses to pitch, the first part of this dissertation tests the following three hypotheses:

1. Mandarin-English bilinguals will implement lists with a rate of declination that is intermediate to monolingual Mandarin and English speakers.

2. Mandarin-English bilinguals will align the peak of two-syllable falling contours in Mandarin (H) and English (H*L) at a point intermediate to Mandarin and English monolinguals.

3. Mandarin-English bilinguals will produce Mandarin and English voiceless bilabial stop consonants with VOT values which are intermediate to monolingual speakers of Mandarin and English.

Even if the tone-sequence framework is validated, however, the problem of how to represent phonological manipulations of pitch range through sequences of H and L tones remains a problem. How, for example, can the register difference between statements and questions heard
during Mandarin speech be captured in Optimality Theory? The second part of this dissertation investigates the representation of register through the medium of Mandarin-English. First, possible mappings of tone to segmental material are explored through testing the tonal association scale. In general, we expect L1 to influence L2. In the L2, English, only accented syllables and phrase edges should surface with tone; in the L1, however, all stressed syllables should emerge with tone. We don’t really expect the exact L1 pattern to emerge in English since Mandarin-English bilingual speakers, presumably, realize that English is not a tonal language. We do, however, expect more tones, and we do expect that since the hierarchy is predicted to be universal, it will account for the ME pattern one way or the other. But there is no intermediate prediction fit into the scale; it’s either all stressed syllables or none (with an exception to the accented syllable), so let’s see what happens.

The second part of this study is therefore guided by the following three hypotheses:

4. A surface contour that corresponds to the English or Mandarin ranking of constraints within the tonal association scale will occur in Mandarin-English:

   (a) If following the Mandarin-like pattern where $\sigma' \leftarrow T$ outranks Dep $[T]$ then a F0 minimum or maximum will occur within the domain of every accented and lexically stressed syllable.

   (b) If following the English-like pattern where Dep $[T]$ falls between $\sigma^* \leftarrow T$ and $\sigma' \leftarrow T$, then a F0 minimum or maximum will occur within the domain of every accented syllable.
The patterns yielded through this exploration are then compared to findings on register differences between statement/question pairs in Mandarin-English. Results are used to determine whether register distinctions correspond to a particular tonal pattern; if so, a sequential representation for contrastive pitch registers will be inferred.

5. A difference in statement/question register will be predictable from tonal pattern.

If results of this study are to be reliably interpreted, it is first necessary to assure that the Mandarin and English speech categories under study are indeed perceptually similar to ME bilinguials. Toward this goal, a preliminary study aimed at determining ME underlying representations of English falling contours in lists and English /p/ was conducted. Details of this preliminary study are presented in full in Chapter 3. Following, Chapter 4 reviews the methodology used to study category interaction and register in Mandarin-English. Chapter 5 describes findings on Mandarin-English category interaction. Chapter 6 reviews the findings on monolingual English and Mandarin-English tonal patterns found in English statements and questions. Chapter 7 discusses results on pitch register and its relationship to tonal patterns in Mandarin-English. Finally, Chapter 8 concludes this dissertation.
CHAPTER 3: PRELIMINARY PERCEPTION STUDY

Chapter 3 describes details of the methodology used to determine ME speakers’ phonological representations of English voiceless bilabial aspirated stops and falling contours in word-lists. The results of the experiment described in this chapter are used to build the hypotheses for the remainder of the study.

3.1. Overview of the perception experiment

During a 10-minute sitting, participants read a consent document (see Appendix A), filled out a questionnaire about their language background and history (see Appendix B), and completed four perception tasks. The goal of the perception experiment was to determine bilingual speakers’ phonological representations of English aspirated voiceless bilabial stops and falling contours in bi-syllabic words with initial stress. For each perception task, participants read directions, participated in a practice session, and then performed in the actual test. Written directions were given to participants in English.

3.2. Participants

Six Mandarin-English bilinguals participated in this experiment. Participants were originally from Beijing but residing in the Washington DC area at the time of the experiment. All participants were undergraduate students enrolled at Georgetown University whose length of residency in the United States was between 2 and 4 years. All participants reported having strong reading and writing skills using the Chinese Pinyin system. Their use of the Pinyin system in the transcription task confirmed that this was true.
3.2.1. Participant Questionnaire

Participant questionnaires were administered to all participants to determine information regarding participant age, previous language exposure, and language use (See Appendix B for questionnaire). The questionnaire was given in English. Its purpose was to confirm that the language of all participants matched, and to determine the average AOA, LOR, amount of L1 and L2-use, phonological representation of pitch in English yes/no questions.

3.2.1.1. Language

Section A of the questionnaire, which was made up of ten questions, was intended to elicit information about language background and dialect. A large variety of Chinese dialects are spoken within the People’s Republic of China, but Standard Mandarin is the official language used in business and educational settings. The standard dialect is based on Beijing Mandarin. Since the effects that speakers’ home dialects may have on the perception or production of Standard Mandarin pitch are unknown, attempts were made to find ME bilingual speakers who spoke only the standard dialect of Mandarin; hence, only speakers from Beijing participated in this study. Additionally, since American English dialects can differ greatly, only ME bilingual speakers who residing in the Washington, DC area participated in the study.

3.2.1.2. AOA and LOR

Section A also aimed to elicit information about AOA and LOR. Individuals between the ages of 19 and 23 were chosen to participate; the average AOA was 18 years of age.
3.3. Perception task 1: Transcription

3.3.1. Phonetic stimuli

A transcription task was designed to examine the following English categories:

(a) /p/
(b) /H*L/

A female, monolingual speaker of English produced eight different word-lists with falling contours (H*L). The final two words in each list were “Janey and Michael.” The initial word in each list was different each time; the words are presented in Table 1. The lists were recorded using a Sony Digital Microphone; speech samples were digitized at 44kHz using Praat Speech Analysis Software (Boersma & Weenink, 2011). The initial word from each word-list was extracted and used as stimuli for the transcription task. Table 3.1 displays each stimulus item. A copy of all items, including the example and practice transcription items can be found in Appendix C.

Table 3.1. Stimulus items produced with a H*L pitch accent.

<table>
<thead>
<tr>
<th>ma.ma</th>
<th>mi.mi</th>
<th>na.na</th>
<th>mo.mo</th>
<th>ma.mi</th>
<th>ni.na</th>
<th>pa.pa</th>
<th>pa.pi</th>
</tr>
</thead>
<tbody>
<tr>
<td>H*L</td>
<td>H*L</td>
<td>H*L</td>
<td>H*L</td>
<td>H*L</td>
<td>H*L</td>
<td>H*L</td>
<td>H*L</td>
</tr>
</tbody>
</table>

Note. Words are transcribed using broad IPA transcriptions. Syllable boundaries are indicated using “.”” Association between pitch accent and syllable is indicated by “|”
3.3.2. Materials

Participants were given a worksheet that indicated the directions of the task in English, one example, one trial, and then eight numbered lines on which to transcribe the test stimuli. A copy of the worksheet can be seen in Appendix D.

3.3.3 Procedures

Participants were directed to listen to words played on a computer and transcribe them using the Chinese Pinyin system. They were further told that not all of the words they were about to hear would necessarily be real words in Mandarin. First, participants read the directions for the task in English. Then, they listened to an example. Next they participated in a trial transcription of the word “Tommy.” Any questions related to the task were asked and discussed following the trial transcription. Following, the real transcription task was given. Each test-word was played a total of 2 times with a one second pause in between repetitions. When a participant indicated having finished transcribing an item, the next word was played.

3.3.4 Results of /p/ stimuli

A total of 24 tokens were analyzed (4 /p/ sounds x 6 speakers). In pinyin, /pʰ/ is written as p and /p/ is transcribed as b. Following from this, three possible transcriptions for English /p/ were considered: p, b, or other. An “other” category was considered in case participants were to transcribe English /p/ using an unpredicted sound. Table 3.2 displays the results of the /p/ transcriptions. English /p/ was unanimously represented as p in the Chinese pinyin system, and so most closely resembles Chinese /pʰ/.
Table 3.2. Results of /p/ transcriptions.

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>b</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

3.3.5 Results of H* over bi-syllabic words

The variables under consideration were “tone” (1.5, 1.3, other) and word-type (reduplicants, and non-reduplicants). Since /H.0/ and /H.L/, represented in pinyin via “1.5” and “1.3” respectively, are closest in terms of acoustic similarity, each of these sounds were considered as a separate level for the variable “tone.” In addition, an “other” category was considered for any other pitches that were transcribed. Some forms were reduplicants, where the second syllable used the same segmental material as the first; others were not. Since Mandarin has a phonological rule that changes the second pitch to a neutral tone in reduplicated forms (Yip, 2002), this was also considered as a possible variable affecting how pitch would be transcribed. A total of 48 tokens were analyzed (8 H*L tokens x 6 listeners). In general, whether a word was a reduplicate form affected the transcription of tone; reduplicants were generally transcribed as 1.5 (or H.0) while non-reduplicants were transcribed as 1.3 (or H.L). One participant transcribed *mama* with two high tones (5.5). Results are indicated in Table 3.3.

Table 3.3. Results of H*L transcriptions.

<table>
<thead>
<tr>
<th></th>
<th>Total forms</th>
<th>H.0</th>
<th>H.L</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduplicant</td>
<td>24</td>
<td>23</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Non-reduplicant</td>
<td>24</td>
<td>0</td>
<td>24</td>
<td>0</td>
</tr>
</tbody>
</table>

65
3.4 Perception task 2: Similarity judgment

3.4.1. Phonetic stimuli

A similarity judgment test was used to examine the following contrasts:

(a) Mandarin /pʰ/ and /p/ versus English /p/
(b) Mandarin /H.L/ and /H.0/ versus English /H*L/

A female, monolingual speaker of Mandarin recorded sixteen word-lists. Her speech samples were digitized at 44kHz using Praat Speech Analysis Software (Boersma & Weenink, 2011). Each word-list contained three items. The final two words were “ta2ta2, ta2ti3” The initial words were also two-syllables long but fell into one of four patterns: (a) H.L tonal pattern on words with syllable-initial /pʰ/, (b) H.L tonal pattern on words with syllable-initial /p/, (c) H.L tonal patterns on words beginning with nasal consonants, or (d) H.0 tonal patterns on words beginning with nasal consonants. The group of words in the H.L tonal pattern (c) corresponded to the words in the H.0 tonal pattern (d) in that they had identical segments. The initial word from each list was extracted and used to prepare similarity judgment stimuli. In addition, the English stimulus items used in the transcription task were used to prepare triadic similarity judgment stimuli.

Each triad was created by placing an English word in between two Mandarin words. All words in a stimulus item either matched in the H.L tonal pattern, but used two Mandarin words with different segmental patterns (initial-pʰ or initial-p), or they matched in segmental makeup, but the two Mandarin words differed in whether they used H.L or H.0 patterns. Half of the stimulus items contained H.L pattern first, and the other half included the H.0 pattern first. An inter-stimulus timing of .5 seconds was placed between each word in the triads. A list of the
similarity judgment stimuli can be seen in Table 3.4. A list of all triadic-similarity-judgment stimuli, including the example and practice stimuli, can be found in Appendix E.

Table 3.4. Triadic-similarity-judgment trials

<table>
<thead>
<tr>
<th>Mandarin - English - Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>maHmaL - maH*Lma - maHma0</td>
</tr>
<tr>
<td>miHmi0 - miH*Lmi - miHmiL</td>
</tr>
<tr>
<td>naHnaL - naH*Lna - naHna0</td>
</tr>
<tr>
<td>moHmo0 - moH*Lmo - moHmoL</td>
</tr>
<tr>
<td>maHmiL - maH*Lmi - maHmi0</td>
</tr>
<tr>
<td>niHna0 - niH*Lna - niHnaL</td>
</tr>
<tr>
<td>paHpaL - paH*Lpa - baHbaL</td>
</tr>
<tr>
<td>baHbiL - baH*Lbi - baHbiL</td>
</tr>
</tbody>
</table>

3.4.2 Materials

A worksheet was designed to record participants’ similarity judgments of triadic similarity judgment trials. The worksheet included directions for the task, one example, one practice trial, and eight spaces for recording similarity judgments for each stimulus item. See Appendix F for the similarity judgment task worksheet.

3.4.3 Procedure

Each contrast was tested by eight similarity judgment trials. Participants listened to each trial twice, spaced 1 second apart, and then circled 1 or 3 on their worksheet, indicating whether
the first or final word was more similar to the center word. First, the participants were directed to read the directions on their worksheet and follow along with the example. Next, participants followed a practice session using “ta1.mi5- ta.mi- ta1.mi3” as a stimulus item. Questions pertaining to the experiment were answered following the trial session. Finally, participants completed the task using real test stimuli.

3.4.4 Results of /p/

The variables under consideration were “response” (1 or 3) and “segment” (/pʰ/ or /p/). Since half of the triadic judgment stimuli included the /p/ first, while the other half included it last, “response” was used to control for whether the order that the tokens were presented affected participants’ responses. The “segment” variable is considered to find out whether one Mandarin segment is perceptually more similar to English /p/ than another. A total of 12 tokens were analyzed (2 words x 6 listeners). Results are shown in Table 3.5. Speakers unanimously chose /pʰ/ as being more similar to English /p/. There was no affect of bias toward choosing either response 1 or 3.

Table 3.5. Results of /p/ similarity judgment task.

<table>
<thead>
<tr>
<th>Total forms</th>
<th>/pʰ/</th>
<th>/p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response 1</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Response 2</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
3.4.5 Results of $H^*$

The variables under consideration were “response” (1 or 3), “tune pattern” (H.L versus H.0), and “reduplicant pattern” (reduplicant or non-repduplicant). The “response” was assessed in order to control for affects stemming from the order of items in the triadic trials. The “tune pattern variable” is meant to assess whether word items are heard as being more similar to a high-neutral or high-low tune in Mandarin. Finally, since Mandarin has a phonological rule that changes the second pitch to a neutral tone in reduplicant forms (Yip, 2002), this was also considered as a possible variable affecting how pitch would be perceived. A total of 48 tokens were analyzed (8 triadic similarity judgment stimuli x 6 speakers). Results are shown in Table 3.6. Speakers unanimously chose /H.0/ as being more similar to English /H*L/. There was no affect of bias toward choosing either response 1 or 3, nor did reduplication patterns affect the results.

Table 3.6. Results of /p/ similarity judgment task.

<table>
<thead>
<tr>
<th>Total forms</th>
<th>H.0</th>
<th>H.L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduplicant</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Non-repduplicant</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Response 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduplicant</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Non-repduplicant</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>
3.5 Discussion and implications

Results of the transcription and similarity judgment tasks imply that the ME phonological representation of English /p/ is /pʰ/. In other words, English /p/ is likely categorized as being the same as Mandarin /pʰ/. For the phonological representation of H*L, however, the transcription and similarity judgment tasks conflicted. The transcribers, in general, transcribed reduplicated forms with a H.0 pattern, or neutral tone, but the non-reduplicated forms were written as having a high tone followed by a low tone. The similarity judgment task, however, revealed that all forms were closer to the H.0 form, regardless of whether they were reduplicated or not. One reason for this may be that Mandarin allows for the neutral tone to occur on reduplicated forms, although the neutral tone is quite restricted in other words; it can only appear on a few morphological-markers. The Mandarin speakers may have transcribed the English words using the closest “real” Mandarin word in their lexicon. When asked to hear three words next to one another, however, the non-word, produced with a neutral tone on the second syllable, was truly the more similar of the sounds. It may be that using a native spelling system to transcribe foreign words is just not an adequate tool for determining phonological representations of words. Following this line of reasoning, the ME phonological representation used for H* sequences in this dissertation will be H.0; further, Chapter 4 only considers reduplicant forms where this discrepancy is not an issue. A more thorough investigation of ME pitch perception in English reduplicant and non-reduplicant words is needed; however, it is beyond the scope of this dissertation.
CHAPTER 4: METHODOLOGY

4.1. Overview of the experiment

During a 40-minute sitting, participants signed a consent form (see Appendix A), filled out a questionnaire about their language background and history (see Appendix B) and completed three production tasks. The purpose of the production tasks was to elicit the targeted forms under study (voiceless bilabial stops, two-syllable falling contours, list declination, statement and question pitch contours). For each production task, participants read directions, participated in a practice session, and then recorded utterances for the actual recording task. Written directions were given to participants in Mandarin for tasks aimed to elicit Mandarin utterances, and in English for those tasks aimed to elicit English utterances.

