FUNDAMENTAL FREQUENCY IN MANDARIN CHINESE AND ENGLISH: IMPLICATIONS FOR SECOND-LANGUAGE SPEAKERS

by

TANYA LYNN VIGER

A dissertation submitted to the Graduate Faculty in Linguistics in partial fulfillment of
the requirements for the degree of Doctor of Philosophy,
The City University of New York

2007

UMI Number: 3283624

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the
Graduate Faculty in Linguistics in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Professor Gita Martohardjono

Date

Chair of Examining Committee

Professor Gita Martohardjono

Date

Executive Officer

Professor Winifred Strange

Professor Robert Vago

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

FUNDAMENTAL FREQUENCY IN MANDARIN CHINESE AND ENGLISH:
IMPLICATIONS FOR SECOND-LANGUAGE SPEAKERS

BY

TANYA LYNN VIGER

Adviser: Professor Gita Martohardjono

This dissertation presents the results of a production experiment designed to investigate differences between the English and Mandarin utterance-level prosodic contours produced by native and non-native speakers. Materials were controlled across languages for sentence length, sentence-level focus, and illocutionary type. A bi-directional design permitted a double comparison of L1 English with L2 English by Mandarin speakers, and of L1 Mandarin with L2 Mandarin by English speakers. The particular focus of the study was on global pitch setting and pitch contours at sentence boundaries. Of interest was whether L2 speakers would appropriately differentiate sentences of three illocutionary types: statements, questions, and *surprise* echo questions, which are typically prosodically disambiguated in both English and Mandarin, Native language data for both English and Mandarin confirmed some widely noted phenomena—the final F0 rise in English yes/no and echo questions, and a global raising of pitch throughout Mandarin yes/no and echo questions. L2 production data for both groups of L2 learners exhibit a striking absence of central aspects of the target language’s utterance-level prosody. In addition, transfer of global L1 prosodic contours into the L2 was entirely absent; there was no sign of transfer of global prosodic strategies from Mandarin into English, or from

English into Mandarin. Some transfer effects were observed, however, with regard to local phenomena within the final syllable. L1-Mandarin speakers compressed illocution-related pitch excursions in English into the final syllable only, in the fashion of pitch excursions that realize Mandarin lexical tones. L1-English speakers imported a final-syllable rise on the final syllable in Mandarin questions. Both participant groups exhibited a considerably narrower F0 range overall in their L2 than in their L1. A range of possible explanations is offered for why transfer was not more extensive and why in particular it was confined to utterance-final position. A tentative conclusion is that local and global prosodic phenomena may differ fundamentally in their linguistic, processing, and neural bases.

Acknowledgements

Completion of this dissertation would not have been possible without the wisdom, support and encouragement of many people. First, I would like to thank the members of my committee for their invaluable comments on the manuscript. The insight and perspective on all matters linguistic and practical given so generously to me by Professor Gita Martohardjono have helped me since the day I started out in the Linguistics Ph.D. program. I owe a debt of gratitude to Professor Winifred Strange, both for her consistent encouragement of my research and for having taught me most of what I know about speech acoustics and experimental design. Professor Robert Vago has also provided valuable comments to help link the experimental results to phonological theory.

This dissertation research was partially funded by the generous support of the National Science Foundation, Division of Behavioral and Cognitive Sciences, Dissertation Improvement Grant #0314356. Part of this dissertation was also published as an article written with Professor Janet Dean Fodor: "Fundamental frequency in Mandarin and English: Comparing first- and second-language speakers", which appeared in Interfaces in Multilingualism (2006) John Benjamins. The author especially wishes to thank the editor of that volume, Professor Conxita Lleó, and two anonymous reviewers for their insightful comments and advice. This dissertation also benefited from audience discussion at the conference "Interaction of Language Components in Bilingual Acquisition" at the University of Hamburg Collaborative Research Center on Multilingualism, April 2005. Sincere thanks also go to Professors Janet Dean Fodor and

Dianne Bradley for their invaluable assistance in shaping the ideas and analyzing the data presented both in the book chapter and in this dissertation.

I would also like to take this opportunity to thank the Graduate Institute of Linguistics at National Taiwan University for hosting me as a visiting scholar from 2002 to 2004, which provided me with the opportunity and laboratory resources to conduct this research. Special thanks are due to Professors Kathleen Ahrens and Wen-yu Chiang for agreeing to sponsor me as a Visiting Scholar. Professor Wen-yu Chiang contributed valuable input throughout this research, particularly in the design of the Mandarin experimental materials. Chiu-yu Tseng of Academia Sinica provided encouragement and countless insightful comments on Mandarin prosody and data analysis. Other Taiwanese colleagues who have made vital and much appreciated contributions to the development of Mandarin experimental materials and to data segmentation and analysis include You-ni Wu, Yun-hsiang Lai and Yun-hsia Bai.

My family, friends and colleagues in New York City and in Taiwan have enriched my life and work immeasurably. Special thanks go to my mother, Mariann Milner, and to my good friends Gina Cogan, Marcia McGibbon, Harriet Taber, Maureen Kealy, Jennifer Jolliffe, Angela Utschig and Cynthia Khatib. Without them, I would not have survived this process. This dissertation is dedicated to the memory of my grandfather, Joseph Rocco Visceglia, whose kindness and unwavering faith in me have sustained me throughout my life.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.0 Summary

Studies of the ways in which knowledge of a first-language (L1) sound system shapes the development of second-language (L2) phonology have primarily focused on learners' abilities to perceive and produce single-sound contrasts in the L2 (Flege, 1995). However, recent accent-rating studies have found that prosody also makes a significant contribution to the perception of a non-native accent in L2 learners' speech. (Anderson-Hsieh & Koehler, 1988; Munro, 1995; Tajima, Port, & Dalby, 1997). Nevertheless, fewer studies have been done on the specific ways in which L2 prosody differs acoustically from that of native-speakers. Chapter 2 of this dissertation will provide a summary of current research investigating characteristics of L2 phonology and situate the current study within that research.

1.1 Statement of the Problem

Adult L2 learners rarely attain native-like competence, however great their motivation and efforts to learn. This is particularly true for the acquisition of L2 phonology. Non-native accents persist even for speakers with advanced knowledge of vocabulary and syntax. For the learner, this can be both a source of frustration and a barrier to intelligibility. More research is needed to investigate the specific acoustic properties contributing to accents that are perceptibly non-native so that instructional materials and methods can be developed to address these barriers to intelligibility.

1.2 Purpose of the Study

The purpose of this study is to make a bi-directional comparison in production of L2 English by Mandarin speakers and production of L2 Mandarin by English speakers. Its particular focus is on production of global pitch setting and pitch contours at phrase boundaries for three different sentence types. This research provides data on the development of second-language prosodic competence. Though the data represent only two languages, they are of interest because the prosodic systems of Mandarin Chinese and English differ substantially, and thus offer a particularly clear view of transfer phenomena or other discrepancies between L1 and L2 performance. This study pairs a tone language with a non-tone language, which provides insight into cross-linguistic differences in lexical, sentential, and discourse-related pitch phenomena for both native and non-native speakers.

1.3 Significance of the Study

No previous research has paired a tone language, such as Mandarin, with a non-tone language, such as English, for an acoustic study of the specific ways in which cross-linguistic differences in use of pitch affect production of prosody in a second language. The present study constitutes a novel investigation of the interaction of two distinct prosodic systems in second-language processing. Descriptive data on specific acoustic properties of L2 prosody will increase understanding of the ways in which pitch (fundamental frequency or f0) is processed and categorized in speech. These data can also be used to improve language learning materials and software by adding a more detailed account of prosody to current pronunciation-shaping paradigms. These data could also be used to develop a perception study to investigate which aspects of L2 prosody most significantly contribute to its perception as non-native. As will be shown, some aspects of prosody present more difficulty for L2 learners than others do. These may require particular attention in second-language instruction, which is not represented in current materials. Relevant differences in the prosodic systems of English and Mandarin are summarized below.

1.4 The Prosodic Systems of Mandarin and English

There are many regional varieties of Mandarin, which to a greater or lesser extent differ from each other lexically, syntactically, and phonologically, as well as prosodically. Thus, the prosodic features of Taiwan Mandarin (Guoyu), the variety spoken by the L1 Mandarin speakers in this study, are not identical to those of other varieties, such as Beijing Mandarin. (Tseng [2004] provides a comprehensive discussion of prosodic differences between Beijing and Taiwan Mandarin). However, the following description of the features of Mandarin prosody that are relevant to this experiment also holds true for Guoyu.

An important prosodic feature of Mandarin is that it has lexical tone, which means that the pitch contour associated with a syllable can distinguish lexical meaning. Mandarin has four tones: Tone 1 is high and level; Tone 2 is high and rising; Tone 3 falls, then rises; and Tone 4 starts high, and then falls low. For instance, the syllable *ma* produced with a high, level tone means *mother*; with a high rising tone it means *hemp*; with a falling-rising tone *horse*; and with a high, falling tone it means *to scold* (Chao, 1968).

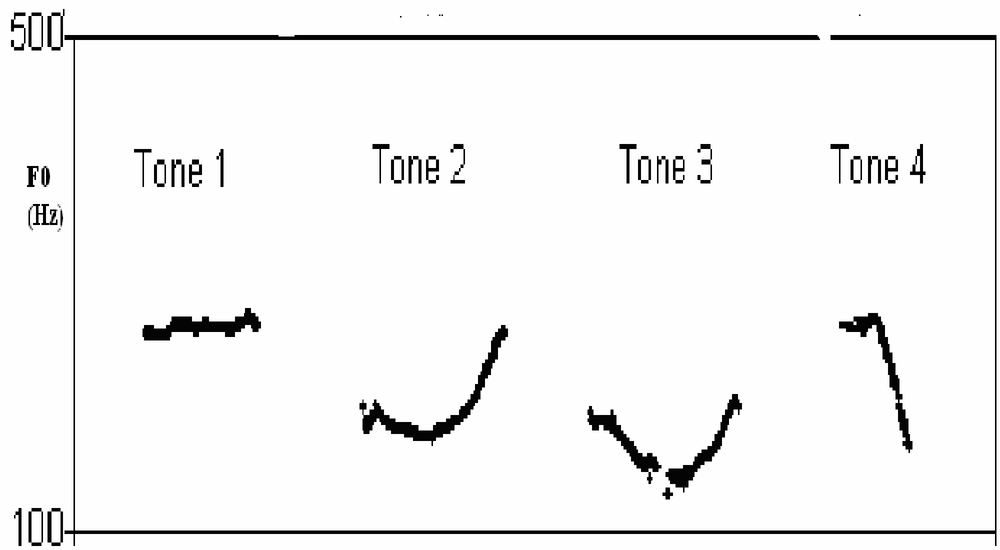


Figure 1: The four Mandarin tones, all produced on the carrier syllable “ma”, spoken by a female native speaker of Guoyu, recorded by the first author.

Tone languages may also use pitch contours at the level of phrases and sentences, to cue the presence of a phrase boundary, to express narrow focus, or to convey utterance-level illocutionary force such as the distinction between a statement and a question. Languages without lexical tone, such as English, use pitch contours primarily or exclusively for this purpose, and not for lexical identification. Pitch contour is therefore relevant to more levels of linguistic representation in Mandarin than it is in English.

According to the Beckman & Pierrehumbert (1986), framework of intonational description, the typical declarative utterance in English is composed of one or more intonation phrases, each with one or more high level pitch accents associated to prominent words. At the end of a sentence, pitch falls to a low final boundary tone. A yes/no or echo question utterance in English typically has one or more low level pitch accents associated with prominent words, and its pitch rises to a high final boundary tone (Pierrehumbert & Hirschberg, 1990). Thus, native English speakers employ pitch accents and boundary tones according to illocutionary utterance-level intent (e.g., question versus statement). Most notably, a low boundary tone occurs at the right edge of a declarative contour, and a high boundary tone occurs at the right edge of an interrogative contour.

English has also been found to exhibit an early high pitch setting (as well as overall pitch expansion) in production of *surprise-redundancy* contours, as reflected in echo questions whose intention is to elicit confirmation (Beckman & Ayers-Elam, 1997). Pitch variation in English also relates to information structure, such as marking topic/focus or new/old information at the discourse level. High utterance-initial pitch setting is associated with *topic jump* (Hirschberg & Pierrehumbert, 1986), i.e., the introduction of a new topic into the discourse.

Mandarin also uses F0 to signal illocutionary force. However, questions and declaratives are distinguished primarily by the global pitch range throughout the utterance, rather than by specific utterance-final boundary tones. Yes/no questions are produced in an overall higher pitch range than statements (Shen, 1990). Echo questions exhibit a higher initial pitch setting than statements do (Chang, 1998; Tseng, 2003) and are produced with an overall expanded pitch range (Peng, et al., 1999). Since these differences between illocutionary types in Mandarin do not involve substantial changes in the lexically-determined pitch contours within an utterance, they largely (though not fully) preserve the pitch variations that realize the lexical tones.

1.5 Summary of Design

This bi-directional production experiment was designed to investigate the differences in the English and Mandarin utterance-and discourse-level prosodic contours produced by native and non-native speakers. Materials were controlled across languages for sentence length, sentence-level focus, and illocutionary type. The bi-directional design permitted a double comparison of L1 English with L2 English by native Mandarin speakers, and of L1 Mandarin with L2 Mandarin by native English speakers.

The particular focus of the study was on global pitch setting and on pitch contours at sentence boundaries. Of interest was whether L2 speakers would appropriately differentiate sentences of three illocutionary types embedded in a mini-discourse: statements, questions, and *surprise* echo questions, which are typically prosodically disambiguated in both English and Mandarin, though in different ways, as noted above. Though these data represent only two languages, they are of more general value because the prosodic systems of Mandarin Chinese and English differ so radically. The comparison between them can provide insight into how L2 learners adjust to a substantial shift in the linguistic functions of prosodic contours from the L1 to the L2.

1.6 Hypotheses

Previous comparisons of L1 and L2 discourse-level prosody in English have shown that L2 speakers demonstrate sporadic use of prosodic markers related to discourse structure, including high pitch at phrase boundaries to link related constituents and the paratone—an expansion of pitch range to signal topic shift (Wennerstrom, 1994; 1998). Moreover, L2 speakers tend to divide utterances into shorter intonational phrases than native speakers do (Hewings, 1995). Gandour, et al. (2003) found that language-specific features of local and global prosody are processed differently by native and non-native listeners. The findings briefly summarized by these researchers suggest that local and global prosody are processed differently by L1 and L2 speakers. Based on these results, the investigator of the present study predicted that deviations from native-speaker norms in L2 will follow from differences in length of the temporal window required to realize f0 contours and the level of representation (lexical, utterance, discourse) that the prosodic event is used to mark. That is to say, L2 speakers have a more limited window of prosodic planning than native speakers do, and will thus be less likely to successfully produce prosodic events that require long-range planning than those that can be realized over a shorter temporal window (i.e. one syllable). From these general predictions follow the specific predictions enumerated below:

1. L2 speakers will set global pitch in L2 at the same height for all sentences in the mini-discourse, whereas in L1, they will use global pitch setting to differentiate the various discourse positions of the three utterance types (statements, questions, and *surprise* echo questions).

2. L2 speakers will confine illocutionary prosody (e.g., declarative falls and interrogative rises in English) to utterance-final syllables. In their production of L1, in contrast, they will demonstrate evidence of longer-range planning to signal differences in illocutionary type.

1.7 Subject Populations

The participants in this study were all adult L2 learners of English or Mandarin, with high intermediate to advanced skills in the L2). To ensure relative homogeneity within subject groups, the native speakers of Mandarin were all Taiwanese and the L1 English speakers were all North American. The North American English speakers had no knowledge of any lexical tone language other Mandarin. For the L1 Mandarin population, it would have been ideal to work with speakers who spoke and understood no Chinese dialect other than Mandarin, but this was not a practical possibility. However, participants were excluded if they had knowledge of Chinese dialects other than Southern Min, the dialect most widely represented in Taiwan. Thus, all L1 Mandarin participants spoke the same variety of Mandarin, all were Mandarin dominant, and all had knowledge of the same local dialect. Though dialectal or ideolectical differences across participants undoubtedly existed in both L1 groups, each group was considered homogenous for the purposes of this study.

In addition, all participants in this study were female. Male and female speakers often differ substantially with respect to their range of fundamental frequency, and it seemed likely that in group data, this might obscure the pitch phenomena of interest. Female speakers were chosen based on the greater ease of pitch tracking represented by female speakers' F0 range. It seems likely, but is not certain, that the findings can be extended to male speakers.

This research was conducted at the Graduate Institute of Linguistics of National Taiwan University. For this reason, there exists a natural asymmetry between the two participant groups; i.e., the L1 English speakers were living in L2 territory, while the L1 Mandarin speakers were not. Thus, the possibility cannot be unequivocally ruled out that Taiwan Mandarin prosodic patterns had a greater impact on the Mandarin pronunciation of the L1 English speakers than North American English patterns had on the pronunciation of English by the L1 Mandarin speakers. For the most part, L2 English had been acquired in Taiwan under the supervision of non-native teachers, whereas the L1 English speakers had received all of their Mandarin instruction from native speakers.

1.8 Overview of the Dissertation

Chapter 2 provides a review of previous research related to L1 and L2 processing of local and global prosodic events, acquisition studies of L2 prosody and a review of models of L2 phonology which might be used to interpret the results of this study. It also includes a review of the range of studies demonstrating that non-native f0 patterns contribute to native speakers' perception of non-native accents. Chapter 3 provides a detailed account of experimental data collection and analysis methods. A detailed presentation of results, including statistical analyses, is given in Chapter 4, and Chapter 5 discusses the implications of the results, limitations of the study, and possible directions for further research.

Chapter 2

Literature Review

2.0 Introduction

This chapter provides a review of previous research comparing L1 and L2 processing of local and global prosodic events and summarizes relevant acquisition studies of L2 prosody as well as studies that demonstrate the contribution of non-native prosody to native speakers' perception of non-native accents in L2 speech. It includes a discussion as to whether current models of L2 phonology may be used to interpret results of studies related to utterance and discourse level prosody.

2.1 Lateralization studies of L1 and L2 prosodic processing

Recent studies have reported differences in the neurological resources underlying L1 and L2 processing, which may predict possible discrepancies in L1 and L2 attainment of local versus global prosody. Poeppel (2003) reviews evidence from a variety of sources indicating that there are two different time frames that characterize prosodic phenomena, and that they exhibit different lateralization patterns: left-hemisphere auditory areas preferentially extract information from short (approximately 20-40-minute) temporal windows, while those in the right hemisphere preferentially extract information from longer (approximately 150-250-minute) temporal windows. Gandour, et al. (2003) found that "the pitch contours associated with Mandarin tones are processed in the left hemisphere (LH) by Chinese listeners, whereas pitch contours associated with intonation are processed bilaterally with a right hemisphere (RH) preference. English listeners, on the other hand, exhibited bilateral activation for lexical tones, [and] right-sided activation for intonation." They interpret these data to suggest that temporal integration windows are linked to language-specific features of local and global prosody, which are processed differently by native and naïve listeners (i.e., listeners who had no knowledge of Chinese). Moreover, behavioral and neuroimaging research demonstrates that native Chinese speakers process Mandarin tones as a linguistic property predominantly in the left hemisphere (Hsieh, Gandour, Wong, & Hutchins, 2001; Klein, Zatorre, Milner, & Zhao, 2001; Wang, Jongman, & Sereno, 2001). The results of PET (positron emission tomography) studies have also shown left hemisphere frontal activation for Mandarin Chinese tones by native Chinese speakers (Hsieh, et al., 2000; Klein, et al., 2001); whereas, for speakers of a non-tone language such as English, the processing of Mandarin tone lies in the homologous right hemisphere frontal regions, which would make its processing similar to the processing of pitch as a non-linguistic stimulus. Current research in L2 prosodic processing seems to point toward independent control of pitch and lexical information, as well as independent control of global and local prosody.

2.2 Acquisition Studies

2.2.1 Acquisition of L2 English Prosody.

The majority of research on L2 learners' mastery of sentence and discourse-level prosody has focused on English prosody as it is produced by non-native speakers from different L1 backgrounds. Generally, the aim has been to identify acoustic features that differ from those of native speakers. For example, McGory (1997) studied the English prosody of Mandarin Chinese speakers' investigating the production of narrow prosodic focus in American English by beginning and advanced L1-Mandarin speakers. Target words were embedded in six different prosodic positions in sentences, which were presented as both statements and questions for speakers to read aloud. McGory's results suggest that L1 Mandarin phonology has an influence on the production of English prosodic patterns: Native Mandarin speakers of English produced a different pitch contour on the stressed syllable of different lexical target items in the same intonation condition. Native English speakers, in contrast, produced the same tonal pattern on the target words in the same intonation condition. This suggests that Mandarin speakers are associating a particular tone shape to a lexical item, rather than to an intonation condition. Mandarin speakers also produced pitch accents on target words that were in post-focus positions, whereas English speakers did not. This again suggests that for Mandarin speakers of English, pitch excursion is relevant on a lexical level, and is assigned to every stressed syllable regardless of utterance-level prominence. Wennerstrom (1998) recorded 18 L1 Mandarin speakers of L2 American English to investigate possible correlations between aspects of intonational proficiency and scores on a global-proficiency speaking test. The four prosodic variables studied were: (a) correct assignment of pitch accent to content versus function words; (b) use of high pitch at phrase boundaries to link related constituents; (c) use of pitch to create contrastive focus; and (d) use of the paratone, which is an expansion of pitch range to signal topic shift. One kind of intermediate phrase boundary that occurs in English is a high phrase accent. When this phrase accent occurs at an intermediate phrase boundary, it indicates there is still more information to come within that intonation phrase. In English, this is known as a *continuation rise*. Results indicated that only use of the paratone proved to be a significant predictor of scores on a global proficiency test. Since Wennerstrom (1998) used the SPEAK test, which employs global ratings, intonation may not have been sufficiently teased apart from other aspects of fluency, such as segmental pronunciation and word choice. Nevertheless, interesting trends of discrepancy between native and non-native intonation emerged in Wennerstrom's acoustic measurements of two other prosodic variables. The greatest difference between native (NS) and non-native speakers (NNS) was the likelihood to use pitch to create contrastive stress; NS had a mean score of 65%, versus NNS, who had a mean score of 30%. However, this variable was subject to considerable inter-speaker variation—some NNS subjects scored as high as 60%. As for utterance-medial boundaries, NS produced between 25 and 29 out of a possible 30, while NNS produced as few as 15. Wennerstrom posits possible transfer from Mandarin, since utterance-initial pitch height, not boundary tone, is a more salient cue for contrasting the illocutionary force of utterances in Mandarin.

Schack (2000) examined English production of one native female Mandarin speaker’s English intonation when producing declaratives, echo questions, and yes/no questions. Generally, the English intonation of this speaker was characterized by "a much denser assignment of tonal targets than that of a native speaker" (p. 2). Also, the speaker seemed to apply the same tonal pattern to words spoken in isolation as those embedded in larger utterances. This would be expected from a speaker whose L1 requires that a tone contour to be assigned to every stressed syllable. Schack’s stimulus materials included declaratives, echo questions, and yes/no questions in order to determine whether the speaker's choice of utterance-level tune assignment would also reflect Mandarin interference. Native English speakers produced the tones assigned to those sentence types by assigning one or more pitch accents to stressed constituents, and then interpolating pitch between the last (nuclear) accent and the boundary tone assigned to mark the phrase edge. There are two distinct tones represented in the stimulus materials: one for declaratives and one for echo and yes/no questions. These are distinguished from one another both by their choice of nuclear pitch accent and their choice of boundary tone.

The Mandarin speaker, in contrast, produced utterance-appropriate boundary tones, but the way in which she realized them suggests Mandarin interference. She confined the interrogation rise to the utterance-final syllable, regardless of the location of the phrase’s nuclear accent, the position of which was varied in the stimulus materials. In Mandarin, the domain of pitch movement is generally confined to the voiced portion of one syllable. Overall, she exhibited much greater pitch excursion within the utterance; a pitch accent seems to have been assigned to every lexically stressed syllable. This was realized as repetition of a low-high-low pitch accent pattern, which was downstepped in pitch throughout the utterance. This pattern was typical of the one she assigned to the same words uttered in isolation; she assigned a rising tone to stressed syllables and a falling tone to unstressed syllables.

2.2.2 Acquisition of L2 Mandarin Prosody

Studies of English learners' acquisition of Mandarin tone have varied widely with respect to experimental task and subject population. Moreover, few researchers have analyzed English speakers' tone production in full utterances, which would provide the most specific evidence for L2 representation of the interaction between tone and intonation in speech production. Still, the studies summarized in this section provide support for the hypothesis that f0 is processed at two distinct levels of representation for English and Mandarin speakers.

Broselow, Hurtig, and Ringen (1987) investigated the question of whether English speakers' ability to perceive Mandarin tones is related to their position in a string of syllables. After a training period, 50 English speakers with no prior knowledge of Mandarin or any other tone language were asked to identify Mandarin tones occurring in syllables presented both in isolation and in sequence. In the training period, each of the four tones was introduced using the carrier syllable *ti*. The subjects were asked to identify the tones of individually presented syllables. This was repeated with a different carrier syllable. Subjects were then asked to identify tones occurring in sequences of two and three syllables. Each tone occurred the same number of times in each position in the sequence. The results of Broselow, et al. show that in isolation, tones 1 (high level) and 4 (high falling) were the easiest to identify, which is to be expected since they are acoustically the most distinct. In sequence, however, correct identification of the fourth tone was far greater in final position than it was in initial or medial position (In final position, a pitch contour resembling the fourth tone would be appropriate in either an English declarative utterance or in a word-final syllable). However, there were cases in which the fourth tone was misidentified as the first tone; this, too, was much more likely to happen in final position. The authors claim that in these cases, English speakers were ignoring the final fall because they perceived only the high starting point of Tone 4, then parsed the fall as part of the intonation contour and, therefore, not relevant to tone identity.

Shen (1989) found similar error patterns involving American English speakers' production of Tones 1 and 4. This experiment required American English speakers who had received less than one year of Mandarin instruction to read a familiar passage aloud in Mandarin. The passages were recorded and analyzed for tonal error patterns. Errors in production of Tones 1 and 4 were found to be the most prevalent. Shen argues that both exhibit pitch contours which also have supra-segmental functions in English. The pitch pattern of Tone 1 is often used to mark a lexically stressed or phrasally-prominent syllable in English, and the pattern of Tone 4 is used to mark a word boundary if the word appears in isolation, or to mark a phrase boundary if that phrase is declarative. However, misproduction of these tones was found to be in f0 height, not f0 contour. Shen concluded that these two tones are more conducive to errors in pitch register, because English speakers will tend to use these familiar pitch contours in the same way they would to produce English intonation. Contrary to prediction, misproduction of Mandarin tones was found to be in F0 height and range, not F0 contour. English speakers produced Tone 1 at a much lower height with respect to their overall pitch range than native speakers did and Tone 4 in a much more compressed pitch range than native speakers. Interestingly, they did not use these tones in Mandarin to serve the same functions as phonetically similar tonal variations in English. They did not use Tone 1 to signal lexical or phrasal prominence in Mandarin, though this is signaled by a high level tone in English; and they did not employ Tone 4 to mark the end of a nuclear accent-final declarative sentence, though this is marked by a falling tone in English.

