2005-08-24

A Comparison of Beijing and Taiwan Mandarin Tone Register: An Acoustic Analysis of Three Native Speech Styles

Richard Christen Torgerson
Brigham Young University - Provo

A COMPARISON OF BEIJING AND TAIWAN MANDARIN TONE REGISTER:

AN ACOUSTIC ANALYSIS OF THREE NATIVE SPEECH STYLES

by

Richard C. Torgerson Jr.

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Language Acquisition

Brigham Young University

December 2005

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Richard C. Torgerson Jr.

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Matthew B. Christensen, Chair
Date	Dana S. Bourgerie
Date	Shawn L. Nissen

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Richard C. Torgerson Jr. in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

____
Date

____
Matthew B. Christensen
Chair, Graduate Committee

Accepted for the Department

____
Ray T. Clifford
Graduate Coordinator

Accepted for the College

____
John R. Rosenberg
Dean, College of Humanities

ABSTRACT

A COMPARISON OF BEIJING AND TAIWAN MANDARIN TONE REGISTER:

AN ACOUSTIC ANALYSIS OF THREE NATIVE SPEECH STYLES

Richard C. Torgerson Jr.

Department of Language Acquisition

Master of Arts

This study investigated the possibility of pitch and tone register differences between native speakers of Taiwan and Beijing Mandarin by means of an acoustic analysis of three speech styles. Speech styles included spontaneous interview, spontaneous descriptive, and controlled read sentential speech. Data analysis included long segments of recorded speech in order to discern any statistically significant pitch register differences between the two dialects. Speech style and read tones were also analyzed.

Results suggest that tones produced in Taiwan Mandarin are in a slightly lower register than those produced in Beijing Mandarin. Surprisingly, speech style was not a significant predictor of pitch register in long segments of recorded speech. Despite a limited sample size, this research effectively promotes the inclusion of sociolinguistic variables such as dialect in the field of tone research.

ACKNOWLEDGMENTS

First of all, I must thank my Father for giving me the opportunity to obtain a good education and helping me to complete this thesis. He has guided me along the way, making the research process worthwhile. My wife Melanie has given me constant support, understanding, insight and patience. My daughter Karen has provided much needed relief with her smile and playful attitude. My parents and parents-in-law have been so supportive and encouraging. I am so grateful for all of them. I am also thankful for each of my committee members; for helping me to remember the primary importance of family, even in the midst of a demanding research project.

Table of Contents

List of Tables

List of Figures

CHAPTER ONE

INTRODUCTION

Accurate descriptive research of the languages and dialects spoken in China is still an open field for sociolinguistic researchers (Norman, 1988; Simmons, 1999). Although there has been a great deal written concerning the linguistic differences between many languages and dialects of China (i.e., Lien, 1986; Rose, 2002; and Yue-Hashimoto, 1986), less research has been specifically focused on dialectal differences between the standard Mandarin Chinese of Beijing and that spoken on the island of Taiwan.[1]

Some differences between the linguistic features of Beijing and Taiwan Mandarin have been noted in the literature. For example, Cheng (1985) has suggested that Taiwan Mandarin syntax has been influenced by Taiwanese (Southern Min). Sanders (1992) indicated that significant semantic differences exist between the use of modal verbs in Beijing and Taiwan Mandarin. More recently, Swihart (2003) provided a sampling of lexical and syntactic differences between the two dialects.

M. Y. Chen (2000) summarized some of the tonal differences between Standard Mandarin Chinese (hereafter *SC*) and Taiwanese, although the specific implications of this language contact were left unmentioned. In his report about dialectal variations of *SC*, Chen (1999) has indicated that pitch contour (or shape) is consistently the most affected of all phonological features. This susceptibility would suggest potential tone contour differences between Beijing and Taiwan Mandarin. Duanmu (2000) has provided several examples of tonal differences between Taiwan and Beijing Mandarin,

---

[1] Languages such as Cantonese and Shanghaiese, which have traditionally been defined by some as Chinese dialects due to a shared written language, are considered distinct languages in this study due to their lack of mutual intelligibility in oral speech.

including the opinion that Taiwan Mandarin is characterized by the frequent use of low tones while Beijing Mandarin uses rising or high tones, a difference of tone register. Furthermore, M.Y. Chen (2000) has reported that “southern [languages including Southern Min] typically have larger tonal inventories than the Mandarin group… Furthermore, sandhi processes take different forms in different dialect (sub)groups….”(p. 2). Thus, it is theoretically possible that tonal dissimilarities could exist between the two *SC* dialects in terms of tone contour, tone register and/or tone sandhi (changes caused by interaction of adjacent tones).

In terms of tone register, Fon & Chiang (1999) have proposed that Taiwan Mandarin speakers produce tones that vary significantly from the standards set by *SC*. Their findings suggest that tones produced by Taiwan Mandarin speakers may be in a lower register. However, because their case study included only one female participant from Taiwan, further investigation is needed to determine the validity of the claim that there are significant dialectal differences between tone register in Beijing and Taiwan Mandarin.

Research Description

I conducted a follow up study of Fon & Chiang’s (1999) case study in order to investigate potential pitch or tone register differences. Pitch register varies from tone register in that it spans over long portions of continuous speech while tone register spans only over single tones. Both pitch and tone are measured acoustically and represented by fundamental frequency (hereafter F0). While pitch is what humans perceive, F0 is considered the most accurate numerical representation of pitch. Hence, both terms will be used interchangeably throughout this study. My analysis of native speech included a descriptive assessment of F0 extracted from longer segments of continuous speech in order to discern any significant pitch register differences between the two dialects of *SC*. In addition, I performed an analysis of a sampling of individual tones produced in read speech in order to test the impact of dialect on tone register in a more controlled setting. Both approaches were made under the premise that if average F0 measures from a long portion of speech revealed significant differences between the two dialects, so would a more controlled analysis of individual tones. Because many prior studies on tone register had been limited to read speech, data collection included three different types of speech: spontaneous interview, spontaneous descriptive, and controlled read sentential speech. The research project was designed to test three null hypotheses:

1) Pitch register differences cannot be predicted by dialect background; or there are no significant pitch register differences between groups of male native speakers of Mandarin Chinese from Beijing and Taiwan.

2) Pitch register differences cannot be predicted by speech style; or there are no significant pitch register differences between groups of male native speakers in three speech styles: read, descriptive and conversational.

3) There are no significant tone register differences between the five tones produced in Taiwan Mandarin and those produced in Beijing Mandarin; or there are no individual tone register differences in read speech between male native speakers of Mandarin Chinese from Beijing or Taiwan.

Thus, while this study looks closely at tone and pitch register, my research hypotheses did not warrant an analysis of tone contour. This concludes chapter one. Chapter two will cover a wide range of topics related to the research hypotheses by means of a thorough review of related literature. Chapter three will introduce the methods used in the present study. Chapters four and five will present results and conclusions.

CHAPTER TWO

REVIEW OF LITERATURE

This literature review is divided into six sections. The first two sections provide important background information concerning tone and the two dialects being investigated. The teaching of Mandarin lexical tone is considered and then readdressed in chapter five. Sections four through six of the review set the stage for the approach of my research study. Each section of the review is summarized below:

1) A brief discussion of the historical developments of the *SC* of Beijing and Taiwan
2) An explanation of Mandarin lexical tone and tone change rules
3) The teaching of Mandarin lexical tone
4) Acoustic analysis of spontaneous speech
5) The perception of Mandarin lexical tone
6) The production of Mandarin lexical tone

A Brief History of Standard Mandarin Chinese (*SC*)

China is a land of many languages. Norman (1988) observes that despite similar writing systems, "Chinese is more like a language family than a single language made up of a number of regional forms" (p. 187). According to the *Language Atlas of China* (1987), the major languages include the official language of *SC*, Cantonese (*Yue*), Shanghaiese (*Wu*), Fuzhou (*Minbei*), Hokkien-Taiwanese (*Minnan*), *Xiang*, *Gan*, *Hakka* as well as a plethora of minority languages. The development of a national language, namely *SC*, is a rather recent event in Chinese history. Throughout China's long history, linguistic standards of varying types and influence have been established, stemming from the capital or other major cities. Various official languages have served as a tool of administrative and/or education centers, only to vanish away before the dawn of a new empire. However, the huge mass of purely local languages has remained a constant throughout these shifts. Even in present day China, most Chinese grow up speaking one or more languages that are phonologically distinct from *SC*, many that are not even mutually intelligible.

Written Chinese has had a tremendous effect on the Chinese spoken in both Taiwan and Mainland China (DeFrancis, 1984). Norman (1988) and Coblin (2000) suggest that a Classical written language, or *wenyan*, served as the only shared linguistic element between the many linguistic and social communities within China until the start of the 20th century. *Wenyan* was a written language, thus, despite its wide acceptance as a tool for commerce and bureaucracy, pronunciation still varied greatly between regions.

According to Chen (1999), the early 1900's was marked by the emergence of what is called the *Baihua* or 'plain language' movement. This vernacular language movement was backed by many scholars, and was designed mainly to unify China through the establishment of a national written language based more on modern speech than on Classical Chinese. However, the manner in which this could best be carried out was widely debated. By 1919, a commission produced the new 'Dictionary of national pronunciation' (*Guoyin zidian*), hoping that standardizing the pronunciation of characters would somehow unify vernacular pronunciation as well. This attempt failed, but some progress was made in literary language planning.

After 1920, the *Baihua* movement had all but eliminated the Classical literary language (*wenyan*), causing it to disappear from the formal language of government and journalism by the 1940's (Norman, 1988). During this time the Nationalist government supported the promotion of a national vernacular (*guoyu),* which was the spoken counterpart of *baihua*. During the late 40's, conflict between the Communists and Nationalists resulted in a large body of Nationalists crossing over from the Mainland to the island of Taiwan. Thus, ever since 1950, the development of *SC* has been divided into at least two dialect groups, *guoyu* in Taiwan and *putonghua* on the Mainland.[2]

The History of Taiwan Mandarin: Guoyu

The Nationalists came from various parts of China, and were far from being a linguistically homogenous group. Thus, the *guoyu* and *baihua* that had been promoted for over twenty years on the Mainland served well as a unifying element to the group. As the Nationalists came to the island of Taiwan, a massive political, social and linguistic change took place for the Southern Min, Hakka and aboriginal people already living there.

Guoyu became the new standard and official language of a people who had just experienced over 50 years of Japanese rule (including language planning). In fact, just before Taiwan's return to China in 1945, it is reported that "By 1944, 71 percent of the local population were proficient in Japanese. The percentage was much higher in the middle and younger generations, with a large proportion of youngsters unable to speak any Chinese at all. Japanese was by all measures the standard language in Taiwan" (Chen, 1999, p. 31). Hence, even before the massive influx of Nationalists in 1949, Taiwan became the target of intense language promotion and planning.

The drastic shift from Japanese to Chinese amidst an already diverse linguistic community has created a complex story of language and dialect shift, language contact, and language maintenance (see Kubler, 1985; Young, 1989). A little over 70% of the

---

[2] The use of the term dialect may be controversial, as some may consider the extent of the differences between Beijing and Taiwan Mandarin to be on a smaller level; making them two varieties of the same dialect, rather than two dialects of the same language.

population of Taiwan is composed of the Southern Min people, with smaller amounts of Mainlanders (15%) and Hakka (12%), and a very small number of aborigines living in the mountain regions (about 2%) (Kubler, 1985; Huang, 1994). Taiwan's unique linguistic environment, as well as its social and political distance from Beijing, has created a *SC* which varies from *putonghua* in all major linguistic subsystems while still remaining mutually intelligible (Norman, 1988; Chan & Tai, 1989).

For over thirty years the Nationalists imposed strict governmental restrictions against speaking anything but *guoyu* on mass media, in schools and in public places (Chiung, 2001). Nevertheless, the Southern Min people maintained a fair amount of their native language as they talked in their homes and amongst friends (Chang, 1996). Chan & Tai (1989) argue that even in Taibei, the capital city of Taiwan, the local Southern Min dialect has influenced *guoyu* (and vice versa). Similarly, Norman (1988) reports widespread use of *guoyu* "in metropolitan areas like Taibei, but in smaller towns and rural areas the Southern Min and *Kejia* [Hakka] dialects are used almost exclusively, despite the fact that almost all people under forty possess at least a minimal working knowledge of the standard language" (p. 248). Of course, this preponderance of non-standard language use has seen revitalization since the late 80's when the martial law of the Nationalist's "Mandarin only policy" was softened in acceptance of a multi-lingual society (Huang, 1997).

The History of Beijing Mandarin: Putonghua

In Mainland China, *putonghua* became the new name of the desired standard language, and it was energetically supported by the Communist government. This encouragement was, however, apparently not as rigid as that taking place in Taiwan.

Chen (1999) reports that in the late 80's only about 50% of the population of Mainland China could speak *putonghua*, while in the early 90's about 90% of the population of Taiwan could speak *guoyu*. These numbers suggest that for many of the people in Mainland China, especially those in rural areas, the establishment of *putonghua* was less effective than the establishment of *guoyu* in Taiwan.

Norman (1988) and Chen (1999) both note that the less educated Chinese tend to have lower proficiency in *putonghua*. Furthermore, Chen (1999) proposes other sociolinguistic factors such as language prestige, homogeneity of local dialects (or languages), and local economies are each major factors affecting the achievement of the standard.

Conclusion

As we have seen in both Taiwan and the Mainland, "… [*SC*] is not in fact identical to any one local dialect in all its details" (Norman, 1988, p. 250), but is a socio-political creation and conglomeration of different aspects of different languages and dialects. Thus, *putonghua* and *guoyu* have become standards, comparable to Standard American and British English. Both dialects share a great deal, and both are fairly high prestige forms. However, there is also a rich array of internal variation and individual differences between the two.

Mandarin Lexical Tone and Tone Sandhi

An Explanation of Mandarin Lexical Tone

Tone has been defined as "the contrastive, or linguistic, functioning of the fundamental frequency at the word or syllable level" (Connell et al, 1983, p. 337). In Mandarin Chinese, a single tone can be combined with a single syllable to create a word, making tone just as significant as the consonant or vowel sound. Mandarin tones have traditionally been described as having four main contours or tone shapes; a high, level tone (first), a rising tone (second), a low or low dipping tone (third), and a high falling tone (fourth). The ‘fifth’ tone is called the neutral tone, primarily because it does not have a consistent contour; it is unstressed and usually follows the rising, level, or falling contour of the last part of the preceding syllable (Norman, 1988; Shen, 1990).

Although intonation and stress are utilized in other non-tonal languages such as English, variation in tone is not usually lexically significant. Kubler et al (1997) attest that, without tone, the 400 basic syllables of Mandarin would be lexically overloaded. With only 400 syllables, words with the same sound and/or same spelling have a great potential to be confused with one another. Because tone is as important as the consonant or vowel sound in a word, an omission or mispronunciation of tone is not only a mistake in pronunciation; it is sometimes an inadvertent production of an entirely different word. For example, production of the morpheme *hao* can mean “good” or “number,” depending on whether you use a low dipping (third) or high falling (fourth) tone. Thus, in addition to learning new vowel and consonant sounds, non-natives must also retain and perceive tones accurately to enable clear communication. Furthermore, recent research has also found that tones are influenced by both voluntary and involuntary phonetic constraints such as speech rate, stress, declination and intonation, making perception and production of tone all the more difficult for the L2 learner (Shen, 1989; Xu, 2001; Shih, 2000).

Two other aspects that must be considered in a discussion of Mandarin lexical tone are pitch register and tone register. As mentioned in the introduction, pitch register is represented by the average pitch and pitch range of one’s voice over a long segment of speech. Some may call this intonation. Tone register is the average pitch and average pitch range over one tone. Tone and pitch register are interrelated and both are influenced by a host of other linguistic variables listed in detail in chapter three. It has been suggested that, when speaking Chinese, native speakers have a wider pitch register than non-native speakers Chinese (G.T. Chen, 1973). This difference in pitch register may be one of the main problems for L2 learners of Mandarin as they attempt to produce tones (Miracle, 1989; Shen, 1989; Q. Chen, 2000).

Tone Sandhi: Rules of tone change

In addition to attaching correct tones to syllables to form words, students must also learn to use tone sandhi rules in speech. Li and Thompson (1981) describe tone sandhi in this way: “A syllable has one of the tones in the language when it stands alone, but the same syllable may take on a different tone without a change in meaning when it is followed by another syllable” (p. 8). In this study, tone sandhi rules are defined as basic changes in tone register or tone contour that facilitate speed and ease in the production of multi-syllabic speech. Obviously, some tone changes are voluntary, and some are not (see Xu, 1997; Xu, 2001).

The most commonly cited example of tone sandhi involves the third tone. While a third tone often has a low or dipping contour as a single syllable, two third tones adjacent to each other often require the speaker to change the first syllable of the phrase to a rising tone. Hung (1990) suggests that this change is based on the syntactic and semantic relationship between the constituents, but this is only one tone change out of many that actually take place. The following are some of the major tone sandi rules as noted in the literature.

First tone rules

This tone may be the least affected in connected or spontaneous speech, but is still subject to slight variation as well as neutralization (Shen, 1990; Lin, 2001). Xu (2001) has shown that the first tone can slant slightly upward or downward based on whether the following tone is a high or low tone. It may be most affected by intonation and/or stress when in the sentence final position (Papousek & Hwang, 1991).

Second tone rules

A second tone changes to a high level tone when preceded by a first or second tone and followed by any other tone other than neutral (Li & Thompson, 1981; Shen, 1990). Hence, the second tone may also be susceptible to other carry-over affects which adjust its shape or register (Xu, 1997).

Third tone rules

The third tone appears to vary more than any other tone in speech (Chao, 1968). A third tone preceding any other tone (excepting the third and the neutral) may lose the rising half of its contour (S. Chen, 1973; Shen, 1990; M.Y. Chen, 2000). In fact, other researchers suggest that the third tone is usually produced as a low tone, rather than with a dipping contour (Tsung, 1987; Xu, 2001; Chen, 2005). As mentioned earlier, a third tone preceding another third becomes a rising tone (Shen, 1990; Duanmu, 2000). A third tone produced three times in succession (e.g. *hao*(3) *ji*(3) *zhong*(3) = ‘quite a few kinds’) can actually become *hao*(2) *ji*(1) *zhong*(3) (Chao, 1968). Chao (1968) also argues that a third tone produced in between a first or second tone and a third tone is usually pronounced as a high level tone. Finally, a third tone produced before a falling tone can become a rising or a falling tone (Shen, 1990).

Fourth tone rules

In contrast, little has been said concerning the fourth tone, except that a fourth tone preceding another fourth tone does not drop all the way, but stops about half way down (Chao, 1968; Shen, 1990).