4.2. Participants

The analysis of this study was structured as a repeated measures design. A group of 18 Mandarin-English bilingual speakers was recorded two times, once speaking English and once speaking Mandarin. Their Mandarin productions were compared to a group of 14 monolingual Mandarin speakers and their English productions to a group of 16 monolingual English speakers. In addition, 10 native speakers of English participated in the study by rating the bilinguals’ degree of foreign accent. While it would have been ideal to have two groups of bilingual speakers, one that spoke in English and the other that spoke in Mandarin, and thus obtain independent samples, it proved impossible. Thus, in this study, I follow Field (2009) who recommends using p-values that are less than 0.01 for non-orthogonal comparisons.

The monolingual Mandarin speakers who participated in this study all came from Beijing, China. In order to qualify as a bilingual participant in this study, and meet the criteria
necessary for the SLM’s rigorous participant selection, the bilingual speakers needed to be originally from Beijing, China as well. Additionally, they had to be residing in one city in the United States, and none of the bilingual speakers could have lived in any other area during their lifetime; otherwise, it was would not be clear whether variations found in bilingual speech stemmed from exposure to different dialects or were simply part of common variation in bilingual speech. Moreover, I searched out participants who were true bilinguals and that did not report speaking any languages other than Standard Mandarin and English. First attempts at finding a group of bilingual participants to fit this description took place in the DC area, but speakers who were originally from Beijing also tended to speak third and even fourth languages in DC. Through much searching, speakers who matched these criteria were finally located near the San Antonio area through contacts in the very large Chinese Association at the University of Texas in San Antonio.

Once a group of bilingual participants were found, place of residency, AOA, LOR, frequency of L1-use and L2-use, native speaker reaction to degree of foreign accent and judgments of phonological representations for English /p/ and H*L were controlled for all bilingual participants. Participants were between the ages of 28 and 46. ME participants arrived in the USA between the ages of 19 and 26. A similarity judgment task (described in Section 3) was used to confirm that all participants perceived English /p/ and /H*L/ as being similar to Mandarin /pʰ/ and Mandarin /H.0/ respectively. Controls for participant selection are described in turn, followed by a summary chart of all participants.

4.2.1. Participant Questionnaire
Participant questionnaires were administered to all participants to determine information regarding age, previous language exposure, and language use (See Appendix B for questionnaire). The questionnaire was given to all ME bilingual and E monolingual participants in English. Mandarin monolingual participants filled out a questionnaire in Mandarin. The purpose of the questionnaire was to confirm that the language and dialects of all participants matched, and to determine the average AOA, LOR, and amount of L1 and L2-use for each bilingual individual.

4.2.1.1. Language and dialect

Section A of the questionnaire, which is made up of ten questions, was intended to elicit information about language background and dialect. Only speakers from Beijing participated in this study. Additionally, since American English dialects can differ greatly, only ME bilingual and monolingual English speakers who resided in the San Antonio, Texas area participated in the study. The monolingual English speakers were born and raised in San Antonio, Texas as well.

Since many educated speakers of both English and Mandarin are expected to study foreign languages in classroom settings during their education, monolingual speakers who had taken a foreign language, but who had never advanced past beginning levels were still chosen to participate in this study.

4.2.1.2. AOA and LOR

Section A also aimed to elicit information about AOA and LOR. Since the processes of category assimilation and dissimilation are thought to be most prevalent in the speech of immigrants with long LORs and low-frequency L1-use, attempts were made to find participants who have resided in the Washington, DC area for a good deal of time; thus, participants with an
LOR between 8 and 14 years were selected to participate in the study. To ensure that these participants began learning as adults, individuals between the ages of 28 and 46 were chosen to participate; the average AOA was somewhere between 18 and 36 years of age.

4.2.1.3. L1-use and L2-use

Section B of the questionnaire, consisting of 10 questions, was used to calculate a score for each participant’s amount of L1-use. To respond to each question, the participants decided how often they participated in different tasks; they marked their response along a 6-point scale where the left indicated “never” and the right was “always.” Each point on the scale was assigned a number (1 through 6), starting with 1 on the left side. A score for L1-use was calculated by averaging the points of all the questions pertaining to the L1 (Mandarin). Bilingual participants who scored 5 or more, on the L1 language-use section were eliminated from the study since they were less likely to have had sufficient exposure to the L2 to have resulted in changes to the L1 and L2 phonetic systems. A score for L2-use was also be calculated by averaging the points of all questions pertaining to the L2 (English). Those who scored less than 2 on the L2 language-use section were also eliminated from this study; this applied to two bilingual participants, bringing the number of bilingual participants down to 16.

4.2.2. Degree of foreign accent

To determine participants’ degree of foreign accent, each speaker was recorded saying a total of 3 English utterances (a question, a statement, and a list). The same utterances used to determine degree of foreign accent were also analyzed as utterances for the main tasks. (To see elicitation method used to elicit utterances, see Section 4.5. To view the utterances, see Appendix G.)
Ten native speakers of American English (five females and five males) were asked to evaluate the three utterances produced by each of the bilingual talkers (ratings were made only of the 16 bilingual speakers whose speech samples were analyzed in this study). All raters were between the ages of 30 and 50, and were living in the Washington, D.C. area during the time of testing. None of the raters were proficient in any language other than English, although they may have had some elementary knowledge of a non-tonal language, such as French or Spanish.

A total of 240 utterances ([16 ME speakers x 4 repetitions x 3 utterances] + [4 E speakers x 4 repetitions x 3 utterances]) were presented over a loudspeaker at a comfortable level. Each English utterance was played separately in three randomized blocks. Each block contained all speakers’ utterances repeated four times. The first 20 utterances played were used to familiarize raters with the range of accents they would hear. These utterances were not analyzed. The remaining 60 utterances in each block were played in randomized orders for each rater.

Raters were told they were listening to Chinese immigrants and American English speakers. They were asked to rate how heavy each accent was using a scale from 1 to 9 where 1 represented no accent at all, and 9 represented an extremely strong foreign accent. An average rating was calculated for each speakers’ sentence.

4.2.3. Similarity judgment task

The same task described in Section 3 was used for bilingual-participant selection in Section 4. Bilingual participants who judged English /p/ and /H*L/ as being similar to Mandarin /pʰ/ and /H.0/ were selected to participate. The reason for this is that the utterances elicited during the speaking tasks were made to test whether English /p/ and Mandarin /pʰ/ or English H*L and Mandarin H.0 interacted at the phonetic level; if a speaker did not perceive these items
as similar, then no interaction between these items would be expected. All 16 bilingual participants whose speech samples were maintained after calculating amount of English exposure, judged English /p/ and /H*L/ as being similar to Mandarin /pʰ/ and /H.0/.

4.2.4. Summary of bilingual participants

Table 4.1 provides a summary of the 16 bilingual participants whose speech was analyzed in this study.

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Education</th>
<th>AOA</th>
<th>LOR</th>
<th>L1-use</th>
<th>L2-use</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME1 Finance Officer</td>
<td>B.A.</td>
<td>32</td>
<td>10</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>ME2 Engineer</td>
<td>B.A.</td>
<td>24</td>
<td>7</td>
<td>3.2</td>
<td>2.8</td>
</tr>
<tr>
<td>ME3 Teacher</td>
<td>B.A.</td>
<td>24</td>
<td>7</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>ME4 Teacher</td>
<td>M.A.</td>
<td>27</td>
<td>7</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>ME5 Student</td>
<td>B.A.</td>
<td>29</td>
<td>10</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>ME6 Engineer</td>
<td>B.A.</td>
<td>40</td>
<td>10</td>
<td>3.2</td>
<td>2.8</td>
</tr>
<tr>
<td>ME7 Scientist</td>
<td>Ph.D.</td>
<td>31</td>
<td>12</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>ME8 Scientist</td>
<td>Ph.D.</td>
<td>32</td>
<td>15</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>ME9 Biomedical Researcher</td>
<td>Ph.D.</td>
<td>31</td>
<td>15</td>
<td>2.8</td>
<td>2.4</td>
</tr>
<tr>
<td>ME10 Biologist</td>
<td>M.A.</td>
<td>32</td>
<td>10</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ME11 Scientist</td>
<td>Ph.D.</td>
<td>30</td>
<td>14</td>
<td>3.2</td>
<td>2.9</td>
</tr>
<tr>
<td>ME12 Finance Officer</td>
<td>B.A.</td>
<td>30</td>
<td>10</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>ME13 Biologist</td>
<td>Ph.D.</td>
<td>31</td>
<td>9</td>
<td>2.1</td>
<td>3</td>
</tr>
<tr>
<td>ME14 Biologist</td>
<td>Ph.D.</td>
<td>32</td>
<td>10</td>
<td>2.8</td>
<td>3</td>
</tr>
<tr>
<td>ME15 Scientist</td>
<td>Ph.D.</td>
<td>32</td>
<td>8</td>
<td>3</td>
<td>2.4</td>
</tr>
<tr>
<td>ME16 Medical Researcher</td>
<td>Ph.D.</td>
<td>39</td>
<td>7</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Note. AOA = Age of Arrival; LOR = Length of Residency; L1-use = Amount of L1-use; L2-use = Amount of L2-use

4.3. Target structures

Five target structures were elicited in Mandarin and English word-lists, statements, and questions: (a) voiceless bilabial aspirated stops (b) word-list declination, (c) two-syllable falling contours, and (d) statement pitch contours, and (e) question pitch contours. Voiceless bilabial stops, declination, and falling contours were elicited via word-lists containing four items. Each
item used to measure declination and falling contours was made up of two-syllable words with nasals, vowels, and liquids so that pitch could be easily measured from the recorded data; items used to study voiceless bilabial stops also had two syllables, but the initial consonant of each syllable began with a /p/. To assess statement and question pitch contours, Mandarin and English statements and questions with initial focus were elicited. The questions and statements corresponded to one another in that all words were the same, and pitch alone identified whether the utterance was a question or statement.

4.4. Materials

4.4.1. English lists

Two sets of four words were constructed. Each word had two syllables, and initial stress. The first set of words was made up of nasals and vowels since these are ideal sounds for measuring pitch. The words were arranged into four lists, each list contained all four words, but the order of words differed across lists. The second set of words contained two items with voiceless bilabial stops and two words made up of nasals and stops. These words were also arranged into two lists, each containing four words in different orders. All lists were placed onto index cards. The six lists can be seen in Appendix H; two example lists are provided here for clarity:

List One: Mimi, Momo, Mama, MaoMao
List Two: Momo, Mimi, MaoMao, Mama

4.4.2. Mandarin lists
Two sets of four words were constructed. Each word had two syllables where the first syllable was associated to a high tone and the second a low tone. The segments used to make up the words matched those of the English list items. Like the English lists, the Mandarin lists were written on index cards in different orders. The Mandarin lists are presented in Appendix H. One example is provided below for clarity; note that it matches the English list in terms of segmental material (the tone is transcribed to the right of the syllable to which it associates; recall that “0” is used to indicate no lexical tone):

List One: miHm0, moHmo0, maHma0, and maoHmao0

4.4.3. English statement task worksheet

A worksheet designed to elicit statements in English was prepared for the participants. The worksheet included directions and examples for the task. A trial section, and the actual task utterances were included. There were seven utterances elicited in the task. The number of words in each utterance ranged from 2 to 5. The stress patterns alternated so as to test for differences in pitch assignment related to differences in the spacing of stressed syllables. The utterances that were elicited can be found in Appendix I. The worksheet is in Appendix J.

4.4.4. English question task worksheet

A worksheet designed to elicit questions in English was also given to the participants. The worksheet included directions and examples for the task, a trial section, and the actual task utterances. The utterances roughly corresponded to the statements on the statement task worksheet; however, the initial name was different. An example of this worksheet can be seen in Appendix K.

4.4.5. Mandarin statement task worksheet
A worksheet designed to elicit statements in Mandarin was prepared for the participants. The worksheet included directions and two examples for the task. Two trials and the actual task utterances were included as well. Four Mandarin statements were elicited. The first two statements had two syllables. The first utterance was used to elicit the tonal pattern H.L and the second, L.H. The last two utterances included three syllables each, with tonal patterns H.0.L and L.0.H. These utterances were expected to have similar phonological structures as short English questions and statements. The Mandarin statements can be found in Appendix L.

4.4.6. Mandarin question task worksheet

A worksheet designed to elicit questions in Mandarin were prepared for the participants. The worksheet included directions and examples for the task. A trial section and the actual task utterances were included. The utterances corresponded roughly to the statements in the statement task. The only difference was that they had a different initial syllable.

4.5. Recording procedures

Data was recorded in a quiet room using a Sony microphone. The participants were asked to perform in three separate recording tasks designed to elicit lists, statements and questions. A native speaker of English, the author of this study, conducted experiments done in English, while a female native speaker of Mandarin conducted experiments done in Mandarin. Half of the bilingual speakers recorded utterances first in Mandarin and then in English; the other half performed the tasks in the opposing order so as to control for language order. The tasks were also given to participants in alternating orders so as to control from any affects of task order.
4.5.1. List task

The list task was designed to elicit lists with falling contours. Participants were given written directions in English during the English task, and in Mandarin during the Mandarin task. Every list task also included two examples of what the participants should do in the task. This was followed by two trial lists where participants practiced a short dialogue before actually recording the true data. Any questions concerning the task were discussed at that time. Discussions with regard to questions were conducted in the language that the task targeted.

A pile of lists written on index cards was set before the participants; the researcher also had a stack of lists. The goal of the task was to get the researcher to put her lists in the same order as the participants’ by reading the lists in response to questions that were heard over a computer. No other words were allowed.

For example, over a computer, the participants would hear, “My list says, ‘Mimi, Mama, Momo, and Nana,’ which order is your first list in?” The computerized question had three purposes: first, it was intended to model list intonation with falling pitch contours, second, it gave an example of how the list item should be pronounced, and finally, it provided a prompt to which the participants could respond. The lists used in this task can be found in Appendix M.

4.5.2. Statement task

The statement task was designed to elicit statements with contrastive focus in utterance-initial position. For the first task, participants were given written directions in English during the English task and in Mandarin for Mandarin tasks. Two examples of were also provided. The examples were followed by two trial utterances where participants practiced a short dialogue
before actually recording the true data. Any questions concerning the task were discussed at that time. Again, discussions about the exercises were held in the target language of the task.

Participants were given a worksheet with short sentences on it. They were told each sentence represented the “true situation.” Next, they heard a recorded utterance that was similar to the one on the note card, but that differed in one word. The participants were asked to correct the recorded utterance by supplying the correct information. Next, they listened to their recorded utterance and underlined which word in the sentence was emphasized. If the participant realized that he or she had forgotten to emphasize the first word, the sentence was re-recorded.

For example, if the participant read the sentence, “Papa worried Manny,” and the recording said “Susan worried Manny,” the participant corrected the recording by responding “PAPA worried Manny.” Next, he listened to his recording and underlined the word “Papa” on the worksheet. A list of the recorded prompts used for this task can be found in Appendix N.

4.5.3. Question task

The question task was designed to obtain questions with contrastive focus in utterance-initial position. Sentences that corresponded with the statements from Task 1 were used. For example, if the statement “Papa worried Manny,” was elicited in Task 1 as a statement, it was also elicited in Task 2, but as a question.

Task 2 followed similar procedures to Task 1. Participants were given a worksheet with sentences on it. These sentences described the actual event. They then heard a recording that said a similar sentence, but one word in that sentence was incorrect. The participants asked a question to clarify what the recorded speaker had intended to say. Next, the participants listened to their recordings while viewing the questions they had just produced. They underlined the word that
was emphasized in the recording. If a speaker realized that he or she had neglected to emphasize the initial word, the utterance was re-recorded. For example, if the speakers read the utterance, “Johnny worried Manny,” and then heard the utterance “Papa worried Manny,” they responded by asking, “PAPA worried Manny?” Then, while looking at a list that said “Papa worried Manny?” they underlined the word “Papa” on the worksheet.

After the recording session, the speakers were asked to listen to their individual recordings in random order. Each sentence was played a total of two times. For each sentence they listened to, they were asked to identify whether it was a statement or a question. This was to assure that the participants had actually formed what they believed to be a question or a statement.