Many questions remain unanswered by Shen's study. It can be argued that all of the pitch contours represented by Mandarin tones have an intonational counterpart in English. The shallow rising contour of Tone is similar to that of an English interrogation or continuation rise when it is realized in phrase or utterance-final position, and the rising contour of Tone 2 is similar to the realization of English narrow focus. Why are these not marked by a compressed pitch range or register difference? Since Shen's reading passage elicited neither narrow-focused constituents nor interrogative utterances, this remains an open question.

Chen (2000) conducted the most comprehensive study to date on English speakers' L2 Mandarin production of tone in connected speech. Chen analyzed L2 production of tones at the word level and compared them to production of tones embedded in phrases. American English speakers who had received at least two years of Mandarin instruction were recorded speaking extemporaneously on their choice of subject. Utterances were cut out of the recorded signal and then randomized. These utterances were presented to native speaker expert judges, who rated them for tonal errors.

Chen (2000) found boundary-marking interference in utterance-final position. English speakers tended to end a declarative with a falling tone or a low-level tone. They were more likely to be correct when producing Tone 3 and Tone 4 in final position. Moreover, they underproduced Tone 1, which is usually overproduced, in the declarative phrase-final position. As discussed in Chapter 1, Tone 1 is the high-level tone, which is acoustically similar to the realization of a lexically stressed acoustically prominent syllable in English declaratives. Chen also found a prosodic interference effect resembling the English continuation rise in the intermediate-phrase final position, calling it the "more will follow intonation". Chen admitted that not enough tokens of that contour could be culled from spontaneous speech to generalize, but the tokens that were available exhibited the following pattern:

1. Guanyu wo muchin de qing kuang …

Target: F (Tone 4, localized on *kuang*)
Produced: R (Tone 2, localized on *kuang*)
Gloss: Regarding my mother's situation,
H* H-

English has the option of realizing a low or a high phrase accent in this position; these data suggest the presence of a continuation rise. In this case, the syllable on which a continuation rise would have been realized in English was pronounced with a rising tone, although its lexically assigned tone is the falling tone. Chen's natural speech data did not provide enough tokens of some sentence patterns to make generalizations about continuation rises and interrogatives. A follow-up study using more controlled materials could specifically elicit different illocutionary types. A closer examination of phrase-final effects could determine whether Chen's results can be generalized over utterance types and L2 speaker groups.

2.2.3 Conclusion

The studies summarized in this section provide evidence that L2 speakers often produce pitch contours in non-target ways in a second language for which f0 is relevant at a different level of representation than it is in their native language. Previous studies have varied too widely in terms of the utterance types they were able to elicit to present a comprehensive representation of emerging L2 prosodic competence for L2 English and L2 Mandarin speakers. The experiment presented in the thesis was designed to control for segmental content, sentence-level focus, and utterance type across languages, with the aim of teasing apart the representation of f0 at the phrase, sentence and discourse levels in the L2 phonological systems of English and Mandarin speakers.

2.3 Mandarin Accent in L2 English and English Accent in L2 Mandarin

2.3.0 Introduction

Three distinct factors can contribute to the perception of non-native accent: listener-related factors, medium-related factors, and speaker-related factors. Listener-related factors include the native listener's hearing ability, experience with accented speech, and expectations for the listening situation. Medium-related factors involve the ambient signal-to-noise ratio and any possible filtering or distorting effects present in the environment. Both medium and listener factors are tangential to the current study, which is a production study concerned with acoustic analysis of non-native pitch contours. Therefore, this paper focuses exclusively on studies involving the speaker-related factor of pronunciation (other speaker-related factors include grammar and usage).

The studies presented in this section provide evidence that the supra-segmental features of duration and fundamental frequency (f0) correlate with native listeners' perception of a non-native accent in Mandarin speakers' production of English. No research has yet emerged on the acoustic correlates of perceived accent in the L2 Mandarin of native English speakers, except for one experiment investigating the effect of perceptual training on Mandarin tone production, which will be summarized at the end of this section.

2.3.1 L1 Mandarin Speakers of English.

Studies have shown that perceived degree of accent does not necessarily correlate with intelligibility. Munro and Derwing (1995) claim that "as yet, there is no indication that reduction of accent necessarily entails increased intelligibility" (p. 19). This study investigated the possible correlation among perception of accentedness, intelligibility, and perceived comprehensibility in the speech of L1 Mandarin learners of English. Ten native speakers of Mandarin Chinese (all advanced ESL speakers with moderate to heavy accents) and two native speakers of Canadian English were recorded while they extemporaneously described the events depicted in a cartoon story. The passages of extemporaneous speech were cut up into utterance-sized tokens, which were then randomized and presented to native listeners. Listeners rated the speakers using three different evaluation metrics: comprehensibility (rated on a 9-point scale); intelligibility (evaluated with a word-by-word transcription of each utterance) and degree of accent (rated on a 9-point scale, 1=no foreign accent; 9=very strong foreign accent).

Results showed that speech that had been rated as heavily accented was often perfectly transcribed. This suggests that not all of the features which contribute to the perception of accent necessarily lessen intelligibility. The results of a follow-up study corroborated earlier findings (Derwing & Munro, 1997). Often heavily accented speech of non-native speakers was just as intelligible as that of native speakers.

Anderson-Hsieh, Johnson, and Koehler (1992) investigated the relationship between raters' assessment of non-native English pronunciation and their actual acoustic deviations from native speaker norms with respect to segmental properties, syllable structure, voice quality, and prosody. Sixty reading passage samples taken from speakers of eleven different L1s, including Mandarin, were rated impressionistically on overall pronunciation by three ESL teachers on a seven-point scale. Prosody was rated separately, on a seven-point impressionistic scale. The authors chose this impressionistic method on the grounds that it is difficult to quantify most prosodic errors in terms of discrete errors. The samples were then acoustically analyzed for sources of deviance from native norms. The authors performed Pearson correlations and multiple regressions in order to correlate deviance in each area of pronunciation with native listeners' ratings. The strongest relationship found was the correlation between the prosodic variable and pronunciation ratings. The errors in prosody considered by this study encompassed “deviations in patterns of stress and intonation as well as in timing, phrasing, and rhythm” (p.21).

Tajima, Port, and Dalby (1997) investigated the temporal properties contributing to Mandarin-accented English. They re-synthesized two-word utterances in Mandarin-accented English to match temporal characteristics of the same utterances recorded by native English speakers. They also temporally distorted the same utterances recorded by native English speakers to match the temporal characteristics of Mandarin-accented ones. Their listening experiment had four sets of stimuli: original Chinese (OC – no manipulation); original English (OE – no manipulation); temporally corrected Chinese (TCC – re-synthesized) and temporally distorted English (TDE, re-synthesized). Listeners were asked to identify the stimulus phrase in a forced-choice identification task with four alternatives: the target phrase plus three phonetically similar distracter phrases. Listeners were always able to distinguish re-synthesized from natural speech. Their results showed significant effects of temporal correction. Temporal corrections resulted in higher rates of identification for TCC than OC and in lower rates of identification for TDE than OE.

Intelligibility of unmodified Chinese phrases was 39%, which increased to 58% after temporal correction. Intelligibility of unmodified English stimuli declined after temporal distortion from 94% to 83%. The authors estimated the effect of temporal distortion as the difference in correct identification between OE and TDE (10.8%) and the difference in correct identification between OC and TCC (19.3%). The effect of spectral differences on intelligibility was calculated as the difference between OE and TCC (35.3%) or the difference between TDE and OC (43.8%). The authors concluded from these figures that while temporal characteristics significantly contribute to intelligibility, spectral characteristics such as vowel formant values make a larger contribution.

Munro (1995) removed the spectral characteristics of Mandarin-accented English using a low-pass filtering technique in order to determine whether untrained listeners could accurately distinguish native from non-native speech in the absence of segmental information. After low-pass filtering, considerable prosodic information remains in the signal, including fundamental frequency (f0) and duration. The second part of this study examined the acoustic differences between native and non-native speech samples in order to determine the source of prosodic deviations from native speaker utterances. Both scripted single-sentence utterances and extemporaneous narratives in English were recorded from 10 native speakers of Mandarin and 10 native speakers of Canadian English. The single sentences represented a variety of English intonation contours: two declaratives with different syntactic structures and two interrogatives, which included one wh-question and one yes/no question.

Twenty untrained native English-speaking listeners were played the filtered samples. They were then instructed to rate each sample on the likelihood that it had been produced by a native speaker of English. The sentences were rated on a four-point scale, and the narratives were rated on a nine-point scale. Overall, listeners assigned significantly higher ratings to samples produced by native English speakers than to those produced by native Mandarin speakers. The 10 listeners also agreed strongly with one another. Acoustical analyses of the samples were performed to determine which properties contributed to the perception of non-native accent. Information remaining in the low-pass filtered stimuli included: intonation (f0 contours); temporal information, such as word and vowel duration; location of stressed syllables; and overall rate of speech, which included location and duration of pauses.

Speaking rate was a source of significant difference across groups. Mean rates were determined to be 3.5 syllables per second for native English speakers and 2.4 syllables per second for native Mandarin speakers. However, the rate of individual non-native speakers' rates overlapped considerably with those of native speakers. Therefore, it was clear that differences in speaking rate could not account for rating differences entirely, and that other sources of possible deviation should be investigated.

The results of fundamental frequency (f0) measurements showed that Mandarin speakers exhibited a continuously falling pitch in the declarative sentence condition, while native English speakers exhibited a flat contour over the first two words, followed by a fall. In the interrogative condition, female Mandarin speakers exhibited a much higher final f0 rise than native English speakers. In addition, a slightly higher pitch overall was found for both male and female Mandarin speakers. These results suggest that fundamental frequency does provide a cue for non-native accent even in the absence of segmental information. Since this study tested only one token of each sentence type, more items are needed to determine whether the contours produced by Mandarin speakers are typical. Longer segments of discourse could also be investigated to determine whether the prosody related to creation of information structure would also have an effect on listeners' perception of differences between native and non-native speech. Subsequent research could also isolate f0 from other prosodic variables to determine whether it alone can create the perception of a non-native accent.

2.3.2 L1 English Learners of Mandarin

Wang, Jongman, and Sereno (2001) investigated English speakers' production of Mandarin tones. This study was designed to evaluate the effectiveness of perceptual training on American English speakers' production of Mandarin tones. Native Mandarin listeners' judgments were obtained before and after perceptual training. It was found that perceptual training significantly improved English speakers' Mandarin tones, as judged by native listeners' ratings. Since this study used tones occurring in isolated syllables as stimulus materials, and the subjects were naïve speakers with no knowledge of Mandarin or any other tone language, these results are tangential to questions regarding L2 learners' development of an L2 Mandarin phonological space.

2.3.3 Conclusion

The studies summarized offer evidence that non-native production of prosodic features such as f0 contour and syllable duration contribute to perception of non-native accent in the speech of L1 Mandarin learners of English. Further research can begin to tease apart the prosodic variables of f0, duration, and amplitude to examine their individual contributions to accented speech, in order to determine whether non-native production of supra-segmental characteristics in Mandarin contribute to native listeners' perception of an English accent in L2 Mandarin connected speech.

2.4 Models of L2 Phonological Representation

2.4.1 Segmental models

Existing models of L2 phonology developed within speech science were designed to describe the acquisition, perception, and production of individual L2 phones/phonemes. Theories such as the Perceptual Assimilation Model (Best, 1995) and Flege's (1995) Speech Learning Model represent the ways in which L2 speech sounds are filtered through native language phonemic categories in the process of forming new L2 categories. These models make predictions about the relative difficulty of distinguishing and producing L2 phonemes based on their acoustic proximity (for SLM) or gestural proximity (for PAM) to existing L1 phonemic categories. They were not designed to make predictions for the acquisition of phrase-, sentence- or discourse-level acoustic events. To date, there exists no unified model for L2 acquisition of prosodic phenomena occurring over larger windows of time.

2.4.2 UG-based hypotheses

More general L2 acquisition hypotheses could be extended to make predictions about the development of L2 prosody. These include the Full Transfer/Full Access hypothesis (Schwartz & Sprouse, 1996; applied to phonology by Archibald, 1998) and several formulations of the Partial Transfer hypotheses, such as the one formulated by Vainikka and Young-Scholten (1994). The Full Transfer/Full Access hypothesis incorporates access to Universal Grammar (UG). According to this hypothesis, L2 grammars consist of principles and parameters provided by UG, with parameters set to the values of the learner's L1 during the earliest stage of L2 acquisition; thus, the initial state of L2 grammar is the end state of L1 grammar. The process of L2 learning can be described as the learner's attempting to reset the parameters to the values appropriate for the L2. Applying this to the acquisition of prosody, the prediction would be that L2 learners in the initial state will transfer the prosodic contours of their native language to their emerging second language phonological system. According to Partial Transfer theories, L2 learners transfer only some of the properties of the L1 into the interlanguage grammar. For example, Vainikka and Young-Scholten (1994) propose that in the development of L2 syntax, learners transfer lexical but not functional categories. For the case of phrasal prosody, a partial transfer hypothesis would require that there be some principled distinction between prosodic phenomena which are transferred into the L2 and prosodic phenomena which are not. However, it must be noted that these theories were developed to model acquisition of L2 syntax, so they may not necessarily apply to phonological development.

2.4.3 Conclusion

Results of the studies presented above, as well as the results which are reported in this dissertation, suggest that it is necessary to develop new models of L2 acquisition to adequately represent the acquisition of L2 phonology at the phrase, sentence and discourse levels, since issues raised with respect to perception, production and linguistic competence are likely to differ in principled ways from those that are relevant to segmental phenomena.

Chapter 3

Method

3.0 Overview

This experiment was designed to investigate the differences between the English and Mandarin utterance-level prosodic contours produced by native and non-native speakers. The materials were controlled across languages for sentence length, sentence-level focus, and illocutionary type. The bi-directional design permitted a double comparison of L1 English with L2 English by Mandarin speakers, and of L1 Mandarin with L2 Mandarin by English speakers. Global pitch setting and pitch contours at sentence boundaries were the particular focus of the study. Of interest was whether L2 speakers would appropriately differentiate sentences of three illocutionary types (statements, yes/no questions, and 'surprise' echo questions) embedded in a mini-discourse structure; these types are typically disambiguated prosodically in both English and Mandarin, although in different ways.

3.1 Hypotheses and Predictions

In general, deviations from native-speaker norms present in L2 speech will follow from processing differences in production of global and local pitch events between L1 and L2 speakers, more specifically:

1. Based on previous research (Hewings 1995; Wennerstrom 1994), which demonstrated L2 speakers' failure to produce target-like global pitch cues related to information structure, it was predicted that both groups of L2 speakers will produce all illocutionary types within mini-discourses at approximately the same pitch height, whereas L1 speakers will use global pitch height to differentiate the three types in terms of their illocutionary type and position in the discourse.
2. Based on previous research, which demonstrated L2 speakers' tendency to confound utterance-level and lexical-level prosodic events (McGory 1997, Broselow et al. 1987), it was predicted that production of utterance-level prosodic events, such as interrogative f0 rises and declarative f0 falls, will be subject to negative L1 transfer. Specifically, L1 English speakers will fail to produce lexically-relevant pitch contours in Mandarin when they occur at phrase boundaries. They will override a phrase-final lexical tone with the pitch contour that is appropriate to the illocutionary force of that utterance in English. L1 Mandarin speakers, in contrast, will confine their utterance-level phrase-final pitch contours in English to the phrase-final syllable, because any single pitch event is restricted to the domain of the syllable in Mandarin.

3.2 Participants

The participants in this study were all adult L2 learners of English or Mandarin, with high intermediate to advanced skills in the L2. A questionnaire was administered prior to recording, which elicited information likely to correlate with spoken language fluency, including age of beginning language instruction, total number of years of instruction, instruction methods, amount of in-country exposure, and amount and context of present daily exposure. All participants had begun second language learning after puberty, had acquired most of their language skills in formal classroom settings, and were university students recruited through postings on National Taiwan University bulletin boards.

To ensure relative homogeneity within subject groups, the native speakers of Mandarin were all Taiwanese and the L1 English speakers were all North American. The North American English speakers had no knowledge of any lexical tone language other than Mandarin. For Mandarin, it would have been ideal to work with speakers who spoke and understood no Chinese dialect other than Mandarin, but this is not a practical possibility. However, participants were excluded if they had knowledge of Chinese dialects other than Mandarin and Southern Min (the dialect most widely represented in Taiwan). Thus, all L1 Mandarin participants spoke the same variety of Mandarin, all were Mandarin dominant, and all had knowledge of the same local dialect. Though dialectal or ideolectal differences across participants undoubtedly exist in both L1 groups, each group is considered homogenous for the purposes of this study.

All the participants in this study were female. Male and female speakers often differ substantially with respect to their range of fundamental frequency, and it seemed likely that in group data, this might obscure the pitch phenomena of interest. Female speakers were chosen based on the greater ease of pitch tracking represented by female speakers' F0 range. It seems most likely, but is not certain, that the findings can be extended to male speakers.

This research was conducted at the Graduate Institute of Linguistics of National Taiwan University. For this reason, there exists a natural asymmetry between the two participant groups; i.e. the L1 English speakers were living in L2 territory, while the L1 Mandarin speakers were not. Thus, the possibility cannot be ruled out that Taiwan Mandarin prosodic patterns had a greater impact on Mandarin pronunciation by the L1 English speakers than North American English patterns had on the pronunciation of English by the L1 Mandarin speakers. For the most part, L2 English had been acquired in Taiwan under the supervision of non-native teachers, whereas the L1 English speakers had received all of their Mandarin instruction from native speakers.

Recordings were made by twenty-four subjects in each experimental group. Subsequently, nine subjects were excluded from each group for the following reasons: missing data (defined as the absence of two or more experimental sentences); language-background information found not to fit inclusion criteria (e.g., exposure to dialects other than Mandarin and Southern Min, or North American English); and poor-quality recordings (e.g., with amplitude too low to extract pitch information in the relevant areas). Ultimately, the productions of fifteen L1 Mandarin and fifteen L1 English speakers were retained for analysis.

3.3 Materials

Experimental materials were designed to elicit three sentence types differing with respect to illocutionary force: declarative, yes/no question, and echo question. All sentences consisted of nine syllables in yes/no questions and eight syllables in declarative and echo questions[1]. The final four syllables of each utterance contained only sonorants, so that an uninterrupted pitch track could be extracted from them. In addition, in the Mandarin experimental sentences, the three syllables preceding the final syllable were all Tone 1 (the high level tone), in order to minimize tonal coarticulation[2] effects. For both languages, each sentence type was represented by three different sub-types whose utterance-final items varied with respect to prosodically relevant properties. In all, there were five tokens of each of the nine sub-types in each language. The Mandarin sentences ended with one of the following: (a) Tone 1, whose high, level contour in phrase-final position would resemble neither an English declarative nor an interrogative; (b) Tone 2, the rising shape of which would resemble the phrase-final contour of an English interrogative, but not a declarative; (c) Tone 4, the falling shape of which would resemble an English declarative, but not an interrogative.

The English experimental sentences ended with (a) final, (b) penultimate and (c) antepenultimate nuclear-stressed syllables, in order to investigate differences between L1 and L2 speakers in the timing of the boundary rise or fall in relation to the nuclear accent. Based on previous research, it was anticipated that the low nuclear accent preceding the F0 rise at the end of question would be produced by native English speakers on the stressed syllable of the final word, and that the rise to the final boundary tone would extend from that accent to the end of the sentence, thus spreading over the final two syllables in the case of the antepenultimate-stress items. Each set of materials included

---

[1] There was one accidental exception in the English materials, whose yes/no question consisted of ten syllables.
[2] Tonal coarticulation refers to the downstep or upstep of pitch created by the presence of falling or rising adjacent tones.

fifteen experimental items, three practice items (two overt and one hidden) and two experiment-final fatigue items.

3.4 Procedure

Participants were presented with two sets of materials to read aloud, one in English and one in Mandarin[3].For all participants, the L1 set preceded the L2 set; this was to ensure that participants were comfortable with the task before moving on to more difficult materials. The Mandarin and English sentences, though matched in some respects as noted above, differed completely in their content, in order to avoid any semantic priming from the first set to the second. Participants were given a ten-minute break between the two language sets. They were instructed to speak the materials as naturally as possible, as if they were talking to a friend. They were also instructed to repeat a sentence until they were satisfied with the accuracy and fluency of their production before they moved on to the next item[4]. The last token of each recorded sentence was chosen for analysis.

Each set of materials included fifteen experimental items, with three covert practice items at the beginning, and two experiment-final fatigue items at the end. Each experimental item consisted of a three-sentence mini-dialogue containing a sequence of: (1) a yes/no question; (2) an affirmative answer to that question; and (3) an echo question soliciting confirmation of that answer. The three sentences in each dialogue were identical in content except for the markers of illocutionary force (initial 'do-support' in the

[3] Mandarin and English experimental sentence sets appear in Appendix A
[4] Mandarin and English instruction sheets are given in Appendix B.

English yes/no question; the final question particle ma in the Mandarin yes/no question), and the linking words (e.g. yes, really) between the sentences.

Each experimental sentence appeared individually on a 4” X 7” card. In order to minimize the possibility that production of one sentence set would be influenced by the semantic content of the preceding set, a filler sentence (The next sentence set is number X/ 這是第 X組句子) was read aloud after each experimental item. The Mandarin sentences were presented to the L1 Mandarin speakers in Chinese characters only. They were presented to the L1 English speakers in Chinese characters with a pinyin runner specifying both segmental content and tone information, in order to avoid processing difficulties due to lexical unfamiliarity. For the same reason, both English and Mandarin sets of materials were accompanied by translations of the first sentence of each set.

The experimental procedure was carried out in the following order:

1. The participant read a description of the task she was about to perform, and signed a consent form, which explicitly granted permission to record the session. Note: all materials related to instruction or consent were presented in the participant’s L1.
2. Participants filled out a language background questionnaire (administered in the participant’s L1).Information solicited by this questionnaire[5] included how much time the participant had spent in an English or Mandarin speaking country. The questionnaire was intended to allow for the possibility of post-hoc comparison between aspects of linguistic experience and production of native-like prosody.
3. Participants read the experimental instructions.
4. Participants read the L1 experimental sentences aloud.

[5] Mandarin and English versions of the questionnaire have been reprinted in Appendix C.

5. Participants took a ten-minute break.
6. Participants read the L2 experimental sentences aloud.

3.5 Design

The design of this experiment is 3 X 3; within each language there are two factors, each of which has three levels. One factor is final word type. For Mandarin, the levels of this factor turn on lexical tone (Tone 1, Tone 2, and Tone 4, all words being monosyllabic); for English, the levels of this factor turn on syllabicity (one-syllable, two-syllable, three-syllable, all words being initial-syllable stressed). The second factor in both languages was the three illocutionary variants of the same five sentences, occurring in a sequence in the experimental presentation, with levels declarative, yes/no question, and echo question. Five different sentences were constructed for each of the three final word type conditions, so for each language; there are a total of the 15 sentence sequences.

In addition, there were two explicit practice items, in order to allow participants to become accustomed to the procedure, and one covert practice item at the start of the main experiment, whose purpose is to protect the first experimental item. Two fatigue items after the fifteenth experimental sentence to protect the final experimental item.

One fixed pseudo-randomization order was administered to all subjects. This order divided the fifteen experimental items into five sub-blocks in the run through the sentence sequences, each sub-block having exactly one example of each final word type. Thus, each Mandarin sub-block contained one Tone 1 sequence, one Tone 2 sequence and one Tone 4 sequence. Each English sub-block contained one one-syllable sequence, one two-syllable sequence and one three-syllable sequence. The sub-blocks were also ordered so that the same item type did not appear consecutively across blocks.

3.6 Data Analysis

Recordings were made in a soundproof booth using a head-mounted microphone and a laptop computer with SoundForge 4.5 software. Input was digitized at 22,050 Hz. with 16-bit resolution. Spectrographic representations and pitch contours for all utterances were generated using the signal processing software Speech Analyzer 2.1. All recordings were segmented into individual syllables, whose boundaries were determined according to the segmentation guidelines outlined in Olive, Greenwood, & Coleman (1992) for North American English, and Tseng and Chou (1999) for Mandarin Chinese.

Two separate sets of pitch data were extracted: an average F0 for each whole syllable; and a sequence of F0 measurements at 20 ms. intervals over the final four syllables. The syllable-average measurements allowed for observation of pitch excursion trends that spread over entire utterances, such as the slope and timing of an illocution-sensitive rise or fall. The 20-ms interval extraction allowed for a fine-grained analysis of pitch excursions within individual syllables, in the areas in which syllable-internal differences between subject groups and languages were anticipated. Mean F0 and duration were normalized across groups using a separately-recorded citation form of the last word of each experimental sentence. Normalization was performed in order to allow abstraction away from speaker- and group-specific variation in F0 register and speech rate. F0 and duration normalization constants can be found in Appendix D.

Chapter 4

Results

4.0 Introduction

After some preliminary observations on general performance, this chapter focuses on differences in production of fundamental frequency, primarily on syllable-average F0 data, with brief mention of dynamic F0 data (sampling every 20ms). Results for English are discussed first, followed by results for Mandarin, in each case beginning with the data from native speakers and then comparing their performance with that of L2 speakers. Finally, hypotheses presented in Chapter 1 are evaluated individually, and some general conclusions are derived from comparison of the two target languages and speaker groups.

4.1 General Performance Data: Speech Rate, Pauses and F0 Range

4.1.1 Speech rate

Some general performance differences were observed, including those in global speech rate. Not unexpectedly, L2 speakers spoke more slowly on average than L1 speakers of the same target language. L1 Mandarin speakers were 25% slower than native speakers in their production of L2 English materials. L1 English speakers were 28% slower in their production of Mandarin materials than native speakers were. A duration adjustment constant was applied to both native and non-native data to compensate for this difference[6].

4.1.2 Pauses

Pausing tended to be more prevalent in non-native than in native speech, which potentially complicates analysis of prosodic factors. Two kinds of pauses occurred in the data: grammatically-licensed pauses at syntactic junctures, and unlicensed pauses which appear to represent hesitations of various kinds. Examples of licensed opportunities to pause in Mandarin and English include clause or constituent boundaries. Generally, it was found that L2 speakers took advantage of licensed opportunities to pause, which native speakers rarely did, and also made a considerable number of unlicensed pauses, which native speakers never did. Comparison of the mean number of both licensed and unlicensed pauses across L2 utterances within a mini-dialogue revealed that both types of pause decreased with successive sentences in a sentence triplet. In a pair wise test comparing the first, second and third utterances of all sentence triplets, third utterances were found to exhibit significantly fewer pauses than first utterances ($F(1,14) = 33.42$, $p<.0001$)[7]. The most plausible interpretation of this is that repetition increases lexical familiarity, and that the source of unlicensed pausing was hesitation caused by lexical access difficulties and/or orthographic opacity. However, hesitation pauses should not confound the analysis of F0 height and contour, since there is evidence that no F0

---

[6] A table summarizing pitch and duration adjustment constants can be found in Appendix D.
[7] A table summarizing these results can be found in Appendix E.

resetting occurs after hesitation pauses; instead, the F0 track resumes at the ending height of the hesitation (Cruttenden, 1997).