Neutral tone rules

In the past, neutral tones (or fifth tones) were believed to be just that, neutral: without contour or standard register. Shen's (1990) acoustic analysis reveals that

> "... except for sentence-initial and sentence-final neutral tones, catathesis generates the tonal value of the neutral tone, i.e., a neutral tone following a falling tone continues to drift down, and it continues to rise following a rising tone... it undulates in connected speech under the effect induced by the interplay of the preceding tone and intonation" (p. 48).

Hence, the fifth tone does not have a single, specific contour; it usually follows the rising, level, or falling contour of the last part of the preceding syllable (Norman, 1988; Shen, 1990).

The Teaching and Learning of Mandarin Lexical Tone

Developing native-like pronunciation of a foreign language is one of the great frontiers of second language acquisition for both learners and researchers. However, most research suggests that while many adult learners struggle indefinitely to develop good pronunciation, young children seem to acquire it quickly and naturally (Archibald, 2001). Research and reason both suggest that learners' tones can improve, but it is still unclear what major factors hinder and help L2 learners acquire native-like tone production.

While the communicative teaching approach is emerging in the Chinese teaching field (Xing, 1997), this approach to learning must also be applied with caution. Different teaching contexts, different learner goals, different learner backgrounds and motivations, even different faculty, administrators and employers demand not just a learner focused, but a learning focused, cafeteria style approach to the teaching of tone (Richards, 2001; Bai, 1996). Bai (1996) has also suggested that this ability to pick and choose from a variety of approaches and methods is one of the great abilities and responsibilities of a good language teacher.

While there may not be one comprehensive approach that can aid students in every aspect of tone acquisition, or throughout the many years of required study, there are many perspectives about how Mandarin tone can be most effectively introduced. Native speakers' acquisition of tone naturally follows a different path than that of non-native learners. For example, Chinese children's phonological development is shaped largely by massive amounts of personalized, meaningful input coupled with contextualized communication between caregiver and child (Gleason, 2001). In contrast, many teachers still attempt to teach tones to adult learners in a matter of weeks using decontextualized listen and repeat drills, even despite recent attempts by some in the Chinese teaching field to "see meaning as the core and foundation of language" (Loke, 2002, p. 65).

Reflecting on some of the special challenges for learners of "truly foreign languages," Jorden and Walton (1987) discuss the tone and pronunciation systems of Chinese. They suggest that Chinese teachers who dedicate several weeks to the pronunciation of single syllables often assume "there will be a direct transfer from work on isolation syllables to the natural utterances of true speech" (Jorden & Walton, 1987, p.

115). Walton's experience with L2 learners suggests there is a difference between ability to produce correct tone from an elicited response, and to produce complete sentences with tone contours intact and tone sandhi rules applied. Jorden and Walton (1987) lament that "pronunciation practice tends to rely much too heavily on mimicking rather than self-generation" and that students do not "internalize articulatory strategies for production" (p. 116).

Chen's (1997) perspective of learners' difficulties with Mandarin tones focuses on transfer and developmental problem areas. He believes tonal errors stem from three problem areas: first, the different pitch range and register 'habits' English speakers already possess, secondly, interference from intonation and other prosodic features of English. Finally, he refers to the inherent difficulty of certain tonal features; features Chinese children also struggle with.

Despite these different theories, researchers have made little progress at coming to a consensus of how to improve the teaching of tones. For example, S. Chen (1973) suggests teaching students to produce the third tone as both a low tone and as a full dipping tone in contextualized tone pairs. This learning task is proposed to address the fact that the third tone has traditionally been taught with a full dipping contour, rather than as a low level tone. While this learning task could help students recognize and produce both allophonic variations of the third tone appropriately, this approach doesn't address other aspects of tone acquisition such as how to help students put this into long-term memory.

Chen (1975) argues that students are generally taught the tones in chronological order; starting with the first high level tone, and concluding with the high falling fourth.

This method supposedly causes students to struggle with differentiating tones accurately, and they don't develop a wide enough pitch range to produce tones in the proper register. Chen suggests that "the four tones should be introduced by pairs with the order of high-low [1st -3rd] and then rising-falling [2nd -4th]" (Chen, 1975, p. 25). This learning order could help students to widen their pitch range when they produce tones, but there is little evidence to warrant such a pattern of introduction. In addition, this proposal only considers tone production on a bi-syllabic level; we have no idea whether the learner will retain this information in connected or spontaneous speech.

In accordance with S. Chen (1973), Tsung (1987) argues that the third tone needs to be taught as a low level tone, rather than as a full dipping contour, in order to eliminate "problems in perception, production and acquisition" (Tsung, 1987, p. 87). The suggested fix is an adjusted representation system that shows the third tone as a low level tone, and an increased focus on the teaching of the 'half-third.' Tsung (1987) provides quite a convincing argument for the teaching of the low-level third tone. However, it is also unclear what the long-term effects of Tsung's adjusted system of representation will be.

Shen (1989) states that traditional beliefs about tone errors are centered on contour errors rather than register errors. However, similar to Chen (1975), she suggests that students are not developing awareness of their own pitch range, nor are they learning how to produce tones in the appropriate tone register. She argues that learners should first perceive and produce sounds in the three levels of pitch in their own voice (high, mid, and low). She further encourages yet another system of representation, this one simplified from five tone levels to only three tone levels (high, mid, and low). It is possible that this approach would have a positive effect on the acquisition of tone, but the method has not been followed up or tested in task-based research. Furthermore, her suggestions are geared towards the introduction of tones, not on later training; how would her method be applied at later stages of learning? Finally, Shen's (1989) suggested changes to widely accepted systems of representation are highly unlikely.

In yet another argument for a new system of representation, Bar-Lev (1991) contends that students should learn one tone at a time with his own double system of representation, designed for ease of pronunciation and long-term effects. Bar-Lev's complaint is that many students are forced to learn de-contextualized tones in isolation, or in minimal pairs, thus, they are unable to produce tones in sentences of their own creation or in spontaneous speech. However, his approach is not shown to work more effectively to help students acquire and produce tones more efficiently in any type of speech. The new system of representation is likely to help beginners more than other systems, but it has not been applied on a wide scale, nor followed up in task-based research.

Lundelius (1992) also focuses on the manner in which tones are presented to the learner. He suggests that tone marks, diacritics or numbers do not help students remember tones. A new system of representation, called "Tonally spelled pinyin" (Lundelius, 1992, p. 95) is presented. Lundelius' system of representation makes use of pinyin, and the method's effectiveness is also supported by his study. However, like his predecessors, Lundelius does not address tone acquisition as a whole, but only the way that tones are represented to the learner. It is doubtful that any of these approaches focused solely on how tones are presented to the learner can solve the major problems learners have with tone acquisition.

In a refreshing turn, Chan (2003) suggests that because of the ready availability of speech processing and analysis technology, students and teachers should be utilizing this software to aid the learning process. With this software, tone can be taught with further attention toward stress, tone context and intonation. Students can compare their speech with native speech as well as receive audiovisual feedback via speech analysis software. It may be that this approach addresses more fully the need for students to improve their pronunciation *throughout* the acquisition process, not just in the first few weeks. However, this approach also runs the risk of being misused. Untrained students and teachers may misinterpret feedback or perhaps neglect more essential activities, such as talking with a native speaker on a regular basis.

Finally, Chen (2005) suggests that there are still many misconceptions being perpetuated in Chinese language teaching materials which affect the teaching of tone. Chen claims that there is still confusion about the difference between the neutral tone and weak stress. The third tone is still being presented as a dipping tone rather than a low level tone. To solve matters, Chen argues that dictionaries should retain the original tones of unstressed syllables and that the third tone should be depicted as a low level tone rather than a dipping one. Chen’s focus on improving some problematic areas of teaching materials is timely, although it is unclear how these improvements will be implemented.

I have shown how several well intentioned proposals to help students learn tone still remain largely untested and unproven (i.e. Chen, 1975; Shen, 1989; Bar-Lev, 1991; Lundelius, 1992). Furthermore, many proposals have been focused solely on the introduction or representation of tone, which is only a small part of the acquisition process. Evidence of what learners are truly acquiring and producing in spontaneous speaking environments is needed if educators are to match tasks with deficiencies appropriately.

Acoustic Analysis of Spontaneous Speech

Acoustic Analysis

In this study, acoustic analysis refers to instrumental analysis performed either by spectrogram, computer, or other devices capable of measuring and depicting pitch by collecting and reporting F0 as well as other aspects of tone (i.e. amplitude, duration, etc.). While pitch is what people hear, F0 is the closest numerical representation of pitch available. Acoustic analyses of F0 were traditionally reported in Hertz (Hz), but recent research has also used semitones and Equivalent Rectangular Bandwidth (ERB) to represent F0 in a manner that is presumably closer to human perception. Lewis (2002) suggests that the ERB scale is reportedly superior to both Hz and semitones in reflecting frequencies below 500 Hz.

Spontaneous Speech

Miller and Weinert (1998) define spontaneous speech as having five key properties. These five properties as a whole[3] serve well as a working definition of spontaneous speech, which is summarized as follows:

1. Produced in real-time, impromptu conversation; no written script.
2. Governed mainly by implicit knowledge of the language.
3. Produced as part of inter-communication between two or more parties in a particular context (i.e. talking with a classmate in the hall, etc.)

[3] The five key properties are almost exactly the same as listed in Miller and Weinert's (1998) book in both format and content; changes were generally stylistic, with an example and "stress" added in numbers 3 and 4 respectively.

4. Involves voice quality, pitch, rhythm, amplitude, stress, etc.
5. Includes "gestures, eye-gaze, facial expressions, and body postures, all of which signal information" (Miller & Weinert, 1998, p. 22).

Yang (1995) argues that "…natural data is complicated, but if the goal is to understand natural speech, then it is important to investigate a level of complexity sufficient to model natural speech" (p. 58). An important distinction between spontaneous and deliberate production is indicated in Bialystok's model of how individuals learn a second language (Bialystok, 1979). Spontaneous production (output) is governed solely by "implicit linguistic knowledge," and deliberate or planned speech is governed by both implicit and explicit linguistic knowledge (Ellis, 1994, p. 357). However, the majority of research on the acquisition of Mandarin lexical tone has been based only on read or planned speech, keeping both native and non-native inherent linguistic knowledge obscured.

Ellis (1994) reiterates the need to analyze natural native and non-native speech, in casual settings, before trying to establish universals of language acquisition. Furthermore, Lakshaman and Selinker (2000) point out that most spontaneous speech studies are focused solely on natives *or* non-natives; rarely are these two groups compared in the same study in similar speech scenarios. Any attempt to measure the abilities of non-natives must be compared to how native speakers are performing in corresponding circumstances. I should note here that my study was originally designed to include non-native speakers, but due to time constraints was limited only to native speakers. It is hoped that this study may serve as a foundation for further investigation into non-native tone production.

Acoustic Analyses of Spontaneous Speech

Many empirical studies that analyze spontaneous speech are focused on the L1 and are performed with children as participants. For example, Tardif et al (1997) analyze parent-child interaction produced by family members in their own homes. To analyze the speech of Mandarin-speaking children, the researchers invited caregivers and children to commence with normal activities while they recorded them for one hour. Dunn and Flax (1996) perform their analysis of spontaneous speech produced by children in 25 minute play sessions with a parent, suggesting that specific language impairment can be detected more accurately in natural contexts. Their analysis suggests that some children with specific language impairment had more spontaneous speech errors than “normal children whether they met psychometric discrepancy criteria or not” (Dunn & Flax, 1996, p. 1).

Liu’s (2001) study focuses on understanding the nature of speech produced by Mandarin Chinese speaking mothers when they speak to their infants vs. other adults. She recorded a total of 32 mothers living in the southern city of Kaohsiung Taiwan. Participants interacted in a sound attenuated room with their infants and an adult in one recording session. Target words were elicited in semi-spontaneous speech by providing 21 pictures/objects with the target bisyllabic words printed on them to be used during the play session.

Liu’s (2001) acoustic analysis included sampling the words produced at 20 kHz, 16 bit resolution and a low pass filter set at 10 kHz. She extracted pitch from five points of each tone, located at the beginning, middle, and end, along with the high and low point of each contour. Results were then pooled to reveal the average pitch height and contour. Liu utilized an autocorrelation (AC) algorithm in CSL speech analysis software to gather

her data, leaving out the first and last 5ms of each tone to avoid miscalculations. When the macros did not extract F0 values that seemed reasonable, analysis parameters were adjusted and pitch measurements were done by hand (Liu, 2001; 56).

Table 2.1 F0 values of the four tones (Liu, 2001; 100)

Table 3.a. Mean F0 (Hz) values of Mandarin tones measured in three positions in AD and ID speech for mothers of 6-8 month-olds.

	Measured position within vowel					
	Initial		Middle		Final	
Tone	ADS	IDS	ADS	IDS	ADS	IDS
1	236.04	289.28	235.49	300.35	234.28	306.61
	(9.28)	(8.70)	(9.06)	(10.44)	(8.74)	(11.10)
2	216.56	257.01	211.16	253.71	226.80	289.88
	(10.12)	(6.36)	(8.49)	(6.57)	(9.27)	(7.06)
3	205.06	256.34	186.22	228.57	188.93	244.74
	(5.27)	(8.23)	(5.14)	(6.57)	(6.01)	(7.81)
4	248.51	329.73	235.06	296.76	219.28	264.44
	(7.50)	(8.14)	(7.90)	(8.71)	(9.33)	(8.39)

Note. **Standard errors in parentheses. ADS: adult-directed speech; IDS: infant-directed speech.**

Liu (2001) found that vowel type and speech type were both significant factors in determining the pitch range and pitch height of each speaker. Infant directed speech had significantly higher mean pitch. Furthermore, Liu's analysis reveals some intriguing pitch movements in the four tones of Mandarin Chinese (see Table 2.1 above). For example, the second tone was produced with a slight dipping contour in both speech styles (see Fon et al, 2004). Since Liu's (2001) speakers were all speakers of Taiwan Mandarin living in Kaousiung Taiwan, this data provides a strong argument that this phenomenon exists in Taiwan Mandarin.

Liu's (2001) study also revealed an intriguing phenomenon relating to speech style and tone range. Despite the significant differences between the tone range in different speech styles, the respective ranges of each tone fit into a very nice hierarchy (see figure 2.1 below, taken from Liu, 2001; 110). Tone 1 consistently had the smallest range, tone 4 the widest, with tones 2 and 3 in-between. Infant directed speech had wider range in each category, but the hierarchy was maintained.

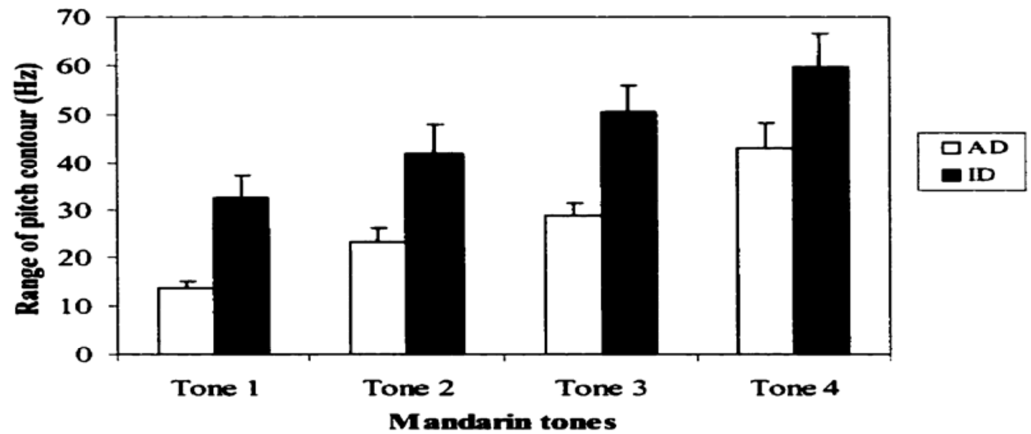


Figure 22.a. Averaged F0 ranges of pitch contour of 4 Mandarin tones (in Hz) in AD and ID speech across mothers of 6-8-month-olds (se in error bars).

Figure 2.1 Average tone range in two speech styles (Liu, 2001; 110)

Acoustic analyses have also been used with other languages, and in other fields; such as sociolinguistics. Yuasa (2001) performed an acoustic analysis of spontaneous speech in order to explore differences in the pitch range of American English and Japanese based on the relationship between speakers; familiar vs. non-familiar. She had eight American and eight Japanese participants speak in their native language in a casual atmosphere (e.g. a living room) for about 10 minutes. Microphones were attached to informants' throats. Although speech was not controlled, Yuasa provided the topic of food for conversation, because food was "an emotionally non-provocative but not uninteresting topic" (Yuasa, 2001, p. 103).

Yuasa's (2001) analysis of F0 did not utilize the mean and Standard Deviation as measures of pitch range because these parameters were used in earlier researcher to analyze only read speech. Spontaneous "speech data … tends to occasionally contain a very large Standard Deviation. This causes the 3.5SDs or 4SDs calculation to sporadically exceed the speaker's whole pitch range" (Yuasa, 2001, p. 26). Instead, Yuasa randomly selected a series of intonation groups from each speaker's conversation, making sure that each group was about one minute long. She extracted the highest and lowest F0s, and then measured median range values that were close to the mean (50^{th} percentile).

Yuasa (2001) found that both male and female Japanese speakers' pitch range was wider when they spoke with someone familiar to them vs. someone who was unfamiliar. She also found that females had a wider pitch range than males using the ERB and Hz scales, but that they had a more narrow pitch than males when she measured pitch range in semitones.

In another socio-phonetic study, Lewis (2002) performed an acoustic analysis to investigate potential interlocutor effects on average pitch, pitch variability and pitch range of young English speaking females' spontaneous speech. These three variables were operationalized by using the median, standard deviation, and 80% range of pitch, respectfully. She analyzed a total of 12 female participants who spoke with four different interlocutors; one male and one female peer, as well as one male professor and one female professor.

> Measurements were made using the autocorrelation method in Praat software (see Boersma, 1993, for a description of the algorithm). A unique combination of settings within the Praat algorithm was created for each subject based on what gave the most accurate results; I experimented with settings until I eliminated/reduced spurious or suspicious measurements, and then these settings were used consistently, i.e. the same settings were used for all measurements made on data for a given subject (Lewis, 2002, p. 47-48).

This method of measuring data has great potential to establish a personalized analysis for each speaker, but could also limit a researcher's ability to make generalizations. Lewis (2002) used different pitch settings for many speakers, but also claims that she used very accurate default settings of the Praat autocorrelation method for the rest. In line with current trends, Lewis used the ERB scale because it is reportedly superior to both Hz and semitones in reflecting frequencies below 500 Hz. Her analysis suggested that interlocutor variables such as gender and status had a significant effect on the pitch measurements made.

Limitations of Spontaneous Speech Data

One of the major limitations of spontaneous speech data is one shared by all linguistic data, the inevitable influence of the observer's paradox. Simply by observing, the researcher is influencing the nature of the spontaneous speech. Because spontaneous speech is typically produced as intercommunication between two or more interlocutors (participants), and influenced by other members of the 'audience' (Bell, 1984), any participant may play a significant role in how communication is carried out.