4.6. Data preparation

Each utterance was digitized at 44 kHz and analyzed using Praat Speech Analysis Software (Boersma & Weenink, 2011). A total of 932 utterances were transcribed, and phonetic data was extracted for examination of VOT of voiceless bilabial stops, declination rate, timing of peak in falling contours, and pitch register between corresponding statements and questions (16 ME speakers x [6 M lists + 4 M questions + 4 M statements + 6 E lists + 5 E questions + 5 E statements] + 14 M speakers x [6 M lists + 4 M questions + 4 M statements] + 16 E speakers x [6 E lists + 5 E questions + 5 E statements]). Procedures for transcription and data abstraction are described in turn.
4.6.1. Transcription

4.6.1.1. Prosodic structure

Digital spectrograms and waveforms were used simultaneously to label phrase, word, and syllable boundaries. Phrase boundaries were marked at the beginning and end of each utterance, and at the end of pauses. Since L2 data has been noted to have greater breaks between word boundaries (Zsiga, 2003), a pause was only counted as a phrase boundary if its duration was greater than or equal to the duration of the shortest syllable in the utterance. Word boundaries were marked between all content and grammatical words, and syllable boundaries within each word were also labeled. In words such as Manny or marry, the exact location of the syllable boundary was not always clear since the center consonant could be part of the coda consonant in the first syllable, part of the onset consonant during the second syllable, or part of both syllables. To determine the location of the syllable boundary for these words, rhyme duration was considered; the syllable boundary was placed during the nasal or liquid segment at a point that made the duration of each rhyme within the word equal, or the duration of the second rhyme exactly one half of the first.

4.6.1.2. Prominence relations

Syllables were labeled as accented, stressed, or unstressed. To determine whether a syllable was stressed or not, syllable duration and vowel quality were used. Stressed syllables were generally longer and had fully articulated vowels. Unstressed syllables had generally shorter duration and reduced vowels.

To determine whether a syllable was accented or not, the focus item in the phrase needed to first be determined. A word was considered to be under focus if (a) it was described as being
emphasized during the subsequent listening task (described in Section 4.5.2 and Section 4.5.3), and it was the intended focus of the elicited utterance. Once the focus item was established, the stressed syllable in that word was labeled as accented.

4.6.1.3. Tone

In order to find out what intonation patterns Mandarin-English speakers use when producing statements and questions, the latter half of this dissertation switched to a qualitative analysis. Three-hundred and forty statement and question utterances were transcribed by determining where speakers associated tones to the prosodic material (16 ME speakers x [5 statements + 5 questions] + 16 E speakers x [5 statements + 5 questions]). The tonal transcriptions were then examined and sorted into groups according to the intonation pattern found in the utterance. For example, statements with LH+H% patterns were all grouped together, while statements showing a LHL+LH% pattern were put into another group. Once grouped appropriately, the different intonation patterns were described, and these were then examined to find out whether a relationship occurred between pitch register and tonal sequence.

To transcribe tone, the F0 of each utterance was extracted via an autocorrelation algorithm (Boersma, 1993) in Praat using a 30 ms Hanning window with a step size of 5ms. The F0 extractions were graphed alongside digital spectrograms and waveforms; the three sources were used simultaneously to transcribe the data. First, the F0 at the start and end of each phrase was extracted. If the F0 at the edge of the phrase fell within the bottom quarter of the pitch range, it was labeled with a L boundary tone; if it fell within the top quarter of the pitch range, it was labeled with a H boundary tone. Those phrase edges that began in the center of the pitch range were marked as toneless. Next, following assumptions of AM, all abrupt pitch changes (labeled
as elbows) were marked with tonal targets, and interpolation between the targets was assumed when there was little or no F0 change (Ladd, 2008; Jun, 2005).

To determine the tonal values of F0 elbows, the pitch range was again divided into quarters. If a F0 elbow fell within the top quarter of the pitch range, a H was transcribed; if it fell in the bottom quarter of the pitch range, a L was assigned. The transcription of mid-level elbows was examined in relation to overall intonation patterns to determine whether it was a L, H or no tone. Generally, however, if a mid-level elbow directly followed a L tone with few or no syllables in between, it was labeled as H, and when a mid-level elbow directly followed a H, it was labeled as L.

4.6.2. Measurement of phonetic data

To compare F0 configurations across speakers, F0 measurements were normalized by converting raw F0 values to z-score; 12 evenly-spaced F0 measurements were taken per syllable in each utterance. Where the autocorrelation algorithm failed to extract a pitch value, the F0 was manually calculated from the waveform; this was done in 6 files. Cases of pitch doubling and pitch halving were also manually corrected in 26 files. The raw F0 values from each list were normalized by converting Hertz to z-score using the following equation (Rose, 1987):

Formula to convert Hertz to z-score:

$$F_{0,\text{norm}} = \frac{(F_{0i} - F_{0,\text{mean}})}{s}$$

‘s’ is one standard deviation above or below the mean F0.

The mean and standard deviation were calculated from the set of raw F0 values measured in each utterance. Z-score values were, therefore, calculated separately for each utterance recorded.
In addition, digital spectrograms and waveforms were used simultaneously to make five phonetic measurements from content words; these measures are described in Table 4.2.

Table 4.2. Phonetic measurements

<table>
<thead>
<tr>
<th>Phonetic Measurement</th>
<th>Description</th>
<th>Phonetic Markers</th>
<th>Abstraction Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timing of syllable onset</td>
<td>Time (in ms) of the start of a syllable onset</td>
<td>Measured at nearest zero-crossing where the glottal pulse began</td>
<td>Initial syllable of all content words</td>
</tr>
<tr>
<td>Timing of syllable offset</td>
<td>Time (in ms) that a syllable ends</td>
<td>Measured at nearest zero-crossing to the decrease in spectral energy, indicating a change from a vowel to a nasal</td>
<td>Initial syllable of all content words</td>
</tr>
<tr>
<td>Timing of elbow</td>
<td>Time (in ms) when a flat F0 changes to a moving F0</td>
<td>Highest z-score value followed by three lower z-score values</td>
<td>All content words with rising or falling contours</td>
</tr>
<tr>
<td>Peak F0</td>
<td>Highest F0 in a given domain</td>
<td>Taken from highest z-score value</td>
<td>All content words with rising or falling contours</td>
</tr>
<tr>
<td>Duration of VOT</td>
<td>Time elapsed between the release of air through lips and the onset of vocal fold vibration</td>
<td>Measured from beginning of the release burst in /p/ tokens to the first positive peak in the periodic portion of the waveform</td>
<td>Every word-initial voiceless bilabial stop</td>
</tr>
</tbody>
</table>

4.6.3. Data preparation

Phonetic measurements were used to prepare the data for visual and statistical analysis. A description of each variable is described in turn.

4.6.3.1. VOT values

A total of 248 VOT values were prepared for analysis of stop consonant category production (14 M speakers x 4 /p/ tokens x 1 language + 16 ME speakers x 4 /p/ tokens x 2 languages + 16 E speakers x 4 tokens x 1 language). These tokens were taken from raw VOT measurements.
4.6.3.2. Declination

To begin preparation of declination data, the first four lists (which contained only nasal and vowel segments) were divided into groups according to the tonal pattern a speaker used since declination can only be compared across lists whose tones and phonological phrasing are the same. Four different patterns emerged; list-items were produced as (a) one phrase with high tones as seen in Figure 4.1; (b) one phrase with falling contours as seen in Figure 4.2, (c) two phrases with falling contours as seen in Figure 4.3, or (d) one phrase with a rising contour in the initial word, but falling contours in the remaining three words as seen in Figure 4.4.

High Tone Declination Pattern

---

**Figure 4.1.** Word-list with one phrase of high tones
Falling Contour Declination Pattern

Figure 4.2. Word-list with one phrase of falling contours

Falling Contour and Phrase Break Declination Pattern

Figure 4.3. Word-list with two phrases of falling contours. Phrase break is indicated through a L% boundary tone symbol.
Rising and Falling Contours
Declination Pattern

![Figure 4.4. Word-list with one phrase of rising and falling contours](image)

Six Mandarin speakers produced the pattern seen in Figure 4.1, and 5 English speakers produced the pattern from Figure 4.4. The remaining Mandarin and Mandarin-English speakers produced lists with a combination of the patterns seen in Figure 4.2 and Figure 4.3. The remaining English speakers used only the pattern seen in Figure 4.2. Since one phrase of falling contours occurred at least once in the majority of monolingual and bilingual speakers’ productions, these items were used to analyze phonetic interaction of discourse-level pitch production. Only 77 declination sets met the criteria. Declination sets were prepared by extracting the peak z-score value from each word-item. Each declination set contained four peak F0 values (converted to z-score).

4.6.3.3. Peak alignment

A total of 224 Peak Alignment values ([8 M speakers x 4 tokens x 1 language] + [16 ME speakers x 4 tokens x 2 languages] + [16 E speakers x 4 tokens x 1 language]) were calculated from the third list-item in lists containing all nasal and vowel segments using the following equation:
Peak Alignment = \[ \frac{[F0 \text{ elbow} - \sigma \text{ onset}]}{[\sigma \text{ offset} - \sigma \text{ onset}]} \]

Measurements were taken only from the third list-item since it was the least likely to have influence from surrounding boundary tones. (Recall that boundary tones can cause tonal crowding whereby nearby tones will shift to allow room for all tones to be pronounced. Since phrase final positions are likely to have boundary tones, the list-items in final position were avoided. Also, some native Mandarin speakers broke lists into two phrases, each consisting of two words, and so the second word in each list was also avoided.) Values of Peak Alignment could only be made from 8 M participants since the other participants produced list-items with high tones; hence, there was no peak in these items to measure.

4.6.3.4. Pitch register

One hundred and sixty graphs of corresponding statement/question pairs were prepared for analysis of pitch register (\([16 \text{ ME speakers x 5 English statement/question pairs}] + [16 \text{ E speakers x 5 statement/question pairs}]\)). Graphs were time normalized by plotting 12 evenly-spaced F0 values taken from each syllable in all statement and question utterances. F0 contours of corresponding statements and questions for each speaker were graphed in the same display so that differences in pitch use could be clearly seen. The graphs were visually labeled according to whether or not each statement and question pair had different statement and question pitch registers. To confirm visual assessment, the average F0 starting from the word adjacent to the focus item continuing to the second to last word was calculated using Praat Speech Analysis Software, and a Repeated-Measures ANOVA was used to examine whether a difference occurred between the two groups’ average F0 during statements and questions.
CHAPTER 5: EFFECTS OF BILINGUALISM ON CATEGORY INTERACTION

This chapter presents results on how language and bilingualism (versus monolingualism) affect the phonetic implementation of phonologically similar categories in Mandarin and English. Section 5.1 discusses findings for phonetic implementation of word-initial /p/. Findings for F0 peak alignment and word-list declination are in Section 5.2 and Section 5.3 respectively. Finally, implications of the results for the models of second language acquisition of phonology and for models of tone and intonation are considered in Section 5.4.

5.1. Category interaction of voiceless bilabial stops /p/

The SLM was designed to predict how bilinguals’ segmental categories interact. To see whether the hypotheses of the SLM are supported for the participants in this study, we begin by looking at whether language and bilingualism affect the duration of VOT in voiceless bilabial stops produced by monolingual Mandarin, monolingual English, and bilingual Mandarin-English speakers. Two hundred and forty-eight VOT tokens, measured from word-initial voiceless bilabial stops, were assessed (14 M speakers x 4 /p/ tokens x 1 language + 16 ME speakers x 4 /p/ tokens x 2 languages + 16 E speakers x 4 tokens x 1 language). VOT values were then analyzed using two mixed design ANOVAs.

The first mixed-design ANOVA compared the dependent variable, duration of VOT, with three independent variables: number of languages (monolingual or bilingual), language (Mandarin or English), and word token (token 1, token 2, token 3, or token 4), where word token stood for the word from which a /p/ was measured. Number of languages was examined as an independent variable since the SLM predicts bilingual VOT values to differ between monolinguals and bilinguals; in particular, bilingual values should be intermediate to
monolingual values in each group. Furthermore, language was examined as an independent variable because the VOT values of Mandarin and English are expected to be different. Mandarin /p/ is generally pronounced with a longer-lagged VOT than English /p/. Finally, since each speaker produced four tokens of the /p/ sound, a repeated factor needed to be considered to make sure that within-subject relations were correctly accounted for.

Language affected duration of VOT, $F(1, 58) = 37.12, p < 0.01$ with an effect size of 0.57. Participants speaking in Mandarin ($M = 79.05$ ms, $SEM = 1.89$, $n = 30$) produced longer VOT values than those speaking in English ($M = 62.75$ ms, $SEM = 1.89$, $n = 32$). Neither the token from which VOT was measured, the number of languages a participant spoke, nor the interaction of language and language number affected VOT values ($p < 0.01$). So, while mean VOT durations of the bilingual groups were intermediate to the monolinguals’ values as seen in Figure 5.1, the standard deviation for the bilingual groups were too large for them to be significantly different from their monolingual counterparts.
There was a good deal of variation in the bilingual VOT values. This variation could have stemmed from different variables. One possibility is that the variation occurred due to a task-related affect since some bilingual speakers began the task in Mandarin while others began in English. If so, a significant difference between the VOT values uttered by bilinguals who began the task in Mandarin versus those who began the task in English would be expected.

Another possibility is that particular words were more prone to having a VOT value of a certain length; for example, the word *papa* may have been uttered with a longer VOT value by all speakers, while *poppy* may have been uttered with a shorter VOT duration. If vocabulary items were the cause of VOT variation in bilingual data, then a significant difference between VOT values of different tokens would be expected. A final possibility is that the variation in
VOT values in the bilingual data was simply random, in which case, no independent variables would be expected to have a significant relation to bilingual VOT values.

To test these different explanations, the variation in bilingual VOT values was analyzed using a second mixed-design ANOVA. The dependent variable was VOT value; the between-group factor was recording order (Mandarin first, English first), and the repeated measures were language (Mandarin or English) and token (token 1, token 2, token 3, or token 4). Results indicated that language affected each participant’s VOT values, $F(1, 14) = 10.81, p < 0.01$ with an effect size of 0.44. Bilingual speakers produced longer VOT values in Mandarin ($M = 77.36, \text{SEM} = 1.723, n = 16$) than in English ($M = 68.94, \text{SEM} = 2.4, n = 16$). However, interspersed at random, in almost every bilingual’s speech, were VOT values that fell within norms of the opposing language. Neither the token from which the VOT value was measured, nor the order of tasks a participant followed when recording the data affected the VOT value ($p < 0.01$). This indicates that the VOT variation in bilingual speech was unlikely caused by the task items or procedures; rather, the variation in VOT appears to be random.

These results do not support Hypothesis 1 which states that bilingual speakers will produce VOT values at a point intermediate to monolingual speakers. Rather, bilingual speakers’ VOT values formed a bimodal distribution. Generally, the bilinguals produced phonetic values within monolingual ranges for the language they were speaking, although some phonetic values were within monolingual ranges for the opposing language. This is a surprising find since it means that the SLM did not make the correct predictions for the bilingual participants’ segmental category interaction. Given the amount of effort that went into selecting bilingual participants who fit the typical population for which the SLM was written, it is unlikely that the participants
in this study were of a special case. A more likely explanation, and one that is brought up in more detail in Section 5.3 is that the SLM doesn’t accurately represent segmental category interaction in bilingual speech.

5.2. Rate of declination

To test whether the rate of declination in bilingual speech is consistent with the hypotheses of the SLM, a mixed design ANOVA was used to analyze 184 peak F0 tokens (4 tokens per list, 14 monolingual Mandarin x 1 language + 16 bilinguals x 2 languages + 16 English monolinguals x 1 language). For each participant, one word list that fit the pattern XX was selected (for more details review Section 4.6.3.2). The dependent variable measured was peak F0. To assess whether word-list declination occurred between F0 peak values within a single list, the F0 peak values within a single word-list were treated as a repeated measure, called word-list position (1, 2, 3, and 4). This value was expected to be significant, assuming that all participants did in fact have at least some declination in their word-list production. Two between-group variables were also compared, language (Mandarin and English) and number of languages (bilingual or monolingual), to find out whether declination rate varied according to language spoken, or whether a participant was monolingual or bilingual.

The position of the peak F0 significantly affected F0, F (3, 58) = 159.13, p < 0.01 with an effect size of 0.85. Peak F0 declined with each word in the list; as expected, this decline in F0 occurred in both languages. The interaction of language and position was also significant, F (3, 58) = 9.09, p < 0.01 with an effect size of 0.25; the peak F0 of list-final words was higher in Mandarin (M = 0.22, SEM = 0.48, n = 30) than in English (M = -0.21, SEM = 0.56, n = 32). This indicates that both languages use declination in lists, but in English, there was a greater drop in
F0 between the third and fourth words. The differences in rate of decline can be seen in Figure 5.2 and Figure 5.3.