4.1.3 F0 range

Both participant groups exhibited a considerably narrower f0 range overall in their L2 than in their L1. Native English speakers' average maximum f0 in English was 269Hz. and their average minimum was 155 Hz., a difference of 114 Hz., while in Mandarin their average maximum was 236.8 Hz. and their average minimum was 180.3 Hz., a difference of only 56.5 Hz. Similarly, native Mandarin speakers' average maximum in Mandarin was 300.1 Hz. and their average minimum was 189.7 Hz., a difference of 110.4 Hz.,while in English their average maximum was 261.8 Hz. and their average minimum was 187.5 Hz., a difference of only 74.3 Hz. The implications of an overall narrower pitch range in L2 are discussed in Chapter 5.

4.2 English materials: Fundamental frequency

Two separate sets of pitch extractions were generated: one extracted f0 at 20 ms. intervals, and the other averaged f0 over an entire syllable at each marked syllable boundary. The 20-ms interval extraction allowed for a finer-grained analysis of pitch excursion within a single syllable, while the syllable-averaged measurements allowed for observation of global pitch setting and rise/declination over the entire utterance and across utterances in each mini-dialogue. F0 values averaged at the syllable level were the dependent variable in three 2-way ANOVAs, which were performed using repeated measures on utterance types within and between participant groups.These were conducted to determine the statistical significance of between-subject differences in global and initial pitch setting of the three illocutionary types (yes/no question (Q), declarative (D), and echo question (E). Six sets of measurements were taken for each of the three conditions (final syllable stress, penultimate syllable stress and antepenultimate syllable stress):

1) pitch value at syllable 1 for Q, D, and E (QP1, DP1 and EP1)
2) pitch value averaged across syllables 1-8 for Q, D and E (QA, DA, EA)

Pitch was the dependent variable measured across group and utterance type. It was anticipated that both in initial pitch values and average pitch values, i.e. the values of: 1) QP1 - DP1; 2) EP1 - QP1; 3) EP1 - DP1; 4) QA – DA; 5) EA – QA and 6) EA – DA would be significantly different across speaker groups, because L1 speakers are predicted to distinguish illocutionary types and discourse position using utterance-initial and global pitch setting, whereas L2 speakers are not. The scores entered for analysis represented differences in f0 between utterance types within subjects. Each of the three conditions is analyzed separately in sections 4.2.1 – 4.2.3

4.2.1 Final syllable stress condition

F0 data for the English materials are presented in this section, averaged by syllable, with pitch and duration normalization across participants and tokens of the item subtypes. Figure 2a represents the three illocutionary types in English produced by native English speakers in the final stress condition. Figure 4a represents the three illocutionary types in English produced by native Mandarin speakers in the final stress condition.

Figure 2a: F0 contours for three English illocutionary types in the final stress condition, produced by native English speakers.

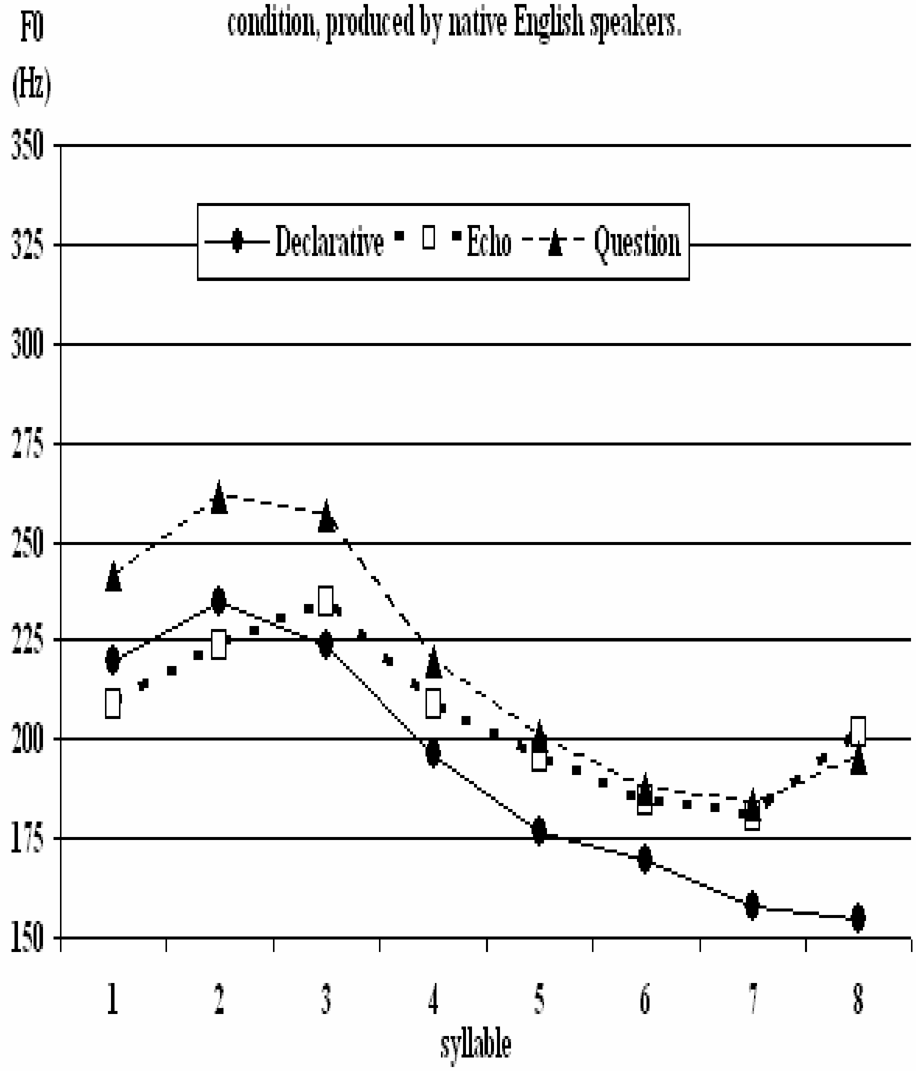


Figure 4a: F0 contours for three English illocutionary types in the final stress condition, produced by native Mandarin speakers.

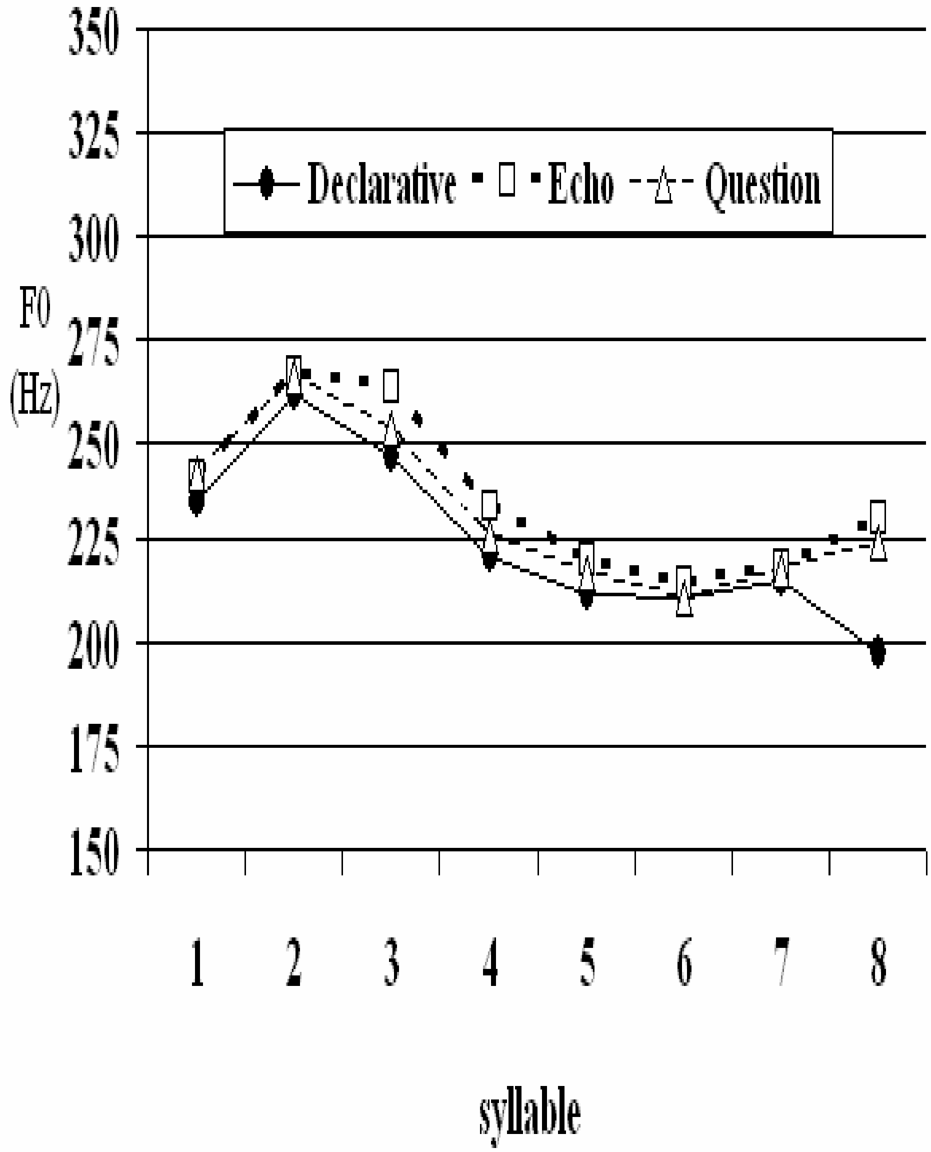


Table 1: *Comparison Of Initial Syllable Pitch Values Across Utterance Types And Speaker Groups In The Final Stress Syllable Condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

SV	SS	DF	MS	F	P Value
Between Subjects					
Group(A)	9490.746	1	9490.746	10.116	***
Error (S*A)	138849.720	148	438.174		
Within Subjects					
Pitch (B)	138719.068	1.138[8]	121886.557	77.672	***
Pitch * Group(B*A)	110977.239	1.138	97494.986	62.138	***
Error (B* sub*w*groups)	264323.426	168.466	1568.998		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated for each condition, the results of which appear in Table 2.

Table 2: *One-Way ANOVA Summary Of The Simple Main Effect Of Pitch In The Final Stress Syllable Condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
QP1_DP1	Between Groups	36907.59	1	36907.589	39.516	.000
	Within Groups	138230.3	148	933.989		
	Total	175137.9	149			
EP1_QP1	Between Groups	76442.34	1	76442.336	70.355	.000
	Within Groups	160805.5	148	1086.524		
	Total	237247.8	149			
EP1_DP1	Between Groups	7118.059	1	7118.059	10.116	.002
	Within Groups	104137.3	148	703.630		
	Total	111255.3	149			

---

[8] Greenhouse-Geisser corrections were performed in all conditions, because Mauchly's sphericity test revealed heterogeneity of covariance.

Significant differences in utterance-initial pitch setting among the three utterance types were found between native and non-native speaker groups in each of the three conditions (QP1-DP1 $F=39.516$, $p<.001$; EP1-QP1 $F=70.355$, $p<.001$, EP1-DP1 $F=10.116$, $p<.01$)

Table 3: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Final Syllable Condition*

Descriptive Statistics

	GROUP	Mean	Std. Deviation	N
QP1_DP1	1	35.877	36.106	75
	2	4.505	23.756	75
	Total	20.191	34.284	150
EP1_QP1	1	-45.364	35.870	75
	2	-.215	29.772	75
	Total	-22.789	39.903	150
EP1_DP1	1	-9.487	25.581	75
	2	4.291	27.438	75
	Total	-2.598	27.325	150

Clearly, native speakers initially set pitch much higher for their English yes/no question sentences than they did for their echo and declarative sentences, whereas non-native speakers made virtually no distinction among utterance types.

Table 4 presents ANOVA summary results for pitch value averaged across syllables 1-8 for Q, D and E (QA, DA, EA) in the final syllable stress condition.

Table 4: *Comparison Of Utterance-Average Pitch Values Across Utterance Types And Speaker Groups In The Final Syllable Stress Condition (QA - DA, EA - QA, EA – DA*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	78.444	1	78.444	.162	.688
Error (S*A)	71701.148	148	484.467		
Within Subjects					
Pitch (B)	36958.424	1.252	29514.415	72.600	***
Pitch * Group(B*A)	25300.542	1.252	20204.615	49.700	***
Error (B* sub*w*groups)	75341.928	185.328	406.533		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 5.

Table 5: *One-Way ANOVA Summary Of The Simple Main Effect Of Pitch In The Final Syllable Condition (QA - DA, EA - QA, EA - DA)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Q_D	Between Groups	13522.61	1	13522.609	52.285	.000
	Within Groups	38277.46	148	258.631		
	Total	51800.07	149			
E_Q	Between Groups	11797.54	1	11797.544	31.752	.000
	Within Groups	54989.76	148	371.552		
	Total	66787.30	149			
E_D	Between Groups	58.833	1	58.833	.162	.688
	Within Groups	53775.86	148	363.350		
	Total	53834.69	149			

Significant differences in average pitch setting among the three utterance types were found between native and non-native speaker groups in two of the three conditions

(Q-D F=52.285, p<.001; E-Q F=31.752, p<.001). It is likely that between-group differences in E-D were not significant because native speakers generally do not use overall pitch height to differentiate echo questions from declaratives.

Table 6: *Means and Standard Deviations For Native (Group 1) and Non-Native (Group 2) Speakers In The Final Syllable Condition*

Descriptives

		N	Mean	td. Deviatior	Std. Error	% Confidence Interval f Mean		Minimum	Maximum
						ower Bound	pper Bound		
Q_D	1	75	5.79283	16.641622	.921609	21.963944	29.621723	-2.9125	84.0875
	2	75	.803286	15.502238	.790044	3.236545	10.370027	23.0125	89.7750
	Total	150	6.29806	18.645414	.522392	13.289794	19.306326	23.0125	89.7750
E_Q	1	75	13.7718	19.974029	.306402	18.367440	-9.176227	55.7625	50.6375
	2	75	.965167	18.551091	.142096	-.303052	8.233385	05.1875	48.1000
	Total	150	.903333	21.171606	.728654	-8.319177	-1.487489	05.1875	50.6375
E_D	1	75	12.0210	22.2830	2.5730	6.8941	17.1479	-33.25	82.80
	2	75	10.7685	15.1713	1.7518	7.2778	14.2591	-17.91	85.18
	Total	150	11.3947	19.0081	1.5520	8.3279	14.4615	-33.25	85.18

4.2.2 Penultimate Syllable Stress Condition

Figure 2b represents the three illocutionary types in English produced by native English speakers in the penultimate stress condition. Figure 4b represents the three illocutionary types in English produced by native Mandarin speakers in the penultimate stress condition.

Figure 2b: F0 contours for three English illocutionary types in the penultimate stress condition, produced by native English speakers.

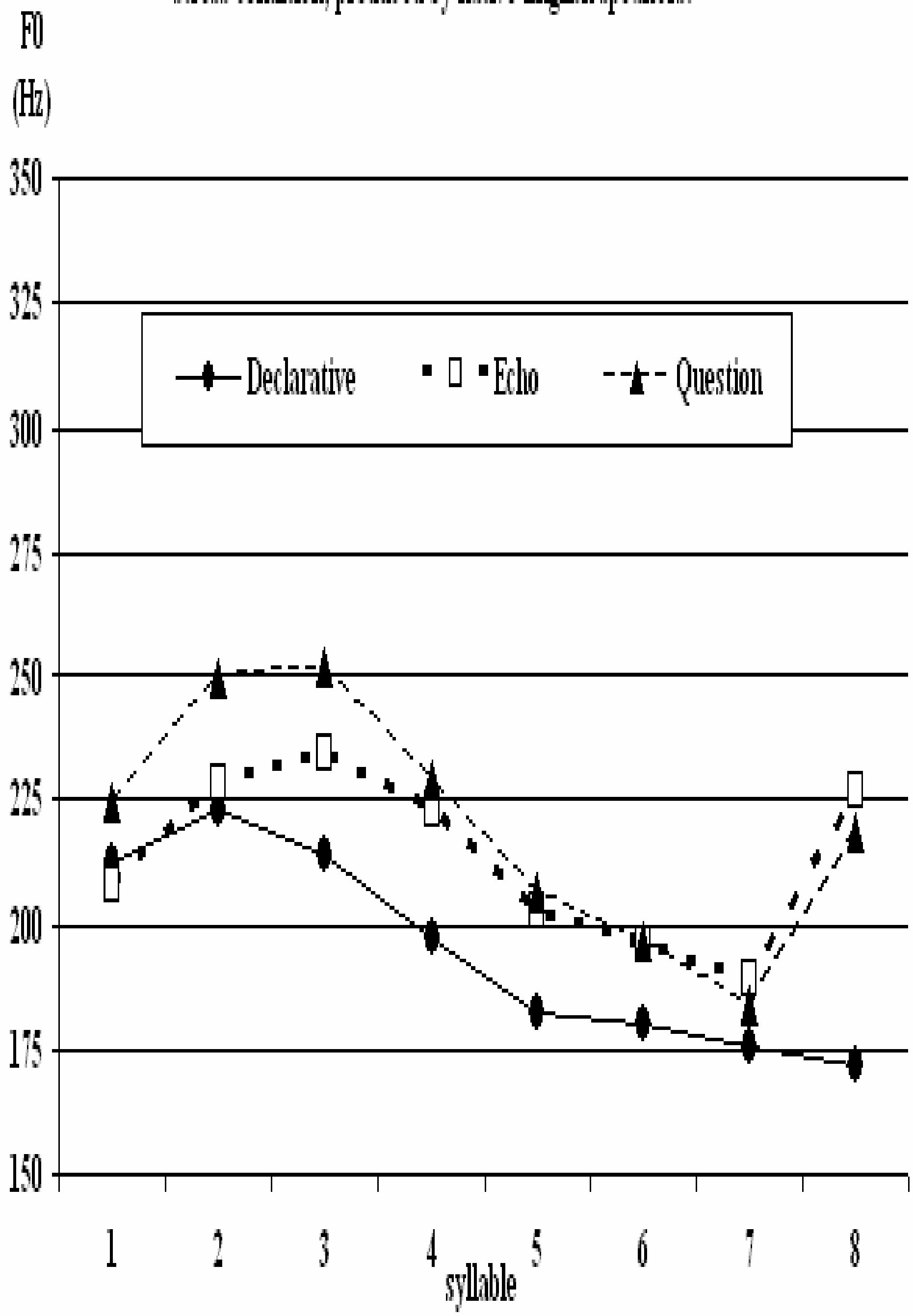


Figure 4b: F0 contours for three English illocutionary types in the penultimate stress condition, produced by native Mandarin speakers.

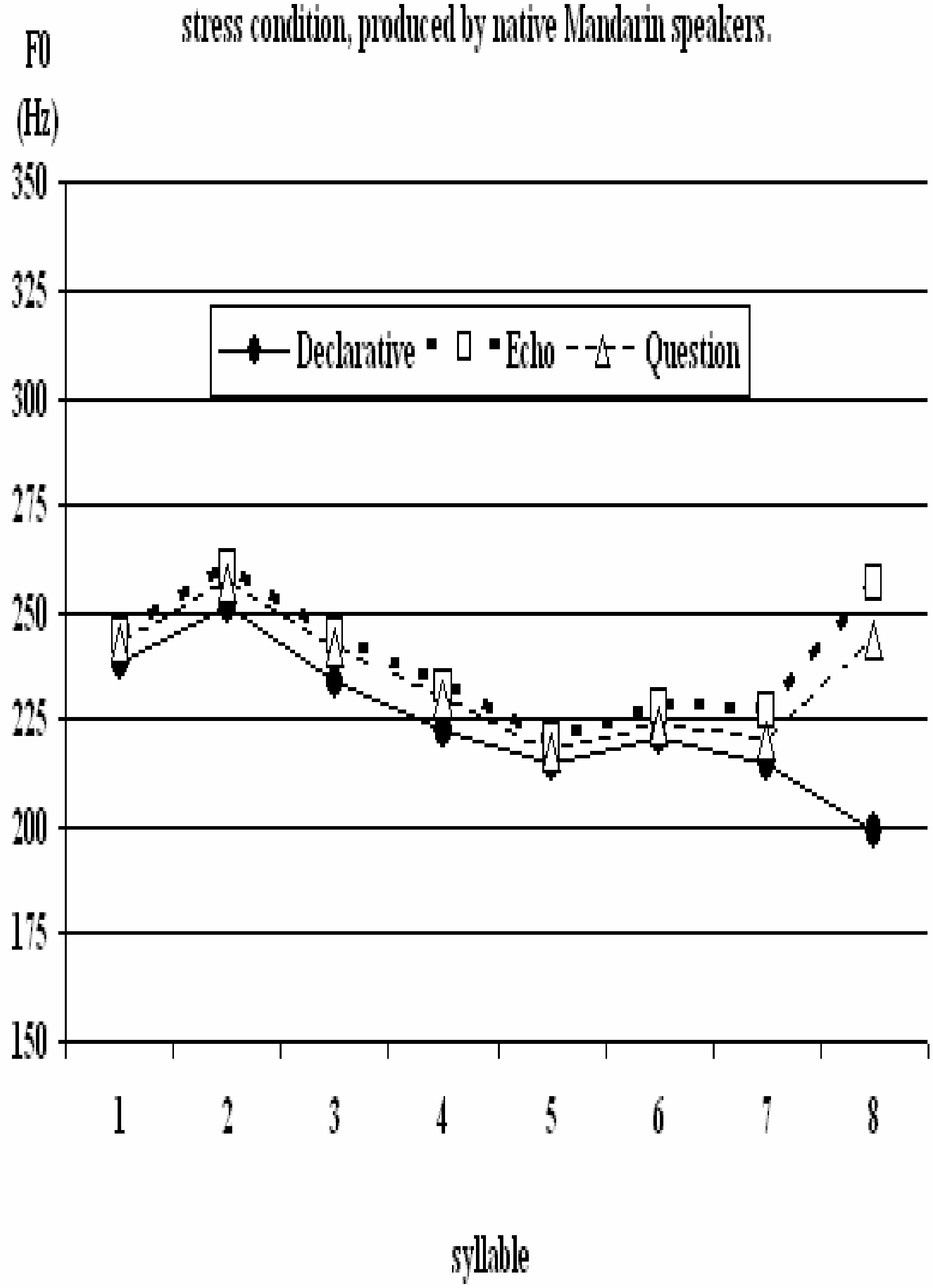


Table 7: *Comparison of Initial Syllable Pitch Values Across Utterance Types And Speaker Groups In The Penultimate Syllable Stress Condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	4299.736	1	4299.736	8.340	.004
Error (S*A)	76298.016	148	515.527		
Within Subjects					
Pitch (B)	34977.698	1.111	31485.803	29.893	***
Pitch * Group(B*A)	68142.508	1.111	61339.702	58.236	***
Error (B* sub*w*groups)	173176.415	164.414	1053.296		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 8.

Table 8: *One-Way ANOVA Summary Of The Simple Main Effect Of Pitch In The Penultimate Syllable Condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
QP1_DP1	Between Groups	24168.11	1	24168.107	43.460	.000
	Within Groups	82303.41	148	556.104		
	Total	106471.5	149			
EP1_QP1	Between Groups	45049.33	1	45049.335	60.641	.000
	Within Groups	109947.5	148	742.889		
	Total	154996.8	149			
EP1_DP1	Between Groups	3224.802	1	3224.802	8.340	.004
	Within Groups	57223.51	148	386.645		
	Total	60448.31	149			

Significant differences in utterance-initial pitch setting among the three utterance types were found between native and non-native speaker groups in each of the three conditions (QP1-DP1 $F=43.460$, $p<.001$; EP1-QP1 $F=60.641$, $p<.001$, EP1-DP1 $F=8.340$, $p<.01$)

Table 9: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Penultimate Syllable Condition*

Descriptives

		N	Mean	td. Deviatior	Std. Error	% Confidence Interval f Mean		Minimum	Maximum
						ower Bound	Jpper Bound		
QP1_DP	1	75	23.581	23.571	2.722	18.158	29.005	-32.8	86.3
	2	75	-1.805	23.593	2.724	-7.233	3.623	-85.8	102.4
	Total	150	10.888	26.732	2.183	6.575	15.201	-85.8	102.4
EP1_QP:	1	75	-28.037	26.064	3.010	-34.034	-22.041	-102.9	29.6
	2	75	6.623	28.398	3.279	8.886E-02	13.156	-143.5	104.6
	Total	150	-10.707	32.253	2.633	-15.911	-5.504	-143.5	104.6
EP1_DP:	1	75	-4.456	23.768	2.745	-9.925	1.013	-80.7	45.1
	2	75	4.817	14.435	1.667	1.496	8.139	-41.1	50.0
	Total	150	.181	20.142	1.645	-3.069	3.430	-80.7	50.0

Table 10 presents ANOVA summary results for pitch value averaged across syllables 1-8 for Q, D and E (QA, DA, EA) in the penultimate syllable condition.

Table 10: *Comparison Of Utterance-Average Pitch Values Across Utterance Types And Speaker Groups In The Penultimate Syllable Condition (QA - DA, EA - QA, EA - DA)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	154.958	1	154.958	100.965	***
Error (S*A)	71388.482	148	482.355		
Within Subjects					
Pitch (B)	30035.712	1.282	23436.211	65.855	***
Pitch * Group(B*A)	13297.368	1.282	10375.646	29.155	***
Error (B* sub*w*groups)	67501.063	189.676	355.876		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 11.

Table 11: *One-Way ANOVA Summary Of The Simple Main Effect Of Pitch In The Penultimate Syllable Condition (QA - DA, EA - QA, EA – DA)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Q_D	Between Groups	7546.445	1	7546.445	32.352	.000
	Within Groups	34522.34	148	233.259		
	Total	42068.79	149			
E_Q	Between Groups	5789.662	1	5789.662	16.859	.000
	Within Groups	50825.84	148	343.418		
	Total	56615.50	149			
E_D	Between Groups	116.218	1	116.218	.321	.572
	Within Groups	53541.36	148	361.766		
	Total	53657.58	149			

Significant differences in average pitch among the three utterance types were found between native and non-native speaker groups in two of the three conditions (Q-D $F=32.352$, $p<.001$; E-Q $F=16.859$, $p<.001$).

Table 12: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Penultimate Syllable Condition*

Descriptives

		N	Mean	Std. Deviation	Std. Error	5% Confidence Interval fo Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Q_D	1	75	23.82983	14.567539	1.682114	20.478147	27.181520	-13.2625	64.5875
	2	75	9.643976	15.946943	1.841394	5.974918	13.313034	-23.7625	96.2750
	Total	150	16.73690	16.803002	1.371959	14.025895	19.447915	-23.7625	96.2750
E_Q	1	75	-7.34500	17.8714507	2.063617	11.4568474	-3.2331526	57.67500	37.93750
	2	75	5.080417	19.1689049	2.213434	.6700522	9.4907811	86.55000	57.50000
	Total	150	-1.13229	19.4928142	1.591582	-4.2772779	2.0126946	86.55000	57.50000
E_D	1	75	16.48483	21.3474543	2.464992	11.5732302	21.3964365	20.27500	85.62500
	2	75	14.72439	16.3651490	1.889685	10.9591140	18.4896717	14.27500	92.57500
	Total	150	15.60461	18.9767748	1.549447	12.5428851	18.6663411	20.27500	92.57500

4.2.3 Antepenultimate Syllable Condition

Figure 2c represents the three illocutionary types in English produced by native English speakers in the antepenultimate stress condition. Figure 4c represents the three illocutionary types in English produced by native Mandarin speakers in the antepenultimate stress condition.