Another more specific weakness of spontaneous speech data is the general lack of control associated with non-read, impromptu conversation. This lack of control makes it very difficult to eliminate intervening variables and thus it is associated with a third limitation; a demand for careful interpretation within context. With so many potential social, cognitive and linguistic variables affecting contextualized speech, researchers must be able to discern both major and minor variables, and analyze the data appropriately.

Finally, just like read or planned speech, spontaneous speech data has only a limited level of generalizability. For example, female native-speaker descriptions of a series of pictures to female non-native-speakers may not reveal anything about how the same individuals would describe the same pictures to male non-natives or to native speakers of either sex. Social factors such as sex, age, relationship between interlocutors, and so on, may play a significant role in shaping the results (Yuasa, 2001; Lewis, 2002).

In light of some of these weaknesses of spontaneous speech data, it is clear that data gleaned from this type of research cannot necessarily replace more controlled data based on explicit memory. However, spontaneous speech data can serve as an important comparison. Despite a lack of control, it resembles casual speech more than read speech. As the following review of literature reveals, much of the research on Mandarin tone perception and production has failed to analyze spontaneous speech at all, neglecting several significant variables influencing language acquisition.

Studies on the L1 Perception of Mandarin Lexical Tone

One cannot acquire ability to produce Mandarin lexical tones without also learning to perceive them. The relationship between perception and production are real, but are not simply causal (Llisterri, 1995; Major, 1994). Simply because of this complex relationship, a review of research regarding the perception of Mandarin lexical tone is necessary. Furthermore, a large body of insightful research has been done in this area which may prove useful in understanding why there is such a discrepancy between L1 and L2 production of tone.

Connell et al. (1983) wanted to understand how much the shape of a Mandarin tone could change before it was recognized as a different tone by native speakers. They artificially modified read monosyllables, then tested native Chinese participants' ability to perceive the original tones. Their findings suggested that tones can withstand a large amount of change without being perceived as different or ambiguous tones. The strength of these results is limited however, partly by their mysterious method of filtering out participants from twenty-eight to ten because of language background, and partly by their exclusive use of monosyllables.

Massaro, Cohen and Tseng (1985) performed a study to better understand the roles of F0 contour and height in native judgments of tone. They had six well-educated native Chinese speakers evaluate prerecorded first and second tones after adjusting the contour and heights of the tones in equal increments. They found that correct recognition depended on both contour and height, and that when one cue was produced in a more ambiguous state the listener would depend more on the other. This study suggests that production of tone in the proper register plays a role in communication. However, some potential drawbacks of this study are 1) the exclusive use of monosyllables, 2) possible conflict of language backgrounds (i.e. there is no indication of where native speaker judges come from, nor where the speaker who produced the tones was from).

Whalen and Xu (1992) conducted four studies about native speakers' tone perception on the suprasegmental level. Utilizing the wonders of modern technology, they altered recorded Mandarin speech and produced it without tone contour. This removal of tone contour enabled the researchers to 'produce' words containing duration and amplitude (volume), but no tone. They eliminated lexical interference by using the same syllable /ba/ for all tests.

The first experiment suggested that natives can easily discern second, third, and fourth tones by their respective amplitudes, despite the lack of tone contour. Even the first tone was recognized 38.5% of the time. When they removed factors related to the respective duration of each tone (e.g. produced a third tone with the shorter duration of a fourth tone); the natives were still able to recognize the same tones at above chance levels (55.3, 69.5, and 92.3% respectively). Apparently, native speakers are able to differentiate monosyllabic tone based solely on amplitude. It is still unclear whether non-native students of Mandarin are sensitive to the same factors. It is also unknown whether this study has direct implications on tones produced or perceived in spontaneous speech.

Whalen and Xu's (1992) last two experiments gauged native speakers' abilities to recognize tone by only part of its contour. Only the first tone was recognized well at all locations on the duration scale. The other three tones were better identified when natives heard "the middle to late portions of the syllable" (Whalen & Xu, 1992, p. 37). The third tone was often mistaken as a second tone if the native heard the later portion of the syllable. Whalen and Xu also claim that low pitch frequency was heard as either second or third tone, suggesting that register may have an impact on tone recognition. These findings raise the question of how different dialectal groups of native speakers, as well as L2 learners, would fare on the same tests.

Shen et al (1993) confirmed an earlier experiment that fundamental frequency (F0) turning point aids in the recognition of second and third tones. Apparently, the shift from falling to rising is the main distinguishing factor of the third tone. They also found that native judgments were not significantly altered by the height of the FO, a result that suggests register may not be vital for tone perception.

Stagray and Downs' (1993) study suggests native speakers of Mandarin perceive more categorical changes in tone, while native speakers of English have a higher differential sensitivity. This higher differential sensitivity allows English speakers to hear the more subtle changes in pitch, yet may not prove helpful in distinguishing the larger categorical changes made in Mandarin tone. Further studies need to be conducted on American students of Chinese to see whether a more categorical perception is developed. This higher differential sensitivity in American students of Chinese may have direct impact on how they learn to produce tones.

Concerning recognition of tone, Fox and Qi (1990) wrote, "It has often been noted in the phonetic literature that the identification of a particular sound can be significantly affected by the surrounding phonetic context" (p. 261). Chen and Cutler's (1997) study revealed that phonological and lexical priming does enhance native speakers' recognition of spoken words in Cantonese. Interestingly enough, semantic priming proved to be the most effective in the audio portion of the study. This study revealed that phonological and semantic association most likely enhances native speaker recognition of vocabulary. This finding corresponds with Lee, Vakoch and Wurm (1996), who suggested Mandarin and Cantonese speakers perceived tones better when they were part of familiar words from their own language background.

This concludes my review of studies investigating native perception of tone. I now turn to studies focused on the production of Mandarin lexical tone by native speakers of *SC*.

Studies on the L1 Production of Mandarin Lexical Tone

The following studies on L1 production of Mandarin tone are divided into two main groups: acoustic and microanalysis.[4]

Acoustic Studies of L1 production

G. T. Chen (1973) conducted an acoustic analysis of read speech to determine the respective pitch ranges of English and Chinese speakers. Chen investigated the pitch range of four native speakers of Midwestern English at the word and sentence level as they read both English and Chinese. G.T. Chen (1973) also recorded four Mandarin speakers who spoke “with no noticeable foreign accent” (p. 161). Chen focused only on F0, leaving duration and amplitude out of the analysis. When speaking their native language, the Chinese had a pitch range 1.5 times wider than the Americans when they spoke English. Even when the Americans spoke Chinese, their pitch range didn’t match the Chinese, although it did have a substantial increase.

For the word level portion of his study G. T. Chen (1973) utilized place-name words that met the following criteria:

“The test word must

1. be identical in meaning.
2. be able to be used in a similar grammatical structure.
3. occur in identical locations in a sentence.
4. have the same number of syllables and similar syllable structure.
5. have similar vowels or main vowels in vowel nuclei.
6. have consonants similar in articulation” (p. 162)

The sentence level portion of the study consisted of ten written responses to conversational questions in both English and Chinese. Participants had two weeks to get familiar with the materials, and Chen conducted several recordings of each test item. He

[4] Microanalysis refers to native speaker judgments of individual tones.

then personally selected the most desirable recordings.

G. T. Chen's (1973) analysis included a measurement of F0 of each vowel at three points. He then found the standard deviation of the pitch of each speaker, as well as the "relationship between the subject's pitch range and his average fundamental frequency" by obtaining the product moment correlation coefficient (G.T. Chen, 1973, p. 166).

G. T. Chen (1973) also utilized a Pitch Range Index (PRI) to indicate relative pitch ranges because there was a positive correlation between the pitch range and average F0 of each speaker. Some of the major weaknesses of Chen's study are 1) the misconception that read speech is equivalent to spontaneous speech (see G.T. Chen, 1973, p. 163), 2) the inclusion of only four participants from each language group, and 3) his method of selecting "the most desirable" of each recording (G.T. Chen, 1973, p. 165). Another potential problem could be the possibility of a transfer of English intonation patterns in the reading of the word list because the words were all English loanwords.

Despite these weaknesses, G. T. Chen's (1973) study still argues quite convincingly that native English speaking students of Chinese have a narrower pitch range than native Chinese sp. Other researchers (i.e. Miracle, 1989; Shen, 1989; Q. Chen, 2000) have also suggested that Americans' comparatively narrow pitch range leads to tonal errors.

In one of the earliest attempts to analyze spontaneous Chinese speech acoustically, Tseng (1981) conducted a series of studies investigating the relationship between intonation and tone, with a special focus on comparing the tones produced in read and spontaneous speech. She found that tones produced in the spontaneous speech of two female natives from Beijing did not correlate well with predictions based on

Chinese phonology, or even with the citation forms recorded in the same study by one of the same participants. She also found read speech to be much simpler than spontaneous, without as much influence from the "interacting/extra-linguistic levels of information" (Tseng, 1981, p. 147). Although she used a very small sample size, and did not match the speech tasks appropriately, Tseng's research suggest that as formidable and complex spontaneous speech may appear, researchers must investigate it before they can develop accurate phonological theories.

In another attempt to understand the nature of Mandarin intonation, Shen (1990) performed a landmark acoustic analysis of six female native speakers of Chinese. Despite previous notions, she provided strong support that intonation and stress really do influence tone contour, which also provided some validation for the educated impressions of Chao (1968) concerning tone sandhi rules. Her design was simple: seventy-two four-word utterances were prepared with all possible tone combinations (except the neutral tone). Thus, the sentences were not natural but "grammatically and semantically acceptable" (Shen, 1990, p. 14). The sentences were written in both statement and question form. After becoming accustomed to the sentences, six female participants from Beijing read them. The read speech was recorded and then analyzed acoustically.

Shen's (1990) pitch traces of the acoustic data showed that sentences in question form were produced in a higher pitch register than sentences produced as mere statements. They also showed what could be considered unusual changes in tone contour (See Shen, 1990, pp. 35-37). These "abnormal" tone contours suggest that native speakers do not always produce tones in citation form.

Some of the weaknesses of Shen's (1990) study are 1) failure to analyze spontaneous speech, 2) small sample size, and 3) lack of attention toward sociolinguistic factors such as sex or possible dialect differences (all participants were from Beijing). Despite these problem areas, Shen's study set a foundation for future studies by strongly suggesting that intonation and stress do affect tone contour, a theory which had not been accounted for in most earlier research on Mandarin tone. Shen (1990) explained,

> Tone is reduced to the paradigmatic components sufficient for signaling lexical oppositions. [Fundamental frequency] is one of the acoustic attributes in the sound wave, which is continuous, dynamic, and context-adjusted. In the phonetic production of tones, while the orosensory targets of tones may be largely independent of context, the motor activity that is required to achieve these targets is likely to depend strongly on the adjacent tones and other suprasegmental influences such as stress and intonation (1990, p. 429).

Thus, native produced tones most likely interact with neighboring tones, stress, and intonation. It is also likely that these variations in tone do not affect native speaker recognition, as long as they are kept within context. Because Shen did not evaluate non-native perception of the tones produced in her study, it is still unclear if L2 learners would be able to perceive the tones accurately.

Acoustic analysis has proven useful in other areas as well. Through acoustic analysis, striking cross-linguistic tendencies have been found to exist in the melodic contours produced by American and Chinese mothers (Papousek et al, 1991). Despite this similarity between Americans and Chinese, significant differences have also been noted in the frequency range of their speech. Papousek and Hwang (1991) assert that Mandarin mothers produce a more complex and rapidly fluctuating frequency than

American mothers. Mandarin speakers (similar to Cantonese speakers) may have higher and lower mean frequencies than English infants the same age (Lee, 1996), and this difference may persist throughout life. It is likely that these differences in higher and lower frequencies between English and Chinese speakers plays a role in molding the perceptive and productive capabilities of infants.

Papousek and Hwang (1991) conducted a study comparing native speech in three simulated social registers – adult conversation, foreign language instruction, and babytalk to infants. Although these contexts were not spontaneous or contextualized, Papousek and Hwang claim real-life context would have only produced even more significant differences. Higher peak frequencies were the norm in both babytalk and foreign language instruction. They found a doubling of terminal rises in babytalk that modified 47.4% of the final first tone syllables, although the other final syllable tones remained unaffected.

Somewhat surprisingly, Papousek and Hwang (1991) also found that the F0 patterns produced in babytalk “exemplified clarification and simplification of global intonation patterns at the cost of tonal variation” (p. 495).   Adult conversation had the most rapid fluctuations in frequency, and the most fused or flat tones. The implications of these findings are not fully understood, but the fact that their acoustic analysis of adult conversation contained the most tone “errors” sheds serious doubt on the present pronunciation target established for L2 learners.

Shen and Lin’s (1991) study of sixteen monolingual Mandarin speakers was focused on tones in connected speech. Specifically, they wondered whether the adjustment of intonation was both anticipatory *and* preservative – in other words – if native speakers would alter their tones both before and after other words to compensate for the respective contours. Shen and Lin's (1991) findings were in the affirmative and they concluded that "acoustic properties of tones in isolated citation form [were no longer] appropriate for use as norms" (p. 421).

In order to better understand contextual tonal variation, Xu (1997) conducted an acoustic analysis of bi-tonal nonsense words read by eight male native speakers of Beijing Mandarin. More specifically, Xu examined tonal variation due to anticipatory and carry-over effects by analyzing the F0 of tone over time. His findings suggested that tones are influenced by both assimilatory and dissimilatory effects. However, Xu found that anticipatory effects were less substantial and were generally dissimilatory, while carry-over effects were greater in magnitude as well as assimilatory in nature.

Xu (1997) had his eight participants read and produce /ma/ four times in isolation, utilizing one word for each of the four tones. Participants then produced nonsense words /mama/in all possible tonal combinations. The nonsense words were produced in the middle of four different carrier sentences (*Wo jiao* ____ *lianxi/lianluo*). *Jiao* occurred as 'tell' (fourth tone) and 'teach' (first tone). The *lian* was either a fourth or a second tone. Xu controlled for speech rate by having participants read to the beeps of a timer.

Xu (1997) utilized a very high level of control and concluded that "while the contextual effects found in this study cannot be considered to be maximal in magnitude, they should be relatively free of the influences of extraneous factors, thereby allowing for relatively straightforward interpretation" (p. 65). The only problem with this conclusion is that by removing "extraneous" factors the researcher could be creating a sterilized sample free from many of the real factors that influence natural, spontaneous speech.

Xu (1997) took measurements of the max and min F0 of each segment, as well as five points of each segment. Each segment consisted of a coda or nucleus of each syllable, thus each /mama/consisted of two segments. In order to understand the effects of both cross syllabic and adjacent tones, Xu used a “set of three-factor repeated-measure ANOVAs” (p. 69). His independent variables were 1) the offset value of the syllable in the carrier sentence immediately preceding the /mama/sequence (high or low), 2) the offset value of the first syllable in the /mama/sequence, and 3) tone of the second syllable (69-70). The dependent variables were the F0 values at vowel onset, one quarter, one half, and three quarters into the vowel, and at the end of a vowel of the second syllable in the /mama/sequence.

Xu’s (1997) findings suggest that F0 contour is influenced substantially by surrounding tones. The greatest influence is from the preceding tone, whose offset value virtually determines the starting F0 of the following tone. A following tone was also found to affect a preceding tone, although this influence was not as strong. Surprisingly, a “low tonal onset raises, rather than lowers, the F0 of the preceding tone” (p. 82).

Fon and Chiang’s (1999) acoustic case study of Taiwan Mandarin raised an important question: what are the differences between tones produced by Mandarin speakers in Taiwan and Beijing? Although their one participant cannot be considered representative, the results suggested by their research are still intriguing. This study has become the primary instigator and catalyst for this study.

Fon & Chiang (1999) had two main goals: 1) to quantify Chao’s five point tonal scale by using an acoustic approach, and 2) to investigate the possibility of tonal discrepancy between Beijing and Taiwan Mandarin. After introducing the problem, Fon

& Chiang discussed both of Chao's systems of representing Chinese tones. They suggest that Chao purposely "adjusted the reference points so that between Scales 1 and 2, the interval is narrowed down to only one semitone while the interval between Scales 2 and 3 is enlarged to three semitones" (Fon & Chiang, 1999, p. 16). They also presented two formulas designed to translate pitch height into Chao's two tonal systems (and vice versa).

Fon & Chiang (1999) analyzed the read speech of one twenty-two year old female student from National Taiwan University. Her parents were both Taiwanese, but they reported that the participant used more Mandarin than Taiwanese on a daily basis. Fon & Chiang only investigated tone in read mono and tri-syllabic words, arguing that tones become so distorted in spontaneous speech that the contours are barely recognizable. Hence, the researchers chose to engage their participant in reading tri-syllabic words selected from a database (CKIP). It is unclear exactly how the words were selected, but Fon & Chiang (1999) did make an effort to avoid "literary words and nonwords" (p. 19). The participant was recorded reading 2107 tri-syllables and 1453 monosyllables (selected from the tri-syllables). The recording sessions consisted of three or four 20-30 minute sessions a day for two days. During the sessions the participant was allowed to rest whenever needed.

They divided tones into different tonal contributions, thus, the number of tonal combinations ranged anywhere from 10 (i.e., 331) to 127 (i.e., 444). They analyzed tones occurring in each position of each tri-syllable to explore contextual variation. Tones were also distributed into word classes, with an overwhelmingly large amount of nouns (86.67%). This distribution was not reflective of the amount of nouns that would be produced in natural speech.

Fon and Chiang (1999) measured two points in tones 1 and 4: the beginning and ending points for tone 1 and the initial *highest* and final *lowest* points. For tones 2 and 3, three points were measured, beginning highest, medial lowest, and the ending highest points. They found that the tonal values were different from the prescriptive ones based on Beijing Mandarin, demonstrating a narrower four-way distinction, lower tonal heights, and more conservative tonal contours. They found the descriptive values of the participants' tones to be 44, 323, 312, and 42.

Fon and Chiang (1999) did not account for linguistic factors such as speech rate, speech style, focus/stress, etc. They ignored the possibility of interlocutor effects, which Gass and Selinker (2001) have described as an adjusting of speech style based on whom we are communicating with. They also failed to limit their results to the female gender, even though their one participant was female. Finally, by analyzing the speech of only one participant they ultimately limited their findings to an ungeneralizable case study.

Despite the fact that this next study is not dealing directly with Mandarin, I include it to shed light on the value of accounting for sociolinguistic variables in tone research. Chiung (2001) conducted an acoustic analysis of thirty Taiwanese (Southern Min) speakers in an attempt to understand the nature of Taiwanese tone 5 (reportedly a low rising tone) as well as major factors affecting its production.

One of the major strengths of Chiung's (2001) study was that he noted the age, sex and region of residence for each participant. Chiung found that at the end of sentences tone 5 was most frequently realized as a dipping tone. It is suggested that this phenomenon stems directly from *guoyu*, which does not have a low rising contour except for that associated with the dipping tone.