**Figure 5.2.** Mandarin declination in word-list with Tone 1 heard over the initial syllable of bi-syllabic words.

**Figure 5.3.** English declination in word-list with H*L contour heard over the initial syllable of bi-syllabic words.
Position, language and number of languages also interacted to create a significant effect on peak F0, $F(3, 58) = 8.44, p < 0.01$ with an effect size of 0.23. Monolingual Mandarin and monolingual English speakers produced significantly different F0 during the final word such that in Mandarin, the peak F0 was much higher than in English, as seen in Figure 14. Bilingual speakers, however, produced the final peak F0 in each language within such varied distances from the peak F0 of the third word that they failed to differ from each other or from either monolingual group. Figure 5.4 displays the average peak F0, converted to z-score for each group. Note that the bilingual groups have a large variation (or standard deviation) while the monolingual groups have a relatively small standard deviation; moreover, the peak F0 in monolingual English speech was considerably lower than monolingual Mandarin speech.

![Figure 5.4. Comparison of final position F0 by monolinguals and bilinguals.](image)

97
Since bilingual speakers did not all produce comparable Mandarin and English lists with the same phonological makeup, statistical analysis of bilingual list declination in each language was not possible since a repeated-measure from the same word-list, produced in each language, could not be found for all participants. Looking at the high standard-deviation in the bilingual group’s word-list declination, in comparison to the large standard deviation in the bilingual group’s VOT durations in Section 5.2, a similar trend appears to be happening. It seems that bilingual speakers produce phonetic values that sometimes fall within the range of monolinguals of the L1, and sometimes within the range of monolinguals of the L2.

This finding does not support Hypothesis 2 which predicts that bilingual speakers will produce lists with a declination rate intermediate to monolingual values in each language. This finding suggests that the SLM makes the wrong predictions, both at the segmental level, and at the tonal level; moreover, it also implies that segments and pitch contours may interact very similarly in bilingual speech since the same general trends that were followed in segmental production also happened in declination production.

5.3. FO peak alignment

The final step in the analysis of bilingual category interaction was to analyze the the F0 peak alignment values measured from pitch contours used to create lexical contrasts in Mandarin (tone), but discourse-level contrasts in English (intonation). To do this, 224 F0 peak alignment values were analyzed using two mixed design ANOVAs ((8 M speakers x 4 tokens x 1 language] + [16 ME speakers x 4 tokens x 2 languages] + [16 E speakers x 4 tokens x 1 language]). Following the same procedures as used for analysis of /p/, the first ANOVA compared monolingual and bilingual peak alignment. Language affected peak alignment,
F (1, 58) = 44.77, p < 0.01 with an effect size of 0.62. Participants speaking in Mandarin (M = .97, SEM = 0.02, n = 30) aligned the elbow of falling pitch contours later than those speaking in English (M = .81, SEM = 0.02, n = 32).

There was no main effect on peak alignment of the number of languages a participant spoke (p > 0.01); however, there was a significant interaction of language and number of languages a participant spoke F (1, 58) = 7.8, p < 0.01 with an effect size of 0.22. Monolingual speakers of Mandarin and English aligned F0 peaks at significantly different locations. The mean bilingual peak alignment values, however, had such large standard deviations that they were not significantly different from each other or their monolingual counterparts. Bilinguals aligned the F0 peak within normal ranges of Mandarin monolinguals and English monolinguals no matter what language they were speaking; this was especially true for bilingual English as seen in Figure 5.5.
Finally, the token from which the F0 peak alignment value was calculated did not affect peak alignment (p < 0.01), indicating that the task item from which the peak alignment value was measured was not the cause of the high standard deviation.

To better understand whether the variation in bilingual F0 peak alignment values was between participants or between tokens, a second mixed-design ANOVA was used to analyze these data. The same independent factors as described for the mixed-design ANOVA in Section 5.2, used to analyze VOT durations, were used here. Thus, while the dependent variable was F0 peak alignment, the between-subject factor was order of tasks (Mandarin first, English first), and the repeated-measures were language (Mandarin or English) and word token (token 1, token 2, token 3, and token 4), referring to the actual word from which the phonetic value was measured.

*Figure 5.5. Mean F0 peak alignment values by monolinguals and bilinguals.*
Language affected each participant’s peak alignment values, F (1, 14) = 15.11, p < 0.01 with an effect size of 0.52. Almost every bilingual speaker aligned the peak of falling contours later in Mandarin (M = .94, SEM = 0.01, n = 16) than in English (M = .84, SEM = 0.01, n = 16).

However, interspersed at random, in almost every bilingual’s speech, were alignment values that fell within norms of the opposing language. When the bilinguals’ alignment values were pulled together, this created a bimodal distribution as seen in Figure 5.6. Neither the token from which the peak alignment value was measured, nor the order of tasks a participant followed when recording the data affected peak alignment (p < 0.01). This indicates that the within-subject variation was unlikely caused by task items or recording procedures.

These findings do not support Hypothesis 3 which states that the bilingual groups will align peaks of two-syllable falling contours at a point intermediate to monolinguals in Mandarin and English. Just as the SLM hypotheses did not accurately predict segmental patterns or
declination patterns, the F0 peak alignment data did not confirm the hypothesis either. This further supports the idea that the SLM is making the wrong predictions about bilingual category interaction; furthermore, once a model is in place that makes the correct predictions for segments, it should work for segmental category interaction and interaction of pitch contours alike. Lastly, since a two-way interaction was found between a similar L1 tonal contour in Mandarin and L2 intonation contour in English, we can confidently state that tone and intonation affect one another (or interact) in bilingual speech, regardless of their linguistic function.

5.4. Summary and discussion

Two major findings are notable from these data. First, interaction of bilingual tone and intonation categories was found; however, the way in which the categories interacted was not predicted by the SLM. Second, findings for how bilinguals implemented segmental categories at the phonetic level paralleled the findings for phonetic implementation of pitch. The implication of these findings for models of bilingual category interaction and models of tone and intonation are discussed in turn.

5.4.1. SLM hypotheses not supported

The SLM predicts that phonetic values of bilingual speech should be intermediate to monolingual norms in both languages. The data reported in this study did not support this hypothesis. Instead, bilingual participants generally produced phonetic values that were within normal ranges of monolingual values for the language they were speaking; occasionally, however, they produced a category that fell into monolingual norms of the opposing language. In other words, bilingual speakers’ phonetic values formed a bimodal distribution. This happened both at the segmental and suprasegmental levels.
These findings indicate that shared L1-L2 speech categories are not actually changing at the phonetic level as suggested by the SLM. Rather, speech categories which have a phonological counterpart in both languages have two possible variants in bilingual speech which correspond to the phonological categories of the L1 and the target language. A better explanation of the data is that the frequency with which one phonetic variant is chosen over the other changes during second language acquisition. It may be that the choice is influenced by the frequency of usage around the bilingual speakers, and, perhaps, by which phonetic variant was most recently heard.

The idea that language users are influenced by input frequencies is not new; in fact it has been suggested for all levels of language (Ellis, 2002; Zipf, 1935). In Flege’s own work, he acknowledges that sensitivity to input frequencies may be an alternative explanation for his data. He dismisses this idea, however, in lieu of the classic SLM explanation. In a study that examined VOT of stop consonants produced by native Italian speakers of English, for example, he found that late learners of English were more likely to have average VOT values that were intermediate to monolingual values while early learners of English were generally able to produce VOT values within native English speaker norms:

The [Native Italian] subjects with an [age of learning] (AOL) of 21 years had spent just over half (56%) of their lives in Canada. Their VOT values may therefore have reflected the overall distribution of VOT values heard in both English and Italian realizations of /p, t, k/. (This assumes, of course, that these subjects did not have separate representations for Italian and English voiceless stops.) Following this line of reasoning, the mean VOT values shown in Fig. 4 may have decreased (become less English-like) as AOL increased because the proportion of all stops heard over the course of a lifetime that were English stops decreased.

We believe, however, that a more likely explanation for the observed AOL effect on VOT is that the likelihood of new phonetic categories being established for L2 vowels and consonants diminishes with increasing AOL, as discussed by Flege (1995).

(Flege, Munro & MacKay, 1995)
Since the study compared aggregated VOT values, it is not possible to discern whether the Italian-English bilinguals produced all VOT values at a point intermediate to monolinguals, or whether they produced roughly 50% of the VOT values like a monolingual Italian speaker, while the remaining 50% were like a monolingual English speaker. (See Section 2.1.4 for more details on SLM statistical design.)

If, however, the correct interpretation of the data is to assume that some of the phonetic values were L1-like, while others were L2-like (as was the case during this study), a clearer connection between learning-mechanism and its effect in second language learning emerges. Recall that the SLM does not explicitly define its learning mechanisms; however, it predicts that learning mechanisms remain intact throughout the lifespan. How might a distributional learning mechanism play out in L2 speech? If speakers really do store frequency-based information taken from the input, then in L2 speech acquisition, we might guess that the storage of phonetic values for a particular shared phonological category would include phonetic variants of the L1 and the L2. A single category could then have a bi-modal distribution within it, and the bilingual speaker simply selects from the database of phonetic variants.

That a bilingual speaker is influenced by distributional frequencies in language input, however, can only be half of the story. If the speakers were selecting a phonetic value at random, we might expect that the phonetic values would be the same during speech in Mandarin and in English. This was not the outcome. Generally, when speaking Mandarin, the ME speakers produced more Mandarin-like phonetic values, although English values slipped in on occasion. When speaking English, more English-like phonetic values were heard, but again, some Mandarin phonetic values entered in too. The exact reason for this is unclear. Maybe this
outcome had to do with recency of exposure. A speaker may be more prone to select a Mandarin-like phonetic value if the last value heard was Mandarin. Such a scenario could be tested by controlling the language a bilingual speaker used, but varying the language that the speaker listened to. For example, if recency of exposure is linked to selection of phonetic value, we might guess that a Mandarin-English bilingual will produce more English-like phonetic values, even when speaking Mandarin, if the directions for a given task are all in English. On the other hand, the same Mandarin-English bilingual will produce more Mandarin-like phonetic values when speaking Mandarin, if the directions for a given task are all in Mandarin.

5.4.2. A parallel in bilingual segmental and supra-segmental production

The second major finding in these data is that bilingual patterns for segmental category production appear to parallel those of pitch production. That is, bilingual speakers’ VOT values formed a bi-modal distribution such that VOT values fell within the range of monolingual speakers from the L1 and L2. The same general pattern was true for pitch. Bilingual category production (where the L1 and L2 phonological representation is identical) can therefore be modeled similarly for pitch as it is for segments. For models of second language acquisition of phonology, this means that a single representation of bilingual category interaction can be made to simulate the acquisition of segments and tones.

5.4.3. Understanding the data from an AM or PENTA perspective

If we stop to consider the two models of tone and intonation (AM and PENTA) explored in this study, we will come out with very different explanations about the bilingual data found in this study. Mandarin-English bilinguals produced F0 peak values of similar Mandarin tone and English intonation contours in such a way that the peak alignment values formed a bimodal
distribution in each language spoken; that is, two separate clusters of values were found for the alignment of the F0 peak in the Mandarin tone, and two separate modes were found for the alignment of the F0 peak in the English intonation contour. From a PENTA perspective, these data must be explained in that bilingual speakers experiment with local pitch targets. Since F0 peak alignment is believed to not be under any control of the speaker, rather it is an incidental by-product of the local pitch targets in a phrase, the variation in F0 peak alignment found in the bilingual data of this study can only be explained in that bilinguals experiment with, or vary their choice of local pitch targets. For a Mandarin speaker using English, this explanation seems reasonable. It makes sense that a second language speaker may not know which local pitch target is needed on a specific syllable, and so the speaker tries out different pitch targets to find out what works. For the Mandarin-English bilingual speaking Mandarin, however, this explanation is at best questionable. Why would learning a second language cause a person to no longer know which local pitch target is needed on a specific syllable? Moreover, why would learning English cause a Mandarin speaker to start using different combinations of lexical tones on a given phrase? It wouldn’t. Mandarin speakers are not changing the lexical tones they are using for a given syllable; they are simply altering the way in which they implement that tone. The tone remains steady, but the alignment used when implementing the tone changes.

The AM theory, on the other hand, provides a better explanation. The distribution in F0 peak alignment values are explained in that Mandarin and English have identical phonological categories which are produced with different phonetic specifications. When learning the L2, speakers acquire a new phonetic variant for a phonological category which has already been formed in their phonological system. This phonetic variant may overlap greatly in terms of the
phonological positions in which it can occur. For example, monolingual English /p/ has two phonetic variants: /p/ and /pʰ/. While /pʰ/ occurs in most phonological environments, /p/ typically appears after /s/. The phonological environments that condition these two phonetic variants for the same phonological category are different. Now enter Mandarin /pʰ/ into the system. The phonological environments under which Mandarin /pʰ/ can occur overlap almost entirely with the environments that English /pʰ/ occurs. Thus, as long as the phonological conditioning is appropriate, the speaker can choose either phonetic variant; a case of free variation (or nearly-free variation) emerges. Social conditioning may also play a role in the selection of phonetic variant chosen since speakers were more likely to choose the phonetic variant that matched the monolingual variant of speakers for the language they were using.

5.4.4. What do these results reveal about tone and intonation structure?

This study has important repercussions for our interpretation of how intonation should be modeled structurally. If tone and intonation should be modeled on separate levels, and with different constructs, then we don’t expect tone and intonation contours to interact in bilingual speech, especially when the bilinguals have such a good command over their two languages. If the Mandarin-English speakers were new to English, it may be argued that the speakers were simply mistaking English intonation for lexical tone; however, since they were well-advanced in their use of English, such an argument is not valid, and so these results suggest that tone and intonation should be represented along a single layer of linguistic structure. This is the representation modeled by tone-sequence models, such as AM.
Chapter 6: Mandarin-English Tone to Syllable Mapping

Chapter 6 describes findings on Mandarin-English intonation of yes/no statements and questions. In Section 6.1, a report of how speakers mapped tone to the prosodic material is given. Section 6.2 describes three patterns for how tones were mapped to syllables in monolingual English and bilingual Mandarin-English. Section 6.3 ends in a discussion about the tonal association scale, and whether aspects of Mandarin tone appeared in Mandarin-English intonation.

6.1. Prosodic structure and syllabic prominence

To assess how tones mapped to syllables using the tonal association scale described in Chapter 2, we must first have an account of the prosodic structure and syllabic prominence for the various utterances produced by participants in this study. All 160 utterances produced by native English speakers were determined to have initial contrastive focus (16 English speakers [5 statements + 5 questions with initial focus]) since (a) speakers selected the initial word as emphasized during their subsequent focus-identification task, and since (b) the initial word was the intended focus in each statement and question assessed in this study. One hundred and fifty-eight utterances produced by Mandarin-English speakers were determined to have initial contrastive focus; the two remaining utterances were short questions whose focus could not be identified by the speaker during the focus-identification task ([16 Mandarin-English speakers x (5 statements + 5 questions with initial focus) – 2 unclear utterances]). The contours heard over these phrases are shown in Figure 6.1. The two speakers produced the same question in a similar manner, and so only one example is shown.
Figure 6.1. Example short question from which contrastive focus could not be identified by Mandarin-English speaker who produced the phrase and then listened to it during a subsequent focus-identification task.

Note that this contour treats the accented syllable and the stressed syllable of the unaccented word as the same, and, consequently, the contrastive focus item could not be retrieved. Since the focus item in the phrase could not be identified, the accented syllable could not be determined using outside measures, and so these examples were removed from the data.

One hundred and thirty-six of the native English utterances were produced with no pauses between words, and were therefore labeled as a single phonological phrase. Twenty-four native English utterances were divided into two phonological phrases, evident from the substantial pause following the focus word as seen in Figure 6.2. All 24 utterances that were produced as two phrases consisted of five syllables; short phrases were never broken into two phrases. Mandarin-English speakers, by comparison, produced long and short phrases as a single phonological phrase in all examples with an exception of two long phrases which were produced by the same speaker.
Figure 6.2. Spectrogram of a long utterance divided into two phonological phrases; the phrase boundary is divided by a substantial pause at the end of the focused element “Manny.”