Figure 2c: F0 contours for three English illocutionary types in the antepenultimate stress condition, produced by native English speakers.

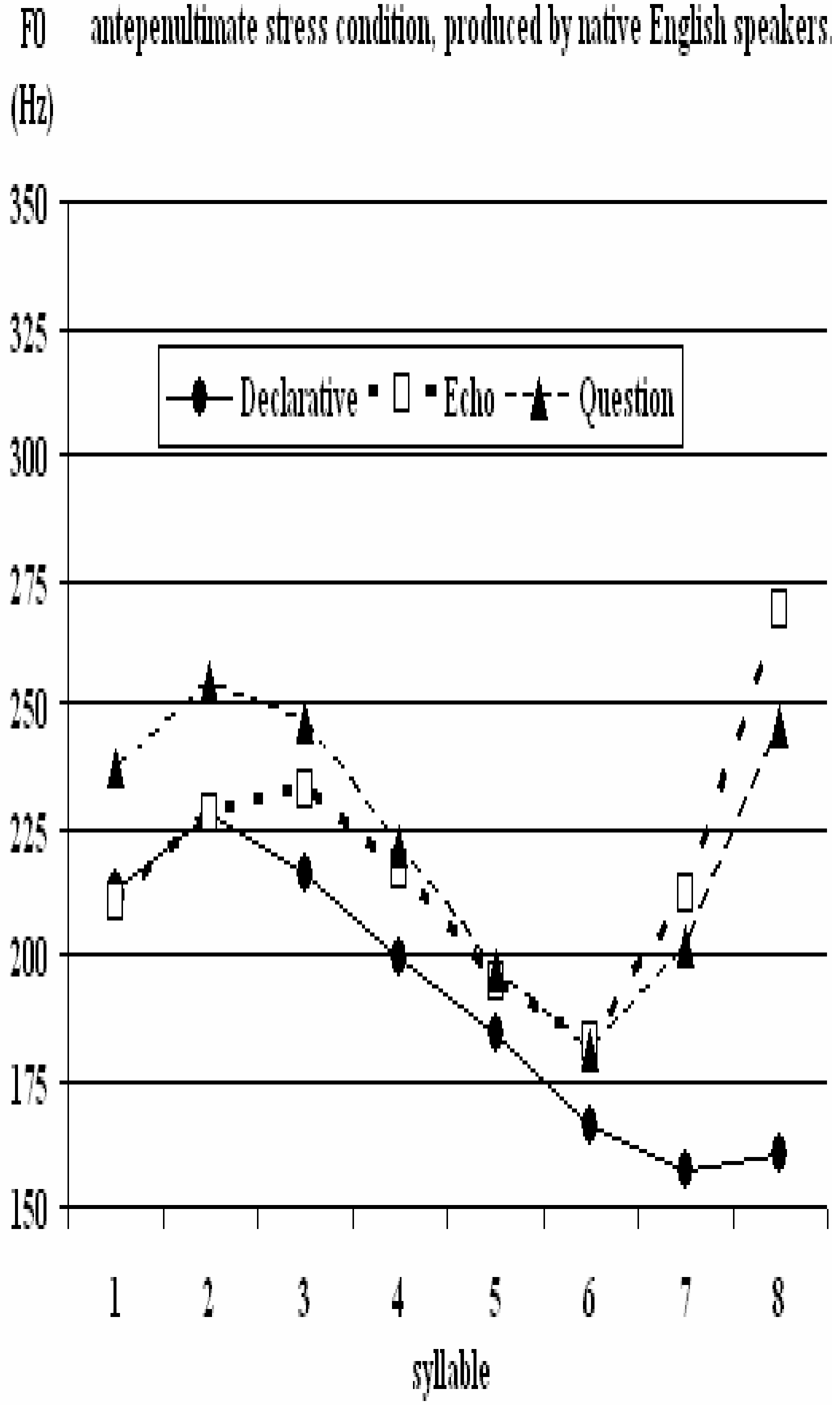


Figure 4c: F0 contours for three English illocutionary types in the antepenultimate stress condition, produced by native Mandarin speakers.

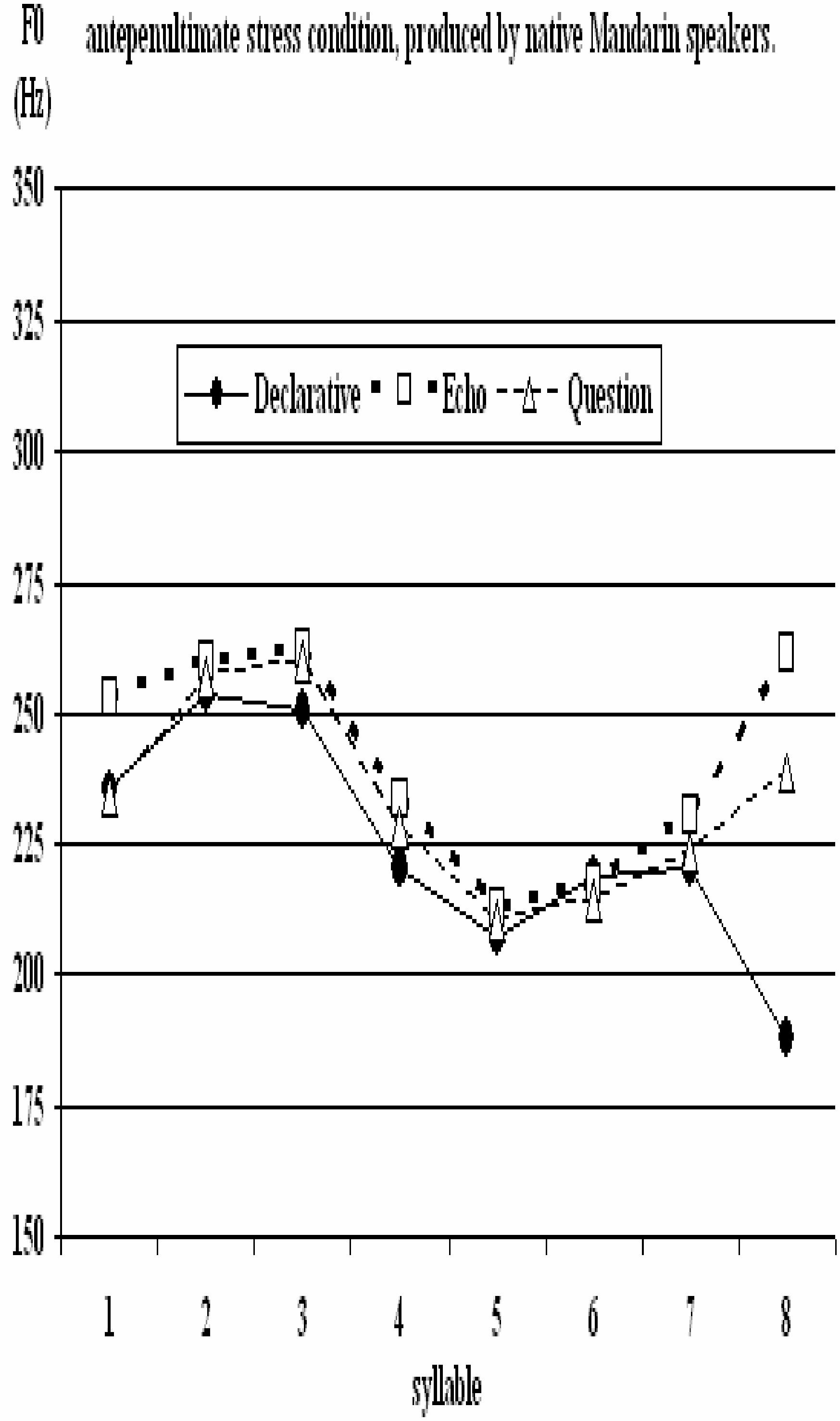


Table 13: *Comparison of initial syllable pitch values across utterance types and speaker groups in the antepenultimate syllable condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	3662.539	1	3662.539	4.595	.034
Error (S*A)	117960.806	148	797.032		
Within Subjects					
Pitch (B)	55218.534	1.186	46558.312	49.071	***
Pitch * Group(B*A)	69257.009	1.186	58395.057	61.547	***
Error (B* sub*w*groups)	166539.924	175.529	948.788		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 14.

Table 14: *One-way ANOVA summary of the simple main effect of pitch in the antepenultimate syllable condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
QP1_DP1	Between Groups	25365.60	1	25365.602	44.566	.000
	Within Groups	84237.70	148	569.174		
	Total	109603.3	149			
EP1_QP1	Between Groups	44807.04	1	44807.042	59.319	.000
	Within Groups	111792.4	148	755.354		
	Total	156599.5	149			
EP1_DP1	Between Groups	2746.904	1	2746.904	4.595	.034
	Within Groups	88470.60	148	597.774		
	Total	91217.51	149			

Significant differences in initial pitch setting among the three utterance types were found between native and non-native speaker groups in two of the three conditions (QP1-DP1 $F=44.566$, $p<.001$; EP1-QP1 $F=59.319$, $p<.001$).

Table 15: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Antepenultimate Syllable Condition*

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
QP1_DP1	1	75	26.792	26.556	3.066	20.682	32.902	-40.6	146.0
	2	75	.784	20.812	2.403	-4.004	5.572	-76.8	67.6
	Total	150	13.788	27.122	2.214	9.412	18.164	-76.8	146.0
EP1_QP1	1	75	-30.628	27.868	3.218	-37.040	-24.216	-160.6	35.1
	2	75	3.939	27.094	3.129	-2.295	10.172	-62.4	128.4
	Total	150	-13.345	32.419	2.647	-18.575	-8.114	-160.6	128.4
EP1_DP1	1	75	-3.836	22.649	2.615	-9.047	1.375	-70.6	64.0
	2	75	4.723	26.126	3.017	-1.288	10.734	-113.3	69.2
	Total	150	.443	24.743	2.020	-3.549	4.435	-113.3	69.2

Table 16 presents ANOVA summary results for pitch value averaged across syllables 1-8 for Q, D and E (QA, DA, EA) in the antepenultimate syllable condition

Table 16: *Comparison Of Utterance-Average Pitch Values Across Utterance Types And Speaker Groups In The Antepenultimate Syllable Condition (QA - DA, EA - QA, EA - DA)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	1821.895	1	1821.895	2.818	.095
Error (S*A)	95669.703	148	646.417		
Within Subjects					
Pitch (B)	39904.180	1.311	30431.728	76.725	***
Pitch * Group(B*A)	21941.064	1.311	16732.696	42.187	***
Error (B* sub*w*groups)	76974.273	194.068	396.636		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 17.

Table 17: *One-way ANOVA summary of the simple main effect of pitch in the antepenultimate syllable condition (QA - DA, EA - QA, EA - DA)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Q_D	Between Groups	15049.86	1	15049.863	62.708	.000
	Within Groups	35520.06	148	240.000		
	Total	50569.92	149			
E_Q	Between Groups	7346.675	1	7346.675	16.633	.000
	Within Groups	65371.64	148	441.700		
	Total	72718.31	149			
E_D	Between Groups	1366.421	1	1366.421	2.818	.095
	Within Groups	71752.28	148	484.813		
	Total	73118.70	149			

Significant differences in average pitch among the three utterance types were found between native and non-native speaker groups in two of the three conditions (Q-D F=32.352, p<.001; E-Q F=16.859, p<.001).

Table 18: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Antepenultimate Syllable Condition*

Descriptives

		N	Mean	Std. Deviation	Std. Error	5% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Q_D	1	75	9.55083	16.73110	1.93194	25.70136	33.40031	-11.412	72.938
	2	75	9.51762	14.14465	1.63328	6.26323	12.77201	-29.488	68.859
	Total	150	9.53423	18.42269	1.50421	16.56189	22.50656	-29.488	72.938
E_Q	1	75	.854500	22.109377	.552971	-12.941406	-2.767594	59.0375	89.2875
	2	75	.142333	19.863936	.293690	1.572057	10.712610	79.2625	57.7125
	Total	150	.856083	22.091681	.803778	-4.420373	2.708206	79.2625	89.2875
E_D	1	75	1.69633	23.666557	.732779	16.251153	27.141514	32.1000	95.6750
	2	75	5.65995	20.236587	.336720	11.003937	20.315968	37.2500	68.8750
	Total	150	8.67814	22.152416	.808737	15.104054	22.252231	37.2500	95.6750

4.3 General Discussion Of Results

As expected, native speaker production of questions exhibited a higher overall F0 than declaratives did, particularly at the end of the sentence. (Mean F0 for syllables 1-8, averaged over all items of each illocutionary type: declarative = 162.3 Hz; yes/no question = 220 Hz; echo question = 232.3 Hz.). The two question types had similar contours, yet there were some interesting differences between them. As is often reported, the echo question exhibited a more expanded pitch range in the sentence-final rise than the yes/no question did. However, the yes/no question type also exhibited an unexpected higher F0 than the echo (and the declarative) at the beginning of the sentence (Mean F0 for the initial syllable, averaged over all items of each illocutionary type: declarative = 214.6 Hz; yes/no question = 234.6 Hz; echo question = 209 Hz). This could be attributed to the initial pitch rising that Hirschberg and Pierrehumbert (1986) associate with *topic jump*, since the yes/no question in each mini-dialogue preceded the declarative and the echo question, so the yes/no question constituted the first mention of the topic of each dialogue.

It should also be noted that both question types had higher F0 than the declarative through the middle part of the sentence, up to the position of the nuclear accent on the last stressed syllable. Comparison of the three stress-pattern conditions suggests that the interrogative and declarative contours are most closely anchored together at that nuclear stressed syllable. The F0 rising seen preceding the nuclear accent in questions compared with declaratives has not been reported previously in the literature. Interestingly, it made the native English interrogative contour more similar than was anticipated to the native

Mandarin contour, which has global F0 rising in the yes/no question condition (see Mandarin results below). However, the two are clearly differentiated by the convergence at the nuclear stressed syllable in the English contour.

As anticipated, sentence-final F0 rising is extended over the last two syllables when the nuclear stress is on the antepenultimate syllable (Figure 2c), but occurs only during the final syllable in the final stress and penultimate stress conditions (Figures 2a, b). This could also explain why the extent of the final rise was greater in the antepenultimate condition than in the other two conditions. The finer-grained F0 data sampled every 20ms, shown in Figures 3a-c below, is valuable here. The different timing of the final interrogative rise across the three stress conditions is apparent. In Figure 3a, for example, one can clearly see the low accent and the rise to the high boundary tone compressed into a single syllable in the final stress condition, in contrast to the antepenultimate stress condition in Figure 3c.

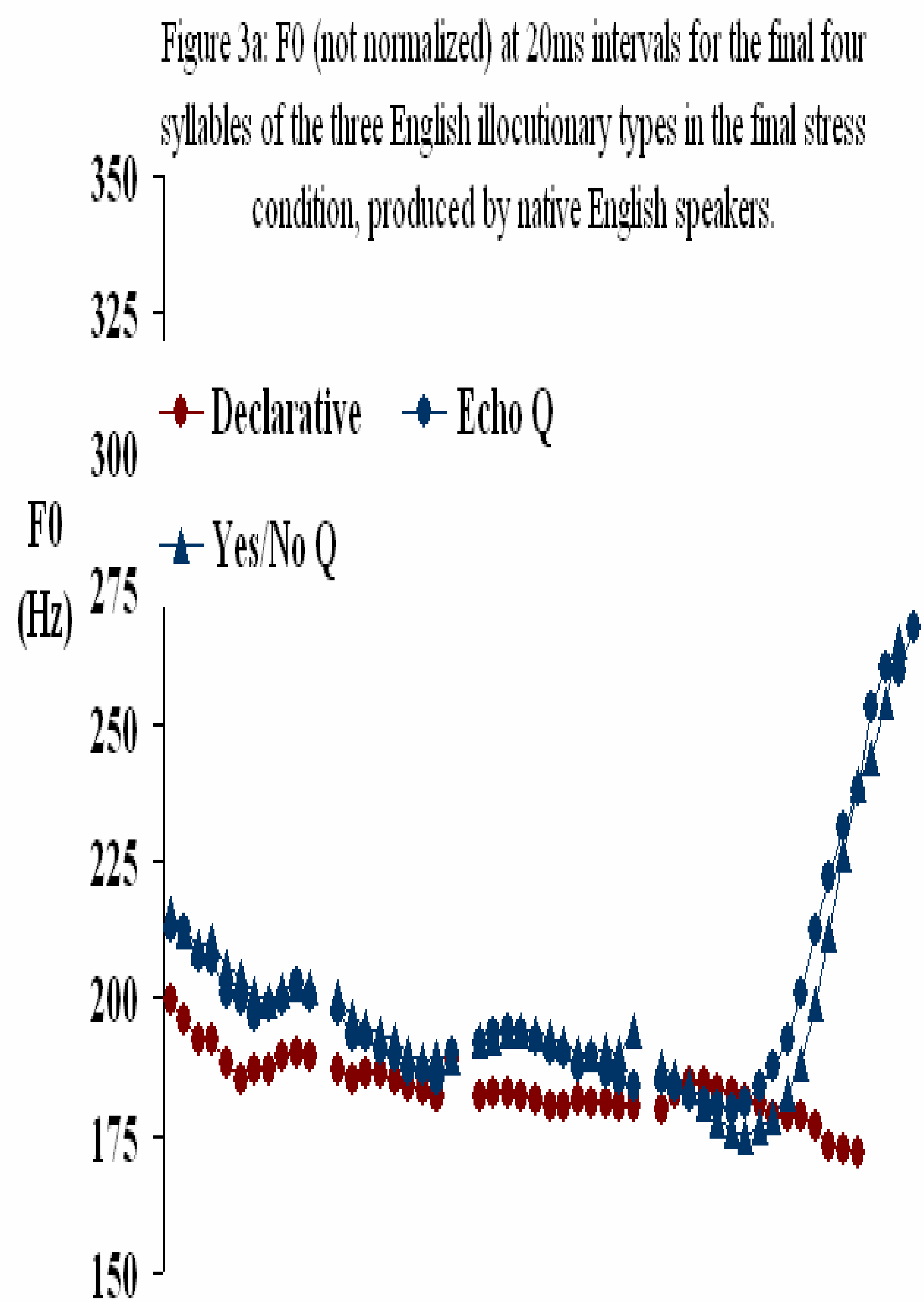


Figure 3a: F0 (not normalized) at 20ms intervals for the final four syllables of the three English illocutionary types in the final stress condition, produced by native English speakers.
350
325
300
275
250
225
200
175
150
F0 (Hz)
Declarative
Echo Q
Yes/No Q

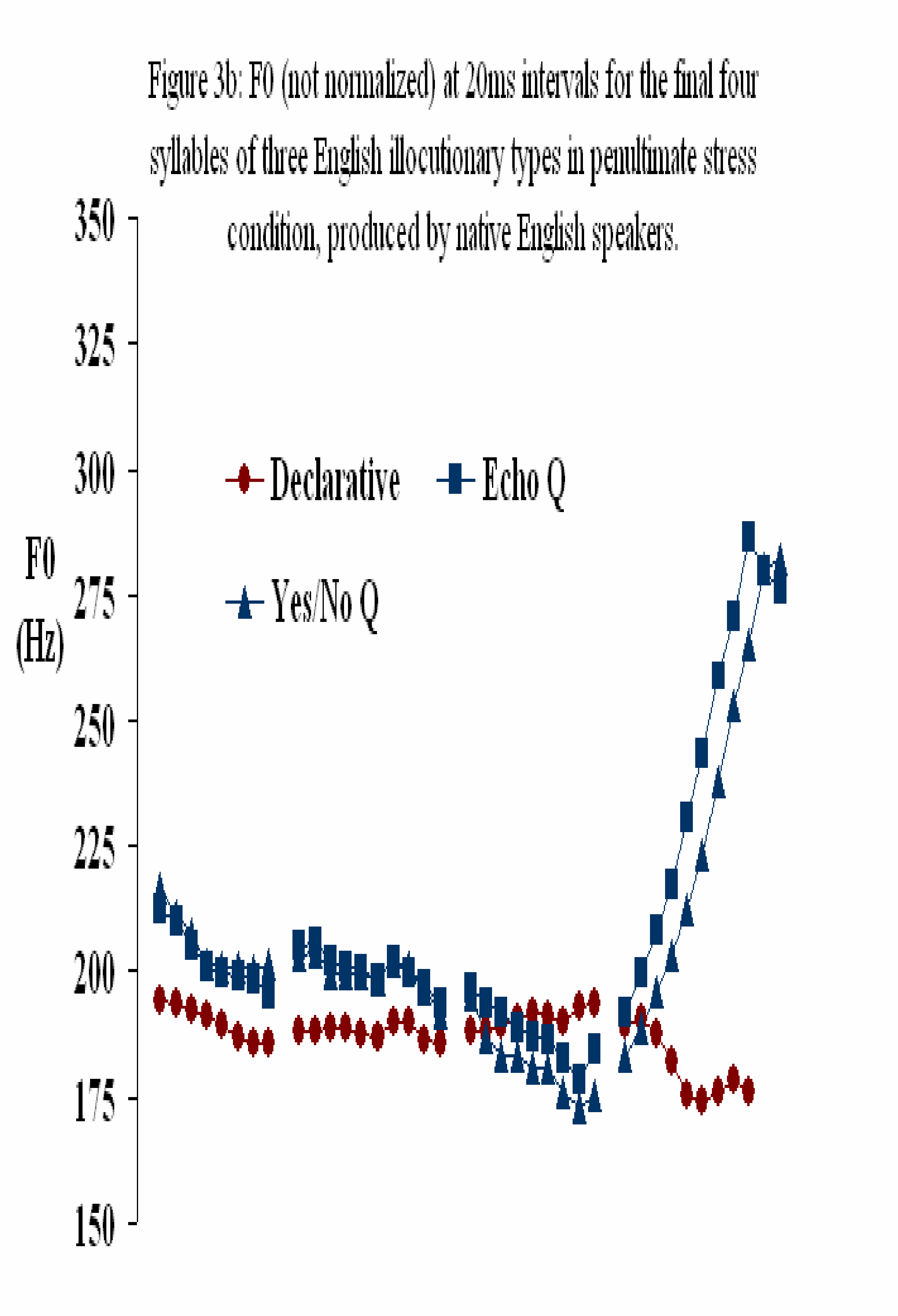
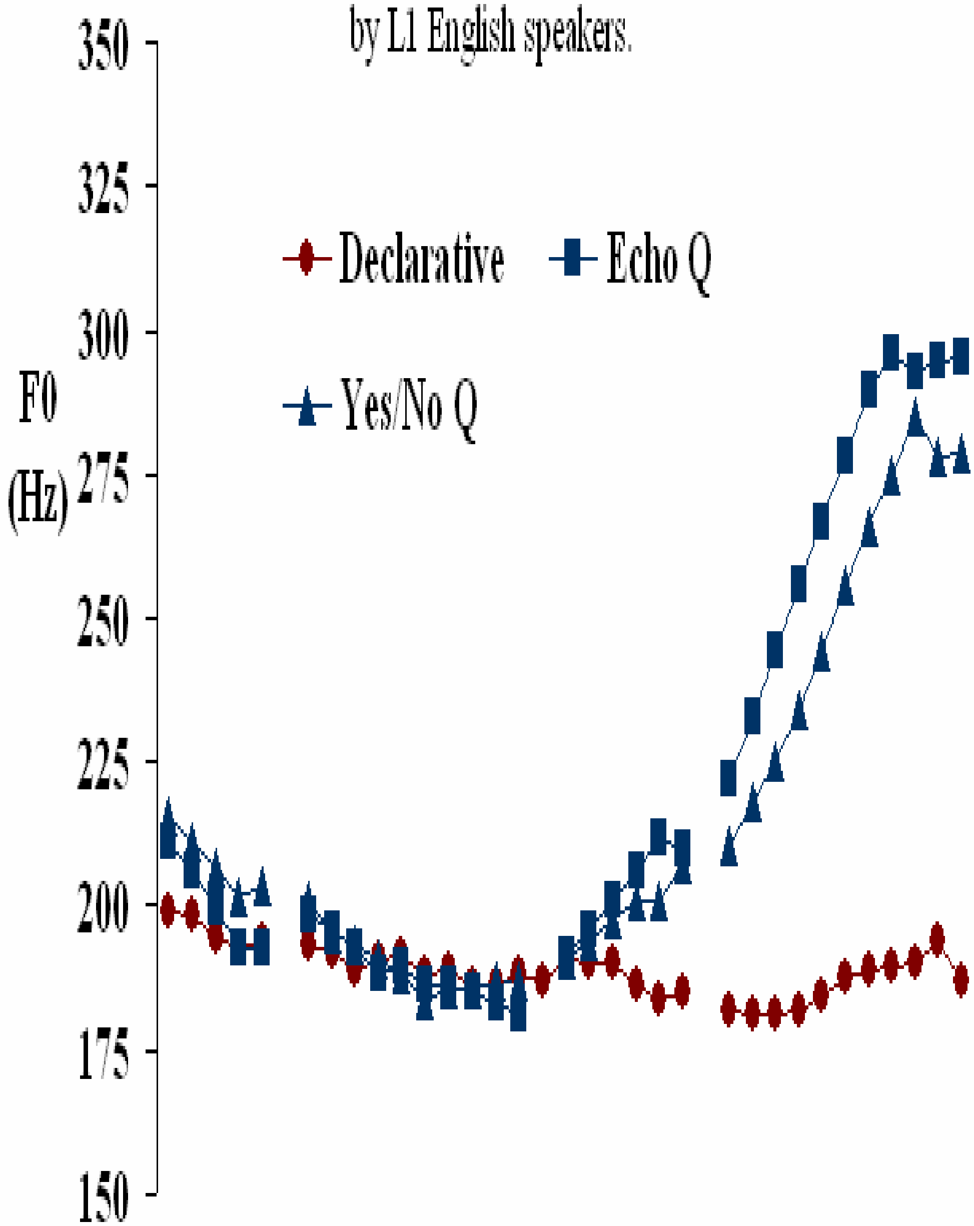


Figure 3b: F0 (not normalized) at 20ms intervals for the final four syllables of three English illocutionary types in penultimate stress condition, produced by native English speakers.
350
325
300
275
250
225
200
175
150
F0
(Hz)
Declarative
Echo Q
Yes/No Q

Figure 3c: F0 (not normalized) for final four syllables of three English illocutionary types in antepenult stress condition, produced by L1 English speakers.



The F0 values in Figures 4a-c do not differentiate between the illocutionary types until the utterance-final syllable. In this respect they resemble neither the native English contours presented above, nor the contours that these same participants produced in their native Mandarin, as further described.

It is clear that these L2 speakers do not use pitch height to distinguish between illocutionary types in English in the same way that native speakers do. Their mean F0 across eight syllables, averaged over participants and item types, was 235.8 Hz. in declarative sentences, 246.3 Hz. in yes/no questions, and 239.7 Hz. in echo questions, a very narrow spread compared with the L1 English speakers, whose mean F0 was 191.6 Hz. in declarative sentences, 220.6 Hz. in yes/no questions, and 212.1 Hz. in echo questions. This failure by the L2 speakers to produce illocutionary differentiation throughout the sentence is not only unlike the native English speakers' performance with respect to global aspects of prosody but also exhibits a very striking absence of transfer from the Mandarin L1. Possible explanations for these findings will be discussed in Chapter 5.