Fon et al. (2004) conducted an analysis of Taiwan Mandarin to determine the true nature of Mandarin tones 2 and 3. Only one female speaker of Taiwan Mandarin was recorded (the same as reported in Fon & Chiang, 1999). Fon et al (2004) reported that the participant's "everyday language is mainly Mandarin," but that she spoke Taiwanese and that her parents were both Taiwanese (p. 251). The participant was recorded reading 222 monosyllables, covering 111 minimal pairs. Only 96 of these monosyllables were analyzed in this study, as the remainder dealt with the 1st and 4th tones. F0 was extracted at three points of each voiced portion; beginning highest, middle concave, and end highest. Figure 2.2 on the next page shows duration and mean F0 measurements of both tone 2 and 3. These findings suggest that both tones have a dipping contour.

In addition, Fon et al (2004) report that tones 2 and 3 differ significantly at the mid and end positions, but not at their starting points. Despite the fact that only one speaker of Taiwan Mandarin was used, the researchers conclude that both *putonghua* and *guoyu* speakers are producing a dipping second tone. This claim demands further investigation involving more speakers from both dialect groups in a variety of speaking tasks.

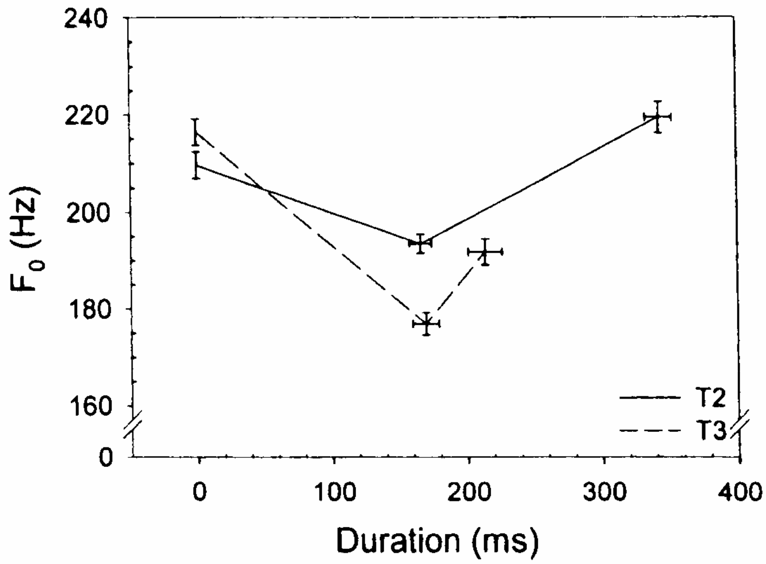


Figure 1: Mean F_0 and duration measurements at three points of 48 minimal pairs contrasting T2 and T3 in Taiwan Mandarin. The vertical and the horizontal bars indicate standard errors.

Figure 2.2 Mean F0 measurements of tones 2 and 3 (Fon et al, 2004; 253)

Microanalysis Studies of L1 production

Li and Thompson (1977) conducted a study of seventeen Mandarin speaking children living in Taiwan. They suggested that children between 1½ - 3 years old master tone before other segmental portions of speech, an argument that has been strengthened by more recent research (Lee, 1996). Li and Thompson claim children had the hardest time producing the rising tone, even confusing the dipping and rising tones until the two-word stage. They attributed this difficulty with the second and third tones to their similar feature: both tones have a rising component. The main weakness of their study was the fact that they personally judged all utterances and recorded the tones they perceived.

In support of Li and Thompson (1977), Hua and Dodd's (2000) study of 126 nursery children in Beijing (aged 1½ - 4½) also suggests tone production is mastered before consonants and vowels. Children were engaged in picture naming tasks. The researchers found tone errors were very rare, but children were only evaluated on words they actually produced (omissions of entire words were not considered). "Only two tone errors were observed in the entire data corpus, and they were produced by the children of the youngest age group" (Hua & Dodd, p. 21). Children from the two youngest groups (aged 1½ - 2½) did occasionally produce tones that should have been altered according to tone sandhi rules. Unfortunately these errors were not well documented in the study.

Native intuition is likely to be quite accurate in the assessment of basic tone skills. It is doubtful, however, that close analysis can be anything but biased as lexical, dialectal, and perceptual differences come into play. I suggest that native speakers would be best utilized providing a holistic analysis (i.e. rate this speakers' tone production on a scale of 1 to 5, 1 being the poorest and 5 being exactly like a native-speaker).

This concludes my review of L1 production of lexical tone. Table 2.2 on the next page presents a brief synopsis of the production studies discussed. Before moving on to review L2 studies, let me first summarize the role that perception and production play in the native acquisition of tone.

The Role of Production and Perception in L1 Acquisition of Tone

Results from L1 research concerning perception and production suggest that early on in their development, Chinese children often confuse the third and second tones and struggle the most with the second tone (Li and Thompson, 1977; Whalen and Xu, 1992; Lee, 1996). Compared to American children, Chinese may develop a wider pitch range early on (G. T. Chen, 1973; Papousek and Hwang, 1991), although the height of this register may also depend on where speakers are from (Fon & Chiang, 1999). Suprasegmental factors such as amplitude and turning point may also play a role in adult native speakers' recognition of tones (Whalen and Xu, 1992; Shen et al, 1993). It is clear that context and connection of spoken form with meaning play a crucial role in the child's ability to produce correct tones (Lee, Vakoch & Wurm, 1996; Chen & Cutler, 1997; Hua & Dodd, 2000). We are still uncertain about children's development of tone sandhi in multi-syllabic speech (Hua & Dodd, 2000).

Table 2.2 Summary of Findings from L1 Production Studies

Source	**Type of Study**	**Participants**	**Basic Findings**
Li & Thompson (1977)	Micro-analysis of spontaneous speech	17 child speakers of Taiwan Mandarin	Chinese children (aged 1.5 – 3 years) mastered tone before consonants and vowels. They struggled most with the rising tone, even confusing second and third tones until the two-word stage.
Tseng (1981)	Acoustic analysis of read and spontaneous speech	2 speakers of Beijing Mandarin	Spontaneous speech does not exhibit the same tonal patterns as read speech. Phonetics and phonology cannot predict how a sentence will be produced.
Shen (1990)	Acoustic analysis of read speech	6 speakers of Beijing Mandarin	Intonation, stress, and tone all interacted mutually in rapid speech.
Papousek & Hwang (1991)	Acoustic analysis of spontaneous speech	6 speakers of Taiwan Mandarin	Adult conversation contained flat and fused tones. Speech directed toward L2 learners emphasized tone more than speech directed toward native infants.
Shen & Lin (1991)	Acoustic analysis of read speech	16 monolingual speakers of Mandarin	In connected speech tones were altered both before and after words to compensate for the respective contour of each word.
Xu (1997)	Acoustic analysis of read speech	8 speakers of Beijing Mandarin	Tones were altered both before and after words, but assimilatory carry-over effects were greater in magnitude, virtually determining the onset of the preceding tone.
Fon & Chiang (1999)	Acoustic analysis of read speech	1 speaker of Taiwan Mandarin	A speaker of Taiwan Mandarin produced unusual and mid-level tones in both isolated and non-isolated syllables.
Hua and Dodd (2000)	Micro-analysis of spontaneous speech	129 child speakers of Taiwan Mandarin	Children mastered tone production before consonants and vowels. Tone sandhi rules seemed to pose somewhat of a problem for a few of the younger children.
Chiung (2001)	Acoustic analysis of read speech	30 speakers of Taiwanese	Found possible influence from *SC* on tone 5 of the Southern Min dialect spoken on Taiwan.
Fon et al (2004)	Acoustic analysis of read speech	1 speaker of Taiwan Mandarin	Found potential similarities between the second and third tones of Taiwan Mandarin.

Studies on the L2 Perception of Mandarin Lexical Tone

As mentioned earlier, Papousek and Hwang (1991) found that when native speakers were engaged in student directed speech; they produced higher pitch frequency changes than when they were engaged in adult conversation or babytalk. Papousek and Hwang (1991) concluded that the pitch frequency changes produced in foreign language instruction “exemplified clarification of tonal patterns at the cost of intonational variation” (p. 495). Thus, L2 students of Mandarin are likely receiving at least some input carrying intensified tones and watered down intonation.

Stagray and Downs (1993) suggested that English speakers without experience in a tonal language have a higher differential sensitivity than native Chinese. This higher differential sensitivity allows English speakers to hear the more subtle changes in pitch, yet this may not prove helpful in distinguishing the larger categorical changes made in Mandarin tone. Further study needs to be conducted on English speaking students of Chinese to see whether a more categorical perception is developed. It may be that English speakers’ capacity to perceive more subtle changes in pitch affects how they produce tones.

Sun (1998) has performed three excellent studies that have helped clarify some of the previous confusion over how American students of Chinese perceive tones. Her first study included fifty students enrolled in an intensive Mandarin language-training program in Beijing. Forty-two of the same group took part in testing again near the end of the semester. Sun investigated participants’ categorical perception of tone utilizing a tone identification task (TIDT). She found an overall strength in the identification of the first and fourth tones. Students who had been introduced to a tonal language in their childhood scored significantly higher in the recognition of second and third tone than did those students without prior tone language experience. Sun did not find that academic level had a significant effect. The twenty-eight native speakers who also participated in the TIDT scored a mean of 98% and above. Errors made by the native speakers when marking the tones of nonsense words were four times higher than their errors with common or uncommon words, showing a possible connection between meaning and tone recognition. 81% of the non-native participants stated that they thought it was easier to recognize the tones of words they knew.

Participants also tended to be more hesitant in the identification of third tone and quicker to identify the fourth. Sun asserted that students' knowledge of the Mandarin tonal system is often lacking. Furthermore, "difficulties in contrasting the four tones persist despite more semesters of instruction" (Sun, 1998, p. 83). Non-final position of tone in polysyllabic words also proved to be more difficult, especially for those without exposure to a tonal language as a child.

Sun (1998) also tested participants' perception of their own tones. Fifty students participated at the beginning of the semester in Beijing, and thirty-six of those took part in the second round of data collection. Sun used students' own production from the reading and repetition portion of earlier production tests as the elicitation instrument. Participants' ability to perceive their own tones increased with higher levels of academic proficiency. Participants were also able to perceive their own first and fourth tones more accurately than their productions of second and third tones. Participants' capacity to recognize their own tones was higher than their ability to recognize native speakers' tones, until the fourth year.

Sun’s (1998) results suggest that American learners do have at least some perceptive parameters that parallel those possessed by native speakers. Tone substitutions varied from participant to participant, and yet the most difficult contrast was between second and third tones. Phonological position affected perception of tone, with higher scores on monosyllabic and the last syllable of polysyllabic words. Sun states that participants had a problem with the phonemic features of tone. It was also found that participants sometimes used the fourth tone as a substitute when they were unsure what tone they heard. While perception of tone did appear to improve through time, lack of control makes it impossible to determine the reasons why. Intriguingly, participants were more familiar with tones they produced in the reading exercise than those produced in the repetition exercise, a phenomenon opposite to that of the native judges. Native speakers perceived the tones produced in the repetition exercise as more authentic.

This concludes my review of L2 perception studies. A discussion of findings occurs after the next section on L2 production.

Studies on the L2 Production of Mandarin Lexical Tone

Naturally, this review does not cover every aspect of tone production. For example, the relationship between non-native speakers’ vocabulary acquisition and their tone production is not considered in depth. See the following for a more specific discussion of how phonology interacts with both spoken and written lexicon (Papagno et al, 1991; Cheung & Hsieh, 1997; Zhou et al, 1999; Lin, 2000; Spinks et al, 2000; Rummer & Engelkamp, 2001).

In an attempt to investigate interference errors, Chiang (1979) made some groundwork observations of potential tonal challenges for L2 learners of Mandarin. He claims that L2 learners' tones are adversely affected by English intonation, yet provides only predictions and speculation to back up his claims. Chiang's focus on the possible influence of English rising and falling intonation on L2 production of tone are intriguing, but his contribution ends there.

Jie's (1988) study of Mandarin tone attrition posed the question of whether an order of attrition exists. Jie concluded that L2 learners' of Mandarin Chinese do not suffer attrition as long as they can remember the lexical item associated with the tone. His study provided an important argument that has also been neglected in the literature; the relationship of general lexical acquisition with tonal acquisition. The major flaw of his study was that it was entirely text-based, that is, participants would read a sentence in Chinese characters and then say the word that correctly filled in the blanks. Tones were not analyzed acoustically, nor were they compared with native production. The most devastating factor was that no tone sandhi rules of speech were accounted for.

In a study supporting transfer from L1, Shen (1989) looked at eight first semester Chinese students. Her microanalysis was designed to determine whether register or contour errors were the most problematic. Four native speakers judged the non-natives' tones, but Shen's acoustic analysis only compared the tones produced by the participant with the worst tones with those produced by a Beijing Mandarin speaker recorded in a prior study (see Shen, 1985). Shen claimed fourth tone errors were the most common, followed by first tone errors. Her results also suggested that the habitual pitch range and suprasegmental production of American English speakers influences tone production. Further studies are needed to verify her claims concerning L1 transfer of register and pitch, as well as the importance of register in natural communication with natives.

Miracle's (1989) study presents contradictory findings, at least at first glance. He found that the second tone was the most problematic, with the third tone in second place. These tonal errors are contradictory of those tones Shen (1989) claimed to be most difficult for learners. The contradiction could be partially explained by the differences in the amount of time their participants had studied Chinese. While Shen's participants only had four months of study, Miracle's ten participants had all engaged in over one year of study (one participant over three years). In partial agreement with Shen's study, Miracle found an equal amount of register and contour errors, strengthening the claim that register errors are being made.

In an attempt to evaluate the effectiveness of two popular systems of romanization (pinyin and gwoyeu romatzyh), McGinnis (1997) investigated the tone production of American and Japanese beginning students of Mandarin. He concluded that gwoyeu romatzyh did not improve pronunciation any more than pinyin. McGinnis recorded students reading familiar dialogues and two native judges assessed tones. He found no significant difference between the performances of the two groups. The Gwoyeu Romatzyh system would not produce significantly better tone production.

McGinnis (1997) analyzed the read speech of 29 American English and 18 Japanese L1 speakers. The independent variables were the two types of romanization used: pinyin and Gwoyeuh Romatzh (GR). The response variables were the percentage of tones produced correctly as judged by two native speaker judges, one from Mainland and one from Taiwan (79% interrater reliability). His findings suggested that there was no significant difference between the students tone production based on the system of romanization they used. More interestingly, McGinnis found that his eight native judges initially chosen to judge the non-native tones had such a low interrater reliability that he was forced to throw out all but two female judges, presumably because they had the highest degree of interrater reliability. This lack of interrater reliability is perhaps one of the most striking features of his study for it suggests that simply selecting a large group of 'native speakers of Chinese,' without any regard to their language background, is an ineffective instrument to assess tone production.

Chen's (1997) study investigated more advanced learners of Chinese; six second year students of Mandarin. A microanalysis of tones was performed by Chen and one other native speaker, who claimed that second and third tones were often confused with each other and that the same mix-ups were found between the first and fourth tones. Chen (1997) stated that the majority of tonal production errors were actually manifested in "*level* tones [not first tones] which do not exist in standard Mandarin at all" (p. 32). This perspective seems to correlate with the register errors reported by Miracle (1989) and Shen (1989). However, as mentioned earlier, Fon and Chiang (1999) have suggested that register "errors" may be produced by native speakers as well.

Sun (1998) engaged participants at four different levels of Chinese instruction in three production tasks: repetition of words spoken by a native, reading of *pinyin* words with diacritics, and oral translation into Chinese. Fifty participants participated in the first study conducted at the beginning of the semester and thirty-nine of those in the second study eleven weeks later.

The repetition task consisted of a simple repetition of the same words participants had heard in the Tone Identification Task (TIDT). The reading task required students to review, and then read aloud 96 common and 96 uncommon words (same words as in

TIDT). Three native speakers, who had scored high in the TIDT, served as judges during each of the three production tasks, marking the tones produced while they were produced. These judges showed less agreement on the nonsense words, indicating that listeners judge correctness of tone on more than just the sounds produced.

Sun's (1998) results suggest that tones are produced more accurately in both common and uncommon words than they are in nonsense words. Correct production on the reading portion was lowest in the di-initial and tri-initial and tri-medial positions. Participants performed best at repeating words just heard, and worst at production dependent solely on long-term memory. Although those in the higher-level classes scored higher in the reading and translation portions, Sun warns that amount of time spent in study will not necessarily resolve problems with tone. First and fourth tones were more easily produced in the short-term memory tasks. Fourth tones were produced less frequently on the translation task, however. This study also suggests that learners have the most difficulty producing the second tone.

Q. Chen's (2000) quantitative and qualitative analysis of L2 Mandarin tone errors is unique because it includes connected speech. His main research question concerns whether there were any "observable error patterns… in sequences of tones, such as in word groups, phrases or sentences" (Q. Chen, 2000, p. 7). He also accounts for the third tone sandhi rule mentioned earlier, namely, that a third tone preceding a first, second, or fourth tone may lose the rising half of its contour (S. Chen, 1973; Shen, 1990). This study expands on Sun's (1998) study in that it includes actual sentences produced by participants with little priming from text. The sample size of forty participants - twenty from two different universities, is also respectable. I also appreciate his consideration of length *and* nature of exposure as well as a holistic evaluation of the participants' general tonal competence.

However, Q. Chen's (2000) study also included a few limitations. First of all, he personally selected participants that had studied Chinese no less than three years and "who were able to make connected speech in Mandarin on a familiar topic with desirable ease and fluency" (Q. Chen, 2000, p. 45). Thus, it is likely that the data collected is representative of only mid to high proficiency students. Furthermore, the twenty participants from the University of Michigan may not be comparable to the twenty students from Brigham Young University because of the nature of language exposure; students from Brigham Young each had about two years of experience abroad. Q. Chen also neglected to account for intervening variables such as participant motivation, goals, etc. Q. Chen also included himself in the group of three native evaluators. Two judges (both from Beijing) were trained to look for Q. Chen's "mid-level alien tone," a phenomenon that was not validated acoustically.

These limitations aside, Q. Chen (2000) set a precedent by analyzing spontaneous speech. Although participants were given a list of topics, they had the freedom to say what they wanted to. Data was elicited without moderation, but data extraction was not randomized; the researcher hand-picked utterances because they were from "a) a natural and continuous speech flow; (b) connected, meaningful and distinct utterances; c) a non-emotional tone, and (d) a normal speed" (Q. Chen, 2000, p. 51).