When utterances were divided into two separate phrases, determining the accented syllable during the second phrase was a difficult task. Recall that the intention of the task was to elicit phrases where only the initial word in an utterance was under focus; to determine whether a word was under focus, two criteria had to be met: (1) speakers described the word as being emphasized during a subsequent listening task, and (2) the word selected by the speaker as the emphasized word was indeed the intended focus for a given task item. When utterances were broken into multiple phrases, an accented syllable presumably occurred in each phrase; however, the procedures used to determine whether a syllable was accented were designed to identify only
a single accented word per utterance. For this reason, utterances which were broken into two phrases were not used to assess the predictions of the tonal association hierarchy.

Table 6.1 and Table 6.2 depict the differences between prosodic structure found in monolingual English phrases and those produced by Mandarin-English speakers. The major difference between the two dialects was in the production of grammatical words such as “to.” Where native speakers of English produced “to” with reduced vowels, the native Mandarin speakers produced them with full vowels as seen in Figure 6.3 and Figure 6.4 respectively.

Table 6.1. Monolingual English prosodic structure

<table>
<thead>
<tr>
<th>Short Phrases: [(Ann)_F won]_i</th>
<th>[(Ma.nny)_F won]_i</th>
<th>[(Ann)_F mar.ried]_i</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ* σ’</td>
<td>σ* σ σ’</td>
<td>σ* σ’ σ’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Long Phrases: [(Ma.nny)_F will mar.ry]_i</th>
<th>[(Ann)_F will need to win]_i</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ* σ σ’ σ’</td>
<td>σ* σ σ’ σ’</td>
</tr>
</tbody>
</table>

Note. Phonological phrases are indicated via [ ]; focus items are marked as ( )F, and syllable boundaries within words are marked via a dot. Syllabic prominence is marked beneath each syllable: accented syllables are indicated via σ*, lexically stressed syllables through σ’, and weak syllables through σ.
Table 6.2. Mandarin-English prosodic structure

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>σ*</td>
<td>σ</td>
<td>σ</td>
</tr>
<tr>
<td></td>
<td>σ’</td>
<td>σ’</td>
<td>σ</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Long Phrases:</th>
<th>[(Ma.nny)F will mar.ry]i</th>
<th>[(Ann)F will need to win]i</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>σ* σ σ’ σ’ σ’</td>
<td>σ* σ’ σ’ σ’ σ’</td>
</tr>
</tbody>
</table>

Note. Phonological phrases are indicated via [ ]; focus items are marked as ( )F, and syllable boundaries within words are marked via a dot. Syllabic prominence is marked beneath each syllable: accented syllables are indicated via σ*, lexically stressed syllables through σ’, and weak syllables through σ.

Figure 6.3. Monolingual English production of the word “to” in the phrase “Ann will need to win.” The first formant frequency is slightly elevated, and the distance between the first and second formant frequencies are relatively large, as expected in the production of a reduced vowel, or shwa. Moreover, the timing of the vowel is very short.
Figure 6.4. Bilingual Mandarin-English production of the word “to” in the phrase “Ann will need to win.” The first formant frequency is low, and there is hardly any distance between the first and second formant frequencies, as expected in the production of a fully articulated high, back vowel, /u/.

The tendency for second language speakers to produce grammatical and content words with equal prominence is not unique to this study; similar reports have been made of second language speech in other studies (Wennerstrom, 1994). The difference in prosodic structure and prominence, however, is expected to cause deviances in the way tone maps to the segmental material in monolingual English and Mandarin English. In particular, we might expect more tones to occur in the Mandarin-English speech since more syllables are treated as prominent.
6.2. Mapping of tone to syllable

A single pattern for mapping tones to prosodic material in monolingual English was found, while two patterns emerged in the Mandarin-English data. Each is described in turn.

6.2.1. Monolingual English

All 16 monolingual speakers of English used the same gross pattern for mapping tones to syllables in statements and questions. The only difference found was in whether long utterances were divided into two phrases, or left as a single phrase, as was discussed in Section 6.1.

6.2.1.1. Statements

One pattern for mapping tones to syllables was found in the monolingual English statements. The native English speakers produced statements with a F0 peak mapped to the stressed syllable of the focused item. A F0 minimum mapped to the right edge of the syllable following the accented one. If the phrase contained more than two syllables, a low, but slightly declining F0 plateau continued to the end of the phrase as seen in Figure 6.5.

![Figure 6.5](image)

*Figure 6.5. Falling F0 contours exemplary of those produced by native English speakers in short and long statements with initial contrastive focus.*

The timing of the F0 peak, indicative of a H autotone, aligned to the end of the vowel gesture which was determined by a decrease in acoustic energy as seen in Figure 6.2. Alignment
of this autotone to the syllable matched results obtained in Chapter 5, where native speakers of English aligned F0 peaks well within the syllable. The F0 minimum, indicative of a L autotone, aligned near the right edge of the following syllable; this is also seen in Figure 6.6.

![Figure 6.6](image)

**Figure 6.6. F0 peak and valley alignment to the syllable during statements produced by native English speakers.**

L consistently mapped to the syllable following the H tone. When the focused item was two syllables, such as in the word *Manny*, the L occurred at the edge of the focus word; in one syllable words such as *Ann*, however, the L occurred at the edge of the following word. Under an AM framework, a H followed by a trailing L tone is typically represented as the complex pitch accent H*L and thus the results of this study confirm earlier findings about contrastive focus in English being marked by a H*L pitch accent (Pierrehumbert 1980; Jun, 2005).

Native speakers of English used the complex falling toneme (H*L) to mark one level of prominence: sentential prominence. Lexical prominence relations, however, were not indicated via tone. Each time a word was under focus, the stressed syllable of that word was produced with a F0 peak. In words that had initial lexical stress, but that were not under focus, however, no peak occurred as seen in Figure 6.7.
Figure 6.7. One level of prominence, sentential prominence, conveyed via tone. The word “marry” with initial lexical stress bears no tone, while the focused element “Manny” involves a H*L contour.

Following the focus word, the remainder of the phrase was produced with a low, flat plateau, and the intonational phrase then ended with a F0 near the bottom of the pitch range. This was treated as a L final boundary tone (%L). Pulling together all of the pieces together, the statements of Monolingual English speakers in this study can thus be summarized as having a H*L-L% pattern in both short and long utterances.

6.2.1.2. Questions

Like in statements, only one pattern for tone-to-text mapping was found in the monolingual English questions. The Native English speakers produced questions with a F0 near the bottom of the pitch range. This was treated as a L initial boundary tone (%L). The F0 remained L through the accented syllable and then began to rise to the right edge of syllable directly adjacent to the accented one. The syllable following the accented one was heard with a higher F0 whether it was part of the focus word or not as can be seen in Figure 6.8.
The timing of the F0 minimum, indicative of a L autotone, aligned to the center of the vowel gesture at the point where the acoustic energy became the strongest. This change in acoustic energy inside the vowel may be from a glottal stop pronounced at the beginning of vowel-initial words; the closure of the vocal folds to their target position during vowel production likely takes time, and until they are in the target position, a slight decrease in acoustic energy can be seen during the vowel. The segmental anchoring position is seen in Figure 6.9. The F0 maximum that followed the L autotone on the accented syllable aligned near the right edge of the following syllable; this is also seen in Figure 6.9.

Figure 6.8. Typical rising F0 contours produced by native English speaker in short and long questions with initial contrastive focus.
Native speakers of English used the complex rising toneme (L*H) to mark one level of prominence: sentential prominence. This was done by aligning the L*H autotone to the stressed syllable of the focus word. Lexical prominence relations, by comparison, are not indicated via tone. This is evident from the F0 during stressed syllables that were not under focus. These words were not indicated with an F0 minima; rather, the F0 continuously accelerated through these syllables toward the edge of the phrase which ended in an F0 near the top of the pitch range as can be seen in Figure 6.10. This pattern was treated as a H final boundary tone (H%).

Figure 6.9. F0 minimum alignment to the syllable during focused elements produced by native English speakers.
Pulling together all of the parts of monolingual English questions, the pattern can be summarized as %L-L*H-H%. Figure 6.11 provides 10 exemplary graphs made from the speech of one of the native speakers of English. Note that the falling and rising contours in statements and questions are similar in the short and long utterances. Also notice the similarity between the patterns used for producing statements and questions: contrastive focus is marked by a complex pitch accent in both types of utterances, and the F0 then continues toward a boundary tone that matches in tonal value to the pitch accent’s second autotone. For example, the H*L pitch accent is followed by a L boundary tone and the L*H pitch accent is followed by a H boundary tone. This pattern results in relatively smooth rising and falling contours which span across the utterance.
Figure 6.11. Monolingual English statements and questions with initial contrastive focus.
6.2.1.3. Assessing the tonal association scale

Using the tonal association hierarchy, we predicted that native speakers of English would map a tone to every accented syllable. The intonation patterns that monolingual English speakers used to differentiate yes/no statements from questions confirmed this hypothesis since every accented syllable did indeed bear tone. Let us now turn to Mandarin-English intonation to see whether the predictions of the tonal association scale continue to hold up.

6.2.2. Mandarin-English Pattern A

Seven of the sixteen Mandarin-English speakers produced statements and questions using very similar intonation contours to one another. The phrases produced by this group of speakers were labeled Mandarin-English Pattern A. The speakers who fell into this group received generally higher scores on the degree of foreign accent task where native speakers of English rated bilingual speakers’ English accents. This indicates that Mandarin-English Pattern A speakers had the overall lower level of English pronunciation. Not surprisingly then, the two question utterances that were removed from the data, due to inability of the speaker to retrieve the focus item during the subsequent listening task (see Section 6.1), were amongst this group. Thus, 68 total utterances were transcribed and grouped together as Mandarin-English Pattern A.

6.2.2.1. Statements

Pattern A Mandarin-English speakers produced statements with a F0 peak aligned to the stressed syllable of the focused item. A F0 minimum aligned to the right edge of the focus item as seen in Figure 6.12.
This pattern used to mark the accented syllable was similar to the H*L pitch accent found in the monolingual English statements; however, two crucial differences amounted. First, the second autotone (L) aligned to the right edge of the focus word whether or not the word was monosyllabic. Recall in monolingual English, on the other hand, the L always aligned to the syllable adjacent to the accented one, whether or not the adjacent syllable was part of the focus item; this was shown in Figure 6.12. This difference in association between autotone and syllable in the Pattern A Mandarin-English statements was treated as a L boundary tone that occurred at the edge of the focus word (L#). Thus in Mandarin-English Pattern A, the focus item was marked with a H tone on the accented syllable and a L tone at the edge of the word (H*-L#). The idea that boundary tones may exist on the edges of focus words in some languages is not totally unique to this study. Similar proposals involving boundary tones on edges of words have also been made for Serbo-Croatian as well (Godjevac, 2005; Zsiga & Zec, to appear).

The second difference between the H*L monolingual English pitch accent and the HL contour found in Mandarin-English Pattern A had to do with F0 peak alignment. In Mandarin-
English, the F0 peak aligned to the accented syllable in two different ways: either the F0 peak occurred at the right edge of the vowel as was seen in monolingual English, or a flat plateau stretched across the entire syllable, ending at the right edge of the syllable as seen in Figure 6.13. Of the 35 utterances that made up the Mandarin-English Pattern A statements, the H tone aligned within the accented syllable 29 times, and to the right edge of the syllable 6 times.

![Figure 6.13](image.png)

*Figure 6.13. Two patterns for F0 peak alignment to the syllable for the H toneme in Mandarin-English: (a) H aligned near the right edge of the vowel; (b) H aligned near the left edge of the vowel and extended to the right edge of the syllable.*

The variation in peak F0 alignment was not surprising; recall in Chapter 5, it was seen that Mandarin-English bilinguals align F0 peaks to the syllable using two different phonetic patterns: one pattern is an early-aligned H found in monolingual English speech, and the second is a late-aligned H found in monolingual Mandarin speech. The variation in alignment was likely due to phonetic interaction between sounds which the bilinguals have categorized as the same in Mandarin and English.

The alternate H tonal alignment pattern in Mandarin-English Pattern A did have an effect on the phonology in one environment. Whenever a late-aligned H occurred during a
monosyllabic word, the L boundary tone on the edge of the focus word (L#) was not present; rather, a short disjuncture occurred following the focus item. This can be seen in Figure 6.14.

![Figure 6.14. Mandarin-English late-aligned H tone followed by a short disjuncture.](image)

Since this disjuncture occurred only in the one specific environment, and because a L always occurred in the other patterns, the disjuncture was treated as an alternate pronunciation of the L#. Such an analysis is not totally unique to this study; ToBI transcribers have noted similar disjuncture and boundary tone mismatches occurring in other languages as well, and so most ToBI transcription systems include disjuncture or tonal marks to indicate when a researcher feels that a boundary tone is present, but the typical tonal events that accompany the boundary tones are not present (Arvaniti & Baltazani, 2005, pg. 107; Beckman, Hirschberg & Shattuck-Hufnagel, 2005, pg. 23; Venditti, 2005, pg. 184).

Unlike monolingual English speakers, Mandarin-English Pattern A speakers marked two levels of prominence, lexical and sentential, through tone. Not all stressed syllables, however, surfaced with tone. Instead, lexical prominence was marked through tone only if a word was multisyllabic, and prominence relations thus needed to be distinguished. For example, in the
word *married*, the initial syllable is stressed while the ultimate syllable is not. To mark the prominence relation between the first and last syllable, a H tone associated to the stressed syllable. In monosyllabic words, such as *need*, there are no other syllables with which to compare prominence, and thus no H tone was inserted on monosyllabic words to indicate lexical stress. This can be seen by comparing Figure 6.15 with Figure 6.16. In Figure 6.15, an extra H is inserted on the stressed syllable of the final word, while in Figure 6.16, the F0 after the accented syllable remains steady and low.

![Figure 6.15](image1.png)

*Figure 6.15. H tone inserted on stressed syllables of multisyllabic words of Mandarin-English Pattern A phrases.*

![Figure 6.16](image2.png)

*Figure 6.16. No H tone inserted on stressed monosyllabic words of Mandarin-English Pattern A phrases.*
Note that the Mandarin-English Pattern A tonal sequence resulted in sentential-level stress having an overall higher pitch range. In phrases with multisyllabic words in the post-focus region, the H*-L# pattern on the focus word appeared to cause a downward adjustment of the pitch range so that subsequent H tones were realized at a slightly lower pitch range. In phrases with monosyllabic words in the post-focus region, the pitch range of the post-focus region remained low to the edge of the phrase. In both scenarios, the focus item was thus realized at an overall higher pitch range than the post-focus region, and thus two levels of stress, sentential and lexical, were easily retrievable from the surface contour.

6.2.2.2. Questions

In contrast to monolingual English speakers, Mandarin-English Pattern A speakers produced a LHL contour over the focus item in questions. In this pattern, the F0 contour began near the bottom of the pitch range. This was treated as an initial L boundary tone (%L). The F0 then rose to a H near the right edge of the accented syllable. A L then aligned to the right edge of every focus item. In bi-syllabic words, this L was realized on the second syllable; in monosyllabic words, the duration of the syllable was extended to accommodate the extra L tone as seen in Figure 6.17. Since the L always occurred at the edge of a focus word, it was treated as a L boundary tone (L#).
In contrast to monolingual English questions, the questions in Pattern A ended in a complex tonal sequence LH. The L tone signaled stress relations between syllables of multisyllabic words. A L occurred on the stressed syllable and the H on the phrase-edge in phrases such as *Ann married?* as is noted in Figure 6.18.

However, a crucial difference occurred between the Pattern A statements and questions. In statements, the boundary tone was a simple tone, L%. Stressed syllables of multisyllabic
words were marked through a H tone. Monosyllabic words in statements, however, did not have a H tone at all. In contrast, for the questions, all final words, even those that were monosyllabic, had a LH pattern. So L occurred on every final word in Mandarin-English Pattern A questions. Compare Figure 6.17 and Figure 6.18. Both are analyzed with a LH%, and so while it is evident that the L is attracted to the stressed syllable of multisyllabic words, it didn’t appear to be inserted especially for these syllables since it occurs on all final words. It’s possible that the position of the stressed syllable is required to carry tone in Mandarin-English Pattern A. It seems, however, that there are two repair mechanisms available to fill the required position if no tone occurs during the input: first, a tone can be moved from one location to another, as was the case with the L tone in questions. Or, if no tone is available to move in that position, a tone can be inserted; this was seen in the statements.