L2 speakers did distinguish the two question types from the declarative by means of a final rise in the questions compared with a fall in the declarative (interestingly, a larger declarative fall than that produced by the native speakers). They also showed a good grasp of the exaggerated quality of the F0 excursion in echo questions: Their final rise is systematically higher in the echo questions than in the yes/no questions. However, a remarkably consistent observation across items is the limitation of the utterance-final rise in the questions to the final syllable, regardless of the position of nuclear stress.

There is no difference in the timing of the rise to the interrogative high boundary tone in the antepenultimate stress items versus the penultimate stress and final stress items. Of special note is the contrast between the 7th syllable in Figure 4c (antepenultimate stress condition) for L2 English, and the 7th syllable in Figure 2c (also antepenultimate stress) for L1 English. In the L2 utterances the declarative and interrogatives are alike at that point, while in the L1 utterances they are already well differentiated. Thus, it appears that Mandarin-native learners of English assimilate the phrasal contours of English to the syllable-locked contours of the lexical tones in their L1. Unlike the lack of transfer observed for F0 in the beginning and mid-portions of the sentence, this syllable-bounded sentence-final pitch excursion plausibly does represent transfer from the L1; that is, an L2 phenomenon (interrogative sentence-final rise) is filtered through the L1 generalization that rises and falls normally within a single syllable.

4.4 Mandarin Materials

4.4.0 Introduction

As for the English data, two separate sets of pitch extractions were generated: one extracted f0 at 20 ms. intervals and the other averaged f0 over an entire syllable at each marked syllable boundary. The 20-ms interval extraction allowed for a finer-grained analysis of pitch excursion within a single syllable, while the syllable-averaged measurements allowed for observation of global pitch setting and rise/declination over the entire utterance and across utterances in each mini-dialogue. F0 values averaged at the syllable level were the dependent variable in three 2-way ANOVAs, which compared utterance types with repeated measures across the two participant groups. This section will present F0 data for the native Mandarin speakers, averaged and normalized for pitch and duration. Figures 5a-5c represent the three illocutionary types in Mandarin produced by native Mandarin speakers in Tone 1, Tone 2, and Tone 4 conditions, respectively. Figures 6a-6c represent the three illocutionary types in Mandarin produced by non-native speakers in Tone 1, Tone 2, and Tone 4 conditions, respectively.

Figure 5a: F0 contours for three Mandarin illocutionary types in the Tone 1 condition, produced by native Mandarin speakers.

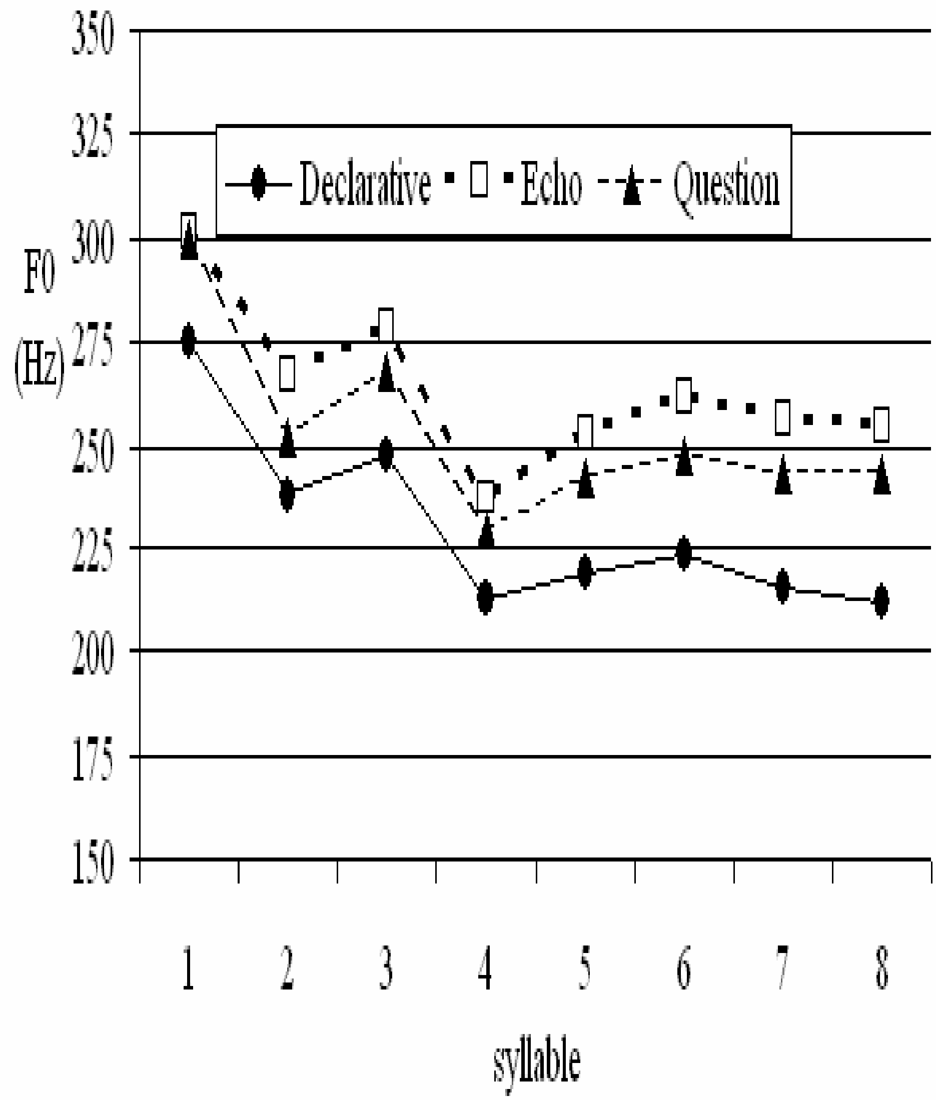
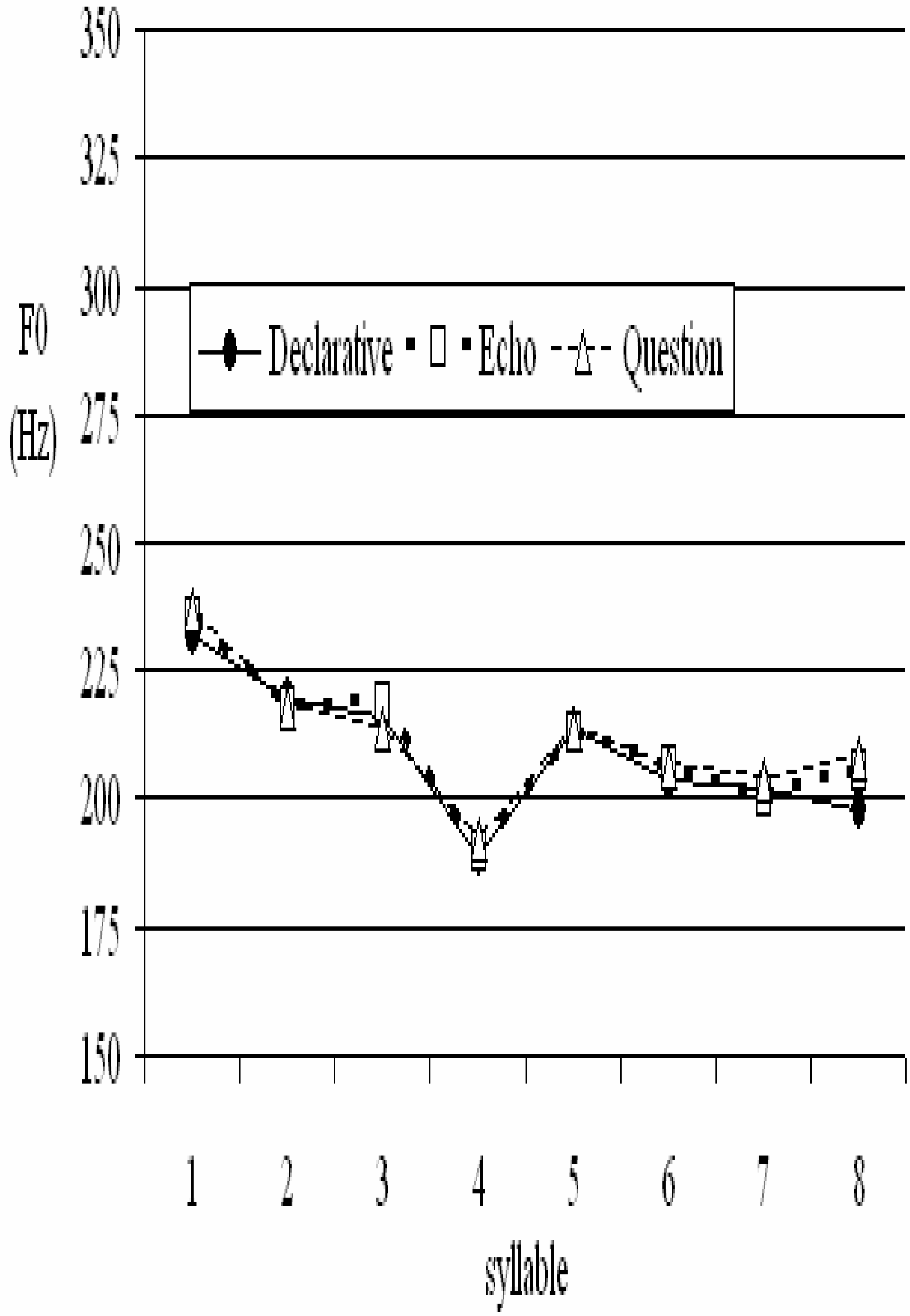


Figure 6a: F0 contours for three Mandarin illocutionary types in the Tone 1 condition, produced by native English speakers.



4.4.1 Tone 1 Condition

Table 19:*Comparison of initial syllable pitch values across utterance types and speaker groups in the Tone 1 condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	48749.237	1	48749.237	74.439	***
Error (S*A)	164266.157	148	1109.906		
Within Subjects					
Pitch (B)	17194.976	1.301	13220.274	16.596	***
Pitch * Group(B*A)	12140.391	1.301	9334.080	11.718	***
Error (B* sub*w*groups)	153337.985	192.497	796.575		

*** p value <.001

An interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 20.

Table 20: *One-way ANOVA summary of the simple main effect of pitch in the tone 1 condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

ANOVA

		平方和	自由度	平均平方和	F 檢定	顯著性
QP1-DP1	組間	22677.658	1	22677.658	36.524	.000
	組內	91892.949	148	620.898		
	總和	14570.607	149			
EP1-QP1	組間	1650.042	1	1650.042	2.382	.125
	組內	02511.576	148	692.646		
	總和	04161.617	149			
EP1-DP1	組間	36561.927	1	36561.927	43.922	.000
	組內	23199.618	148	832.430		
	總和	59761.546	149			

Significant differences in utterance-initial pitch setting among the three utterance types were found between native and non-native speaker groups in two of the three conditions (QP1-DP1 36.524, $p<.001$, EP1-DP1 $F=43.922$, $p<.001$).

Table 21: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Tone 1 Condition*

描述性統計量

		個數	平均數	標準差	標準誤	均數的 95% 信賴區		最小值	最大值
						下界	上界		
QP1-DP:	1	75	27.243	29.6674	3.4257	20.417	34.069	-34.7	129.6
	2	75	2.651	19.0169	2.1959	-1.724	7.027	-50.4	55.9
	總和	150	14.947	27.7296	2.2641	10.473	19.421	-50.4	129.6
EP1-QP1	1	75	8.695	30.6586	3.5402	1.641	15.749	-81.1	100.6
	2	75	2.061	21.1031	2.4368	-2.794	6.917	-60.2	52.6
	總和	150	5.378	26.4400	2.1588	1.112	9.644	-81.1	100.6
EP1-DP1	1	75	35.937	36.4048	4.2037	27.561	44.313	-56.7	198.5
	2	75	4.713	18.4268	2.1277	.473	8.952	-50.0	65.7
	總和	150	20.325	32.7448	2.6736	15.042	25.608	-56.7	198.5

Clearly, native speakers used initial pitch setting to distinguish between interrogative and declarative utterance types, whereas non-native speakers made virtually no such distinction.

Table 22 presents ANOVA summary results for pitch value averaged across syllables 1-8 for Q, D and E (QA, DA, EA) in the Tone 1 condition

Table 22: *Comparison Of Utterance-Average Pitch Values Across Utterance Types And Speaker Groups In The Tone 1 Condition (QA - DA, EA - QA, EA - DA)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	58609.089	1	58609.089	108.586	***
Error (S*A)	79882.755	148	539.748		
Within Subjects					
Pitch (B)	14923.486	1.330	11218.251	33.827	***
Pitch * Group(B*A)	11265.596	1.330	8468.549	25.536	***
Error	65293.262	196.882	331.636		
(B* sub*w*groups)					

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 23.

Table 23:*One-Way ANOVA Summary Of The Simple Main Effect Of Pitch In The Tone 1 Condition (QA - DA, EA - QA, EA – DA)*

ANOVA

		平方和	自由度	平均平方和	F 檢定	顯著性
Q-D	組間	22263.936	1	22263.936	96.010	.000
	組內	34320.106	148	231.893		
	總和	56584.042	149			
E-Q	組間	3653.932	1	3653.932	10.615	.001
	組內	50943.845	148	344.215		
	總和	54597.777	149			
E-D	組間	43956.817	1	43956.817	108.586	.000
	組內	59912.066	148	404.811		
	總和	03868.883	149			

Significant differences in utterance-average pitch among the three utterance types were found between native and non-native speaker groups in each of the three conditions (Q-D F=96.010, p<.001; E-Q F=10.615, p<.01, E-D F=108.586, p<.001). Thus, it appears that native speakers use average pitch to distinguish among all three utterance types, while non-native speakers maintain approximately the same average pitch for all three utterance types.

Table 24: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Tone 1 Condition*

描述性統計量

		個數	平均數	標準差	標準誤	均數的 95% 信賴區		最小值	最大值
						下界	上界		
Q-D	1	75	26.1906	18.66274	2.15499	21.8967	30.4845	-8.80	97.34
	2	75	1.8245	10.74650	1.24090	-.6480	4.2971	-34.07	25.39
	總和	150	14.0076	19.48740	1.59114	10.8635	17.1517	-34.07	97.34
E-Q	1	75	10.498	22.3085	2.5760	5.366	15.631	-50.1	56.1
	2	75	.627	13.8116	1.5948	-2.550	3.805	-20.6	30.8
	總和	150	5.563	19.1423	1.5630	2.475	8.651	-50.1	56.1
E-D	1	75	36.6891	25.61791	2.95810	30.7949	42.5832	-3.70	103.94
	2	75	2.4520	12.38326	1.42990	-.3972	5.3011	-30.41	36.67
	總和	150	19.5705	26.40278	2.15578	15.3107	23.8304	-30.41	103.94

4.4.2 Tone 2 Condition

Figures 5b and 6b illustrate the three utterance types produced by native and non-native speakers in the Tone 2 condition.

Figure 5b: F0 contours for three Mandarin illocutionary types in the Tone 2 condition, produced by native Mandarin speakers.

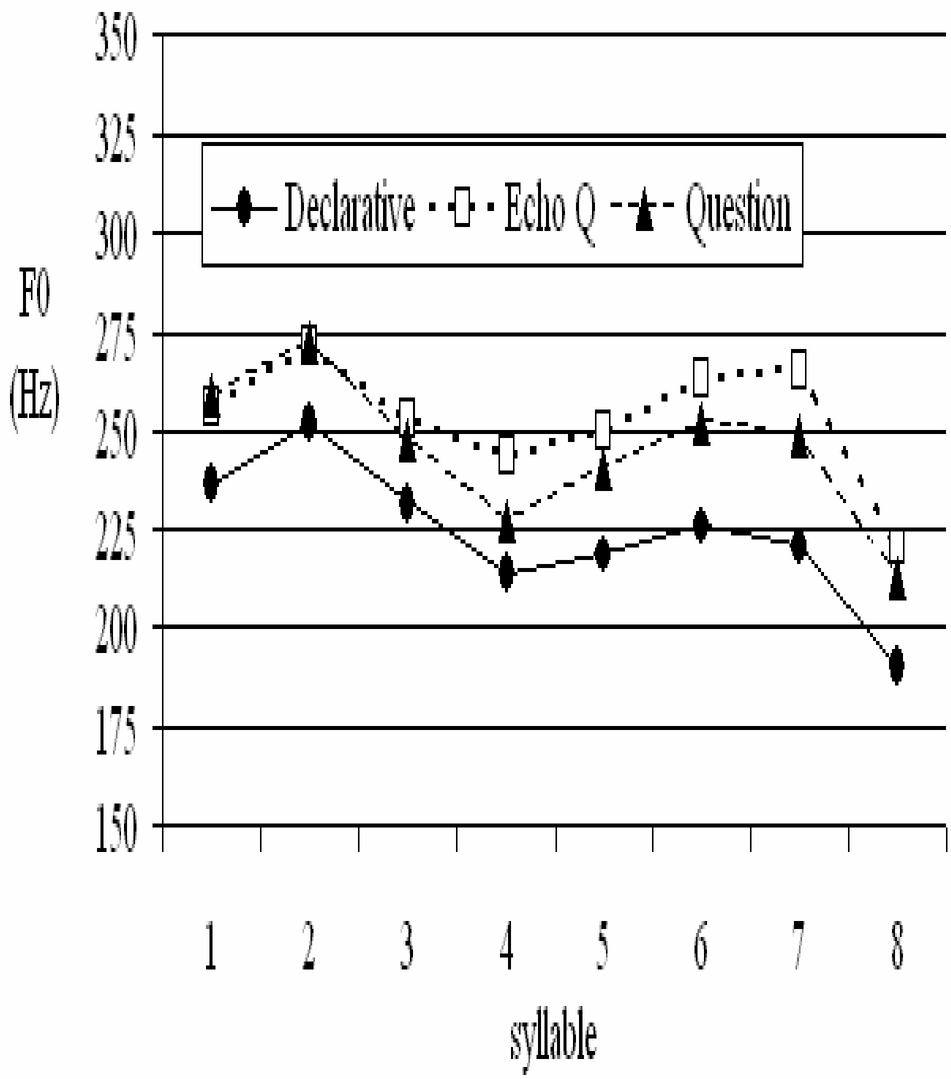


Figure 6b: F0 contours for three Mandarin illocutionary types in the Tone 2 condition, produced by native English speakers.

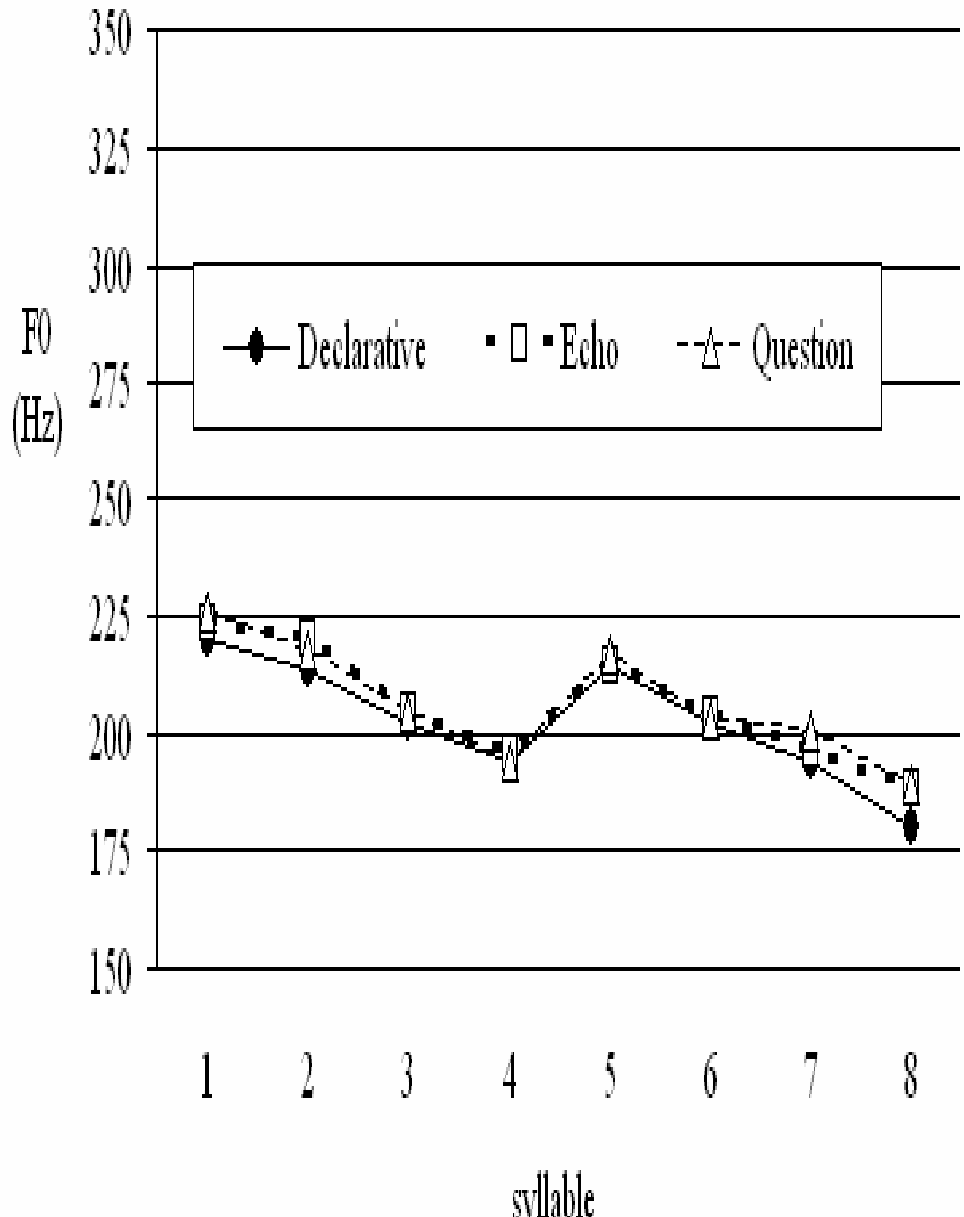


Table 25: *Comparison Of Initial Syllable Pitch Values Across Utterance Types And Speaker Groups In The Tone 2 Condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	23040.889	1	23040.889	19.939	***
Error (S*A)	171020.629	148	1155.5435		
Within Subjects					
Pitch (B)	17962.244	1.246	14419.886	13.794	***
Pitch * Group(B*A)	8523.849	1.246	6842.850	6.546	.007
Error (B* sub*w*groups)	192715.940	184.357	1045.339		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 26.

Table 26: *One-Way ANOVA Summary Of The Simple Main Effect Of Pitch In The Tone 2 Condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
QP1_DP1	Between Groups	14125.14	1	14125.142	17.549	.000
	Within Groups	119122.5	148	804.881		
	Total	133247.6	149			
EP1_QP1	Between Groups	158.929	1	158.929	.202	.654
	Within Groups	116348.6	148	786.139		
	Total	116507.6	149			
EP1_DP1	Between Groups	17280.67	1	17280.667	19.939	.000
	Within Groups	128265.5	148	866.659		
	Total	145546.1	149			

As in the tone 1 condition, significant differences in utterance-initial pitch setting among the three utterance types were found between native and non-native speaker groups in two of the three conditions (QP1-DP1 F=17.549, p<.001, EP1-DP1 F=19.939, p<.001).

Table 27: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers in the Tone 2 Condition*

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
QP1_DP1	1	75	23.299	33.532	3.872	15.584	31.014	-33.1	189.4
	2	75	3.891	22.032	2.544	-1.178	8.960	-48.6	104.6
	Total	150	13.595	29.905	2.442	8.770	18.419	-48.6	189.4
EP1_QP1	1	75	1.42	29.49	3.41	-5.36	8.21	-82	78
	2	75	-.64	26.51	3.06	-6.73	5.46	-117	63
	Total	150	.39	27.96	2.28	-4.12	4.90	-117	78
EP1_DP1	1	75	24.721	36.408	4.204	16.345	33.098	-98.7	175.1
	2	75	3.255	20.193	2.332	-1.391	7.901	-51.8	65.2
	Total	150	13.988	31.254	2.552	8.945	19.031	-98.7	175.1

Table 28 presents ANOVA summary results for pitch value averaged across syllables 1-8 for Q, D and E (QA, DA, EA) in the Tone 2 condition.

Table 28: *Comparison Of Utterance-Average Pitch Values Across Utterance Types And Speaker Groups In The Tone 2 Condition (QA - DA, EA - QA, EA - DA)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	40442.628	1	40442.628	80.252	***
Error (S*A)	74583.578	148	503.943		
Within Subjects					
Pitch (B)	10750.604	1.364	7878.986	29.834	***
Pitch * Group(B*A)	5964.653	1.364	4371.421	16.552	***
Error (B* sub*w*groups)	53331.445	201.941	264.094		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 29.

Table 29: *One-Way ANOVA Summary Of The Simple Main Effect Of Pitch For The Native Speaker Group In The Tone 2 Condition (QA - DA, EA - QA, EA - DA)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Q_D	Between Groups	11751.25	1	11751.251	65.790	.000
	Within Groups	26435.22	148	178.616		
	Total	38186.47	149			
E_Q	Between Groups	4324.058	1	4324.058	14.052	.000
	Within Groups	45542.12	148	307.717		
	Total	49866.18	149			
E_D	Between Groups	30331.97	1	30331.971	80.252	.000
	Within Groups	55937.68	148	377.957		
	Total	86269.65	149			

As in the tone 1 condition, significant differences in utterance-average pitch among the three utterance types were found between native and non-native speaker groups in each of the three conditions (Q-D F=65.790, p<.001; E-Q F=14.052, p<.001, E-D F=80.252, p<.001).

Table 30: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Tone 2 Condition*

Descriptives

		N	Mean	td. Deviation	Std. Error	% Confidence Interval f Mean		Minimum	Maximum
						ower Bound	pper Bound		
Q_D	1	75	0.72133	15.118467	.745730	17.242890	24.199777	14.5750	61.9750
	2	75	.019167	11.343043	.309782	.409369	5.628964	34.4875	33.9000
	Total	150	1.87025	16.008905	.307122	9.287360	14.453140	34.4875	61.9750
E_Q	1	75	.951833	20.476256	.364394	5.240675	14.662992	29.0125	78.2125
	2	75	.786333	14.005607	.617228	-4.008731	2.436064	37.8000	42.6250
	Total	150	.582750	18.294052	.493703	1.631173	7.534327	37.8000	78.2125
E_D	1	75	30.67	25.23	2.91	24.87	36.48	-33	97
	2	75	2.23	10.93	1.26	-.28	4.75	-33	45
	Total	150	16.45	24.06	1.96	12.57	20.34	-33	97

4.4.2 Tone 4 Condition

Figures 5c and 6c illustrate the three utterance types produced by native and non-native speakers in the Tone 4 condition.

Figure 5c: F0 contours for three Mandarin illocutionary types in the Tone 4 condition, produced by native Mandarin speakers.