Overall, only 56% of the analyzable syllables were judged as correct. "Target tone, target tone position in a word, length and nature of exposure to a Mandarin-speaking environment, and the level of general tonal competence all turned out to be important to subjects' tonal performance in connected speech" (Q. Chen, 2000, p. 82). Q. Chen's perception of an alien mid-level tone is possible, but is not backed up by acoustic evidence. In addition, several studies suggest that native speakers may produce tones in the same range (i.e., Shen, 1990; Fon & Chiang, 1999).

This concludes the review of L2 production studies. The following table provides a brief synopsis of the L2 tone production studies discussed.

Table 2.3 Summary of Findings from L2 Production Studies

Source	Type of Study	Participants	Basic Findings
Jie (1988)	Cross-sectional and Longitudinal Microanalysis or read speech	40 former students of Mandarin with experience abroad	L2 learners of Mandarin did not suffer from tone attrition as long as they could remember the lexical item associated with the tone.
Shen (1989)	Microanalysis of read speech	8 American beginning students of Mandarin	L2 learners struggled most with register, not contour. Tonal errors were most commonly fourth tone errors, followed by the first tone.
Miracle (1989)	Acoustic Analysis of read speech	10 American intermediate students of Mandarin	Learners struggled the most with second and third tones and had the same amount of register errors as contour errors.
Chen (1997)	Microanalysis of spontaneous speech	6 American advanced students of Mandarin	Learners confused second and third tones with each other, as well as first and fourth. Register errors were the most striking.
McGinnis (1997)	Longitudinal Microanalysis of read speech	29 American and 18 Japanese students of Mandarin	Suggested no significant difference between the students tone production based on a study of pinyin or Gwoyeuh Romatzh (GR). Found very low agreement between eight native judges.
Sun (1998)	Longitudinal and Cross-sectional Microanalysis of read speech	50/39 students of Mandarin on study abroad	Second tone is the most difficult. Tone production was constrained by memory capacity.
Q. Chen (2000)	Cross-sectional Microanalysis of spontaneous speech	40 American students of Mandarin from two Universities	Found alien mid-level tones in connected speech. Found great variation in individual learners' tonal errors.

The Role of Production and Perception in L2 Acquisition of Tone

The following is a summary of findings gleaned from studies focused on the perception and production of tone by L2 learners. First, both practical experience and reliable research suggest that L2 learners are not internalizing strategies for tone production (Jorden & Walton, 1987; Sun, 1998; Q. Chen, 2000). Secondly, register errors are likely a problem for L2 learners, (Shen, 1989; Miracle, 1989; Chen, 1997) although there have been no studies validating whether these errors produce misunderstanding or even irritation in native speakers. Third, L2 learners often confuse the second and third tones and seem to struggle the most with the second tone (Miracle, 1989; Sun, 1998). Fourth, childhood introduction to tone language may improve recognition of second and third tones, and likely enhances production of tones in some contexts (Sun, 1998). Fifth, connection of spoken form with meaning plays a crucial role for non-natives in the perception and production of tone (Sun, 1998). Finally, we are uncertain about L2 learners' development of tone sandhi in multi-syllabic speech (Q. Chen, 2000).

Connections and Contrasts: L1 and L2 Acquisition of Mandarin Tone

The research discussed has provided a rich and complex assortment of information, but corresponding results come down to six main points. First, many of the studies reviewed suggest that L1 and L2 learners often confuse the third and second tones with each other and struggle the most with the second tone (Li and Thompson, 1977; Miracle, 1989; Whalen and Xu, 1992; Lee, 1996; Snow, 1998; Sun, 1998; Q. Chen, 2000). Secondly, compared to Americans, Chinese may develop a wider pitch range and higher tone register early on (G.T. Chen, 1973; Miracle, 1989; Shen, 1989; Papousek and

Hwang, 1991; Chen, 1997). Third, intonation, stress, and tone all interact in connected speech, making the use of monosyllabic, decontextualized tones obsolete and misleading when applied to judgments of connected or spontaneous speech (Chao, 1968; Shen, 1990; Xu, 1997). Fourth, suprasegmental factors such as amplitude and turning point may also play a role in adult native speakers' recognition of tones (Whalen and Xu, 1992; Shen et al, 1993). Fifth, context and connection of spoken form with meaning play an important role in both L1 and L2 learners' ability to perceive and produce correct tones (Lee et al, 1996; Chen and Cutler, 1997; Sun, 1998; Hua and Dodd, 2000). Finally, the realities of tone sandhi in multi-syllabic speech are not fully understood (Shen, 1990; Hua and Dodd, 2000).

Summary and Conclusions

Prior studies on native and non-native production of Mandarin lexical tone have suffered from at least one of four major design flaws: 1) exclusive use of text-induced or planned speech, 2) evaluation of tone based solely on microanalysis (vs. holistic analysis or acoustic analysis), 3) an inaccurate/unclear conception of what the target tones actually are, and 4) avoidance of the relationship between lexical development and tone acquisition.

Each drawback becomes problematic in our path to understanding non-native acquisition of lexical tone. First of all, previous research endeavors have sought to determine what communication skills L2 learners have acquired based *solely* on text-induced and/or planned speech, and is incapable of providing a thorough understanding of learners' *implicit* linguistic knowledge (Bialystok, 1979; Ellis, 1994; Lakshmanan & Selinker, 2001). Secondly, many studies have employed only native judges to evaluate and discern each learner produced tone (i.e. McGinnis, 1997; Sun, 1998; Q. Chen, 2000). This method is problematic because human judgments have been found to be much too subjective compared to acoustic analyses (see Tseng, 1981; Shen, 1990; McGinnis, 1997).

Finally, most researchers have utilized an oversimplified concept of tone, neglecting its interaction with neighboring tones, intonation, stress, etc. This oversimplification has challenged the entire premise of studies such as Jie's (1988) study, which did not even account for tone sandhi. Although many of the complexities of tone variation in natural native speech have been documented for many years in both impressionistic and acoustic research (see Chao, 1968; Tseng, 1981; Shen, 1990; Xu, 1997), many researchers have not accounted for these findings when attempting to assess non-native tones. Previous findings and opinions concerning tone register need to be confirmed by larger acoustic studies of spontaneous and more controlled speech (see Duanmu, 2000, Fon & Chiang, 1999; Liu, 2001).

It is evident that tone production varies depending on syntax, semantics, context, stress, and other prosodic elements (Connell et al, 1983; Hung, 1990; Shen, 1990; Papousek and Hwang, 1991). From a theoretical and empirical standpoint, there is likely a difference between the tones natives and non-natives produce spontaneously and their pronunciation in more planned speech (Bialystok, 1979; Chao, 1968; Tseng, 1981; Ellis, 1994; Lakshmanan & Selinker, 2001). Only when we come to a greater understanding of the social, cognitive, and linguistic factors influencing spontaneous native and non-native speech can we then move on to the next step; researching and designing appropriate tasks that facilitate learning and acquisition of tone (Ellis, 2000). After my review of the literature, it has become clear that effective pedagogical suggestions concerning tone cannot be made until a more accurate description of native tone production becomes available.

Research Hypotheses

In order to determine whether there are significant tone register differences between native speakers of Taiwan Mandarin and Beijing Mandarin, I conducted an acoustic analysis of spontaneous conversation, spontaneous descriptive and read speech. I developed three null hypotheses to be tested, which are as follows;

1) Pitch register differences cannot be predicted by dialect; or there are no significant pitch register differences between groups of male native speakers of Mandarin Chinese from Beijing and Taiwan.

2) Pitch register differences cannot be predicted by speech style; or there are no significant pitch register differences between groups of male native speakers in three speech styles: read, descriptive and conversational.

3) There are no significant tone register differences between the five tones produced in Taiwan Mandarin and those produced in Beijing Mandarin; or there are no individual tone register differences in read speech between male native speakers of Mandarin Chinese from Beijing or Taiwan.

Purpose

The purpose of this research was to test claims made by Fon and Chiang (1999), who suggested that differences in pronunciation between speakers of Beijing Mandarin and Taiwan Mandarin may also extend to tone register. Previously, judgment of non-native tone register has ignored possible differences influenced by dialectal dissimilarity amongst native Chinese. For example, Q. Chen's (2000) study of non-native tone production utilized three native judges (one from Shanghai and two from Beijing) to judge tones of students, many of whom learned most of their Chinese in Taiwan. Thus, Q. Chen's conclusion that there are "more cases of longer level-tone sequences in the speech of the first 20 subjects, who are all students of Brigham Young University with church missionary experience in a Chinese speaking community" (Q. Chen, 2000, p. 148) could stem partially from this dialectal difference. Perhaps these *alien* tones reported by Q. Chen are acquired, in part, from the pronunciation patterns of native speakers of Taiwan Mandarin.

Furthermore, most prior acoustic analyses of tone production have been based on deliberate or planned speech. An important distinction between spontaneous and deliberate production is indicated in Bialystok's model of how individuals learn a second language (Bialystok, 1979). Spontaneous production (output) is governed solely by "implicit linguistic knowledge," and deliberate or planned speech is governed by both implicit and explicit linguistic knowledge (Ellis, 1994, p. 357). Because the majority of research on the acquisition of Mandarin lexical tone has been based on deliberate or planned speech, the implicit linguistic knowledge of Chinese speakers has been obscured. Thus, this study contributes much to the field in terms of describing what the target pronunciation is.

Potential Impact

Regardless of whether the null hypotheses can be rejected or not, this study enlightens further research on dialectal tonal differences and sheds light on the importance of examining spontaneous speech before coming to conclusions about linguistic systems (Ellis, 1994; Lakshmanan & Selinker, 2001). However, if the null hypotheses are rejected and a significant dialectal difference is found, my study will encourage researchers to investigate more thoroughly what tonal differences exist between Beijing and Taiwan Mandarin. Possible discrepancies between tones prescribed by *SC* and those produced by *SC* speakers from different regions will demand further attention. In addition, future assessments of non-native tone production of Mandarin will be more likely to take into account the nature of the input students receive as they converse with or listen to native speakers from different dialectal or social backgrounds.

This concludes chapter two, and the review of literature. In the next chapter the methods of data collection and analysis are described in detail.

CHAPTER THREE

METHODOLOGY

This chapter explains the methods and reasoning used to collect, extract, and analyze the data of the present study. Nine native speakers of *SC* were recruited, six of whom were speakers of Beijing or Taiwan Mandarin. All participants were recorded in a single walled sound attenuating recording booth speaking in three speech styles; spontaneous interview, spontaneous descriptive, and controlled read speech. 10% and 90% F0 as well as median F0 and 80% pitch range were analyzed as acoustic correlates of pitch register. A sampling of individual tones was compared to determine any significant tone register differences. This analysis included min, max, median and 100% F0 range.

Participants

Collecting representative speech data of the two dialects involved two phases. The first phase consisted of contacting participants that were publicly listed as being from the target locations (Beijing or Taiwan). These potential participants were also asked if they knew any male native speakers of Chinese that may be interested in participating, without mention of dialect or region. While I was fully aware that I only needed participants from the two regions, I made a concerted and successful effort to not bring up the topic to potential participants. The second phase, which occurred after the recording process, involved an assessment of demographic information to determine whether participants were representative speakers of one of the two dialects being analyzed. All participants included in the study fit into the following description: 1) Born and raised in Beijing or Taiwan (14 years or more), 2) participants from Taiwan spoke

Mandarin (*guoyu*) and Taiwanese (*taiyu*), participants from Beijing spoke Mandarin (*putonghua*), and 3) participants' parents both spoke the desired dialect (See Appendix I).

First, let me discuss the procedures of phase one in more detail. Native Chinese speaking participants were recruited through email, phone, personal contacting and networking. Potential participants were asked if they would be interested in participating in a study designed to increase understanding of the Chinese language and improve the teaching of non-native students of Chinese. Names for recruiting over email and phone were collected from forty of the most common Chinese surnames selected from http://www.zhongwen.com. The first thirty of these forty surnames were then entered into the search option of Brigham Young University's online directory. The online directory lists the hometown of each student, if they so desire. Only those who were listed publicly as being from Beijing or Taiwan were contacted. All students with one of the thirty common surnames and who listed themselves as being from these locales were contacted by phone and/or email. I also personally invited individuals thought to be from the target communities. Potential participants were asked if they knew other males that may be willing to participate, without any mention of dialect, place of birth, or residence.

A total of nine male native speakers of Mandarin Chinese were recruited to participate in recordings (See Table 3.1 on the next page). One participant's data was entirely lost due to a computer error and two participants were found not to be representative speakers of either dialect group.[5] Hence, three native speakers from Beijing and three from Taiwan were analyzed for this study. All six participants were aged 26-33, and were current students at Brigham Young University in Provo, Utah.

[5] Subject 4 and 5 were from Linyi and Haarbin.

Table 3.1 Participant Demographic Information

Participant	Location & Hometown	Duration at Hometown	Age	Languages or dialects spoken	Where attended High School	Year arrived in U.S.
1	Taiwan Jidong	16	30	*Guoyu* *Taiyu* Japanese	Taiwan Taibei	2004
2	Taiwan Taibei	30	30	*Guoyu* *Taiyu*	Taiwan Taibei	2004
3	Taiwan Jiayi	20	26	*Guoyu* *Taiyu*	Taiwan Jiayi	2004
4	China Linyi	19	24	*Putonghua* *Shandonghua*	China Linyi	2001
5	China Haarbin	18	26	*Putonghua*	China Haarbin	2003
6	China Beijing	15	27	*Putonghua*	China Suzhou	1999
7	China Beijing	29	33	*Putonghua*	China Beijing	2001
8	China Beijing	25	30	*Putonghua*	China Beijing	2001

Data Collection Procedures

All recordings were done in a sound attenuated booth with a pair of Sennheiser MKH 40 P4B microphones. Audio was processed and saved at a 16 bit resolution at a sampling rate of 44.1 KHz (Participant one was first recorded at 32 bit resolution, and then saved at 16 bit to minimize file space).

Upon arrival participants were welcomed briefly, and then invited to read the consent form, which was presented in simplified or traditional characters (See Appendix D). All participants read the consent form, which included an option to depart if they did not wish to participate. Confidentiality was maintained as names were not recorded at any time. The recordings were not available to the general public, but were analyzed only by the researcher.

Participants then entered the recording booth and were seated in front of a lap-top computer. Each participant was asked if they knew how to go through a Power Point presentation and informed that if they had any problems or questions with the presentation the researcher would be available just outside the recording booth to answer questions.

Participants were then given an opportunity to practice going through the presentation by viewing the first few slides (See appendix A to see the text of the presentation in English or Chinese). After a welcome page, the presentation requested participants to count to ten slowly in their regular voice while the researcher adjusted the volume. The next slide asked them to count again, this time fast. Finally, participants were to ask any questions they had.

The researcher conducted all demographic question and answer sessions, following the same basic wording and format for each participant (See Appendix C). Some slight variations were made in the researcher's speech to make the participants more comfortable. For example, I used the Taiwan Mandarin *nali* (那里/哪里) with participants from Taiwan, and the Beijing Mandarin *nar* (那儿/哪儿) with those from Beijing. In addition, my follow up questions or brief comments after participant responses sometimes varied. After the conversation, I informed participants that the rest of the study would be conducted on their own, and then exited the booth and was seated at an external computer. I should also note that I was present throughout the duration of all recordings, and partially visible through the window of the booth. Participants felt free to ask questions, but I also made sure not to make eye contact and that my movements were minimal so as not to distract.

The PowerPoint presentation contained a set of picture description tasks (See Appendix B). The picture description portion consisted of six slides, each slide containing a series of four progressive pictures published by Byrne (1967). Participants were given specific instructions to describe the story being depicted as they would to a native speaker. Responses were generally produced immediately, as before seeing the actual set of pictures a window popped up saying "Describe this series of pictures: begin now!" Then the set of pictures appeared, and remained until the participant decided to move on to the next slide.

The first set of pictures was designed to serve as a test run, and some participants did have some confusion as to how to describe the pictures. For example, participant 8 nearly completed his description, and then started over again from the beginning. Because some participants struggled with their first attempt at the task, my random selection of the descriptive portion included the description of two slides out of five (slides 2-6). At the conclusion of the picture description, participants were invited to take a 2-3 minute break, including an invitation to exit the booth and use the restroom, get a drink, etc. None of the participants took advantage of this opportunity, but instead commenced with the final portion of the study.

The reading section (see Appendix H) utilized a series of two syllable nonsense words occurring in all possible tonal combinations within four different carrier sentences (see Xu, 1997; reviewed earlier). My sentences were exactly the same as those used in Xu's (1997) study of contextual tonal variation; *Wo jiao* ***mama*** *lianxi/lianluo* (我教/叫_ _ 練習/聯絡). The bolded portion consisted of all four tonal combinations, except perhaps the first-first combination, which may have been produced first-neutral because of the common pronunciation of the *mama* characters (媽媽, ‘mom’).

However, there are three major differences between my methodology and that used by Xu (1997); first of all, participants only read the same sentence three times (instead of six). Secondly, speech rate was not controlled for by having participants read to the beeps of a timer. Both of these changes were made to prevent participant fatigue and anxiety. The final change was the inclusion of a randomized tone sequence, rather than a sequential presentation. The randomization was performed in order to increase validity.

Extraction of F0

All F0 measurements were extracted manually utilizing Adobe Audition 1.5. In addition, all speech from the demographic interview was carefully edited by the researcher using Adobe Audition 1.5 to remove any speech produced by the researcher, as well as any questions or English produced by the participants. Data collected from the descriptive and read portions was left as is, except for the removal of any clicking created from interaction with the computer. Like Xu (1997), I ensured that individual tones extracted from the read portion included the vowel portion only and that before making my selection at the initiation of periodic sound energy, the “waveform of each utterance was displayed on the computer screen in such a way that each glottal cycle was shown clearly” (p. 66).

Pitch measurements included 10%, 90%, median pitch, and 80% (10%-90%) F0 as a measure of average pitch range (Yuasa, 2001; Lewis, 2002). Limiting the analysis of pitch range to 10%-90% of the measured pitch is considered the most accurate for natural speech as it excludes pitch levels that are at the extreme lows and highs. F0 measurements found at very high and low levels is generally found to be spurious data. An example of this can be found in the histogram of F0 from participant one located below. The values below 62.5 and above 125 are most likely spurious data.