Sentential level stress was realized in questions through a LHL pattern. The LH pattern, needed to distinguish lexical stress relations, mapped to the stressed syllable of the focus word; a L edge tone was added to the right edge of the focused element. In monosyllabic words, such as ANN, the initial vowel was extended to accommodate the LH tones; The L edge tone then aligned to the right edge of the focus word. In bi-syllabic sequences, the LH aligned to the stressed syllable, and the L edge tone aligned to the right edge of the unstressed syllable.

Typically, after the LHL pattern on focused elements, the pitch range for the remainder of the phrase shifted downward slightly. This allowed for two levels of stress to be distinguished during questions as well. Sentential level stress was heard at a higher pitch range; it was also indicated with a LH pattern on the stressed syllable. Lexical stress was indicated with a L tone on the stressed syllable of multisyllabic words; monosyllabic words that were not under contrastive
focus were not treated as prominent; nor were L tones typically mapped to these words (unless the monosyllabic word occurred in sentence-final position).

Figure 6.19 gives an overview of 10 statement and question utterances produced by one of the Mandarin-English Pattern A speakers. The unique patterns found in Mandarin-English Pattern A utterances, in comparison to monolingual English utterances are that a L boundary tone occurs on the right edges of focus words in statements and questions; two levels of pitch are marked through tone, sentential and lexical, and the final boundary tone in questions is LH rather than H as seen in English. The combination of the L boundary tone on the edge of words and the LH boundary tone on the edge of question phrases, causing lowering of the pitch range in the post-focus region in Mandarin-English Pattern A examples.
Short Phrases:

<table>
<thead>
<tr>
<th>Statement 1</th>
<th>Statement 2</th>
<th>Statement 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann won</td>
<td>Manny won</td>
<td>Ann marred</td>
</tr>
<tr>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 1</th>
<th>Question 2</th>
<th>Question 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann won</td>
<td>Manny won</td>
<td>Ann married</td>
</tr>
<tr>
<td>L</td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
</tbody>
</table>

Long Phrases:

<table>
<thead>
<tr>
<th>Statement 4</th>
<th>Statement 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manny will marry</td>
<td>Ann will need to win</td>
</tr>
<tr>
<td>H</td>
<td>H L H L</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 4</th>
<th>Question 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manny will marry</td>
<td>Ann will need to win</td>
</tr>
<tr>
<td>L H L L H</td>
<td>L H L H</td>
</tr>
</tbody>
</table>

Figure 6.19. Mandarin-English Pattern A intonation contours in statements and questions.

130
6.2.2.3. Assessing the tonal association scale

The tonal association scale predicted that one of two patterns would emerge in Mandarin-English speech; the first pattern would mark sentential-level stress by associating a tone to every accented syllable. The second pattern would mark sentential and lexical level stress by associating a tone to every stressed syllable. The results of Pattern A speakers’ utterances were not predicted via the tonal association scale. While participants in Pattern A did mark lexical and sentential prominence via tone, there was a focus on prominence relations that was not predicted. Lexical prominence was marked through tone only if a word was multisyllabic, and prominence relations thus needed to be distinguished. For example, in the word *married*, the initial syllable is stressed while the ultimate syllable is not. To mark the prominence relation between the first and last syllable, a H tone associated to the stressed syllable. In monosyllabic words, such as *need*, there were no other syllables with which to compare prominence, and thus no H tone was inserted on monosyllabic words to indicate lexical stress. Further discussion of this point is taken up in Section 6.3.1.

6.2.3. Mandarin-English B

The remaining 9 Mandarin-English participants produced intonation contours on English statements and questions using a slightly different pattern than those in Mandarin-English Pattern A. The speakers who fell in this group tended to have lower ratings from native speakers in terms of their degree of foreign accent; this indicates that these speakers may have been slightly more advanced in their English pronunciation skills.
6.2.3.1. Statements

Statements in Mandarin-English Pattern B followed the same general patterns as those just described in Mandarin-English Pattern A, and so the discussion of these phrases will be kept relatively short. Just as Mandarin-English Pattern A speakers produced a H tone on the accented syllable in statements, and a L tone on the right edge of the focus word, so did Mandarin-English Pattern B speakers. This can be seen in Figure 6.20. This pattern was treated as a H*-L#, just as was the case for Mandarin-English Pattern A.

![Figure 6.20. Mandarin-English Pattern B H tone next to a L boundary tone on the edge of the focus word.](image) 

Also similar to what was found for Mandarin-English Pattern A statements, the alignment of the H tone to the accented syllable was varied: in some sequences, the F0 peak occurred well within the syllable, while in others it occurred at the right edge of the accented syllable. Of the 43 statements which fell into the Mandarin-English Group B phonological pattern, 30 aligned the H early in the syllable, and 13 aligned the H to the right edge of the initial syllable.

The Mandarin-English Pattern A and Pattern B statements diverged with respect to the way tone was used to mark levels of prominence. While Pattern A speakers marked two levels of
stress via tone, Pattern B speakers marked only one: sentential-level stress. This was evident in that bi-syllabic words in sententially unstressed positions, such as marry in the phrase MANNY will marry, did not bear tone; rather a flat plateau near the bottom of the pitch range followed these words as seen in Figure 6.21.

![MANNY will marry.](image)

*Figure 6.21. Mandarin-English Pattern B production of the phrase “Manny will marry.” No H tone is inserted on the stressed syllable of post focus words.*

### 2.2.3.2. Questions

Questions in Mandarin-English Pattern B were a mix of what was seen in the monolingual English and Mandarin-English Pattern A variations. Focused items surfaced with a LH contour as was seen in monolingual English; however, in Mandarin-English Pattern B, the L always aligned within the accented syllable, and the H aligned to the right edge of the focus word. The H was thus treated as a boundary tone on the edge of the word (H#).
In short questions, the need for high pitch range to signal sentential-level stress disappeared once speakers began using tones to contrast sentential prominence relations instead of word-level and sentential relations as seen in Figure 6.22.

![ANNE won?](image)

*Figure 6.22. Mandarin-English Pattern B production of the short question “Ann won?” The post-focus region has a higher pitch range than the focus word.*

In long questions, a F0 plateau was heard near the center of the pitch range. A LH ending was also heard in Mandarin-English Pattern B questions. Just as was the case in Pattern A, the L occurred on the stressed syllable of the final word, and the H at the edge of the phrase (as was seen in Figure 6.18.

An overview of Mandarin-English Pattern B is given in Figure 6.23. Note that the Pattern B speakers, like Pattern A speakers, also inserted tones at the right edge of focus words. The tones, however, differed slightly in Mandarin-English Pattern B. In statements, the boundary tone was L, while in questions, it was H. The right edge of questions in Mandarin-English Pattern B, like in A, also used a complex LH tone. The L surfaced on the stressed syllable of
multisyllabic words; however, in statements, no tones occurred on the stressed syllables of multisyllabic words.

Finally, note that the difference in boundary tones on the edges of words between Mandarin-English Pattern A and Pattern B affected the post-focus pitch range. In Mandarin-English Pattern B questions, the post-focus pitch range was near the middle of the pitch range in long questions, and then ended near the top of the pitch range; statements, on the other hand, ended in a L F0 plateau.
Short Utterances:

<table>
<thead>
<tr>
<th>Statement 1</th>
<th>Statement 2</th>
<th>Statement 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann</td>
<td>&lt;2&gt;</td>
<td>won</td>
</tr>
<tr>
<td>L</td>
<td>L</td>
<td>H</td>
</tr>
</tbody>
</table>

Question 1

<table>
<thead>
<tr>
<th>Question 1</th>
<th>Question 2</th>
<th>Question 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann</td>
<td>won</td>
<td>Manny</td>
</tr>
<tr>
<td>L</td>
<td>L</td>
<td>H</td>
</tr>
</tbody>
</table>

Long Phrases:

<table>
<thead>
<tr>
<th>Statement 4</th>
<th>Statement 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manny</td>
<td>will</td>
</tr>
<tr>
<td>M</td>
<td>H</td>
</tr>
</tbody>
</table>

Question 4

<table>
<thead>
<tr>
<th>Question 4</th>
<th>Question 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manny</td>
<td>will</td>
</tr>
<tr>
<td>M</td>
<td>L</td>
</tr>
</tbody>
</table>

Figure 6.23. Mandarin-English Pattern B intonation contours in statements and questions.
6.2.3.3. Assessment of the tonal association scale

The pattern found in Mandarin-English Pattern B statements and questions were predicted by the tonal association scale in that a tone did map to every accented syllable in the phrase. This is further discussed in Section 6.3.

6.3. Discussion

6.3.1. Review of findings

Chapter 6 explained and discussed three variations in the surface intonation patterns of monolingual English and Mandarin-English yes/no statements and questions with initial contrastive focus. First, the patterns of monolingual English were described. Monolingual English speakers used early-aligned complex pitch accents (H*L and L*H) on the accented syllable to mark sentential-level stress; statements ended in a H boundary tone (H%) at the edge of the phrase, and questions began with a L boundary tone (%L) and ended in a H boundary tone (H%). Mandarin-English Pattern A speakers used a H tone to mark two levels of stress, sentential and lexical; the insertion of a H tone, however, only occurred on the stressed syllables of multisyllabic words which were not under contrastive focus. In addition, differentiating the focus word from the rest of the phrase, a L boundary tone (L#) occurred at the right edge of the focus item. Finally, statements ended in a L boundary tone (L%) and questions began with a L boundary tone (%L) and ended in a LH boundary tone (LH%).

The last pattern described was that of Mandarin-English Pattern B speakers. Pattern B speakers marked sentential-level stress through inserting a H tone on the accented syllable in statements and a L tone on the accented syllable in questions. All focus words ended with a boundary tone; in statements, the tone was L (L#) while in questions it was H (H#). Finally,
statements ended in a L boundary tone (L%) and questions with a LH boundary tone (LH%). The L autotone from the complex LH% aligned to the stressed syllable of the final word in each question example.

Three major points from these data should be noted. First, the tonal association scale was only partially upheld; Mandarin-English Pattern A inserted tones on stressed syllables of multisyllabic words. This was not an anticipated result following the tonal association scale. Second, these data have shown that boundary tones can occur at the edges of words in addition to larger phrases. Third, Mandarin-English intonation contours did not appear to be influenced at the phonological level by the Mandarin lexical tone system; however, it is possible that the unique patterns that occurred in Mandarin-English may be related to patterns of Mandarin intonation. Since research on Mandarin intonation from an AM perspective is so limited, however, it is difficult to assess whether this is the case. Each point is discussed in turn.

6.3.2. The predictions of the tonal association scale

To begin, it was pointed out that the hypotheses of the tonal association scale tested in this study were only partially upheld. Using the tonal association scale, it was hypothesized that speakers would assign tones to prominent syllables using one of two patterns; either they would assign a tone to every accented syllable or to every stressed syllable. While all speakers placed a tone on the accented syllable in each phrase, no pattern emerged where a speaker produced every stressed syllable with a tone. This point is especially apparent in the longer phrases, such as Ann will need to win. The word need was never pronounced with a tone, although it was always produced with a full vowel in each example, suggesting that it was stressed.
While no pattern emerged where all stressed syllables carried tone, speakers who used Pattern A did, however, produce tones on stressed syllables in multisyllabic words, even when these syllables were not accented. Compare statements 3 and 4 in Figure 6.19, for example. H occupies lexically strong syllables in words whose syllables have a strong-weak relation, such as marry. The speakers who used Pattern A did not mark all stressed syllables; rather, they used tone to distinguish prominence relations. That is to say, strong-weak relationships were differentiated through tone; monosyllabic words, such as need, had no weak counterpart, and so tone was not used. This outcome is not predicted by the tonal association scale.

One possible solution may be to add an additional constraint into the hierarchy that requires tones to associate to prominent syllables in strong-weak relations. Such a constraint is defined in 1.

(1) $\sigma'_{[\text{rel}]} \rightarrow T$:

A tone must associate to the stressed syllable of multisyllabic words.

If $\sigma'_{[\text{rel}]} \rightarrow T$ outranked $\sigma' \rightarrow T$ as shown in 2, we could get the correct result to emerge for Mandarin-English.
This would not be without consequence to our overall understanding of prosodic systems. The insertion of such a constraint would suggest that another domain to which tonal assignment is sensitive is present and active in languages. This domain would be somewhere between the foot and phrase, and interestingly, be sensitive to relational comparisons.

6.3.3. Boundary tones on the edges of words

In addition to adaptations to the tonal association scale, the data in this study has shown a need for constraints on boundary tones as well. Gussenhovern (2004) also makes a similar point. In the three patterns found in the English data, boundary tones surfaced on the beginning and ending of utterances as well as on the right edges of focus words. If boundary tones occur on the right edges of words in some dialects, presumably they too can occur on the left edges of words. These prosodic positions must also be captured in our constraint hierarchy.

When writing constraints about how tones associate to structural positions within an utterance, the same logic used to help us make typological predictions about syllable structure,
may provide a good starting point for learning about boundary tones cross-linguistically. Syllable structure is captured via two markedness constraints: Onset and NoCoda (Prince & Smolensky, 1993). The positive onset constraint captures that fact that the least marked syllables include consonants in the onset position, while the negative NoCoda constraint captures that the coda position is more marked typologically.

In monolingual English and Mandarin-English we saw that all phrases ended in a tone at the right edge of a phrase, and so this may be a less-marked position. A constraint such as FinalBoundary (T%) defined in 8 may therefore be needed.

(8) T%

A tone must associate to the right edge of the phrase.

On the opposing end, we saw that a boundary tone on the left edge of the intonational phrase was possible, although less common. Perhaps this is our less-marked position and should be captured by the constraint defined in 9.

(9) No%T:

A tone cannot associate to the left edge of the phrase.

To describe the different patterns, I will use T% to show a final boundary tone on the edge of an intonational phrase, and %T to signify the initial boundary tone on the left edge of an intonational phrase.
If the markedness constraints (T% and No%T) outrank the faithfulness constraints (Max [T] and Dep [T]), we predict a typological system that allows only final boundary tones: T%.

This is shown in Table 6.3.

Table 6.3. OT tableau showing a marked %T-T* input will emerge as an unmarked T*-T% syllable under the ranking markedness >> faithfulness.

<table>
<thead>
<tr>
<th>/%T-T*/</th>
<th>T%</th>
<th>No%T</th>
<th>Dep</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>→T*-T%</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>%T-T*</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>%T-T*-T%</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>T*</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
</tr>
</tbody>
</table>

If, however, the faithfulness constraints outrank the markedness constraints, we predict an intonational system that allows four possible patterns of boundary tones on intonational phrases to emerge: T*-T%, T*, %T-T*, %T-T*-T% as shown in Table 6.4.
Table 6.4. Four OT tableaux which demonstrate $T^*-T\%$, $T^*$, $%T-T^*$, $%T-T^*-T\%$ patterns for boundary tones on edges of intonational phrases will emerge when faithfulness $>>$ markedness.

If we mix the rankings of the faithfulness and markedness constraints, we predict two additional systems to emerge typologically. One allows $T^*-T\%$ and $T^*$ and the other allows $T^*-T\%$ and $%T-T^*-T\%$. These systems are characterized in Tables 6.5 and 6.6.
Table 6.5. OT tableaux which demonstrate $T^*-T\%$ or $T^*$ phrases emerge from ranking $\text{Dep-}[T] >> T\%$ and $\text{No}\%T >> \text{Max}[T]$.

<table>
<thead>
<tr>
<th>/\T^*/</th>
<th>Dep</th>
<th>T%</th>
<th>No%T</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T^*-T%$</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rightarrow T^*$</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/\T^*-T%/</th>
<th>Dep</th>
<th>T%</th>
<th>No%T</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T^*-T%$</td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>$\rightarrow T^*$</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/%T-T^*-T%/</th>
<th>Dep</th>
<th>T%</th>
<th>No%T</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rightarrow T^*-T%$</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>%T-T^*-T%</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6.6. OT tableaux which demonstrate $T^*-T\%$ or $%T^*-T\%$ phrases emerge from the ranking $\text{Dep-IO} >> \text{Onset and NoCoda} >> \text{Max-IO}$.

<table>
<thead>
<tr>
<th>$T^*$</th>
<th>Max</th>
<th>NoT</th>
<th>T%</th>
<th>Dep</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rightarrow T^*-T%$</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>$T^*$</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$%T^*-T%$</th>
<th>Max</th>
<th>NoT</th>
<th>T%</th>
<th>Dep</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rightarrow %T^*-T%$</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>$T^*-T%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rightarrow T^*-T%$</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Crucially, we rule out systems that permit initial boundary tones but not final boundary tones. This prediction holds up across all of the languages represented in Jun (2005). Still, further research focused on boundary tones in cross-linguistic data should continue to be conducted in order to confirm whether this prediction is correct. Use of the constraints T% and NoT, however, may help us to direct research on intonation systems in the future.