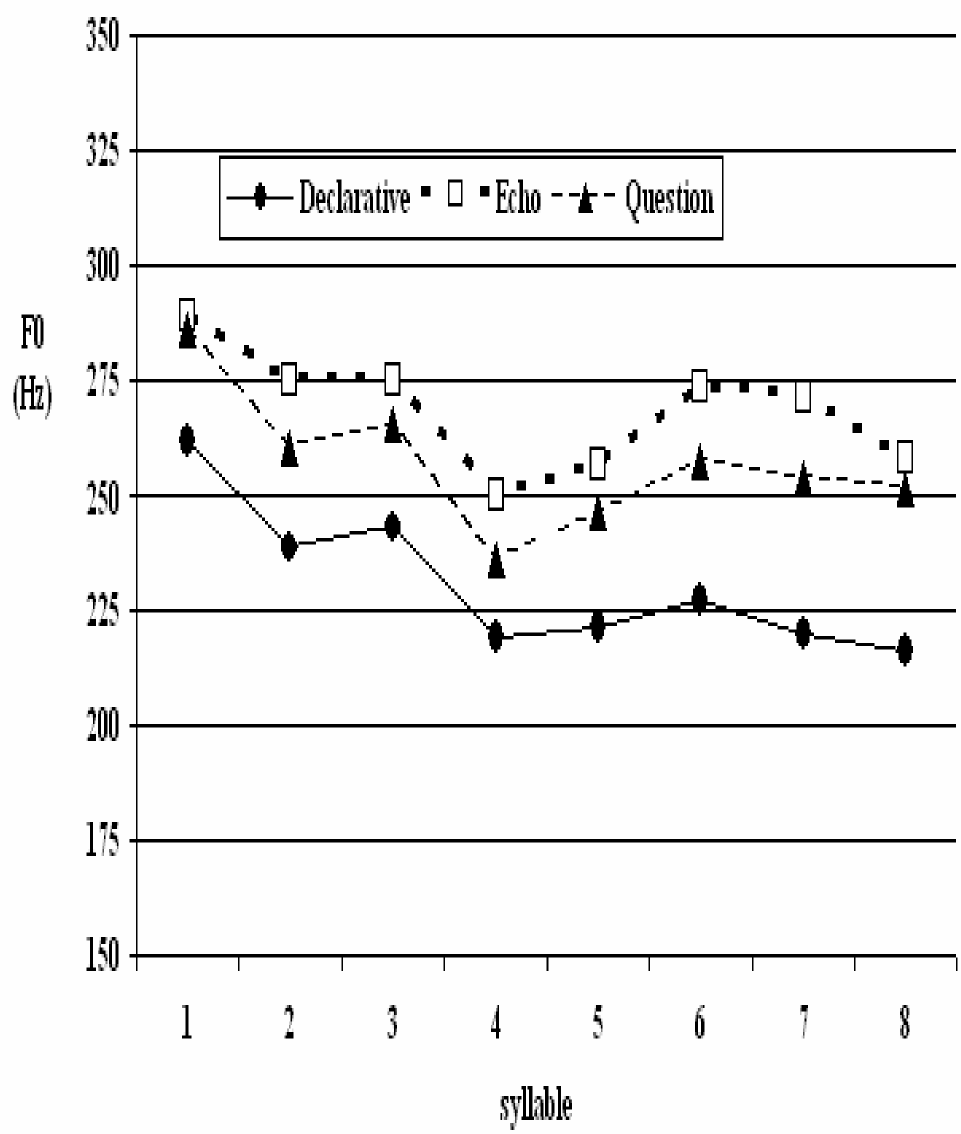


Figure 6c: F0 contours for three Mandarin illocutionary types in the Tone 4 condition, produced by native English speakers.

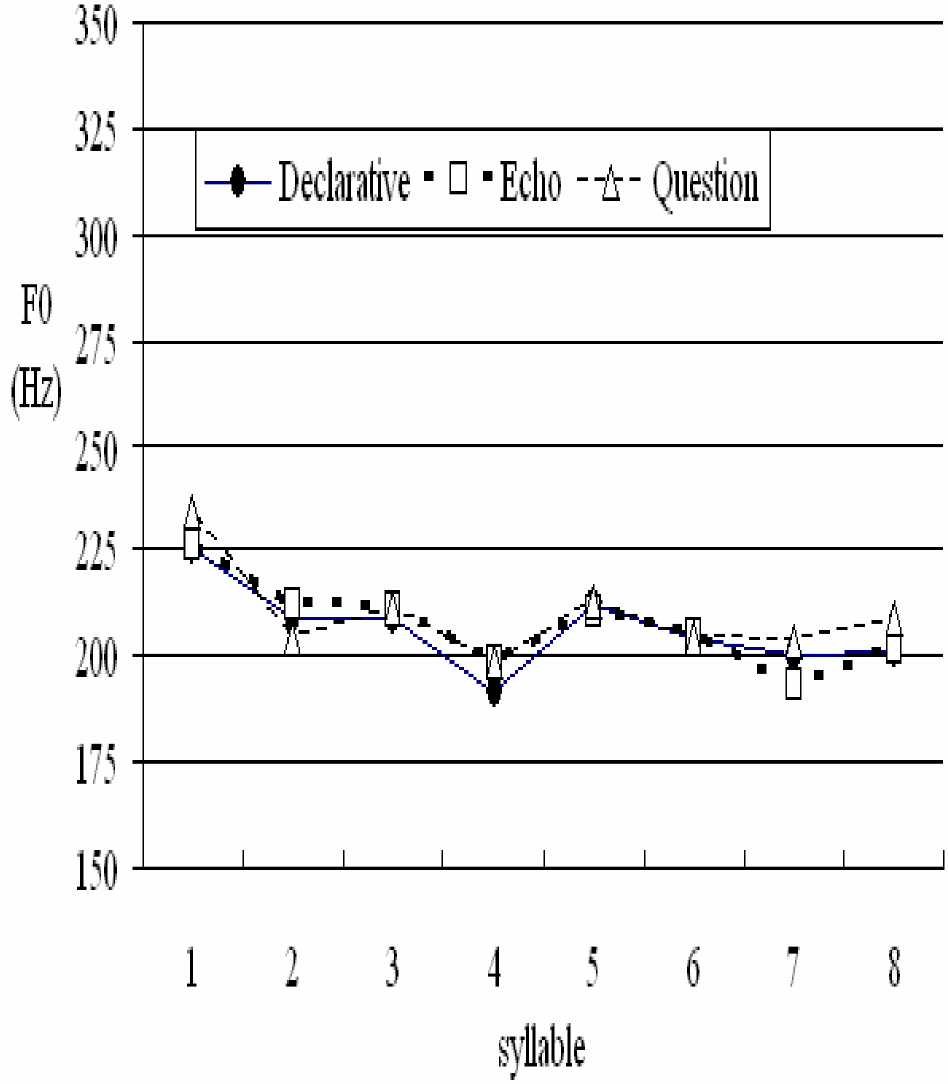


Table 31: *Comparison Of Initial Syllable Pitch Values Across Utterance Types And Speaker Groups In The Tone 4 Condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	587680.491	1	587680.491	55.401	***
Error (S*A)	156975.140	148	1060.643		
Within Subjects					
Pitch (B)	353356.190	1.231	28729.188	28.031	***
Pitch * Group(B*A)	8747.767	1.231	7108.126	6.935	.006
Error (B* sub*w*groups)	186675.830	182.139	1024.907		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 32.

Table 33: *One-Way ANOVA Summary Of The Simple Main Effect Of Pitch In The Tone 4 Condition (QP1 - DP1, EP1 - QP1, EP1 - DP1)*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
QP1_DP1	Between Groups	17278.52	1	17278.520	21.494	.000
	Within Groups	118972.9	148	803.871		
	Total	136251.5	149			
EP1_QP1	Between Groups	6159.370	1	6159.370	8.524	.004
	Within Groups	106946.7	148	722.613		
	Total	113106.0	149			
EP1_DP1	Between Groups	44070.37	1	44070.368	55.401	.000
	Within Groups	117731.4	148	795.482		
	Total	161801.7	149			

In this case, significant differences in utterance-initial pitch among the three utterance types were found between native and non-native speaker groups in each of the three conditions (QP1-DP1 F=21.494, p<.001; EP1-QP1 F=8.524, p<.01, EP1-DP1 F=55.401, p<.001). We see, however, that the effect size is not nearly as large when native speakers use utterance-initial pitch to distinguish between yes/no and echo questions.

Table 33: *Means And Standard Deviations For Native (Group 1) And Non-Native (Group 2) Speakers In The Tone 4 Condition*

Descriptives

		N	Mean	td. Deviatio	Std. Error	% Confidence Interval f Mean		Minimum	Maximum
						ower Bound	Jpper Bound		
QP1_DP	1	75	29.311	34.891	4.029	21.283	37.338	-66.8	170.1
	2	75	7.845	19.757	2.281	3.300	12.391	-27.8	71.1
	Total	150	18.578	30.240	2.469	13.699	23.457	-66.8	170.1
EP1_QP	1	75	5.965	29.459	3.402	-.813	12.743	-108.3	94.8
	2	75	-6.851	24.029	2.775	-12.379	-1.322	-84.7	43.2
	Total	150	-.443	27.552	2.250	-4.888	4.003	-108.3	94.8
EP1_DP	1	75	35.276	32.191	3.717	27.870	42.682	-15.4	156.9
	2	75	.995	23.552	2.720	-4.424	6.414	-112.5	63.9
	Total	150	18.135	32.953	2.691	12.819	23.452	-112.5	156.9

Table 34 presents ANOVA summary results for pitch value averaged across syllables 1-8 for Q, D and E (QA, DA, EA) in the Tone 4 condition.

Table 34:*Comparison of utterance-average pitch values across utterance types and speaker groups in the tone 4 condition (QA - DA, EA - QA, EA - DA)*

SV	SS	DF	MS	F	P VALUE
Between Subjects					
Group(A)	74530.170	1	74530.170	117.974	***
Error (S*A)	93498.909	148	631.749		
Within Subjects					
Pitch (B)	19221.011	1.383	13900.667	41.295	***
Pitch * Group(B*A)	11224.416	1.383	8117.517	24.115	***
Error (B* sub*w*groups)	68886.717	204.646	336.615		

*** p value <.001

A significant interaction effect between speaker group and pitch was found, so simple main effects were calculated, the results of which appear in Table 35.

Table 35: *One-Way ANOVA Summary Of The Simple Main Effect Of Pitch In The Tone 4 Condition (QA - DA, EA - QA, EA - DA)*

ANOVA

		平方和	自由度	平均平方和	F 檢定	顯著性
Q_D	組間	22231.246	1	22231.246	81.861	.000
	組內	40192.792	148	271.573		
	總和	62424.038	149			
E_Q	組間	7625.713	1	7625.713	21.675	.000
	組內	52068.652	148	351.815		
	總和	59694.365	149			
E_D	組間	55897.628	1	55897.628	117.974	.000
	組內	70124.182	148	473.812		
	總和	126021.809	149			

Table 36: *Means and standard deviations for native (group 1) and non-native (group 2) speakers in the Tone 4 condition*

描述性統計量

		個數	平均數	標準差	標準誤	均數的 95% 信賴區		最小值	最大值
						下界	上界		
Q_D	1	75	7.764833	1674522	4442068	2.894645	2.635022	-27.1125	107.2125
	2	75	3.416667	7511437	1259651	1.173132	5.660201	-24.9000	31.4875
	總和	150	5.590750	4683481	6712336	2.288370	3.893130	-27.1125	107.2125
E_Q	1	75	1.778000	6536464	7312878	5.335790	7.220210	-84.1875	70.5750
	2	75	2.482167	0056424	3862922	5.244414	.280081	-29.5625	34.9625
	總和	150	4.647917	0158269	6342854	1.418547	7.877286	-84.1875	70.5750
E_D	1	75	39.54	29.457	3.401	32.77	46.32	-19	119
	2	75	.93	8.941	1.032	-1.12	2.99	-28	31
	總和	150	20.24	29.082	2.375	15.55	24.93	-28	119

4.5 General Discussion of Results

Here, as for the L1 English, pitch setting is generally higher for questions than for declaratives. Unlike English, this is true for both question types even at the beginning of the sentence. (Mean F0 for initial syllable, averaged over all items of each illocutionary type: declarative = 257.8 Hz; yes/no question = 281.6 Hz; echo question = 282.4 Hz.). Through the remainder of the sentence, F0 is somewhat higher for echo questions than for yes/no questions, and it is higher for yes/no questions than for declaratives (mean F0 across all syllables, averaged over all items of each illocutionary type: declarative = 228.2 Hz; yes/no question = 251.8 Hz; echo question = 261.9 Hz.. This differentiation is more constant throughout the sentence than it is in the L1 English utterances, in accord with descriptions in the literature as noted in Chapter 3. Moreover, these L1 Mandarin speakers show little sign of a topic-shift effect in the form of sentence-initial F0 rising in the yes/no question (the first sentence of mini-discourse) above that of the declarative and echo question. It might have been expected, since Tseng (2003) has reported downdrift and pitch resetting effects produced within larger discourse units in Mandarin by Mandarin native speakers. It is possible that in the context of the current study, topic shift manifests itself (very mildly compared with English) in the fact that the initial syllable of the Mandarin utterances is the only one in which the f0 for the yes/no question is as high as that for the echo question.

It is clear that the L1-English learners of Mandarin failed utterly to produce the F0 distinctions that distinguish illocutionary type in Mandarin. Nor did they employ the contours of their native English, with raised F0 for interrogatives through the mid-part of the sentence and a strong final interrogative rise. Apart from a hint of a difference at the end of the sentence (discussed further), there is no systematic difference at all between questions and declaratives with respect to their contour or their F0 range. (Initial syllables, averaged over participants and item types: declaratives 225.2 Hz, yes/no questions 232.2 Hz, echo questions 228 Hz. Mean F0 across all syllables, averaged over all items of each illocutionary type: declaratives = 207.1 Hz; yes/no questions = 208.9 Hz; echo questions = 207.8 Hz.).

Thus, L2 speakers' productions were apparently driven entirely by lexical tone shapes (which were common to the three illocutionary types in each dialogue). A detailed analysis of the L2 lexical tones will not be given here, but it is worth noting that they were approximately correct, though clearly non-native to the ear. Syllable tone shapes were similar to those of the native speakers, though produced within a much more compressed pitch range than native speakers'.

It might have been anticipated that L2 speakers exhibit L1 transfer by overriding a required lexical rising tone (Tone 2) at the end of a declarative and producing instead a falling tone more like Tone 4; and that they might replace a lexically required falling Tone 4 with a rising Tone 2 at the end of an interrogative. The syllable-average data in Figures 6a-f show little evidence of this, since they do not show within-syllable trajectories, but the 20ms.-sampled F0 data in Figures 7-9 reveal interesting dynamic patterns within the final (8^{th}) syllable and also within the *ma* particle of the yes/no questions. (These are the 4^{th} and 5^{th} syllables respectively in the graphs that follow, since these show just the final four syllables of the sentence, plus the *ma* particle in the yes/no questions). Figures 7 and 8 show group mean pitch measurements (not normalized), extracted at 20 ms. intervals over the final four syllables of three Mandarin illocutionary types. These measurements were averaged across participants and tokens of the item subtypes: Figures 7a-c display Mandarin utterances produced by L1 Mandarin speakers in the Tone 1, Tone 2, and Tone 4 conditions, respectively. Figures 8a-c provide the same data for L1 English speakers.

Figure 7a: F0 for the final four syllables (plus "ma" for yes-no questions) of Tone 1 final utterances, by L1 Mandarin speakers.

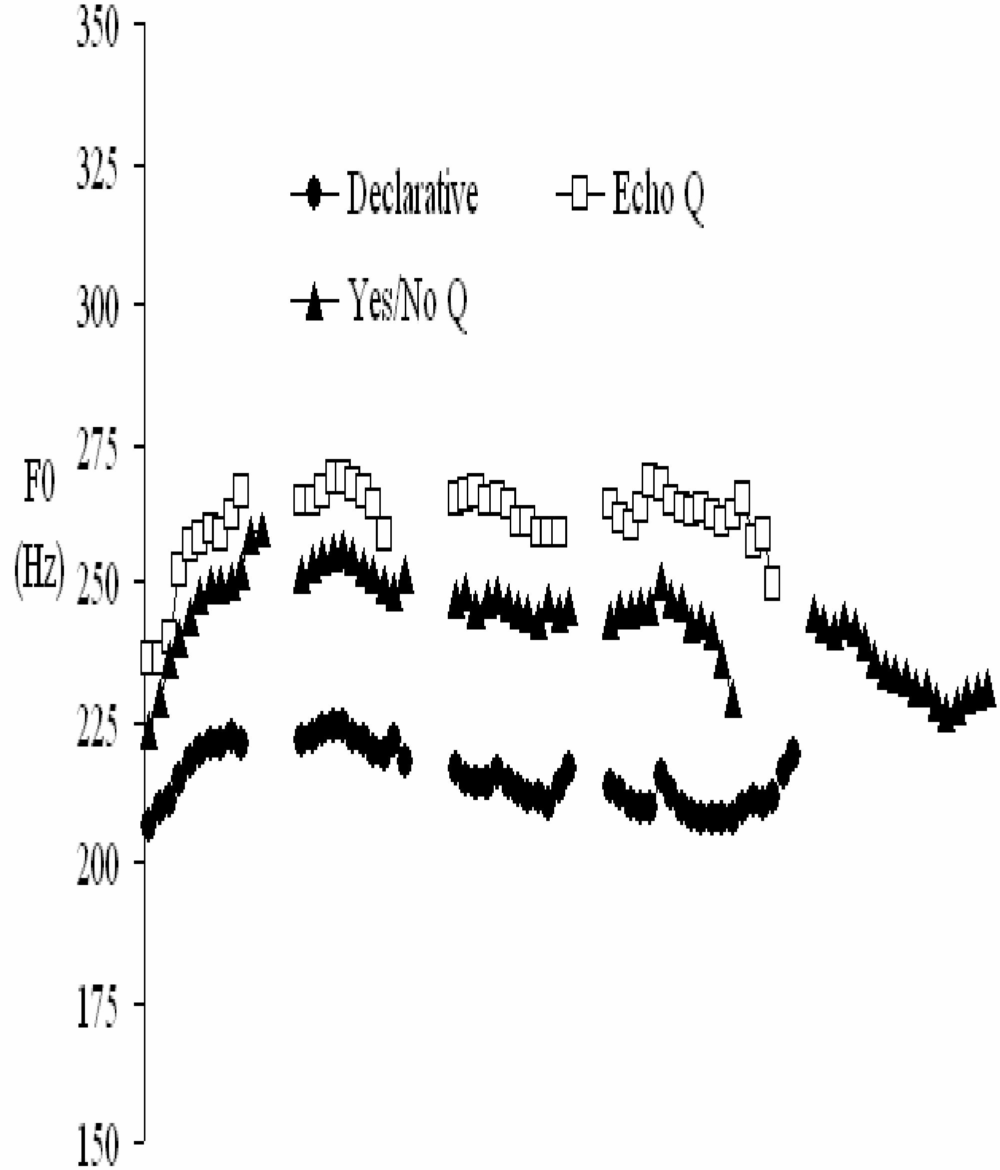


Figure 7b: F0 for the final four syllables (plus "ma" for yes-no questions) of Tone 2 final utterances, by L1 Mandarin speakers.

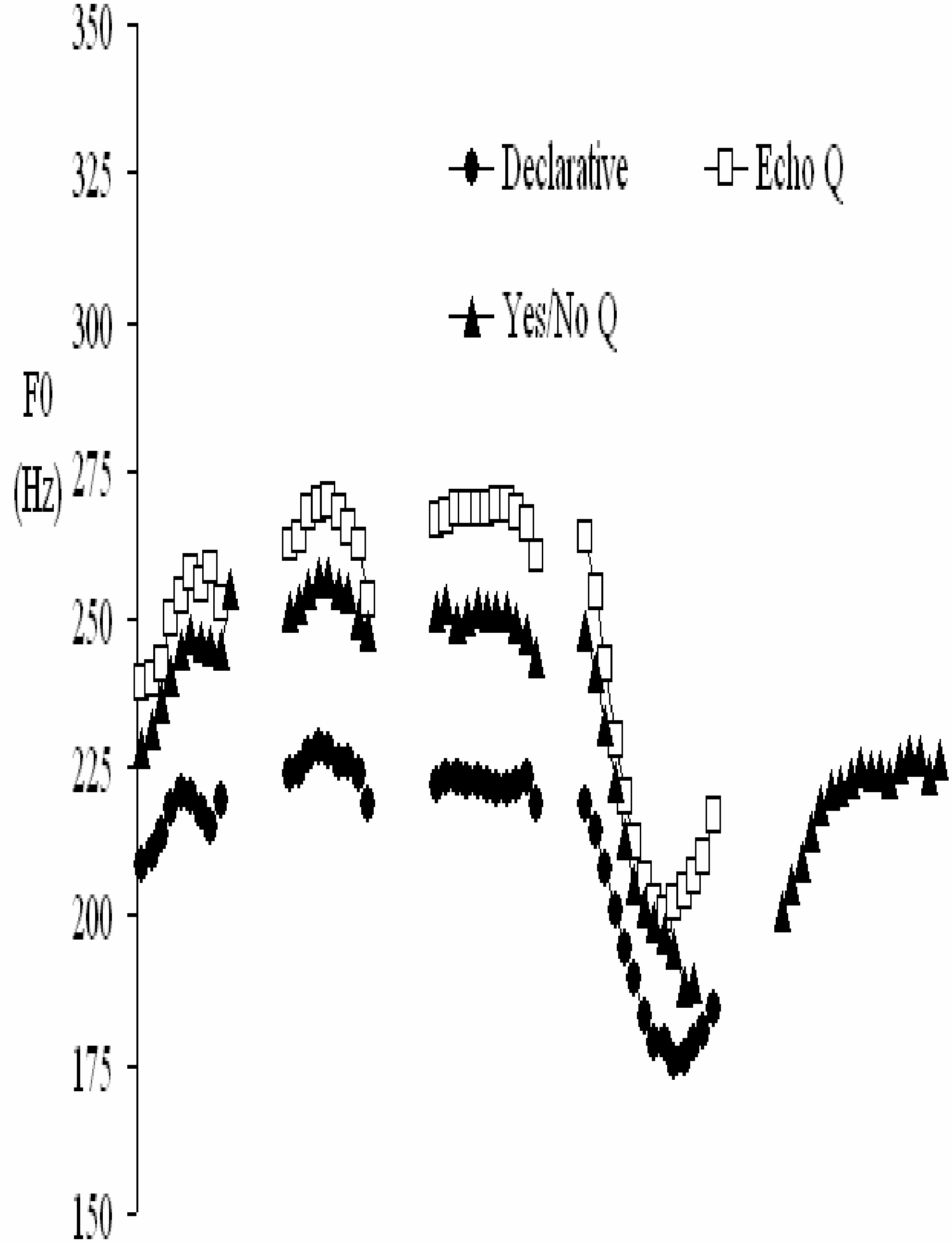


Figure 7c: F0 for the final four syllables (plus "ma" for yes-no questions) of Tone 4 final utterances, by L1 Mandarin speakers.

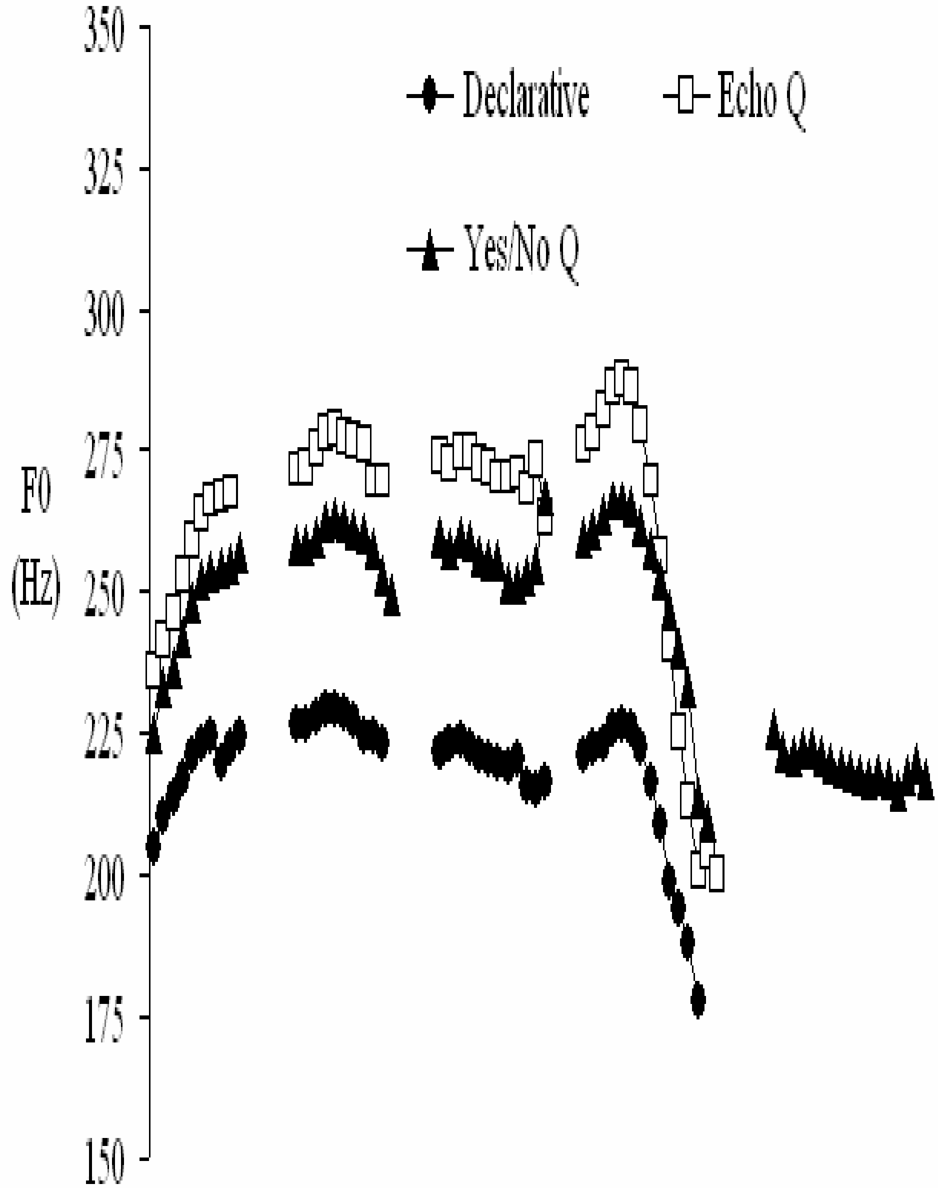


Figure 8a: F0 for the final four syllables (plus "ma" for yes-no questions) of Tone 1 final utterances, by L1-English speakers.