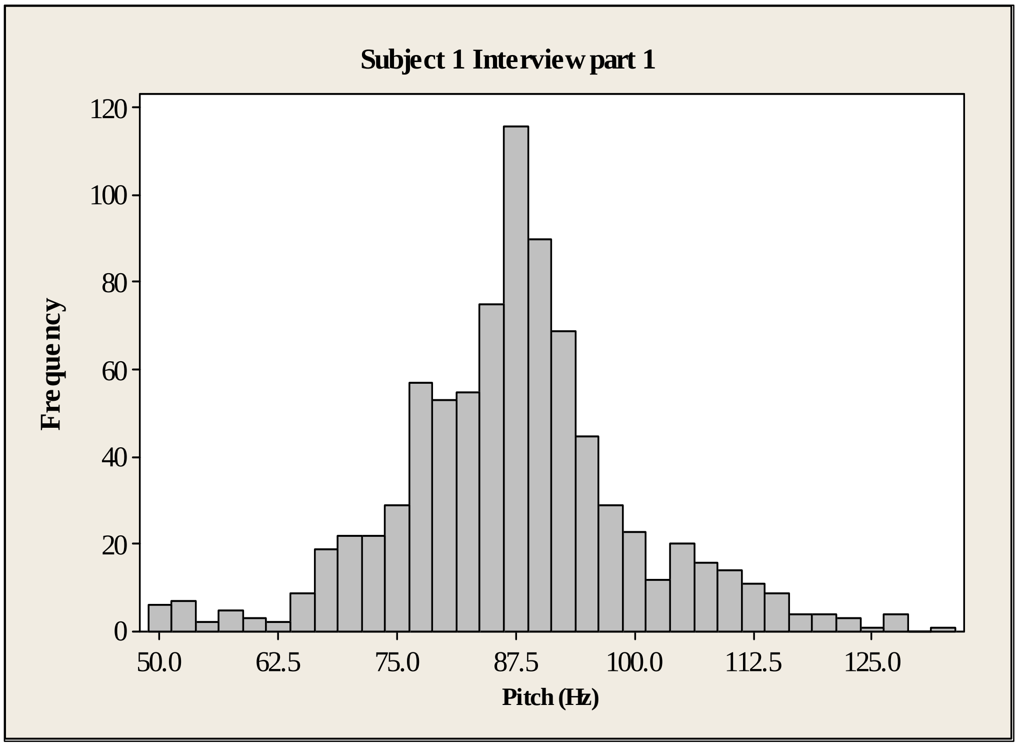


Figure 3.1 Histogram of F0 values from part one of participant one interview
Total duration: 14.7 s Total number of frames: 977 (474 voiced) Time step: 15 ms

All audio files were opened and analyzed in Praat version 4.3.18 (Boersma, 1993). Long segments of each participant's three speech types were opened up in a separate edit window using the autocorrelation method (AC) and standard settings for intonation analysis. These standard settings also meant a time step of 15 ms. The pitch ceiling was set at 300 Hz and pitch floor at 50 Hz, and these settings were not adjusted. Short clips of read tones were opened up with standard settings, but this time utilizing a cross-correlation (CC) method and the very accurate box checked, creating a time step of 5 ms. I utilized the CC method because I was dealing with such a short time window; all individual extracted tones were less than 200 ms (See Boersma, 1993). Utilizing the Praat command 'extract visible pitch contour,' I was able to obtain all necessary descriptive statistics in Hz, semitone, and ERB scales. This data was then saved in MINITAB or Microsoft Excel for further statistical analyses. Statistical analyses are described in detail in chapter four when I present my results.

Research Variables

My analysis was based on a combination of methods used in Yuasa (2001), Lewis (2002), Liu (2001), and Xu (1997). The variables for my first two hypotheses are listed below in two categories; nominal independent variables and continuous dependent variables (Tables 3.2 and 3.3).

Table 3.2 Independent Variables for Dialect and Speech Style ANOVA

Independent Variables	Level 1	Level 2	Level 3
Dialect	Taiwan Mandarin	Beijing Mandarin	--
Speech Type	Interview	Descriptive	Read

Table 3.3 Dependent Variables for Dialect and Speech Style ANOVA

Dependent Variables	Measurement
Average Pitch	Median F0
Pitch Range	80% (10% - 90%) F0
Min Pitch	10% F0
Max Pitch	90% F0

To address my third hypothesis concerning the impact of dialect and tone on read tones, I first conducted a multiple analysis of variance (MANOVA), with the P level set to $p < .05$, in order to determine which variables would be significant predictors of tone register. Because individual participants could not be extricated from dialect spoken, I established dialect and participant as covariates, and ran the MANOVA with tone, dialect and participant as independent variables. The four dependent variables were min, max, median and 100% pitch range in the second /ma/ of each selected /mama/sequence. Any significant independent variables would then be used in an ANOVA to ensure the accuracy of results and to provide necessary statistical information not provided in the preliminary MANOVA.

Research Controls

In addition to these variables, it was recognized that there were many other social, psychological, and linguistic variables that needed to be controlled for. Each of these areas was addressed as follows:

Major Social Factors

Interlocutor effects

In order to control for interlocutor effects, or "adjustments in speech style based on whom we are talking with" (Gass & Selinker, 2001), each participant interacted with the same researcher during the demographic interview. With such a small potential number of participants it was deemed unwise to attempt to stratify the data further by introducing another interlocutor. However, the descriptive and read portions of the study were recorded with the participants in relative isolation[6] and with specific instructions to speak as they would to a native speaker.

Speech style

Labov (1984) explained that speech style includes "any consistent… [set of]

[6] The subject could be seen by the researcher through the recording booth window.

linguistic forms used by a speaker, qualitative or quantitative, that can be associated with a… [set of] topics, participants, channel, or the broader social context." To explore three different styles my participants engaged in three different speech styles to account for potential differences between them (Llisterri, 1995). The three styles chosen are widely used forms of speech in language research; spontaneous conversation during a demographic interview, spontaneous descriptive speech during a picture description task, and read speech in a series of carrier sentences.

Language/Dialect

The potential influence of dialect was at least partially controlled for by stratifying my sample based on place of birth and residence as well as dialect spoken (Beijing vs. Taiwan). Parents' spoken languages and dialects were also noted. The decision to stratify participants due to dialectal background was based on the general lack of recognition of the potential impact of dialect many of the studies reviewed earlier.

Gender

Common sense and research strongly suggest that gender has an impact on how people communicate (Major, 1997; Chan, 1997; Yuasa, 2001; Lewis, 2002). It is possible that tonal differences exist between male and female speakers of Mandarin Chinese. Although it would have been ideal to have over 20 male and 20 female participants from each dialect group, due to temporal limitations, only male participants were recruited. The plus to the *exclusive* use of male speakers was that I was not forced to compare an insufficient number of male and female speakers acoustically.

Major Psychological Factors

Time for planning afforded

There is a difference between speech produced after planning, and that which comes immediately and spontaneously (Ellis, 1994). I controlled for planning time in the interview and descriptive portions of the study, but not during the read portion. The time for planning afforded during the interview was based on natural rules of conversation. When I asked a question, the participant responded as soon as they could think of the response. During the picture description task, each series of pictures was preceded by a pop-up window stating: describe this series of pictures, begin now!

Affective filter

Research participants often have at least some anxiety. I maintained a comfortable and relaxed atmosphere by welcoming them warmly, smiling, laughing with the participants, and ensuring them that they could ask me any questions during the recording process. Furthermore, by sitting outside the booth, yet still within partial sight, I helped participants to feel comfortable while not creating greater anxiety or confusion by leaving them totally alone to perform the description and reading tasks. Finally, none of the tasks included a time limit or a timer; participants were free to speak at whatever rate and/or duration desired.

Controls of Major Linguistic Factors in Long Speech Segments

In my first analysis of longer segments of speech, all linguistic factors listed below were controlled for simply by obtaining 1-3 minutes of data from each participant in each speech style. This method has been utilized by several other researchers in order to obtain an average pitch and pitch range while not needing to scrutinize or eliminate the seemingly endless array of variables that influence natural speech (Yuasa, 2001; Lewis, 2002). By obtaining such long segments of recorded data, the relative impact of any one variable becomes insignificant.

Phonetic/Phonologic environment

The following is simply a summary of the linguistic effects controlled for that can create phonetic/phonological changes to tone; assimilatory carry over affects (Xu, 1997), dissimilatory anticipatory affects (Xu, 1997), F0 peak delay (Xu, 2001), focus (Xu, 1999), articulatory constraints such as maximum speed of pitch change (Xu, 2002), speech rate (Chao, 1968; Shen, 1990), loudness/amplitude (Whalen & Xu, 1992), vowel/consonant sound (Xu, 1997), downdrift/declination (Shih, 2000), tone sandhi (Chao, 1968), intonation (Ho, 1977; Xu, 2001), and word stress (Moore, 1993).

Syntactic/Semantic effects

There are also at least three possible grammatical or semantic effects on tone, such as position in sentence (Ho, 1977; Shih, 2000), relationship between words (Hung, 1990), and sentence type (Shen, 1990).

Controls of Major Linguistic Factors in Analysis of Read Tones

Tone measurements of the read /mama/portions were made by extracting the min, max, average pitch range and median pitch of a sample of tones. The tone extraction here did vary from that of the longer portions of speech used to address my first two research hypotheses. Because this was a more controlled analysis of short audio clips, Xu's (1997) method of extracting min, max and pitch range from 100% of the range of pitch measurements was used. I personally checked each pitch window to ensure that there were no abnormalities or spurious data (See Appendix G for a few examples of Praat pitch windows). All measurements appeared to be free from any problems. Median pitch was still measured rather than mean pitch to ensure an accurate average of speech. Raw data from each tone was extracted automatically utilizing the ‘extract visible pitch contour’ command in Praat.

In the preliminary tone analysis, major linguistic effects were controlled for by extracting all tone samples from the exact same tone context. For example, all first tones were taken from second-first pairs found in each participants recording corresponding with the first paragraph of the seventh slide in the PowerPoint presentation (see figure 3.2 below).

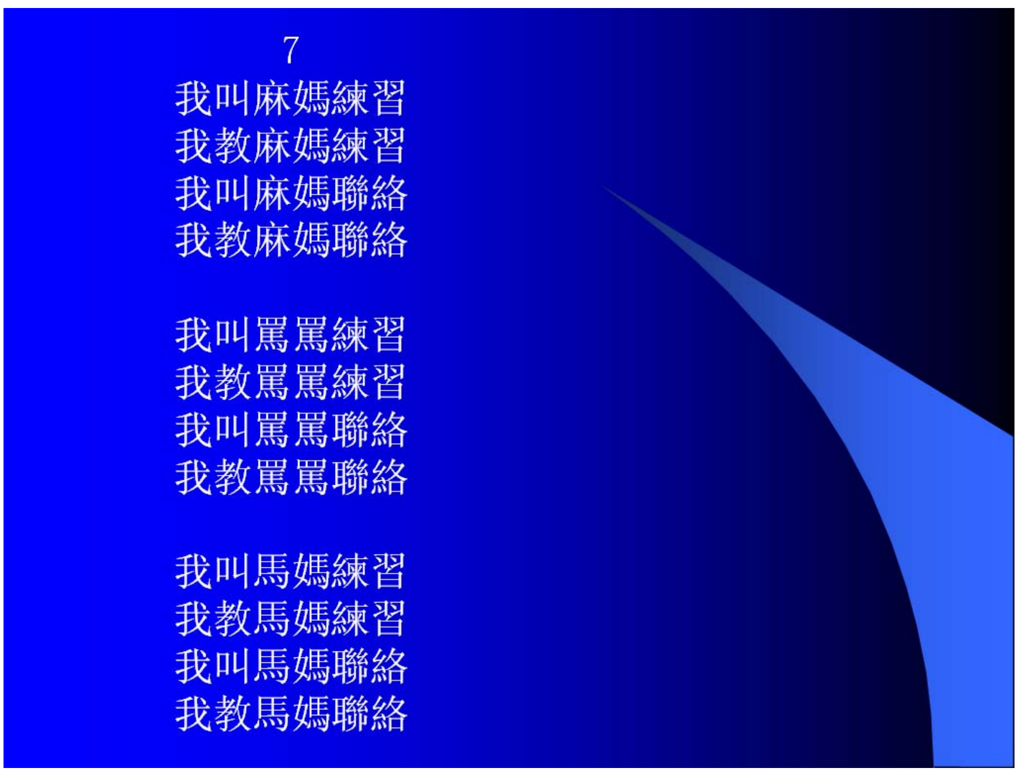


Figure 3.2 Slide 7 from reading elicitation portion of PowerPoint presentation

Thus, sentence position and the influence of neighboring tones were at least standard. Focus, word stress and tone sandhi problems were all avoided by analyzing only the second syllable of the bi-syllabic target word. Vowel and consonant effects were standardized by utilizing only the /ma/ sound. Potential grammatical effects on tone were controlled for in the following manner; sentence position (Ho, 1977; Shih, 2000) was controlled for by looking only at mid-sentence bi-syllables. Syntactic or semantic relationships between words (Hung, 1990) were avoided by utilizing nonsense words. Finally, only declarative utterances were included to control for sentence type (Shen, 1990).

Because each tone pair was produced four times in the two different carrier sentences, I was able to select the second, third and fourth attempts in order to avoid disfluencies (participants sometimes struggled more with the first line of a paragraph). There were only two exceptions to this; Participant 1 tone 3: the first three tones were taken because the very last instance had suffered from data corruption. Participant 6 tone 4: the last two productions were corrected and produced again without any intervention from the researcher. I included the two corrected versions as they were what the participant intended to produce.

In addition, because the presentation included at least two sets of the tone pair in the exact same context, I was able to select the second attempt made by each participant. This ensured that every production analyzed would not be the participants' first attempt to produce it. The precise tones and their location in the PowerPoint presentation are listed on the next page in table 3.4 (See also Appendix H for full reading list):

Table 3.4 Location of Tone and Total Number of Tone Productions

Tone	**Tone Context**	**Location**	**Total**
1	Following a second tone /ma/	Last 3 lines of slide 7 paragraph 1	18
2	Following a first tone /ma/	Last 3 lines of slide 16 paragraph 2	18
3	Following a first tone /ma/	Last 3 lines of slide 17 paragraph 3	18
4	Following a first tone /ma/	Last 3 lines of slide 10 paragraph 2	18
5	Following a first tone /ma/	Last 3 lines of slide 2 paragraph 1	18

This preliminary tone register analysis included a total of ninety tones. Three instances of each tone from each participant; thus 15 total tones from each participant, 45 from each dialect group, and 9 total instances of each tone from each dialect group.

This concludes chapter three. The following chapter presents the statistical analyses of the data as well as the corresponding results.

CHAPTER FOUR

RESULTS

While chapter three introduced the methods of data collection and analysis, chapter four presents the statistical analyses and findings of my study. Below is a summary of my three research hypotheses;

1) Pitch register differences cannot be predicted by dialect.

2) Pitch register differences cannot be predicted by speech style.

3) There are no significant tone register differences in read speech between the five tones produced in Taiwan Mandarin and those produced in Beijing Mandarin.

The following two sections address the research hypotheses in this order; impact of dialect and speech style on pitch register, impact of language and tone on read tone register.

Impact of Dialect and Speech Style on Pitch Register

As mentioned in chapter three, I first conducted a general linear model ANOVA to test the impact of dialect and speech style on median, 10%, 90%, and 80% range of pitch. Dialect was found to be a significant predictor of 90% pitch F (1, 17) = 8.86, $p < .05$, and median pitch F (1, 17) = 9.29, $p < .01$. Speech style was not found to be a significant predictor of *any* pitch measurement. Dialect was not found to be a significant predictor of 10% pitch or 80% pitch range, although it did reach a marginal $p < .1$ with both. It was also noted that participant 7's measurements from the interview showed up as unusual for range, with a large standardized residual of 2.81 R, and for 10%, with a standardized residual of -2.07.

Instead of assuming that this unusual pitch behavior would not alter my results, I decided to try the two-way ANOVA again, this time with participant 7's interview data removed. The results of the ANOVA were very similar, with dialect showing up as a likely predictor of median and 90%, but this time dialect also proved to be a significant predictor of 10%; F (1, 16) = 8.34, $p < .05$. However, it was also noted that after eliminating the unusual data from participant 7, participants 1 and 3 both showed up as having "unusual measurements" in interview and read speech, respectively. After a thorough review of the data collected, and considering that the results of my analysis were not thrown off by my including the data, it was determined that it would be best to continue analysis with all participants, and with all data. This decision was also made under the pretence that "unusual measurements" are easily obtained when you have six participants speaking in three different speech styles. It is also possible that the pitch settings in Praat were inappropriate for the specific speech task or specific participant, but any modifications of individual settings would eliminate my ability to compare the data collected.

To ensure that my groupings of speech into three styles were not misleading, I also collapsed speech style into two main groups; read vs. spontaneous and ran the two-way ANOVA again. Speech style still showed up as insignificant; hence, three speech styles were maintained. It appeared that speech style was not a significant predictor of pitch register effects in long segments of continuous speech (See Figure 4.1 below).[7]

[7] Figure 4.1 contains a boxplot. Boxplots are generally used to analyze sample distributions. In this study, all boxplots display mean and median values, as well as the min and max measurement represented by the top and bottom of each box. Outliers, if any, are also indicated by an asterisk (*).

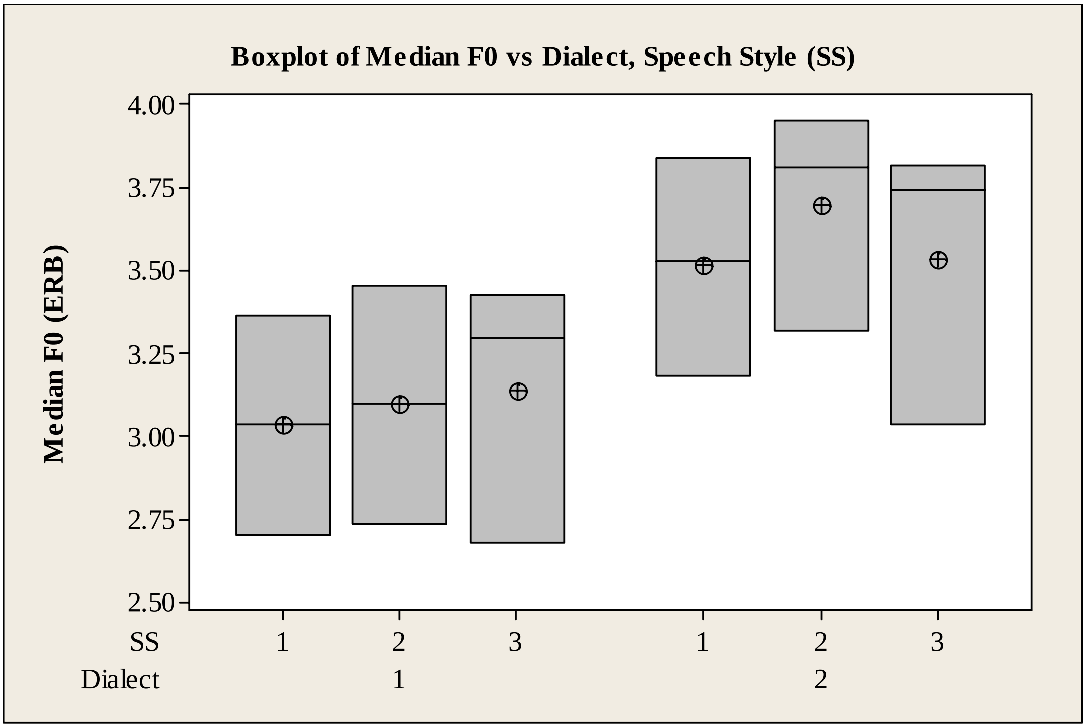


Figure 4.1 The effects of dialect and speech style on median F0. Note: Horizontal bars indicate the median, circles indicate the mean. Dialect 1 = Taiwan Mandarin, Dialect 2 = Beijing Mandarin. SS: 1 = Interview, 2 = Descriptive, 3 = Read.