While the constraints T% and NoT may help us to focus our research efforts on boundary tones at phrase edges in order to better understand typologies of tone and intonation, they alone will not capture the full range of structural typologies. These constraints only make predictions about the edges of the whole phrase; however, the phrase breaks down into smaller components, and these too, presumably, have positive and negative constraints. Is there a set of constraints, for example that require a final boundary tone to surface on focus words, and prohibit a tone from occurring at its left edge? While this may be a possibility, there simply is
not enough evidence in this study to merit the proposal of such constraints. More work needs to be done to find out more about boundary tones on word edges cross-linguistically.

6.3.4. Does Mandarin-English intonation resemble Mandarin lexical tone?

The final notable finding in this study is that Mandarin-English intonation seems to be affected by Mandarin lexical tone at the phonetic level; however, this is not the case at the phonological level. In Chapters 5 and 6, for example, it was seen that Mandarin-English speakers use a varied F0 alignment pattern that draws from both the L1 and L2. At the phonological level, however, less similarity between the Mandarin lexical tone system and the Mandarin-English intonation system was found.

For example, one unique finding for Mandarin-English was that a boundary tone occurred on the right edge of the focus word in statements and questions. This pattern cannot be attributed to the native English pattern since boundary tones do not occur on edges of words in the dialect observed (or in other native dialects for that matter). Moreover, the lexical tone system in Mandarin does not insert tones on the edges of words. Whether this boundary tone pattern came from Mandarin intonation, however, is unclear since there is so little work done on Mandarin intonation from an AM perspective. There is some evidence, though, that may suggest the Mandarin-English pattern had borrowed from Mandarin. Recall that in Mandarin-English Pattern A, speakers marked two levels of stress through tone, sentential and lexical. Due to a L boundary tone inserted on the right edge of all focus words in statements and questions, sentential-level stress was realized at an overall higher pitch range than post-focus words. A similar description of the Mandarin intonation system is made by Xu who notes that focus items in Mandarin statements and questions are heard at a higher pitch range, and with pitch expansion
(1999). It is possible that these similarities arise because both Mandarin and Mandarin-English require a tone on the edge of every focus word. The data in this study, however, cannot confirm whether or not this is the case. More work investigating Mandarin intonation from an AM perspective, however, is needed.

Another finding in the Mandarin-English data was that relational stress was marked via tone. In Mandarin, almost every syllable is stressed, and almost every syllable bears tone. And so the application of lexical tone is not sensitive to relational stress in this language. The tonal pattern that emerged in Mandarin-English Pattern A cannot therefore, be attributed directly to the Mandarin lexical tone system. Mandarin speakers have no L1 experience in marking lexical stress, and it may be that they are using, at the lexical level, the pattern that English speakers use at the sentential level. They don’t insert extra tones on monosyllables because there is no need to mark one syllable as more prominent than another. When they realize a need to mark relational stress, marking one syllable in the word as more prominent, they may appropriate a pitch accent to mark the stressed syllables.

Another surprising finding in the Mandarin-English data was that the right edge of questions ended in a LH boundary tone rather than a simple H tone as was seen in monolingual English. The Mandarin speakers correctly associated rising contours with questions– this is generally taught in English classes. However, when the focused item came early in the phrase, as it was in these data, the Mandarin-English speakers had trouble creating a smooth contour across the whole phrase, and instead they tended to treat the intonation on the focused item and the phrase-level intonation separately. They inserted a particular tonal pattern on the focused word, and then inserted the rise at the end of the phrase. The independence of the two
specifications is particularly obvious in ling utterances where the pattern on the focus word and the pattern at the end of the phrase were separated by a long mid-level plateau. This is, in effect, a limited effect of Mandarin phonology on Mandarin-English. Since Mandarin speakers have not practiced producing tonal patterns that stretch out across a phrase, and instead are accustomed to having a very dense tonal pattern, they insert tones piece by piece rather than tying them all together into one whole contour.

The last notable finding in the Mandarin-English data was the two alternate realizations for the L boundary tone on the edge of a word. After late-aligned H tones, a short pause occurred, while after early-aligned H tones on monosyllabic words, a fully-pronounced L boundary tone emerged. There is no current research on Mandarin tone that suggests L tones can have alternate realizations in Mandarin. Again, however, the research on Mandarin intonation from an AM perspective is quite limited. Certainly, due to the rich lexical tone system, if boundary tones occurred on edges of some words within a Mandarin phrase, it is likely that there would not be room left over on the segmental material for the tone to be realized. It is possible, then, that this alternate realization of the L boundary tone in Mandarin-English stemmed from the Mandarin language. Further research on Mandarin intonation, however would be needed to confirm whether this is the case. If boundary tones do occur on the edges of focus words in Mandarin, one place where the tone could be realized on the segmental material would be in focus items where the second syllable is unstressed and does not bear lexical tone. For future studies on Mandarin intonation then, it may be worth investigating whether a L tone occurs on the right edge of words such as malma when placed in contrastive focus position.
Chapter 7: Pitch Register in Mandarin-English Statements and Yes/No Questions

Chapter 7 presents results on post-focus pitch register in monolingual English and Mandarin-English yes/no statements and questions. Section 7.1 discusses findings on post-focus pitch register between corresponding statements and questions. Section 7.2 turns to the relationship between post-focus pitch register and the tonal patterns reported in Chapter 6. Finally, implications of the results for models of tone and intonation are discussed in Section 7.3.

7.1. Pitch register

To begin analysis of post-focus pitch register in English yes/no statements and questions, graphs of corresponding statements and questions were made by overlaying time-aligned pitch contours for statements and their question counterparts on top of one another. Since it was thought that differences between statement and question post-focus register would be easier to view in longer phrases, only utterances that were five syllables long were analyzed. This made a total of sixty-four graphs of corresponding statements and questions ([16 Mandarin-English speakers x 2 statement/question pairs] + [6 English speakers x 2 statement/question pairs]) to be analyzed.

Graphs were analyzed visually first to determine whether a difference in pitch register emerged between corresponding statements and questions. Three patterns for post-focus pitch register emerged from this.

7.1.1. Monolingual English

All 32 corresponding statements and questions produced by monolingual speakers of English fell into one pattern by where a low plateau occurred in the post-focus region of statements, and a continuously rising F0 contour occurred during the post-focus region of
questions. Statements and questions were thus recognized with a low post-focus pitch range in statements and a high one in questions as seen in Figure 7.1.

Figure 7.1. A divergence in pitch range emerges between English statements and questions as produced by monolingual speakers of English.

7.1.2. Mandarin-English low pitch register in statements and questions

The second pattern that emerged included no difference in post-focus pitch register between statements and questions as seen in Figure 7.2 and Figure 7.3. This was evident in the substantial overlap between the post-focus region in statements and questions. Fourteen of the Mandarin-English corresponding statement and question pairs fell into this pattern.
To confirm the visual assessment, a repeated-measures T-test was used to analyze the difference between the average F0 of plateaus in statements versus questions. The post-focus F0 plateau was not significantly different in the statements and questions $t(14) = -1.01, p < 0.01$. The F0 of post-focus plateaus in statements and questions overlapped.

7.1.3. Mandarin-English low pitch register in statements and high pitch register in questions

The final pattern included 18 Mandarin-English corresponding statement and question pairs. In this pattern, questions had a high post-focus pitch register while statements had a low one. This pattern can be seen in Figure 7.4 and Figure 7.5.
To confirm the visual assessment, a repeated-measures T-test was used to analyze the difference between the average F0 plateau in the statements versus questions. The post-focus F0 plateau was significantly different in the statements and questions $t(18) = -4.21, p < 0.01$. Statements had lower F0 during the post-focus plateaus than did questions.

7.2. Register and tonal pattern
Next, to determine whether post-focus pitch register was related to a particular tonal pattern, post-focus pitch register was compared to Mandarin-English tonal pattern found in Chapter 6 (Mandarin-English Pattern A and Mandarin-English Pattern B). Recall that Mandarin-English Pattern B speakers ended the focus word with opposite values of tone; that is, in statements, the focus word, or the word following, ended in a L tone, while in questions, they ended in a high tone. In Mandarin-English Pattern A, however, all focus words ended in a L boundary tone.

Again, only long utterances were used in this analysis since it was easy to compare the post-focus pitch register in corresponding statements and questions when more syllables followed the focus word. The variables compared are shown in Table 7.1.

<table>
<thead>
<tr>
<th></th>
<th>Difference in pitch register</th>
<th>Overlap in pitch register</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandarin-English Pattern A</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Mandarin-English Pattern B</td>
<td>18</td>
<td>0</td>
</tr>
</tbody>
</table>

Note. Table compares number of corresponding statement/question pairs that fall within each group.

In the corresponding statements and questions produced by Mandarin-English Pattern B speakers, a difference in post-focus register occurred 100% of the time. No difference in register, however, occurred between corresponding statements and questions as produced by those Mandarin-English speakers in Pattern A.

Looking closely at the tonal patterns found in Chapter 6, it could be seen that Mandarin-English Pattern B speakers did not end focus items with similar tones; however, Mandarin-English Pattern A speakers inserted a L boundary tone at the right edge of focus words in both
statements and questions. Incidentally, the presence of this L boundary tone on the right edge of contrastive focus items seemed to be the key pattern that could predict whether a difference between statement and question register would emerge. The L boundary tone appeared to cause the post-focus pitch range to move downward in such a way that the post-focus pitch register overlapped in both statements and questions. These findings support Hypothesis 5 which states that pitch register will be predictable from tonal pattern. A discussion of what this implies for tone-sequence models of tone and intonation follows.

7.3. Summary and discussion

The major finding in Chapter 7 is that a significant relationship between post-focus pitch register and tonal pattern occurred. But what caused this relationship? One possibility is that some Mandarin-English speakers differentiated statements from questions using a high and low pitch register, while others did not. Moreover, the transfer of pitch register to differentiate statements from questions in English coincidentally coincided with the tonal pattern a Mandarin-English speaker used.

The more likely possibility is that some speakers produced a tonal sequence that included a L boundary tone on the right edge of all contrastive focus words. When articulated in fluid speech, this L tone, in addition to the anticipation of the upcoming L tone, created a low F0 plateau, heard as a low pitch register, in both statements and questions. On the other hand, some speakers produced a L tone near the edge of focus words in statements, while questions had a H tone near the edge of the focus word. In these examples, the opposing tonal values at the edges of post-focus words created a divergence in pitch range that, when realized in quick speech was heard as mid or low F0 plateaus, or high and low pitch register. Statements were heard with a
lower overall pitch in the post-focus region, while statements were heard with a higher one. The post-focus “pitch register,” however, was simply an incidental bi-product of the tonal pattern; pitch register in and of itself was not the actual phonological pattern.

So what does this mean for tone-sequence models? How should we account for pitch register in statements and questions? I believe the answer is that once the tonal sequences are correctly accounted for, the facts about pitch register will come for free. While an account of the Mandarin intonation system is beyond the scope of this dissertation, what this study can offer is evidence that, register is derived from the tonal pattern, not orthogonal to it. If in English this relationship occurs, there is a chance that the same could hold true in other languages, such as Mandarin.
CHAPTER 8: CONCLUSION

8.1. Review of the study

Through examination of Mandarin-English tone and intonation, this study answered five research questions:

(1) Do speech categories interact in bilingual speech?
(2) Do pitch and segments interact similarly in bilingual speech?
(3) Does L1 tone interact with L2 intonation in Mandarin-English speech?
(4) Does the tonal association scale predict Mandarin-English tonal patterns?
(5) Is register predictable from tonal pattern in Mandarin-English?

Answers to these questions, and their implications for models of tone and intonation follow.

8.1.1. How do speech categories interact in bilingual speech?

Chapter 5 described the results of a phonetic study which examined phonetic values heard in segments versus those in pitch contours to see if bilinguals produced phonetic values intermediate to monolingual values. The SLM hypotheses were not confirmed. Rather, bilingual speakers produced phonetic values of phonologically identical L1/L2 categories similarly to speakers of the language for which they were speaking. A bi-modal distribution, however, was found in bilingual speech; it revealed that while speakers generally produced phonetic values that were within the norms of monolingual values for the language they were speaking, sometimes phonetic values entered in that fell into the range of values normal for monolinguals of the opposing language. It is suggested that a better model of bilingual category interaction would thus treat phonological categories as having different phonetic variants. These variants are
selected for actual speech according to the phonological environment first, and then the social environment second.

8.1.2. Do pitch and segments interact similarly in bilingual speech?

The results in Chapter 5 not only confirmed that the SLM made the wrong predictions about the way bilingual categories interact, but it also revealed a similar pattern between the way segmental categories interact in bilingual speech, and the way that pitch contours interact. In both cases, bilingual speakers produce L1 and L2 phonetic variants of phonologically identical categories in both languages that they speak; however, when speaking the L1, more L1-like phonetic values are produced and when speaking the L2, more L2-like phonetic values emerge. It is therefore suggested that a model that represents segmental category interaction in bilingual speech should work dually for the interaction of intonation contours in bilingual speech.

8.1.3. Does L1 tone interact with L2 intonation in Mandarin-English speech?

Chapter 5 also discussed the results about the interaction of tone and intonation at the phonetic level in Mandarin-English speech after examining phonetic peak alignment to the syllable in English and Mandarin falling contours heard over multisyllabic-words with initial stress (mama). Monolingual English speakers were found to consistently align the F0 peak well within the initial syllable, whereas monolingual speakers of Mandarin aligned the peak to the right edge of the syllable. Bilingual speakers produced phonetic alignment values with a bi-modal distribution in both languages. Most of the phonetic values fell within the norms of monolingual speakers for the language the bilinguals were speaking; however, some values which fell within norms of the opposing language were also found. This happened in both their L1 (Mandarin) and their L2 (English). It is suggested that this finding supports a model of
phonology that represents tone and intonation along the same level of linguistic structure since similar phonological contours in the L1 and L2 interact despite one contour being used for lexical purposes in the L1 and the other for discourse-level purposes in the L2.

Chapter six assessed the interaction of Mandarin tone and Mandarin-English intonation at the phonological level. Even though interaction was found at the phonetic level, there were no direct phonological transfers from Mandarin tone into Mandarin-English intonation. It was suggested that the more surprising tonal patterns that were found in Mandarin-English intonation may have stemmed from Mandarin intonation; however, since so little research has been conducted on the Mandarin intonation system from an AM perspective, whether this is true or not remains unclear.

8.1.4. Does the tonal association scale predict Mandarin-English tonal patterns?

Chapter 6 discusses the predictions of the tonal association scale in comparison to the surface contours that emerged in monolingual English and Mandarin-English statements and questions. Three intonation patterns were found for mapping tones to the prosodic material in English statements and questions. In the first pattern, the monolingual English pattern, tones associated to the accented syllable, the syllable following the accented one, and the phrase edge in all utterances; in questions, an initial boundary tone at the left edge of the phrase was also found. In the second pattern, Mandarin-English Pattern A, tones associated to the accented syllable, the right edge of the focus word, the stressed syllable of multisyllabic words, and the phrase edge in all utterances; in questions a boundary tone was also heard on the left edge of the phrase. Finally, the third pattern, Mandarin-English Pattern B, was realized with tones associated to the same general positions as were heard in Pattern A; however, a H tone aligned to the right
edge of focus words in questions, while a L tone aligned to the right edge of the focus word in statements. These typologies were not predicted by the tonal association scale since no pattern emerged where a tone associated to every stressed syllable in a phrase, and since a pattern emerged by where stressed syllables of multisyllabic words were realized with tone.

It is suggested that the tonal association scale, as written, may be too simplistic to capture the set of facts which are needed to describe typologies of tone and intonation systems. It is suggested that a set of constraints that map tones to prosodic material using positive and negative constraints may be a more fruitful way to begin making better typological predictions about the types of tonal systems which are out there. However, this will be a long-term project that will take the minds of many researchers looking at a large variety of languages.