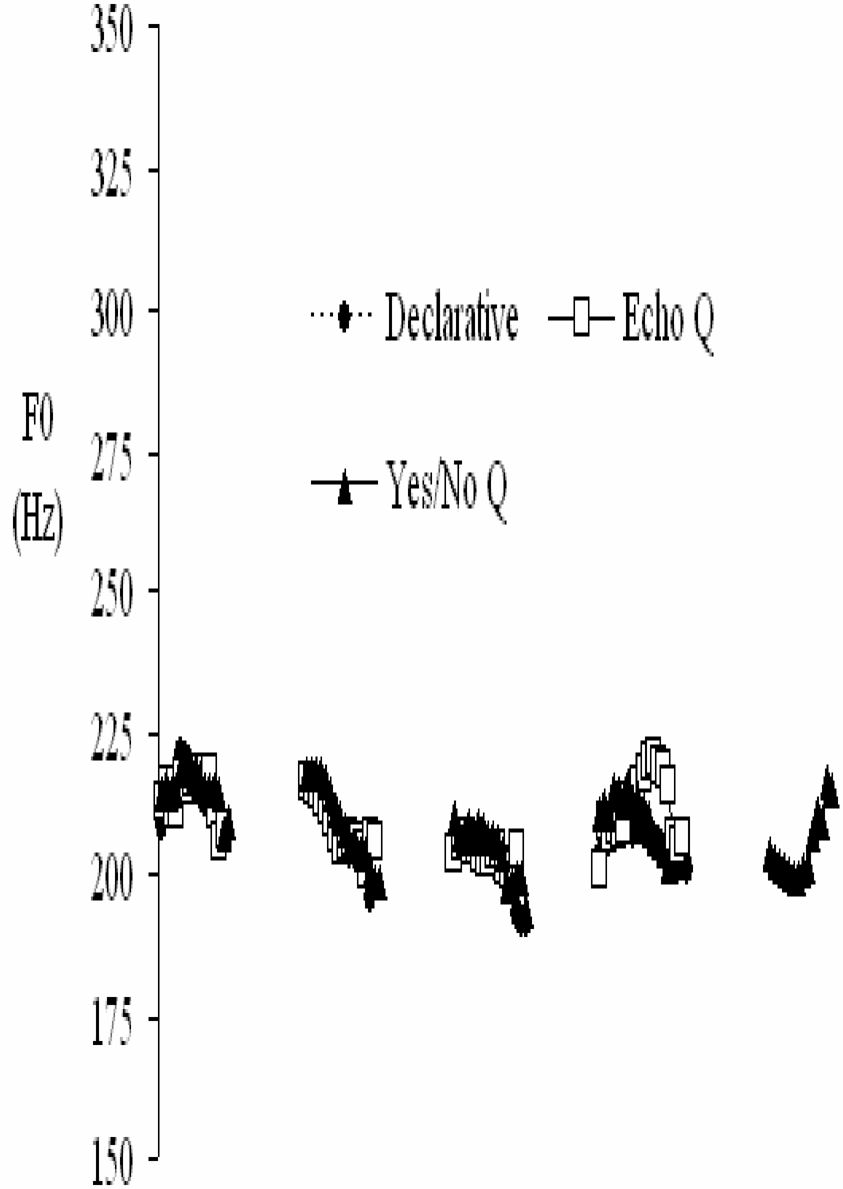
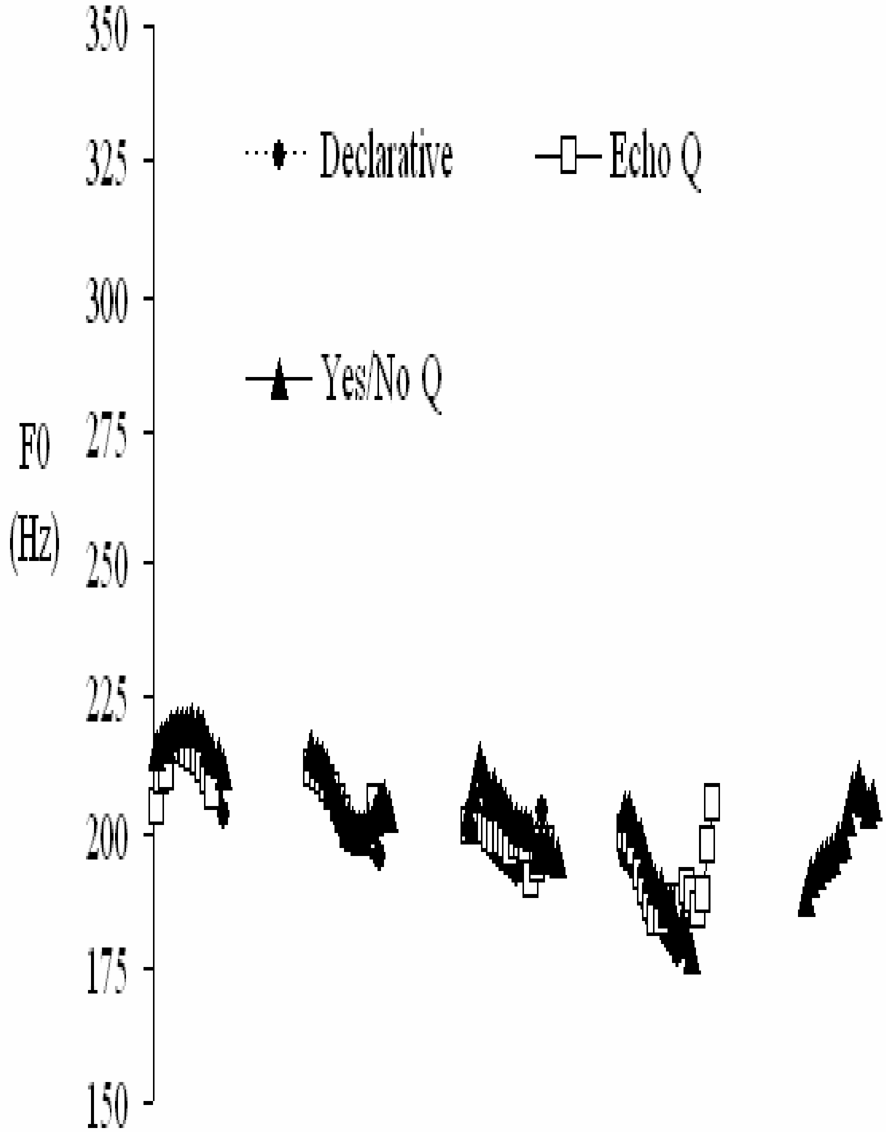


Figure 8b: F0 for the final four syllables (plus "ma" for yes-no questions) of Tone 2 final utterances, by L1-English speakers.



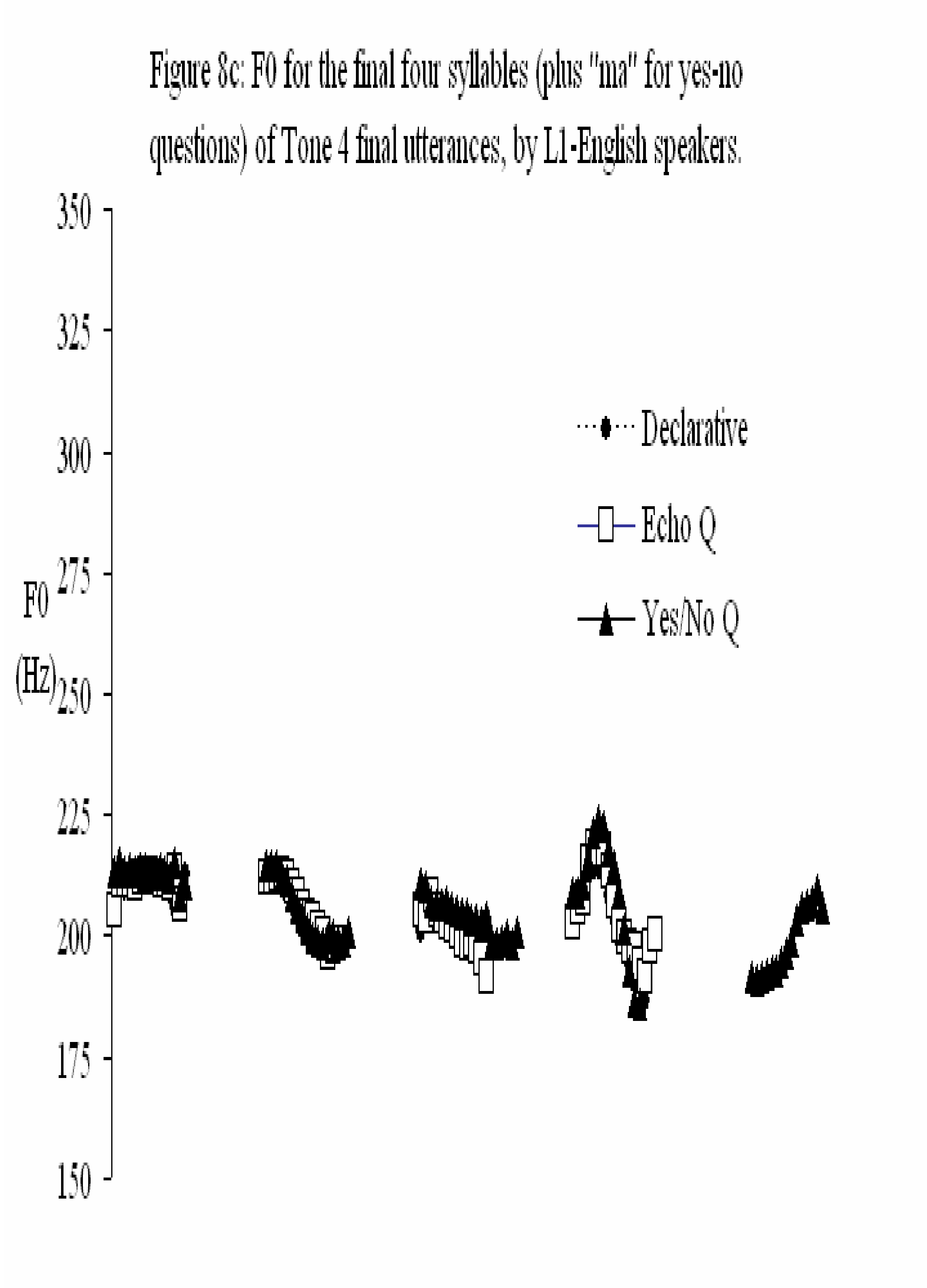


Figure 8c: F0 for the final four syllables (plus "ma" for yes-no questions) of Tone 4 final utterances, by L1-English speakers.
F0 (Hz)
350
325
300
275
250
225
200
175
150
Declarative
Echo Q
Yes/No Q

In all three tone conditions, L2 participants produced a final rise in the yes/no question, which was confined to the *ma* particle. This is not true of L1 Mandarin speakers, whose productions were more tailored to the tone that preceded *ma*. They performed tonal coarticulation from the endpoint of the previous syllable to rise or fall to a level pitch on the *ma* particle. In general, native Guoyu Mandarin speakers start the *ma* particle at the ending pitch of whichever tone preceded it, then rise or fall to a steady mid-level tone (W.Y. Chiang, personal communication). In the echo question, which has no *ma* particle, there appears to be some tendency for the L2 speakers to extend the last syllable of the sentence and to produce a slight rise in F0. In short, it seems possible that the native English speakers were transferring, at least partially, the English interrogative final rise to the final syllable of a Mandarin interrogative sentence to the particle *ma* if it is present and otherwise to the utterance-final syllable of echo questions. This F0 rise does not seem to extend backward into preceding syllables as it does in some cases in L1-English (see preceding data), but this may be for lack of a clearly located nuclear accent to initiate it. This final rise is apparent for Tones 2 and 4, while for the Tone 1 echo question the rise does not extend to the end of the syllable.

Thus, some phenomena in the L2 productions of Mandarin could be interpreted as transfer of English question intonation. But more striking is the lack of transfer of major sentence and discourse-level prosodic features.

4.6 Evaluation of hypotheses

4.6.1 Hypothesis 1

Hypothesis 1 proposed that L2 speakers would produce global pitch in L2 at the same height for all sentences in the mini-discourse, whereas in L1 they would use global pitch setting to differentiate the various discourse positions and illocutionary force of the three utterances. Hypothesis 1 was confirmed. For Mandarin production of native English speakers, just as for English production of native Mandarin speakers, there was a remarkable absence of the expected broad rising of F0 for questions compared with declaratives, even though this is a characteristic that is common (details aside) to both languages.

4.6.2 Hypothesis 2

Hypothesis 2 proposed that L2 speakers would confine illocutionary prosody (e.g., declarative falls and interrogative rises in English) to utterance-final syllables. In their production of L1, in contrast, they would demonstrate evidence of longer-range planning to signal differences in illocutionary type. Hypothesis 2 was also confirmed. Native Mandarin speakers tended to limit their utterance-final rise in questions to the final syllable, regardless of the position of nuclear stress within the sentence. There is no difference in the timing of the rise to the interrogative high boundary tone in the antepenultimate stress items versus the penultimate stress and final stress items. Thus, it appears that L1 Mandarin speakers of English assimilate the phrasal contours of English to the syllable-locked contours of the lexical tones in their L1. This syllable-bounded sentence-final pitch excursion plausibly does represent transfer from the L1.

In Mandarin, native English-speakers tended to produce a final rise in the yes/no question, confined to the *ma* particle. In echo questions, which have no *ma* particle, there appears to be some tendency for the L2 speakers to extend the last syllable of the sentence and to produce a slight rise in F0. In short, it seems possible that native English speakers were transferring, at least partially, the English interrogative final rise to the final syllable of a Mandarin interrogative sentence: to the particle *ma* if it is present, and otherwise to the utterance-final syllable of echo questions. Statistical analyses of these results were not generated, since they were not nearly as striking as L2 speakers' absence of central aspects of the target language's utterance-level prosody, as well as their overwhelming lack of transfer of global L1 prosodic contours into the L2.

4.7 Summary of Findings

Native language data for both English and Mandarin confirm some widely noted phenomena—the final F0 rise in English yes/no and echo questions, and a global rising of pitch throughout Mandarin yes/no and echo questions. L2 data exhibit a striking absence, in the productions of both groups of L2 learners, of central aspects of the target language's utterance-level prosody. In addition to this failure to match target language norms, a resounding lack of transfer of global L1 prosodic contours into the L2 was observed; there was no sign of transfer from Mandarin into English, or from English into

Mandarin. Transfer effects, however, were observed with regard to local phenomena within the final syllable. L1-Mandarin speakers compressed illocution-related pitch excursions in English into the final syllable only, in the fashion of pitch excursions that realize Mandarin lexical tones. L1-English speakers imported a final-syllable rise on the final syllable in Mandarin questions.

In many ways, the prosodic markings of illocutionary force in L1 English and L1 Mandarin are more similar to each other than either is to its L2 counterpart. Oddly, this implies that transferring the L1 prosody into the L2 might have produced a more native-like effect than the lack of positive transfer[9] that actually occurred. Chapter 5 will discuss various possible explanations for why positive transfer was not more extensive, and for the striking discrepancies observed between production of global and local prosodic phenomena.

---

[9] 'Positive transfer' refers to transfer of an L1 grammar feature which is compatible with the L2 grammar, as opposed to 'negative transfer', which refers to transfer of an L1 grammar feature which is incorrect for the L2.

Chapter 5

Discussion and Conclusion

5.0 Introduction

These data show that L2 speakers fail to produce the global pitch setting cues that native speakers use to differentiate illocutionary sentence types and to link utterances in discourse. Brazil (1997) proposes relative pitch onset level as the most salient cue for interpretation of English prosodic units larger than the individual utterance. Brazil describes this unit as a *pitch sequence*: a stretch of consecutively lower pitch settings embedded within a discourse, which serve to delimit longer sections of speech. Tseng (2003) has reported similar downdrift and pitch resetting effects produced within larger discourse units in Mandarin by Mandarin native speakers.

Pickering (2004) has also demonstrated L2 speakers' failure to produce larger units of prosody. In a study of 10 L1 Mandarin speakers' production of English academic lectures, Pickering found that while native speakers showed evidence of a hierarchical system of prosodic units to create semantic cohesion within paragraphs and signal aspects of information structure, L2 speakers were unable to manipulate pitch setting to create paragraph-level intonational units. No equivalent study has yet investigated whether this is also true in non-native Mandarin production.

In these data, both groups of L2 learners failed to produce global pitch setting, which is a central aspect of each target language's discourse-level prosody. In addition, while some negative transfer of local phenomena (e.g., the timing of rises and falls) was observed, there was no sign of global pitch setting transfer from Mandarin into English, or from English into Mandarin. Various possible explanations for these findings will be discussed in this chapter: (a) on-line processing difficulties; (b) lack of robust positive input in both natural language and instructional materials; and (c) possible confounding effects of a slower speech rate, which may precipitate the need for L2 speakers to divide utterances into a larger number of prosodic units. Finally, a range of hemispheric lateralization hypotheses will be discussed. These hypotheses predict that differences in neurological processing between L1 and L2 may contribute to difficulties in L2 attainment of global prosody.

5.1 On-Line Processing Resources

These data suggest that L2 learners may sacrifice global prosody to focus on local prosody. One possibility is that the processing resources needed to concentrate on accurate production of local phenomena, (e.g., segmental contrasts, lexical stress), have depleted the resources that might have been allocated to control of global phenomena. It is also possible that the additional demands of on-line processing presented by the L2 production task may have obscured participants' underlying prosodic competence. It could be that even for global prosody, they had the correct principles in their acquired L2 grammar but were unable to deploy them on-line due to resource limits or due to the lexical decoding demands presented by a task which required them to read aloud.

Perception is generally found to be less demanding than production, so further research would be useful to investigate how well L2 learners at this level perform in a perceptual task; for example, a listening task requiring discrimination between the three illocutionary types based on recordings from which lexical and syntactic cues have been excised by low-pass filtering[10].

5.2 Insufficient evidence in the L1 input

There is some evidence that L1 speech input provides learners with a less clear model for global than for local prosodic phenomena, especially if, as is sometimes suggested, global prosody is more graded, or exhibits more optionas than lexical stress and tone distinctions (Bolinger, 1988; van Bezoojien, 1995). The across-the-board pitch range difference between questions and declaratives in Mandarin is a distinction that is inherently difficult to detect except in an extended interaction with the same speaker. In fact, Chang (1998) found that even native speakers experienced difficulty in distinguishing Mandarin echo questions from statements when they were presented in isolation in a forced-choice perception task.

Related to the optionality/gradability of some phrase-level prosodic phenomena, it has been suggested that when L2 speakers are uncertain, they prefer to be neutral rather than to risk making a positive error (Mennen, 1998). This offers a possible explanation both for the limited prosodic transfer observed and for the general reduction in f0 range in L2 speech documented in Chapter 4. Neutrality is easier to achieve in prosody (even some L1 speakers tend toward monotone pronunciation) than it is for segmental contrasts, where small differences in production have considerable effect on perception. Thus, in the case of prosody, L2 speakers' lack of confidence with respect to target patterns may manifest itself in the use of a narrow pitch range, in order to avoid the risk

[10] I thank an anonymous reviewer of an earlier version of this manuscript for this suggestion.

of producing noticeably incorrect pitch variation. Pickering's (2004) study showed that the absence of clear pitch sequence structuring in L2 English data was exacerbated by the participants' overall narrower pitch range. Shen (1989) also proposed that many L1 English speakers' lexical tone errors in Mandarin might be attributed to their reduced pitch range in L2 production.

5.3 Lack Of Explicit Instruction

A different but very likely possibility is that many L2 learners know only what their teachers or textbooks tell them, and that few language teachers are attuned to many of the details of utterance- and discourse-level prosody. The rise in pitch at the end of an English question is often mentioned, but the proper timing of that rise probably is not, nor the global rising of pitch in Mandarin questions. These facts are not widely known among linguists, and so it would not be surprising if they have not made their way into mainstream language pedagogy[11]. To substantiate this pedagogic explanation will require investigation of individual differences and of how they relate to the mode of acquisition: teacher-dependent in a classroom setting, or in more natural interactive contexts. Intervention studies, in which prosodic facts and exercises are emphasized in instruction, could also be informative.

[11] Some EFL textbooks and teacher training volumes written by linguists (e.g. Wong [1987], Pennington [1996]) provide more comprehensive information on prosody, but possibly these works have not found their way into the majority of EFL classrooms. Moreover, no mention of global prosody is made in any of the Mandarin Chinese textbooks the author could find.

5.4 Speech Rate

Phonological production was also found to be slower for these L2 speakers than for native speakers. Speech rate is a factor known to influence prosodic phrasing even in native language production: the slower the speech, the more prosodic breaks there tend to be, so the fewer syllables (mora, etc.) there are in each prosodic phrase. It is possible, therefore, that an L2 speaker needs to divide sentences into two or more prosodic phrases even if they are short enough for a native speaker to encompass within a single prosodic phrase (see discussion of Hewings [1995] in Chapter 2).

Prosodic phrase boundaries are associated with pitch modulations, often a sharp resetting of the F0 from low to high or vice versa, though language-specific details vary. Thus, the slower speed of L2 speech could bring about, indirectly, complex changes in the F0 contour. These data clearly confirm the L1-L2 difference in speech rate, but they do not contain any decisive evidence of an L2 tendency to divide sentences into two prosodic phrases. This would be expected to show up in increased duration of the syllable preceding the break, but no consistent locus of a duration increase was found in instrumental analysis of these recordings.

5.5 Hemispheric Lateralization Hypotheses

Recent FMRI studies point in the direction of differences in the hemispheric lateralization of L1 and L2 processing, which may especially relate to a possible discrepancy between the L2 attainment of local versus global prosody. Two major hypotheses have emerged to explain hemispheric lateralization: a cue dependent hypothesis and a hypothesis which proposes a hybrid of cue and task dependency. Each hypothesis will be described separately, followed by a discussion of which model is compatible with these data.

5.5.1 Cue Dependent Hypothesis

The cue dependent hypothesis claims that speech processing involves neurobiological mechanisms specialized for particular properties of the speech signal (e.g., duration of the acoustic signal). However, these mechanisms do not consider the linguistic function of the cue in question. Poeppel (2003) reviews evidence from a variety of sources, which indicate that there are two different time frames that characterize prosodic phenomena, and that they exhibit different lateralization patterns: left-hemisphere auditory areas preferentially extract information from short temporal windows, while those in the right hemisphere preferentially extract information from longer temporal windows.

5.5.2 Task- And Cue-Dependent Hybrid Hypotheses

Task-dependent hypotheses claim that language-specialized neural mechanisms are employed in speech processing; these mechanisms are chosen according to the specific demands of the linguistic task Hybrid hypotheses propose interaction between the left and right hemisphere during both processing of linguistic and other kinds of acoustic signals. Two such models will be discussed.

One model (Gandour, et al., 2004) proposes that complex sounds are first processed in the right hemisphere: If linguistic processing is required, the stimulus is then transferred to domain-specific areas in the left hemisphere. Gandour, et al. (2003) found that "the pitch contours associated with Mandarin tones are processed in the LH (left hemisphere) by Chinese listeners, whereas pitch contours associated with intonation are processed bilaterally with a RH (right hemisphere) preference. English listeners, on the other hand, exhibited bilateral activation for lexical tones, [and] right-sided activation for intonation" (p.321). They interpret these data to suggest that temporal integration windows are linked to language-specific features of local and global prosody, which are processed differently by native and naïve listeners (i.e. listeners who had no knowledge of Chinese). Friederici and Alter's (2004) dual-pathway model claims that acoustic signals that have either a lexical or syntactic function are processed in the left hemisphere, but that acoustic signals having utterance- or discourse-level functions, such as intonation are processed in the right hemisphere.

5.5.3 The Current Data

In the absence of FMRI data, only general claims can be made to connect these data to lateralization hypotheses. Moreover, these hypotheses were formulated to model perception of linguistic acoustic signals, so extension of their predictions to interpret production data must be mitigated by the confounding factors mentioned earlier in this chapter. Bearing these stipulations in mind, these data appear consistent with Poeppel's cue-dependent hypothesis. L2 speakers produced acoustic signals with both lexical and utterance/discourse-level functions; however, they confined their productions to syllable-length windows of time. L1 Mandarin speakers compressed illocution-related pitch excursions in English into the final syllable only, in the fashion of pitch excursions that realize Mandarin lexical tones. L1-English speakers imported a final-syllable rise on the final syllable in Mandarin questions. Moreover, there was no sign of global pitch setting transfer from Mandarin into English, or from English into Mandarin, which suggests the absence of long-range intonational planning, rather than a separation of lexical and discourse-related functions.

5.6 Limitations of the Study

The experimental task in this study requires participants to read pre-scripted materials aloud. Acoustic differences between the prosody of read speech and spontaneous speech have been debated (Cruttenden, 1997); and possible effects of these differences have been taken into account in discussion of the results. However, in order to allow for reliable, systematic cross-item comparisons of the prosodic phenomena of theoretical interest, the experimental materials had to be strictly controlled for utterance type, number of syllables, and segmental content. Thus, completely scripted experimental materials were necessary; it was not possible to work with spontaneous speech corpora, or even elicited utterances. To ensure the most natural output possible under such constraints, the experimental sentences were arranged in mini-dialogues (as question, declarative answer, and echo question, all with the same propositional content), and the subjects were encouraged to use an everyday speaking style.

The order of presentation of materials may also have influenced the results. Participants first performed the task using materials in their first language, followed by a break, then the second language materials. This presentation method was chosen so that it would be maximally clear to the participants how the sentence sets were patterned, so that they would be acclimated to the task before they attempted the more challenging second language materials. None of the items in one language were translation equivalents of those in the other language, or semantically related to them in any way. Nevertheless, it must be acknowledged that this strategy encourages native language transfer more than the reverse order would have done. In future research, two different presentation orders could be used and their outcomes compared.

5.7 Conclusion

In summary, the results of this study suggest that a combination of on-line processing difficulties, lack of robust positive input, and possible differences in neurological resources underlying L1 and L2 processing contribute to difficulties in L2 attainment of global prosody. Neurolinguistic investigation of the independent control of global and local prosody seems a promising avenue to pursue for future L2 production studies.

Appendix A: Experimental Materials[12]

Set 1: Mandarin materials for L1 English speakers[13]

PRACTICE ITEMS

P1

火車常常 遲 到 很 久 嗎 ？
huo che chang chang chi dao hen jiu ma
(Is the train often late?)[14]
是。 火車常常 遲 到 很 久。
Shi. huo che chang chang chi dao hen jiu
真的 ？ 火車常常 遲到 很 久 ？
zhen de huo che chang chang chi dao hen jiu

P2
他 講 故事 給 大家 聽 嗎 ？
ta jiang gu shi gei da jia ting ma
(Does he tell stories for everyone to hear?)
是。他 講 故事 給 大家 聽。
shi ta jiang gu shi gei da jia ting
真的 ？ 他 講 故事 給 大家 聽 ？
zhen de ta jiang gu shi gei da jia ting

[12] In the materials set presented to participants, the pinyin gloss appears directly under the relevant syllable, and tone information appears directly above the pinyin. Discrepancies in the number of characters between morphographs and their pinyin glosses make this alignment extremely difficult to reproduce here.
[13] Mandarin materials presented to L1 Mandarin speakers were presented in the same order, but without glosses and translations.
[14] An English translation of the first sentence of each set appeared opposite the Mandarin sentence. Since the semantic content of the two remaining sentences in each set was almost identical to the first, no translations of those were necessary.