Because speech style was not found to be a significant variable, I conducted a one-way ANOVA with dialect as the only independent variable. The one-way ANOVA results produced the same basic findings as the two-way ANOVA; 10% pitch was marginal, $p < .1$, pitch range was not predicted at all, but 90% and median pitch were predicted well; F (1, 16) = 7.98, $p < .05$ and F (1, 16) = 8.34, $p < .05$, respectively.

Impact of Dialect and Tone on Read Tone Register

As mentioned in chapter three, a MANOVA was first conducted to determine the most significant variables. The results of the MANOVA were as follows. Dialect and tone both showed up as significant predictors, with dialect at $p < .05$, and tone at $p < .001$. Participant was not a significant predictor. Thus, the MANOVA results allowed me to proceed and test my original research hypothesis with dialect and tone as independent variables in a two-way ANOVA (P level: $p < .05$). Dependent variables were min, max and median F0, as well as 100% F0 range. Dialect and tone were both found to be significant predictors of each dependent variable (See Table 4.1 below).

Table 4.1 Two-Way ANOVA: Min, Max, Median, Range vs. Dialect, Tone

Two-Way ANOVA P results for Language			
Min	Max	Median	Range
.023	.000	.001	.003
Two-Way ANOVA P results for Tone			
Min	Max	Median	Range
.000	.000	.000	.000

The next several pages consist of a series of figures depicting the nature of the results. Figures 4.2 and 4.3 represent the impact of dialect on F0 measurements, which were all significantly higher in the Beijing Mandarin group. Figures 4.4 and 4.5 depict the impact of tone on F0, suggesting that each tone has a respective average and range. Figures 4.6 and 4.7 represent the combined influence of dialect and tone on F0, revealing that participants from Beijing produced all five tones in a higher and wider tone register than the Taiwan Mandarin speakers.

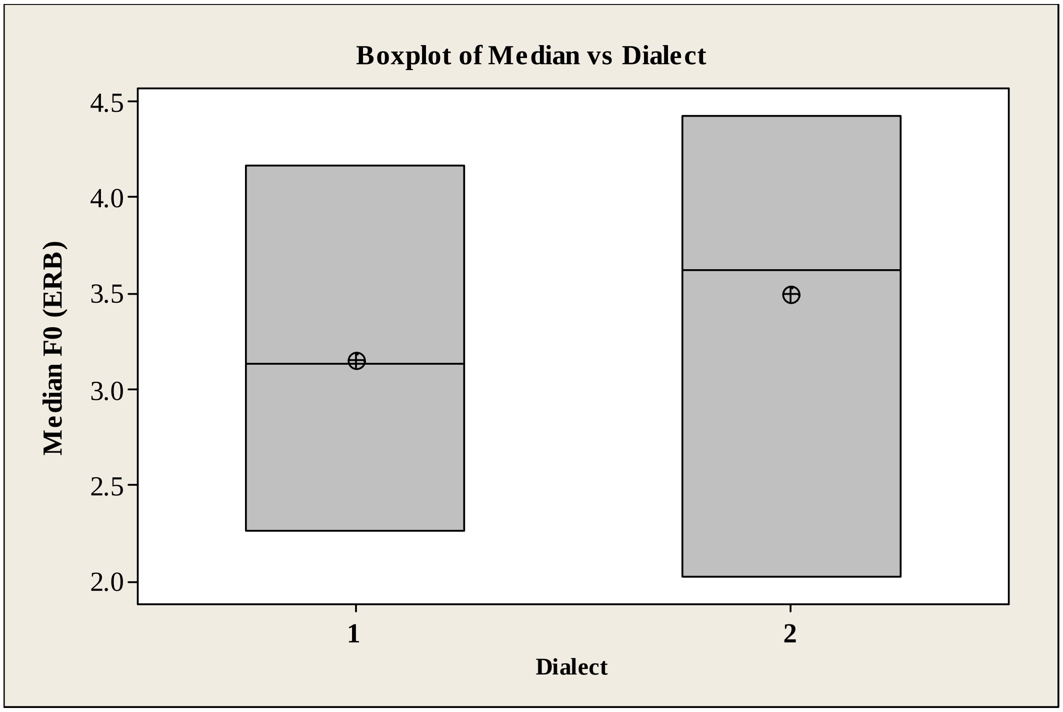


Figure 4.2. Effect of Dialect on Median F0. Note: The horizontal bars indicate median values, circles indicate the mean. Dialect 1 = Taiwan Mandarin, Dialect 2 = Beijing Mandarin.

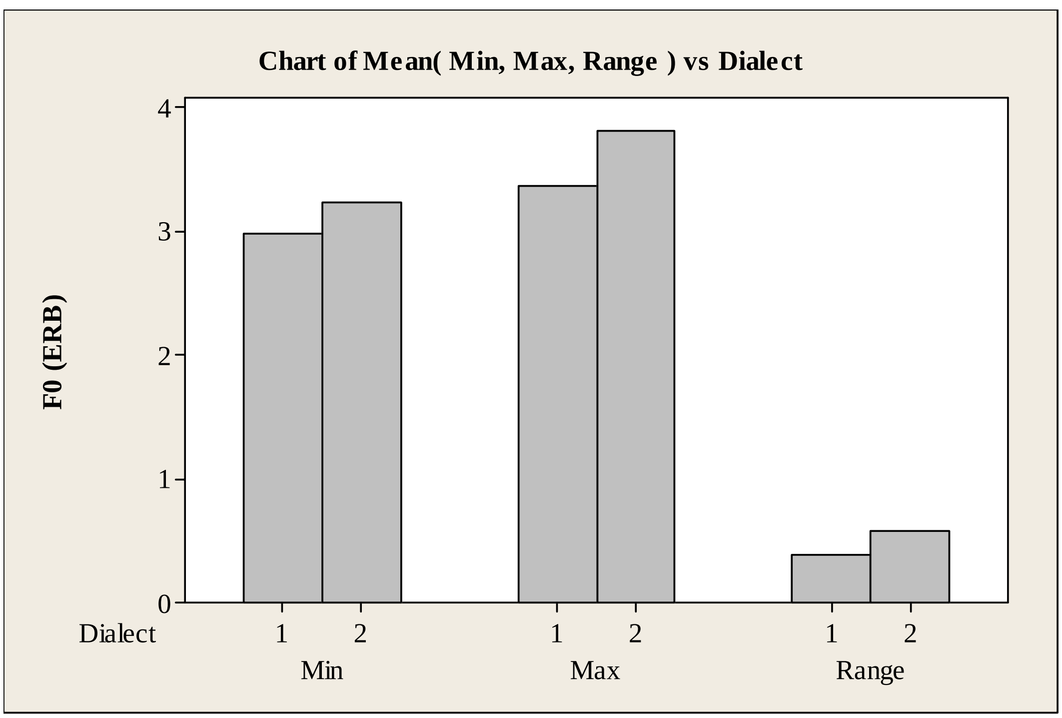


Figure 4.3 The effects of Dialect on min, max, and range of F0. Note: Dialect 1 = Taiwan Mandarin, Dialect 2 = Beijing Mandarin.

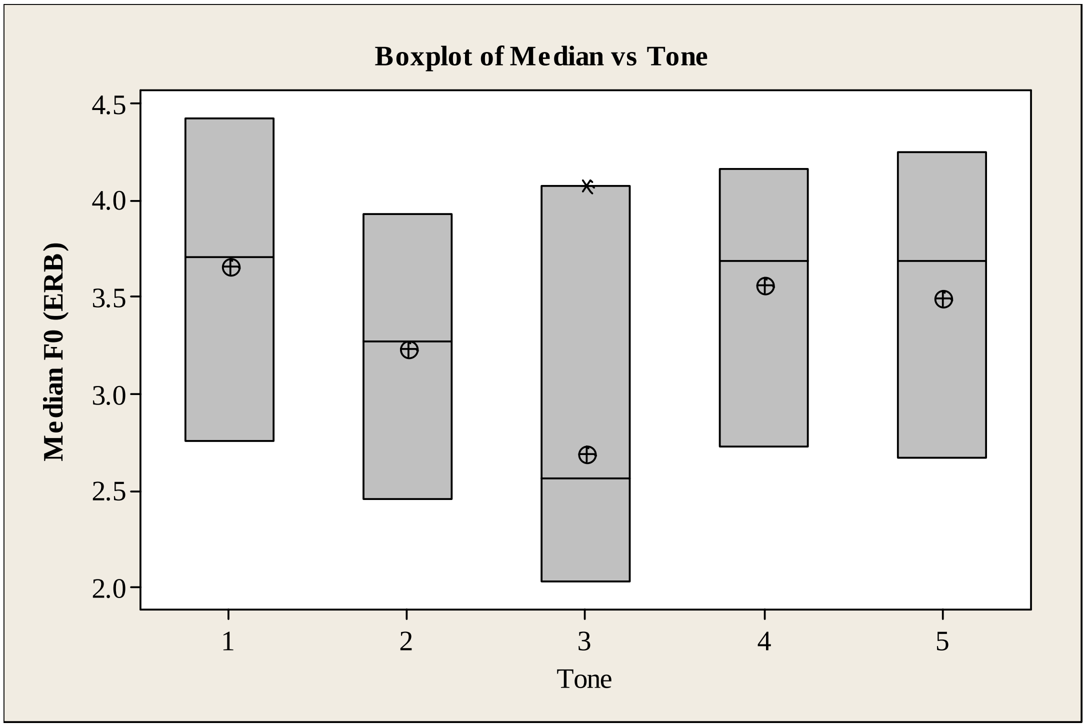


Figure 4.4 The effects of tone on median pitch. Note: The horizontal bars indicate median values. The asterisk (*) indicates a pitch outlier.

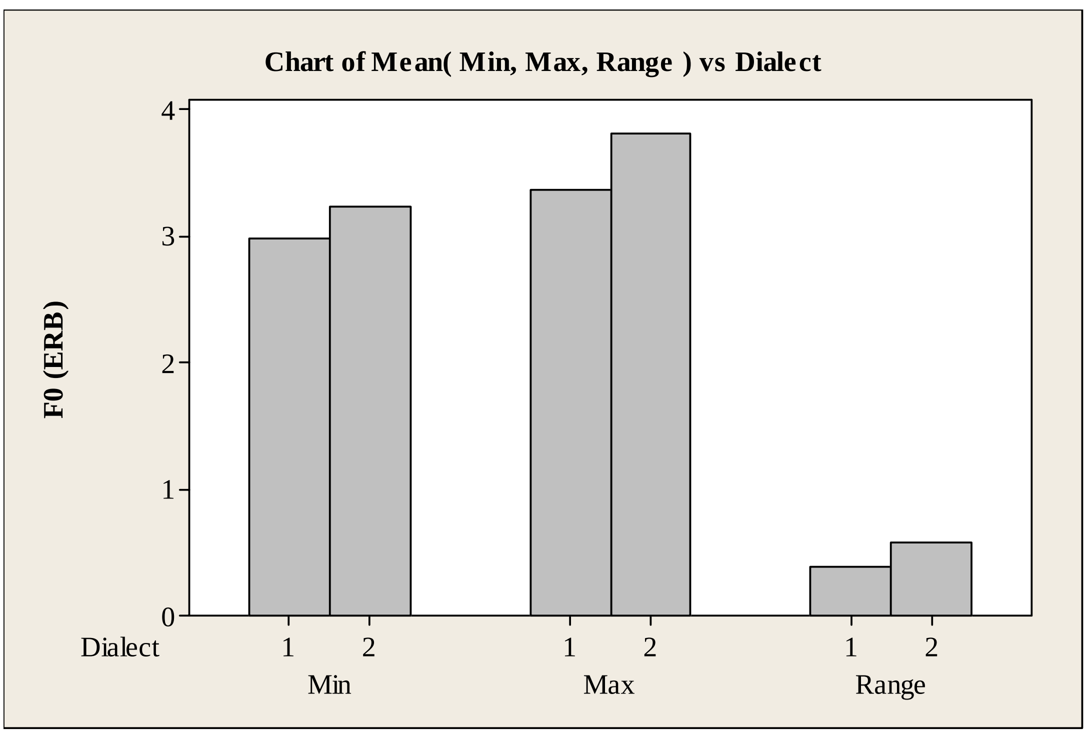


Figure 4.5 The effects of tone on min, max, and range of F0. Note: The horizontal bars indicate median values.

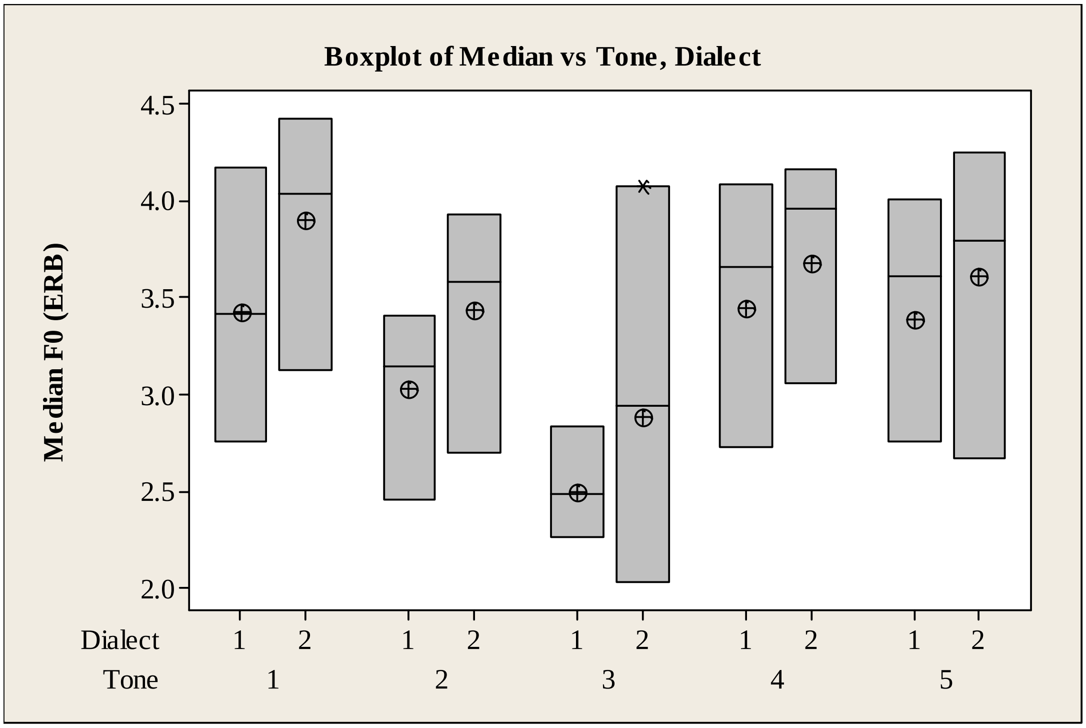


Figure 4.6 The effects of dialect and tone on median F0. Note: Circles indicate the mean, horizontal lines the median. Dialect 1 = Taiwan Mandarin, Dialect 2 = Beijing Mandarin. The asterisk (*) indicates a pitch outlier.

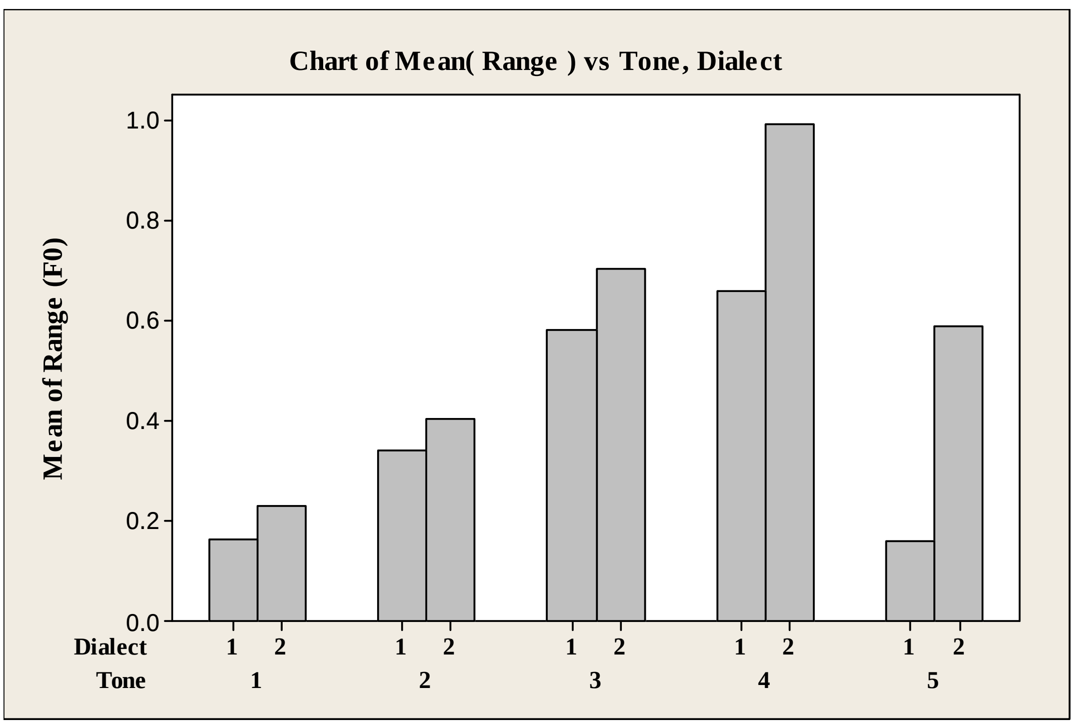


Figure 4.7 The effects of tone and dialect on mean pitch range measurements. Note: Dialect 1 = Taiwan Mandarin, Dialect 2 = Beijing Mandarin.

Discussion of Results

Discussion of the Impact of Speech Style and Dialect on Pitch Register

Speech style did not appear to be a significant predictor of any pitch measurement. Thus, pitch register differences cannot be predicted by speech style and my second null hypothesis is maintained. It is possible that analyzing such long portions of speech, or even the exclusion of the top and bottom 10%, veiled the impact of speech style. Perhaps a close analysis of comparable tones from each speech style would reveal significant differences.

In terms of dialect, these data also suggest that while a speaker's minimum pitch or pitch range is not predicted well by dialect background, at least two aspects of pitch register are predicted well; namely, 90% and median pitch. It is possible that the similarity between minimum F0 values is due to the fact that such long portions of speech were recorded and that pitch values tend to remain on the lower end in continuous speech (Lewis, 2002). In addition, the relative range of pitch may not necessarily equal pitch register. In other words, a speaker may produce tones in a lower pitch register on average, but still have just as wide of a relative range as someone with someone who speaks in a higher pitch register. Thus, it is still possible that Taiwan Mandarin tones occur in a slightly lower register. At this point results are not conclusive, but will be further discussed as they are compared with results from the next section; an analysis of individual tones produced in more highly controlled read data.