8.1.5. Is register predictable from tonal pattern in Mandarin-English?

Finally, Chapter 7 provides the results about the relationship between pitch register in statements and questions and their tonal patterns. A relationship between post-focus register and tonal pattern is found. Those patterns which use a L boundary tone at the edge of focus items in statements and questions do not have a difference in register emerge, while those which have opposing tonal values during or after the focus word get a difference in sentence-type register. It is suggested that register may be an incidental consequence emerging from the way tones associate to the prosodic material. When the proper tonal patterns are found for a particular system, the facts about register should therefore come for free. No added structural level is needed to account for register.
8.2. Future research on Mandarin pitch register from a tone-sequence framework

This dissertation has focused narrowly on Mandarin-English in order to explore the linguistic representation and structure of tone and intonation. Let us not forget, however, that the original question leading up to this study dealt with a divide between linguists and their choice of model for representing tone and intonation. While most linguists use a tone-sequence model for the representation of tone and intonation, some, particularly those who focus on East-Asian tonal languages, such as Mandarin, use contour interaction models since the facts of the languages which they examine have proven to be difficult to account for from a tone-sequence framework. A case in point is the Mandarin Chinese pitch register. In statements the post-focus pitch register is realized at the bottom of the pitch range, while in questions, it is realized near the top of the pitch range. This system allows for all lexical tones to be realized, and at the same time, the sentence-type register can be deduced. How can such a system be represented within a tone sequence model?

The insights gained from examining Mandarin-English may open a new path for studying intonation, and in particular pitch register, in Mandarin. Mandarin is a tonal language by where lexical tones are heard on almost every syllable across a phrase. The presence of lexical tone makes it very difficult to examine intonation simply because the tonal tier is so dense. This study, however, has given us insights toward two areas that may transpose into the study of the Mandarin tonal system as well: (a) post-focus pitch register is related to tonal pattern, and (b) where most Mandarin-English focus words had a L boundary tone at the edge of word in statements, late-aligned H tones were not followed by a L tone in monosyllabic words; rather, a disjuncture was heard. From these facts, it was argued that pitch register is an incidental
bi-product of tonal pattern, and that L boundary tones on the edges of monosyllabic focus words are left to float when they are preceded by a late-aligned H tone since, the alignment of the H tone left no segmental material on which the L boundary tone could be realized.

These facts may help us toward research in Mandarin. If in Mandarin, a different in post-focus pitch register occurs between statements and questions, it is possible that it too stems from a particular tonal pattern that occurs on the edges of focus words. Moreover, since in Mandarin we know that lexical tones fill the segmental material in most syllables, we might expect that a boundary tone on the edges of contrastive words would not be able to associate to the segmental material in most cases. If these boundary tones do exist, is there anywhere they would likely surface? Words that end with a toneless syllable may provide the perfect context for analyzing whether boundary tones do occur in these positions. In particular, since reduplicant forms, such as ma1ma1 (1 stands for the H tone in this example) surface without lexical tone on the second syllable, these words placed in contrastive focus may surface with boundary tones on their right edges. If so, we will have evidence that boundary tones exist on edges of some words in Mandarin; moreover, these tones, if present, may be the cause of the low and high post-focus pitch register which emerges in statements and questions of Mandarin.
Appendix A

Consent for perception study only:

Consent to Participate in Research

Discourse-level Meaning and Phonetic Restructuring

You are being asked to participate in a research study. The purpose of this research is to better understand the relationship between Mandarin and English pitch when spoken by Mandarin-English bilingual speakers. Your participation in this study will take approximately 15 minutes. The following outlines the procedures of this study:

1. You will be given a questionnaire to fill out about your language history. The questionnaire will be conducted in English.
2. You will be led through a transcription task where you will be asked to write down words that you hear using the Standard Chinese Pinyin System.
3. You will be led through a perception task where you will be asked to decide whether words you hear are similar or not.

By participating in this study, you will be helping linguistic researchers to better understand similarities and differences between Mandarin, English, and Mandarin-English pitch. This is beneficial information for not only researchers, but also language teachers who may be able to use this information to improve English pronunciation courses.

There are no known risks to participating in this study. The only personal information that will be collected is contact information and background information about your foreign language studies. No personally-identifiable information will be made public. Data files will not contain any personally identifiable information. Only the researcher will have access to the data, which will be kept in a password-protected computer.

You may contact Kimberly Teague at XXX-XXX-XXXX any time you have questions about the research, or your rights as a research participant.

Your participation in this research is voluntary, and you will not be penalized or lose benefits if you refuse to participate or decide to stop.
Consent for full experiment:

Consent to Participate in Research

Discourse-level Meaning and Phonetic Restructuring

You are being asked to participate in a research study. The purpose of this research is to better understand the relationship between Mandarin and English pitch when spoken by Mandarin-English bilingual speakers. Your participation in this study will take approximately 40 minutes. The following outlines the procedures of this study:

1. You will be given a questionnaire to fill out about your language history. The questionnaire will be conducted in English.
2. You will be led through a transcription task where you will be asked to write down words that you hear using the Standard Chinese Pinyin System.
3. You will be led through a perception task where you will be asked to decide whether words you hear are similar or not.
4. You will be given a set of index cards with Mandarin and English lists on them. You may read through the cards silently first.
5. You will be led through a speaking task where you will be recorded reading the lists.
6. You will be given a worksheet with Mandarin and English sentences on them. You may read through it silently.
7. You will be led through a listening task where you will be asked to listen to a recorded utterance, and then respond to a question; your response will be recorded.
8. You will underline the word that differs between the listening and recording task.
9. You will listen to all of your recorded sentences and tell the researcher whether they are questions or statements; you’ll also explain if a word was emphasized.

By participating in this study, you will be helping linguistic researchers to better understand similarities and differences between Mandarin, English, and Mandarin-English pitch. This is beneficial information for not only researchers, but also language teachers who may be able to use this information to improve English pronunciation courses.

There are no known risks to participating in this study. The only personal information that will be collected is contact information and background information about your foreign language studies. No personally-identifiable information will be made public. Data files will not contain any personally identifiable information. Only the researcher will have access to the data, which will be kept in a password-protected computer.

You may contact Kimberly Teague at XXX-XXX-XXXX any time you have questions about the research, or your rights as a research participant.

Your participation in this research is voluntary, and you will not be penalized or lose benefits if you refuse to participate or decide to stop.
Appendix B

Participant Questionnaire

Section A

1. Age: _______

2. Where are you from (include city and country)? ____________________________

3. Have you lived anywhere besides Beijing or Washington, DC? If so, please specify where and how long you lived there. __________________________________________

4. Do you speak or understand any Chinese dialects other than standard Mandarin? ________________

5. How many years did you study English before moving to the USA? _________________________

6. Were your teachers in China native speakers of English? _______________________________

7. What dialect of English did you formally study (British/American/other)? _______________

8. How many years have you lived in the USA? ______

9. Current occupation: ___________________________

10. Do you speak any languages other than Mandarin or English? If so, please list your speaking level in each of them (beginning, intermediate, advanced). We are not interested in your reading and writing level, only your speaking level. ________________________________

Section B

Directions: Write the number that best corresponds to how often you do the following:

0 = never 1 = rarely 2 = sometimes 3 = often 4 = always

CHINESE

1. I search Chinese websites on the internet.

2. I talk on the phone in Chinese.

3. I watch TV or movies in Chinese.

4. I speak to my roommates (or family if you live with your family) in Chinese.

5. I go out with Chinese people during the weekends and nights.

6. I use online video conferencing or Skype in Chinese.

7. I speak Chinese during my breaks at school or work.
8. I do my shopping with Chinese people at Chinese stores.


**ENGLISH**

10. I search English websites on the internet.

11. I talk on the phone in English.

12. I watch TV or movies in English.

13. I speak to my roommates (or family if you live with your family) in English.

14. I go out with American people during the weekends and nights.

15. I use online video conferencing or skype in English.

16. I speak English during my breaks at school or work.

17. I do my shopping with English people.

18. I listen to English music.

19. When I speak, native speakers misinterpret me because of
   a. My grammar __________
   b. My Vocabulary __________
   c. My pronunciation of consonants and vowels __________
   d. My intonation __________

20. When I speak, I
   a. Avoid certain words __________
   b. Avoid using special intonation __________
   c. Avoid certain grammatical structures __________

**Section C**

Directions: Circle the answer that you think is most appropriate.

1. My English pronunciation is
   a. Generally clear
   b. Sometimes Clear
   c. Not clear at all
2. When I speak with native speakers of English I am
   a. Easily understood
   b. Sometimes difficult to understand
   c. Very difficult to understand

3. I speak English at a pace that is
   a. Close to American English
   b. Sometimes too fast
   c. Sometimes too slow

4. I feel I can use intonation to convey emotions
   a. Similar to how they are expressed in English
   b. Sometimes differently than in English
   c. Very differently from American English

5. I can pronounce
   a. Most English sounds without difficulty
   b. Many English sounds, but have difficulty with a few sounds
   c. Some English sounds, but have difficulty with most

6. My English pitch (intonation) is
   a. Close to American English
   b. Sometimes different from American English
   c. Very different from American English

7. When I speak English, I feel that
   a. Grammar is my weakest point
   b. Vocabulary is my weakest point
   c. Pronunciation of consonants and vowels is my weakest point
   d. Intonation is my weakest point

Appendix C

List of utterances used for Degree of Foreign Accent Rating Task

1. Mama, Momo, Nana, Nini
2. Manny will marry?
3. Anne will need to win.
Appendix D

List of items transcribed in Transcription Task

Example:
1. talta1

Trial:
1. Tommy

Task Items:

1. Mama
2. Mommy
3. Mimi
4. Nana
5. Nina
6. Momo
7. Papa
8. Poppy
Appendix E

Transcription Task

Directions: You will hear a total of eight words. Each word will be played two times. Write each word using the Standard Chinese Pinyin system. The words are not necessarily real words in Mandarin!

Example One: 
If you hear [listen for example 1]
You write  
ta1.ta1

Trial One:  

Trial 1.  ________________

Example Two:  
If you hear [listen for example 2]
You write  
tao1.tao1

Trial Two:  

Trial 2.  ________________

Task Items:  

1.  ________________

2.  ________________

3.  ________________

4.  ________________

5.  ________________

6.  ________________

7.  ________________

8.  ________________
Appendix F

List of Similarity Judgment Triads:

Example 1:

1. ta2ta2     ta2ta2     ta1ta1

Trial 1:

1. ta1mi5     Tommy     ta1mi3

Example 2:

2. tao2tao2   tao2tao2   tao1tao1

Trial 2:

2. ma1mi5     Mommy     ma1mi3

Task Items:

1. ma1ma5     Mama     ma1ma3
2. ma1mi5     Mommy     ma1mi3
3. mi1mi3     Mimi     mi1mi5
4. na1na3     Nana     na1na5
5. ni1na5     Nina     ni1na3
6. mo1mo3     Momo     mo1mo5
7. pa1pa5     Papa     ba1ba5
8. ba1bi3     Poppy     ba1bi3
Appendix G

Similarity Judgment Task

Directions: You will hear three words in a row. Decide which word, the first or the last, is most similar to the word in the middle. Circle “1” if the first word is more similar; circle “3” if the last word is more similar.

Example One:

You hear [listen for example 1]

You circle “1.”

Example. 1  _____  3

Trial One:

Trial. 1  _____  3

Example Two:

You hear [listen for example 1]

You circle “1.”

Example. 1  _____  3

Trial Two:

Trial. 1  _____  3

Task Items:

1. 1  _____  3

2. 1  _____  3

3. 1  _____  3

4. 1  _____  3

5. 1  _____  3

6. 1  _____  3
Appendix H

English lists:

1. Mimi, Mama, Momo, and Nana
2. Mama, Nana, Mimi, and Momo
3. Momo, Mimi, Mama, and Nana
4. Nana, Momo, Mama, and Mimi
5. Papa, Poppy, Mama, and Mimi
6. Poppy, Mama, Mimi and Papa

Mandarin lists:

1. mi1mi5, ma1ma5, mo1mo5, and na1na5
2. ma1ma5, na1na5, mi1mi5, and ni1ni5
3. ni1ni5, mi1mi5, ma1ma5, and na1na5
4. na1na5, mo1mo5, ma1ma5, mi1mi5
5. pa1pa5, pa1pi5, ma1ma5, mi1mi5
6. pa1pi5, ma1ma5, mi1mi5, pa1pa5
Appendix I

Elicited Utterances:

1. Ann won.
2. Manny won.
3. Anam won.
5. Ann.
6. Manny will marry.
7. Ann will need to win.

1. Ann won?
2. Manny won?
3. Anam won?
4. Ann married?
5. Ann?
6. Manny will marry?
7. Ann will need to win?
Appendix J

List Task

Directions: In front of you, you will see a stack of cards. Choose one card from the pile and read it to yourself silently. Next, you will hear a girl talking. She will tell you the order that her list is on, and then ask you a question. Respond to her question using the list on your card.

Example One:

You choose:
Molly, Manny, Peter, and Polly

You hear:
I got the list that says “Manny, Molly, Polly, and Peter.” Which order is your list in?

You say:
Mine says “Molly, Manny, Peter, and Polly.”

Trial One:

Sally, Peter, Anna, and Susan

Answer to Trial One:

You hear:
I got the list that says “Peter, Sally, Anna, and Susan.” Which order is your list in?

You say:
Mine says “Peter, Sally, Anna, and Susan.”

Example Two:

You choose:
Molly, Peter, Manny, and Polly

You hear:
I got the list that says “Manny, Molly, Polly, and Peter.” Which order is your list in?

You say:
Mine says “Molly, Peter, Manny, and Polly.”

**Trial One:**

Sally, Susan, Anna, and Peter

**Answer to Trial One:**

You hear:

I got the list that says “Peter, Sally, Anna, and Susan.” Which order is your list in?

You say:

Mine says “Sally, Susan, Anna, and Peter.”

---

Appendix K  

**Correction by Statement Task**

Directions: You will read a series of seven sentences. Each sentence represents the “true” scenario. You will hear a girl describe the situation; however, she will say something different from the “true” scenario. First, correct her sentence, and then underline the word that was different.

**Example One:**

You read:

Mary worried Sammy.

You hear:

Tommy worried Sammy.

You say:

MARY worried Sammy.

You underline:

Mary worried Sammy.

**Trial One:**

Lisa is helping Susan.
Answer to Trial One:

You read:
Lisa is helping Susan.

You hear:
John is helping Susan.

You say:
LISA is helping Susan.

You underline:
Lisa is helping Susan.

Example Two:

You read:
Sammy married Anam.

You hear:
Carl married Anam.

You say:
SAMMY married Anam.

You underline:
Sammy married Anam.

Trial Two:

Mary is taking Minnie.

Answer to Trial One:

You read:
Mary is taking Minnie.

You hear:
John is taking Minnie.

You say:
MARY is taking Minnie.

You underline:
Mary is taking Minnie.
Asking for Clarification Task

Directions: You will read a series of seven sentences. Each sentence represents the “true” scenario. You will hear a girl describe the situation; however, she will say something different from the “true” scenario. First, ask whether she meant to say something else, and then wait for her response. Then, underline the word that was different.

Example One:

You read: Mary worried Sammy.

You hear: June worried Sammy.

You say: JUNE worried Sammy?

You hear: Sorry. MARY worried Sammy.

You underline: Mary worried Sammy.

Trial One:

Lilly is helping Susan.

Answer to Trial One:

You read: Lilly is helping Susan.

You hear: Matt is helping Susan.

You say: MATT is helping Susan?

You hear: Sorry. LILLY is helping Susan.

You underline: Lilly is helping Susan.
Example Two:

You read:
Minnie married Carl.

You hear:
May married Carl.

You say:
MAY married Carl?

You hear:
Sorry. MINNIE married Carl.

You underline:
Minnie married Carl.

Trial Two:

Jill is taking Susan.

Answer to Trial One:

You read:
Jill is taking Susan.

You hear:
Matt is taking Susan.

You say:
MATT is taking Susan?

You hear:
Sorry. JILL is taking Susan.

You underline:
Jill is taking Susan.
Appendix M

List of recorded sentences

Each sentence will be read as both a statement and a question.

**English sentences:**
1. Anne won
2. Manny won
3. Anam won
4. Anne married
5. Anne
6. Manny will marry
7. Anne will need to win

**Mandarin sentences:**
1. 妈妈说叫
   ma1 ma5 shuo1 “jiao4.”
2. 妈妈说玩
   ma1 ma5 shuo1 “wan2.”
3. 妈妈说丢碗
   ma1 ma5 shuo1 “diu wan3.”
4. 妈妈说打猫
   ma1 ma5 shuo1 “da3 mao1.”


Clements (1979)