HIDDEN PRACTICE ITEM

P3
姐姐 正在 看 漫 畫 書 嗎 ？
jie jie zheng zai kan man hua shu ma
(Is older sister reading a comic book?)
是。 姐姐 正在 看 漫 畫 書。
Shi. jie jie zheng zai kan man hua shu.
真的？ 姐姐 正在 看 漫 畫 書？
zhen de jie jie zheng zai kan man hua shu

Experimental Sentences

BLOCK 1

EP1 (T2)
每 天 早上 葳葳 撈 魚 嗎？
mei tian zao shang weiwei lao yu ma
(Does Weiwei catch fish every morning?)
是。每 天 早上 葳葳 撈 魚。
Shi. mei tian zao shang weiwei lao yu.
真的？ 每 天 早上 葳葳 撈 魚？
zhen de mei tian zao shang weiwei lao yu

E2 (T1)

妹妹 看 著 媽媽 摸 貓 嗎？
meimei kan zhe mama mo mao ma
(Is little sister watching Mama pet the cat?)
是。妹妹 看 著 媽媽 摸 貓。
Shi. meimei kan zhe mama mo mao
真的？ 妹妹 看 著 媽媽 摸 貓？
Zhen de meimei kan zhe mama mo mao

E3 (T4)
沒 人 發現 妞妞 捏 鹿 嗎？
mei ren fa xian niuniu nie lu ma
(Didn't anyone notice Niuniu pinching the deer?)
是。沒 人 發現 妞妞 捏 鹿。
Shi. mei ren fa xian niuniu nie lu
真的？ 沒 人 發現 妞妞 捏 鹿？
zhende mei ren fa xian niuniu nie lu

BLOCK 2

E4 (T1)
那 本 書 叫"貓 咪 安安 嗎？
nei ben shu jiao mao mi anan ma
(Is this book called "An An the Cat?")
是。 那 本 書 叫"貓 咪 安安。
Shi. nei ben shu jiao mao mi anan
真的？ 那 本 書 叫"貓 咪 安安 ？
zhen de nei ben shu jiao mao mi anan

E5 (T4)
麵 煮 好 時 媽媽 撈麵 嗎？
mian zhu hao shi mama lao mian ma
(When the noodles are cooked, does Mama take them out of the pot?)
是。麵 煮 好 時 媽媽 撈麵。
Shi. mian zhu hao shi mama lao mian
真的？ 麵 煮 好 時 媽媽 撈麵？
zhen de mian zhu hao shi mama lao mian

E6 (T2)
牛 不 走 時 葳葳 拉 牛 嗎？
niu bu zou shi weiwei la niu ma
When the cow won’t walk, does Wei Wei drag the cow?
是。牛 不 走 時 葳葳 拉 牛。
shi niu bu zou shi weiwei la niu
真的？ 牛 不 走 時 葳葳 拉 牛？
zhen de niu bu zou shi weiwei la niu

BLOCK 3

E7 (T4)
那 本 書 叫 貓 咪 安夢 嗎？
nei ben shu jiao mao mi anmeng ma
Is this book called “An Meng the Cat”?
是。 那 本 書 叫 貓 咪 安夢。
shi nei ben shu jiao mao mi anmeng
真的？ 那 本 書 叫 貓 咪 安夢？
zhen de nei ben shu jiao mao mi anmeng

E8 (T1)

蓮花 開 時 妞妞 撈 花 嗎？
Lian hua kai shi niuniu lao hua ma
(When the lotus flowers bloom, does Niuniu pick them?)
是。蓮花 開 時 妞妞 撈 花。
Shi. Lian hua kai shi niuniu lao hua
真的？ 蓮花 開 時 妞妞 撈 花？
zhen de lian hua kai shi niuniu lao hua

E9 (T2)
種 地 瓜 時 媽媽 挖 泥 嗎？
zhong di gua shi mama wa ni ma
是。 種 地 瓜 時 媽媽 挖 泥。
Shi. zhong di gua shi mama wa ni
真的？ 種 地 瓜 時 媽媽 挖 泥？
zhen de zhong di gua shi mama wa ni

BLOCK 4

EP10 (T1)
過 山洞 時 妞妞 彎 腰 嗎？
guo shan dong shi niu niu wan yao ma
(Does Niuniu duck when she enters a cave?)
是。過 山洞 時 妞妞 彎 腰。
shi guo shan dong shi niu niu wan yao
真的？ 過 山洞 時 妞妞 彎 腰？
zhen de guo shan dong shi niu niu wan yao

EP11 (T2)
那 本 書 叫 貓 咪 安妮 嗎？
nei ben shu jiao mao mi anni ma
Is this book called “Anni the Cat”?
是。那 本 書 叫 貓 咪 安妮。
Shi. nei ben shu jiao mao mi anni.
真的？ 那 本 書 叫 貓 咪 安妮？
zhen de? nei ben shu jiao mao mi anni

EP12 (T4)
離開 家 時 媽媽 拎帽嗎？
Li kai jia shi mama ling mao ma
(When she leaves the house, does Mama take her hat?)
是。 離開 家 時 媽媽 拎帽。
Shi. Li kai jia shi mama ling mao
真的？ 離開 家 時 媽媽 拎帽？
zhende? Li kai jia shi mama ling mao

BLOCK 5

EP13 (T2)
姐姐 看 著 葳葳 摸 驢 嗎？
jie jie kan zhe wei wei mo lü ma
(Is big sister watching Wei Wei pet a mule?)
是。 姐姐 看 著 葳葳 摸 驢。
Shi. jie jie kan zhe wei wei mo lü
真的？ 姐姐 看 著 葳葳 摸 驢？
Zhende? jie jie kan zhe wei wei mo lü

EP14 (T4)
工 人 看著 妞妞 挖 路 嗎？
gong ren kanzhe niuniu wa lu ma
(Are the workers watching Niu Niu dig the road?)
是。 工 人 看著 妞妞 挖 路。
Shi. gong ren kanzhe niuniu wa lu.
真的？ 工 人 看著 妞妞 挖 路？
Zhende? gong ren kanzhe niuniu wa lu?

EP15 (T1)
他 撒嬌 時 都 拉 媽媽 嗎？
ta sa jiao shi dou la mama ma
When he acts up, does he always pull on his mother?
是。他 撒嬌 時 都 拉媽媽。
Shi. ta sa jiao shi dou la mama
真的？ 他 撒嬌 時 都 拉媽媽？
zhende? ta sa jiao shi dou la mama?

BLOCK 6

EP16 (T3)
做 鬼 臉 時 葳葳 捏 臉 嗎？
zuo gui lian shi wei wei nie lian ma
(When she makes a face, does Wei Wei pinch her face?)
是。做 鬼 臉 時 葳葳 捏 臉。
Shi. zuo gui lian shi wei wei nie lian.
真的？ 做 鬼 臉 時 葳葳 捏 臉？
Zhende? zuo gui lian shi wei wei nie lian

EP17 (T3)
妹妹 看著 安安 摸 馬 嗎？
mei mei kan zhe an an mo ma ma
(Is little sister watching An An pet a horse?)
是。 妹妹 看著 安安 摸 馬。
Shi. mei mei kan zhe an an mo ma
真的？ 妹妹 看著 安安 摸 馬？
Zhende? mei mei kan zhe an an mo ma

EP18 (T3)
看 小 字 時 媽媽 瞇 眼 嗎？
Kan xiao zi shi ma ma mi yan ma?
When she reads small characters, does Mama squint?
是。看 小 字 時 媽媽 瞇 眼。
Shi. Kan xiao zi shi ma ma mi yan
真的？看 小 字 時 媽媽 瞇 眼？
Zhende? Kan xiao zi shi ma ma mi yan

FATIGUE ITEMS

P4
他 不 想 去 上 數 學 課 嗎？
ta bu xiang qü shang shu xue ke ma
(Doesn't he want to go to math class?)
是。 他 不 想 去 上 數 學 課。
Shi. ta bu xiang qü shang shu xue ke
真的？ 他 不 想 去 上 數 學 課？
zhen de ta bu xiang qü shang shu xue

P5
外婆 要 去 市 場 買 菜 嗎？
wai po yao qü shi chang mai cai ma
(Does Grandma want to go to the market to buy vegetables?)
是。 外婆 要 去 市 場 買 菜。
Shi. wai po yao qü shi chang mai cai
真的？外婆 要 去 市 場 買 菜？
zhen de wai po yao qü shi chang mai cai

Set 2: English materials presented to L1 English speakers[15]

PRACTICE ITEMS

P1

Does Brenda want to study abroad?
Yes. Brenda wants to study abroad.
Really? Brenda wants to study abroad?

P2

Does Barbara sing in the shower?
Yes. Barbara sings in the shower
Really? Barbara sings in the shower?

HIDDEN PRACTICE ITEM

P3

Does he want to eat an ice cream cone?
Yes. He wants to eat an ice cream cone.
Really? He wants to eat an ice cream cone?

BLOCK 1

EP1(2-syll)

Did Monica eat a raw onion?
Yes. Monica ate a raw onion.
Really? Monica ate a raw onion?

E2 (1-syll)

Does the kitchen have a yellow wall?
Yes. The kitchen has a yellow wall.
Really? The kitchen has a yellow wall?

[15] The English materials presented to L1 Mandarin speakers were presented in the same order, accompanied by a Mandarin translation of the first sentence of each set opposite each English sentence. Since the semantic content of the two remaining sentences in each set was almost identical to the first, no translations of those were necessary

E3 (3-syll)

Does Susanna have an enemy?
Yes. Susanna has an enemy.
Really? Susanna has an enemy?

BLOCK 2

E4 (2-syll)
Did Marianna win some money?
Yes. Marianna won some money.
Really? Marianna won some money?

E5 (3-syll)

Does the cloud look like an animal?
Yes. The cloud looks like an animal.
Really? The cloud looks like an animal?

E6 (1-syll)

Does William have to learn a new role?
Yes. William has to learn a new role.
Really? William has to learn a new role?

BLOCK 3

E7 (3-syll)

Does the T.V. have an aerial?
Yes. The T.V. has an aerial.
Really? The T.V. has an aerial?

E8 (1-syll)

Does this train run on a main line?
Yes. This train runs on a main line.
Really? This train runs on a main line?

E9 (2-syll)

Did Jerry see a wooly llama?
Yes. Jerry saw a wooly llama.
Really? Jerry saw a wooly llama?

BLOCK 4

EP10 (1-syll)

Does Mary drink only well known wine?
Yes. Mary drinks only well known wine.
Really? Mary drinks only well known wine?

EP11 (3-syll)

Did Lily say she saw Marian?
Yes. Lily said she saw Marian.
Really? Lily said she saw Marian?

EP12 (2-syll)

Did Ann go to see a new lawyer?
Yes. Ann went to see a new lawyer.
Really? Ann went to see a new lawyer?

BLOCK 5

EP13 (1-syll)

Did the woman buy a yellow lime?
Yes. The woman bought a yellow lime.
Really? The woman bought a yellow lime?

EP14 (2-syll)

Do the polls show an early winner?
Yes. The polls show an early winner.
Really? The polls show an early winner?

EP15 (3-syll)

Does Joanna know a millionaire?
Yes. Joanna knows a millionaire.
Really? Joanna knows a millionaire.

FATIGUE ITEMS

P4

Does his mother live in Iowa?
Yes. His mother lives in Iowa.
Really? His mother lives in Iowa?

P5

Does her brother Jim drive a red car?
Yes. Her brother Jim drives a red car.
Really? Her brother Jim drives a red car?

Appendix B: Instructions

English version

This is an experiment about the speech characteristics of native English-speaking learners of Mandarin as a second language.

Your participation in the experiment will require approximately one hour of your time, and you will be compensated 400 NT for your participation.

If you are tired, you can ask to take a break at any time. I have scheduled a ten-minute break between the English and Mandarin sentence sets.

If at any point in the experiment, you no longer wish to participate, you are free to stop the recording at any time.

First, you will be asked to complete a questionnaire about your language background.

Then, you will be asked to read sets of sentences aloud, first in English and then in Mandarin.

Each sentence set will consist of three sentences, each of which will appear on a separate card. The first will be a yes-no question, the second will be an affirmative answer to that question and the third will be a question which seeks confirmation of the affirmative answer, as if you were surprised to hear that the second sentence is really true.

Example:

1. Did Harry buy a purple balloon?
2. Yes. Harry bought a purple balloon.
3. Really? Harry bought a purple balloon?

Please read each set in the most natural way possible, as if you were talking to a friend.

Between sentence sets, you will see a card to read aloud which says, in the relevant language:

The next sentence set is number X.

Feel free to practice any sentence or sentence set as many times as you like until you feel comfortable with it. Recording will continue through the whole session, but I can remove the practice parts afterward. The last repetition of each sentence is the one I will save. When reading the Mandarin materials, be sure that you understand every word in the sentence before you begin reading it aloud. An English translation of the first sentence of each set appears on the card opposite that sentence. Feel free to consult the translation or to ask me if you have any questions about the meaning of any word that appears in a sentence.

The Mandarin materials will be written in Chinese characters. The Pinyin equivalent of each word, which includes tone information for each word, appears on the line underneath. Please read whichever is more comfortable for you.

Thank you for your participation.

Mandarin version:

這個實驗是為了探討母語是中文的人在學習英文時的語言表現，所以要請你唸一些中文句子和英文句子。

這個實驗將會花掉你一個鐘頭的時間; 為了感謝你全程的參與, 在實驗結束後, 我們將會提供四百元的車馬費.

你有權隨時要求休息或放棄繼續實驗, 只需事先通知主持人.

在實驗開始前, 我們將會先詢問一些有關你學習語言的經驗, 例如你的母語為何? 你會說台語, 客語或其它語言嗎? 等等…

本實驗將分為兩大部分, 第一部分是中文組句子, 第二部分是英文組句子, 在唸完第一部分後, 我們將休息十分鐘再繼續完成第二部分。中文組每部分都有二十一組句子. 每組有三個句子: 第一個句子是問句, 第二個是這個問句的回答, 第三個是反問句. (反問句並不是問句, 而是對前面所講的句子表示驚訝, 想要再進一步求證, 所以再問一次的句子) 例子:

姐姐遲到了嗎?

是, 姐姐遲到了.

真的? 姐姐遲到了?

在開始念實驗句以前, 我們會先有兩組練習句, 讓你熟悉實驗的過程, 之後再正式進入實驗的句子. 進入正式的實驗句後, 請用最自然的方式和正常的速度讀出所有的句子, 在句組與句組中將會有一頁封面如: **這是**

第二組句子. 也請你務必將這頁的句子讀出來. 在讀每一個句子時, 你可以不止唸一次, 每次都可以念到你認為滿意為止.

在二十一組句子都唸完後, 還有另外的二十一組句子. 每個句子都是如下面的句型: "我說 X 的時候，表示句子結束了。", 請你同樣用最自然的方式和正常的速度讀出這些句子, 並且念到你認為滿意為止.

中文組句子唸完後, 將會有十八組英文句子, 所有的實驗過程都將和前面的中文組句子相同，只是語言換成英文。當你讀英文句組時，若遇到不懂意思的句子時，你可以看英文句子旁邊的中文翻譯或詢問實驗主持人，以確定句意。若遇到不會唸的單字時，請先詢問實驗主持人該單字如何唸，再繼續唸整個句子。謝謝你熱心的參與.

Appendix C: Questionnaire

English version

Tanya Viger Date______________
Graduate School and University Center, CUNY

Language background questionnaire

1. Name ____________________________ e-mail __________________________

2. Please list the languages you know. For each language you know, rate your fluency in speaking, listening, reading and writing.

Please circle a number on the scale from 1-7, where 1= poor and 7= excellent

Language Speaking Listening Reading

Writing

English	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
Mandarin	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7

3. Your current age_________

4. Age at which you began learning Mandarin___________

5. Length of time spent studying Mandarin (years, months) ______________

6. Length of time spent in a Mandarin-speaking country. (Please provide dates)________

7. Current use of Mandarin (average number of hours per day) _____________

8. Please put a mark on the scale to indicate approximately what percentage of your teachers were native speakers of Mandarin.

0%--------------------------------------50%--------------------------------------100%

9. Please rate the methods you have been using to learn Mandarin on a scale from1-7.

Please circle a number on the scale where rating 1= I almost never do this and 7= I do this very often.

(a) Practice/conversation at home with family members (e.g. parents, spouse, partner)

1 2 3 4 5 6 7

almost never very often

(b) Practice/conversation with friends who are not native speakers of Mandarin

1 2 3 4 5 6 7

almost never very often

(c) Practice/conversation with friends who are native speakers of Mandarin

1 2 3 4 5 6 7

almost never very often

(d) Mandarin classes at a Mandarin center

1 2 3 4 5 6 7

almost never very often

(e) Mandarin classes at a bushiban

1 2 3 4 5 6 7

almost never very often

(f) Watching/reading/listening to Mandarin on TV, the movies, the radio or the computer

1 2 3 4 5 6 7

almost never very often

(g) Other (please describe)

Questionnaire, Mandarin version

語言經驗調查表 日期___________

姓名___________網址______________

請列出所有妳會的語言, 並依聽說讀寫的能力圈出下列號碼:

(評量值 1-7, 1= 差 7= 母語者程度)

語言種類	聽	說	度	寫
國語	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
英文	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
台語	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7

妳現在幾歲?

妳幾歲開始學英文?

妳學英文多久?

妳曾經在英語系國家住過多久? (請寫出明確的時間)

再妳學習英文時, 來自英文系國家的老師佔百分之幾?

0%--50%--100%

妳學習英文的方法是什麼? (評量值 1-7 1=極少 7=最多)

方法 \ 年齡	1-8歲	9-12歲	13-18歲	19歲 以後
跟家人學習	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
跟台灣朋友練習	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
跟外國朋友練習	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
在學校學習	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
在補習班學習	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
再家自己看電視或是電影學	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7

其它的方法 (請描寫):

妳現在每 天花多少時間是用英文?

再何種情況下使用?

Appendix D: Duration and pitch normalization constants

English Materials

DURATION CONSTANTS

Citation form in each of three item types …		**MA**	**EN**
	One-Syll	410.7	323.9
	Two-Syll	228.9	177.1
	Three-Syll	196.7	154.5
	Citation Grand Mean	**278.8**	**218.5**
	L2 vs. L1, proportional increase	**0.2759**	

Adjustments are:	Lower L2 value by dividing by	1.138
	Raise L1 value by dividing by	0.879

PITCH CONSTANTS

Citation form in each of three item types …		**MA**	**EN**
	One-Syll	233.5	191.5
	Two-Syll	247.7	197.1
	Three-Syll	242.4	195.7
	Citation Grand Mean	**241.2**	**194.8**
	Difference L2 minus L1	**46.44**	

Adjustments are:	Lower L2 value by subtracting	**23.22**
	Raise L1 value by adding	**23.22**

Mandarin Materials

DURATION CONSTANTS

			EN	MA
Citation form in each of three item types …	Tone 1		327	253
	Tone 2		304	241
	Tone 4		298	246
		Citation Grand Mean	**309**	**246**
		L2 vs. L1, proportional increase	**125.61**	

Adjustments are:	Lower L2 value by dividing by	**1.1281**
	Raise L1 value by dividing by	**0.8865**

PITCH CONSTANTS

			EN	MA
Citation form in each of three item types …	Tone 1		202.2	269.2
	Tone 2		185.3	213.9
	Tone 4		222.2	269.8
		Citation Grand Mean	**203.1**	**251.0**
		Difference L2 minus L1	**47.8**	

Adjustments are:	Lower L2 value by subtracting	**23.9**
	Raise L1 value by adding	**23.9**

Appendix E: 1-way ANOVA Output: L2 Pause Data: MA-EN

Summary Data:
Grand Means: 1st utterance = 2.60   2nd utterance = 2.30   3rd utterance = 1.51
(maximum 15)
Percentages: 17.3%   15.4%
10.1%
ANOVA Overview

Omnibus	$F(2,28) = 22.08$, $p<.001$
Q vs. D	$F(1,14) = 3.57$, $.05<p<.10$ (trend)
D vs. E	$F(1,14) = 24.28$, $p<.001$
Q vs. E	$F(1,14) = 33.42$, $p<.0001$

Q = 1st; D = 2nd; E = 3rd

Omnibus

SV	SS	DF	MS	F	P-value
S	119.74	14	8.55		
A	9.53	2	4.77	22.08	.00000173
S*A	6.05	28	.22		

Pairwise Test A: First versus Second Utterance

SV	SS	DF	MS	F	P-value
S	95.45	14	6.82		
A	.67	1	.67	3.57	.0798
S*A	2.61	14	.19		

Pairwise Test B: Second versus Third Utterance

SV	SS	DF	MS	F	P-value
S	71.23	14	5.09		
A	4.71	1	4.71	24.28	.0002233
S*A	2.61	14	.19		

Pairwise Test C: First versus Third Utterance

SV	SS	DF	MS	F	P-value
S	75.82	14	5.42		
A	8.92	1	8.92	33.42	.0000476
S*A	3.74	14	.27		

References

Anderson-Hsieh & Koehler, K. (1988). "The effect of foreign accent and speaking rate on native speaker comprehension" *Language Learning* (38), 561-611.

Anderson-Hsieh, J. Johnson, R. & Koehler, K. (1992). "The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody and syllable structure," *Language Learning* (42), 529-555.

Archibald, John. (1998). *Second language phonology*. Amsterdam: John Benjamins.

Beckman, M. E. & Ayers-Elam, G. (1997). "Guidelines for ToBI Labeling, version 3.0." *Manuscript and accompanying speech materials, Ohio State University*. Department of Linguistics: The Ohio State University.

Beckman M. E. & Pierrehumbert, J. (1986). "Intonational Structure in Japanese and English." *Phonology Yearbook* (3) 255-309

Best, C. T. (1995). "A direct realist view of cross-language speech perception." In *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, Strange, W. (Ed.), Timonium, MD: York Press.

Bolinger, D. (1988. Intonation and its uses: melody in grammar and discourse. London: Edward Arnold.

Brazil, D. (1997). The communicative value of intonation in English. Cambridge: Cambridge University Press.

Broselow, E., Hurtig, R. & Ringen, C. (1987). "The perception of second language prosody," in Ioup, G. & Weinberg, S.H. (eds.) *Interlanguage Phonology: the Acquisition of a Second Language Sound System*, Cambridge, MA: Newbury House Publishers.

Chang Y. (1998). "Les indices acoustiques et perceptifs des questions totales en Mandarin parle de Taiwan." Cahiers de Linguistique Asie Orientale (27) 51-78.

Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press

Chen, Q. (2000). *Analysis of Mandarin Tonal Errors in Connected Speech by English-speaking American Adult Learners: A Study at and above the Word Level*, Department of Instructional Psychology and Technology: Brigham Young University, Salt Lake City, Utah

Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press

Derwing, T. & Munro, M. (1997). "Accent, intelligibility and comprehensibility: evidence from four L1s" *SSLA* (19), 1-16.

Flege, J.E. (1995). "Second-language speech learning: Theory, findings and problems." In *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Strange, W. (Ed.), Timonium, MD: York Press, 233-277.

Friederici, A.D. & Alter, K. Lateralization of auditory language functions – the dynamic dual pathway model (2004). Brain and Language 89, 267-276.

Gandour, J. Dzemidic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L. Satthamnuwong, N., Lurito, J. (2003). "Temporal integration of speech prosody is shaped by linguistic experience." Brain and Language 84 318-336.

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., Li, X., & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage*, 23, 344-357.

Hewings, M. (1995. "Tone choice in the English intonation of non-native speakers." International Review of Applied Linguistics, 33(3), 251-265.

Hirschberg, J. and Pierrehumbert, J. (1986). "The intonational structuring of discourse." In Proceedings of the 24th AnnualMeeting ofthe Association for Computational Linguistics, New York, NY, 136-144.

Hsieh, L., Gandour, J., Wong, D., and Hutchins, G. (2001). *Functional circuits engaged in speech perception depend on language experience*. Brain and Language 76 (3): 227-252.

Klein, D., Zatorre R. J., Milner B., Zhao V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. NeuroImage. 13, 646-653.

McGory, J. (1997). "The Acquisition of Intonation Patterns in English by Native Speakers of Korean and Mandarin.." Unpublished doctoral dissertation, Ohio State Dissertations in Linguistics. Department of Linguistics: The Ohio State University

Mennen, I. (1998). "Can language learners ever acquire the intonation of a second language?" In Proceedings of the ESCA workshop on speech technology in language learning (pp. 17–20). Marholmen, Sweden: International Speech Communication Association.

Munro, M. (1995). "Nonsegmental factors in foreign accent," SSLA (17), 17-34

Munro, M. & Derwing, T. (1995). "Foreign accent, comprehensibility and intelligibility in the speech of second language learners" *Language Learning* (45), 73-97.

Olive, J.P., Greenwood, A. and Coleman, C. (1992). "Acoustics of American English Speech: A Dynamic Approach", New York: Springer.

Peng, S, Chan, M.K.M, Tseng, C.Y, Huang, T, Lee, O.J. and Beckman, M.E. () "A Pan-Mandarin ToBI" in Jun. S. ed. Prosodic Typology and Transcription: A Unified Approach: Oxford University Press

Pickering, L. (2004). "The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse" *English for Specific Purposes* (23) 19-43.

Pierrehumbert, J. & Hirschberg, J. (1990). "The meaning of intonation contours in the interpretation of discourse," in Cohen, P., Pollock, J. and Pollack, M. eds. "Plans and Intentions in Communication." Cambridge, MA: MIT Press 271-312.

Pennington, M. C. (1996). *Phonology in English language teaching: An international approach*. New York: Addison Wesley Longman.

Poeppel, D. (2003). "The analysis of speech in different temporal integration windows: cerebral lateralization as asymmetric sampling in time," Speech Communication 41: 245-255.

Schack, K. (2000 "Comparison of intonation patterns in Mandarin and English for a particular speaker" *University of Rochester Working Papers in Linguistics*, Vol. Spring (2000 (1), 24-55.

Schwartz, B.D. and Sprouse, R.A. (1996). "L2 cognitive states and the Full Transfer / Full Access Model." Second Language Research 12: 40-72

Shen, X. (1989). "Toward a register approach to teaching Mandarin tones." Journal of the Chinese Language Teachers Assoc. (24), 27-47

Shen X. (1990). *The Prosody of Mandarin Chinese*. Berkeley, CA: University of California Press.

Tajima, K., Port, R. & Dalby, J. (1997). "Effects of temporal correction on intelligibility of foreign-accented English," Journal of Phonetics (25) 1-24

Tseng C.C. (2004). "Prosodic Properties of Intonation in Two Major Varieties of Mandarin Chinese: Mainland vs. Taiwan." In Proceedings of ICSLP 2004.

Tseng, C.Y. (2003). "Towards the organization of Mandarin prosody: Units, boundaries and their characteristics" In Proceedings of ICPhS, Barcelona.

Tseng, C.Y. and Chou, F. (1999). "A prosodic labeling system for a Mandarin speech database." InProceedings of ICPhS, San Francisco.

Vainikka, Anne & Martha Young-Scholten. (1994). Direct access to X¡¦-theory: Evidence from Korean and Turkish adults learning German. In T. Hoekstra and B. Schwartz (eds.), Language acquisition studies in generative grammar. Amsterdam: John Benjamins, pp. 265-316.

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language & Speech*, 38, 253–265.

Wennerstrom, A. (1994. “Intonational meaning in English discourse: A study of non-native speakers” Applied Linguistics 15(4) 399-420.

Wennerstrom, A. (1998). "Intonation as cohesion in academic discourse: a study of Chinese speakers of English" *Studies in Second Language Acquisition* (20), 1-25.

Wang, Y., Jongman, A. and Sereno, J. (2001). “Acoustic analysis of American listeners’ Mandarin tone productions before and after training” *Poster presentation at 142nd Meeting of the Acoustical Society of America*.

Wong, R. (1987). *Teaching pronunciation: Focus on English rhythm and intonation*. Englewood Cliffs, New Jersey: Prentice-Hall Regents.

Xu, Y. (2002) “Contextual tonal variations in Mandarin”*Journal of Phonetics* (25), 61-85