Discussion of the Impact of Tone and Dialect on Tone Register

Both tone and dialect have a significant impact on tone register in read speech. These findings suggest that we can reject the third null hypothesis; that dialect is a significant predictor of tone register. These findings correspond with Fon & Chiang (1999), who argued that Taiwan Mandarin speakers produce tones in a lower register.

Overall Results

Results from both sets of analyses suggest three main findings in relation to my original research hypotheses: First, that pitch register differences can be partially predicted by dialect background in long portions of continuous speech, thus we can partially reject the null hypothesis. Secondly, pitch register differences cannot be predicted by speech style, suggesting that the null hypothesis was correct, and that speakers tend to maintain the same averages in different styles of speech. Finally, there are significant tone register differences in read tones produced by Taiwan Mandarin speakers and those produced by Beijing Mandarin speakers. Dialect and tone were both significant predictors of tone register.

Null hypothesis number two deserves more discussion. Why wasn't speech style a significant predictor of pitch register? It is likely that individuals' average pitch and pitch range is not affected by speech style. However, this is not to say that speech style is not a factor that must be accounted for. It is possible that speech style affects tone contour, tone register, rate of speech, or other aspects of Chinese tone production, but that this impact cannot be discerned by a holistic view of pitch measures.

Due to the difference between the data used to investigate my first two research hypotheses and the data extracted from individual tones, I avoided comparing individual tone data directly with the data collected from the longer measurements of speech. Although this approach was tempting, comparing data sets with different durations, and that were obtained with different settings, may have led to erroneous conclusions.

Therefore, the only alternative was to compare the findings from the previous section on read speech with the analyses performed for my first two research questions. Three main differences surfaced. First of all, min pitch is predicted by dialect more efficiently in read speech. Secondly, dialect predicts pitch range in the small sample of read tones, but not overall. Finally, my results suggest that tone predicts pitch measurements more efficiently than dialect does. Thus, in terms of detecting potential dialectal tone register differences, it may be that controlled analysis of read speech is more effective than a holistic analysis. It is also possible that the dialectal register differences occur in read speech and not in other types of speech.

In addition to the main research findings, it was interesting to note potential differences between the production of the neutral fifth tone (see figure 4.7 above). If measurements were accurate, and the speakers truly did intend to produce the neutral tone, it would appear that there may be neutral tone differences between the two dialects that cannot be solely attributed to tone register. It is also possible that this difference was caused by the text, which some may have interpreted and produced as two high level tones rather than a high level and then a neutral tone. This concludes chapter four, and the results of my analyses. The next chapter concludes the study.

CHAPTER FIVE

CONCLUSION

This final chapter presents an overview of the theoretical contributions and pedagogical implications, a discussion of limitations and suggestions for further research, as well as conclusions based on the study.

Theoretical Contributions

This study argues successfully for a closer look at native tone production. It has been shown that a significant dialectal tone register difference exists. There may also be dialectal differences between Beijing and Taiwan Mandarin speakers in terms of tone contour and tone sandhi. Future studies should include a more careful assessment of the language and dialect background of participants.

As mentioned in the introduction, perhaps my largest contribution to the study of Mandarin tone register was a unique socio-phonetic approach to pitch that had previously only been applied to non-tonal languages (Yuasa, 2001; Lewis, 2002). I also accounted for social and psychological variables more openly than prior tone research. By analyzing longer segments of both spontaneous and read speech, one can control for many social, linguistic, and paralinguistic variables without attempting to eliminate them. By analyzing both pitch register and tone register, I have investigated the issue holistically and on a minute scale, bringing the two fields of sociolinguistics and acoustic phonetics a little closer together.

In addition, my inclusion of spontaneous and more controlled forms of speech sets a precedent. Many prior studies of Mandarin tone production have investigated only one type of speech, and most have looked at read speech. By analyzing spontaneous and read speech, I was able to provide a more thorough and holistic view of tone register.

I have encouraged a more thorough review of relevant literature by covering the historical, pedagogical, productive and perceptive aspects of tone acquisition in native and non-native speech. Finally, this study has directly addressed serious deficits in the validity and reliability of prior studies. While findings from only six participants should not be over generalized, they are much more valuable than an analysis of data collected from only one participant (see Fon & Chiang, 1999).

Pedagogical Implications

As mentioned in chapter two, development of native-like pronunciation is one of the last great frontiers of second language acquisition for both researchers and learners. As our understanding of native tone production increases, so will our ability to teach non-natives. Current trends in the Chinese teaching field, which are largely communicative, help address many of the needs of learners with respect to the acquisition of tone (Xing, 1997). However, this approach to learning must also be applied with caution. A learning focused, “diagnostic multifaceted approach” to the teaching of tone is more desirable than one specific bandwagon approach (Bai, 1996). I also agree with Bai (1996), who suggested that as teachers become skilled diagnosticians of their students and the various methods, they are better able to “adjust their teaching methods to maximize learning” (p. 84).

One such useful approach to the teaching of tone may be found in Chan’s (2003) suggestion to utilize speech analysis software to help students compare their multi-syllabic and sentential speech with native speech, as well as receive audiovisual feedback via the software. As discussed in chapter two, most prior suggestions on the teaching of tone have been solely focused on the representation of tone and/or the introduction to tone (see S. Chen, 1973; Chen, 1975; Tsung, 1987; Shen, 1989; Bar-Lev, 1991; Lundelius, 1992). Few have focused on the big picture, as Chan (2003) did, and come up with a comprehensive approach that can aid students throughout the many years of required study and application. In my review of the literature I failed to find an approach emphasizing the importance of regular opportunities to communicate with native speakers, or the crucial role of input and negotiated meaning in tone acquisition.

Remember that Jorden and Walton (1987) lamented that “pronunciation practice tends to rely much too heavily on mimicking rather than self-generation” and that students do not “internalize articulatory strategies for production” (p. 116). Students need to continue to study, practice and approach tone *throughout* their study of Chinese, not just the first few weeks or even the first year. The acquisition process could be greatly facilitated by including assessment of tone perception and production in the upper levels, and providing further training in areas including tone produced on the multi-syllabic level and sentence level. Mid to high proficiency students could be trained how to utilize stress and intonation more effectively.

Chen’s (1997) belief that some L2 tonal errors stem from the different pitch register English speakers already possess is probably accurate, but the tonal ‘errors’ must also be compared more objectively with the native speech of the specific dialect group(s) that learners have been influenced by. This study has suggested that it is possible that learners who interact solely with speakers of Beijing Mandarin will acquire slightly different tones than those who communicate regularly with speakers of Taiwan

Mandarin. A greater understanding of native speech should be preeminent in our efforts to prepare to teach non-native learners. Evidence of what L1 and L2 speakers are producing in a variety of complex speaking environments must be obtained before we assess error (Yang, 1995).

Finally, I make an argument for task based research (Ellis, 2000). Prior studies, such as Shen (1989), have made direct pedagogical suggestions based purely on linguistic findings. Task based research helps bridge the gap between linguistics and language teaching by testing the actual methods and tasks that linguistic deficiencies seem to require. When a linguistic phenomenon, such as the one investigated in this study, is understood theoretically and established by empirical evidence, how should we proceed? If for example, it were an established fact that Taiwan Mandarin speakers' tones were each in a lower register than Beijing Mandarin tones, what would that mean about the teaching of tone? Should we then allow students two tracks of learning; one group who wanted to study Taiwan Mandarin, one who hoped to go to Beijing? Should we then teach each group to produce tones in different registers? Not necessarily. Even if it was an established fact, it would be simply a minute detail of tone production, a detail which should be taken in stride with other aspects of tone and language acquisition. While a fact such as this would suggest that researchers trying to assess L2 tone production or perception must take into account the nature of the input non-natives are receiving, specific methods and approaches to teaching should not be recommended based solely on linguistic research. As Richards (2001) has argued, a teaching method cannot be chosen unless we know about the specific context of the language program, with all the inherent factors and players involved. Task-based research, or investigating the effects of specific tasks in specific contexts, can further help educators to match tasks with deficiencies appropriately (Ellis, 2000).

Limitations of the Study and Suggestions for Future Research

While this study does shed light on what holistic and acoustic analysis of multiple speech types offers to our understanding of tone register, my results are not comprehensive. Native speaker perception of tone register was not addressed directly. Due to temporal constraints, only three male native speakers from each dialect group were analyzed. Non-native speech was not collected or analyzed. Further studies with more participants from each of these groups are needed. As tone contour is not addressed in my research, closer investigation of potential dialectal differences in tone contour is also needed.

My methods of data collection may have been limited. First of all, participants interacted with a non-native speaker of Mandarin. It is possible that my pitch register influenced theirs (see Lewis, 2002). It may be useful for other studies to investigate the impact of interlocutor on pitch measurements. Secondly, my demographic interview was probably not the best way to elicit spontaneous conversation. While it controlled for topic quite well, it did not allow for sufficient data production. If more speech from the interview could have been elicited, then longer segments of the descriptive and read portions could have been compared. Future research could utilize a variety of other, perhaps more relevant, speech styles. Finally, the read portion of my study was problematic. Participants probably needed a set of more natural sentences, or perhaps the reading list could have been prepared with exact replicates in succession rather than in random order to enable greater ease in production.

In addition, I did not account for speech rate or loudness, both of which are acoustic correlates of F0. Would rate of speech prove to be a better predictor of median pitch than dialect? It is possible. While the duration of each recording probably minimized the effect of speech rate and loudness substantially, they are still both intervening variables and could have been accounted for. Secondly, my preliminary analysis of individual tones was just that; preliminary - limited to an analysis of tones produced in only one tonal context. Further analysis of all the tones produced, or at least a larger subset of tones, would be fruitful. Despite these limitations, this research still strongly suggests that there is a dialectal tone register difference between Beijing and Taiwan Mandarin. Furthermore, this study effectively encourages future acoustic analysis of pitch and tone register produced in multiple speech styles.

Conclusions

This study has addressed potential dialectal pitch and tone register differences between native speakers of Taiwan and Beijing Mandarin through an acoustic analysis of three forms of continuous speech; spontaneous interview, spontaneous descriptive, and read sentential speech. Two types of F0 measurements were analyzed; F0 extracted from over one minute long segments of recorded speech, and that extracted from individual tones produced in read speech.

Results coincide with previous findings reported by Fon & Chiang (1999), that Taiwan Mandarin tones are produced in a lower register than those produced in Beijing Mandarin. Thus, this study strongly suggests that the difference does exist, and that future research concerning Mandarin lexical tone must take this and other potential dialectal differences into account.

References

Archibald, J. (2001). Second Language Acquisition. *Contemporary Linguistics: An Introduction, Fourth Edition*. Eds. O'Grady, W., Archibald, J., Aronoff, M., Rees-Miller, J. Boston: Bedford/St. Martin's.

Bai, J.-H. (1996). Some thoughts on the notion of better methods. *Journal of the Chinese Language Teachers Association*, 31, 73-86.

Bar-Lev, Z. (1991). Two innovations for teaching tones. *Journal of the Chinese Language Teachers' Association*, 26, 1-24.

Bell, A. (1984). Language style as audience design. *Language in Society* 13, 145-204.

Bialystok, E. (1979). An analytical view of second language competence: A model and some evidence. *Modern Language Journal*, 63, 257-262.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise-ratio of a sampled sound. *Institute of Phonetic Sciences, University of Amsterdam*, *Proceedings*, 17, 97-110.

Byrne, D. (1967). *Progressive Picture Compositions: Pupil's Book*. London: Longman.

Chan, M.K.M. (2003). The digital age and speech technology for Chinese language teaching and learning. *Journal of the Chinese Language Teachers Association*, 38, 49-86.

Chan, M.K.M. (1998). Gender Differences in the Chinese Language: A Preliminary Report. *Proceedings of the Ninth North American Conference on Chinese Linguistics*, 2, Ed., Lin, H. Los Angeles: GSIL Publications.

Chan, M.K.M., Tai, J.H.Y. (1989). A Critical Review of Norman's *Chinese*. *Journal of the Chinese Language Teachers Association*, 24, 43-61.

Chang, M.-Y. (1996). *Language use and language attitudes among Taiwanese elementary school students in native language instruction programs: A study on language maintenance, language shift, and language planning in Taiwan.* Unpublished doctoral dissertation, Indiana University.

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

Chen, C.-Y. (2005). Proposed modifications in teaching materials for Mandarin phonology. *Journal of the Chinese Language Teachers' Association*, 40, 79-98.

Chen, G.-T. (1973). The pitch range of English and Chinese speakers. *Journal of Chinese Linguistics*, 2, 159-171.

Chen, G.-T. (1975). Tone pair. *Journal of the Chinese Language Teachers' Association*, 10, 25-27.

Chen, J., Li, A., Wang, X. (2004). Contrastive study on prosodic aspects for standard and regional accented Chinese. *Proceedings of the International Symposium on Tonal Aspects of Languages: With emphasis on tonal languages* (TAL, Mar 2004). Pages 32-36.

Chen, H-C., & Cutler, A. (1997). Auditory Priming in Spoken and Printed Word Recognition. *Cognitive Processing of Chinese and Related Asian Languages*. Ed. Chen, H.-C. The Chinese University of Hong Kong: The Chinese University Press.

Chen, M.Y. (2000). *Tone Sandhi: Patterns Across Chinese Dialects*. Port Chester, New York, USA: Cambridge University Press.

Chen, P. (1999). *Modern Chinese History and Sociolinguistics*. Cambridge: Cambridge University Press.

Chen, Q. (1997). Toward a sequential approach for tonal error analysis. *Journal of the Chinese Language Teachers Association*, 32, 21-39.

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.* Unpublished doctoral dissertation, Brigham Young University.

Chen, S. (1973). The third tone and see-saw pairs (A field technique). *Journal of the Chinese Language Teachers Association*, 8, 145-149.

Cheng, R. L. (1985). A comparison of Taiwanese, Taiwan Mandarin, and Peking Mandarin. *Language*, 61, 352-377.

Cheung, H., Hsieh, L. (1997). Learning a new verb in Mandarin Chinese: The effects of affectedness and phonological shape. *Journal of Chinese Linguistics,* 25, 21-39.

Chiang, T. (1979). Some interferences of English intonations with Chinese tones. *International Review of Applied Linguistics in Language Learning*, 17, 245-250.

Chiung, W. T. (2001). Tone change in Taiwanese: Age and geographic factors. *U. Penn Working Papers in Linguistics,* 8, Proceedings of the 25th Annual PLC.

Coblin, W. S. (2000). A Brief History of Mandarin. *Journal of American Oriental Society*, 120, 537-552.

Connell, B.A., Hogan, J.T., Rozsypal, A.J. (1983). Experimental evidence of interaction between tone and intonation in Mandarin Chinese. *Journal of Phonetics,* 11, 337-351.

DeFrancis, J. (1984). *The Chinese Language: Fact and Fantasy***.** Honolulu: University of

Hawaii Press.

Duanmu, S. (2000). *The Phonology of Standard Chinese*. New York: Oxford University Press.

Dunn, M., Flax, J. (1996). The use of spontaneous language measures as criteria for identifying children with specific language impairment: An attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research*, 39, 643.

Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. (2000). Task-based research and language pedagogy. *Language-Teaching-Research*, 4, 193-220.

Fon, J., Chiang, W.-Y. (1999). What does Chao have to say about tones? – A case study of Taiwan Mandarin. *Journal of Chinese Linguistics*, 27, 13-37.

Fon, J., Chiang, W.-Y., Cheung, H. (2004). Production and perception of the two dipping tones (Tone 2 and Tone 3) in Taiwan Mandarin. *Journal of Chinese Linguistics*, 32, 249-281.

Fox, A., Qi, Y.-Y. (1990). Context effects in the perception of lexical tone. *Journal of Chinese Linguistics*, 18, 261-284.

Gass, S.M., Selinker, L. (2001). *Second Language Acquisition: an Introductory Course, 2nd Ed*. Mahwah, New Jersey: L. Erlbaum Associates.

Gleason, J.B. (2001). *The Development of Language*. Boston University: Allyn and Bacon.

Hu, M. (1991). Feminine accent in the Beijing vernacular: a sociolinguistic investigation.

Journal of the Chinese Language Teachers Association, 26, 49-54.

Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *The Journal of Child Language*, 27, 3-42.

Huang, C.-M. (1997). Language education policies in Taiwan: From nationism to nationalism. Unpublished doctoral dissertation, University of Washington.

Huang, X.F. (1994). *Language, Society, and Ethnic Ideology: The Studies of Sociolinguistics in Taiwan* (2nd ed.). Taipei, Taiwan: Crane.

Hung, T.T.N. (1990). Syntax-phonology interaction in Chinese tone sandhi: An overview. *Journal of the Chinese Language Teachers Association*, 25, 43-59.

Jorden, E.H., Walton, A.R. (1987). Truly foreign languages: Instructional challenges. *Foreign Language Instruction: A National Agenda*. Ed. Lambert, R.D. Newbury Park, California: Sage Publications, Inc.

Jie, Z. (1988). *Attrition patterns in learners' use of Mandarin Chinese tones*. Unpublished masters thesis, Brigham Young University.

Juffs, A. (1990). Tone, syllable structure and interlanguage phonology: Chinese learners' stress errors. *International Review of Applied Linguistics in Language Teaching*, 28, 99-118.

Kubler, C. C. (1985). The Development of Mandarin in Taiwan: A Case Study of Language Contact. Taibei, Taiwan: Student Book Co., Ltd.

Kubler, C. C., Yung-O B., Henrichson, G. C., Walton, A. R., Wong, M. M., Wu, W.-L., & Yu, C. (1997). *NFLC Guide for Basic Chinese Language Programs*. Columbus, Ohio: The Ohio State University.

Labov, W. (1984). Field Methods of the Project in Linguistic Change and Variation.

Language in Use. Eds. Baugh, J. & Sherzer, J. Prentice-Hall.

Ladd, D.R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Ladofoged, P. (2001). *A Course in Phonetics, Fourth Edition*. Orlando, Florida: Harcourt, Inc.

Lakshmanan, U., Selinker, L. (2001). Analysing interlanguage: how do we know what learners know? *Second Language Research*, 17, 393-420.

Language Atlas of China. (1987). Compiled jointly by Australian Academy of the Humanities and Chinese Academy of Social Sciences. Hong Kong: Longman.

Lee, T. H.-T. (1996). Theoretical issues in language development and Chinese child language. *New Horizons in Chinese Linguistics*. Eds. Huang, C.-T. J., & Li, Y.-H. A. Dordrecht, Netherlands: Kluwer Academic Publishers.

Lee, Y.-S., Vakoch, D. A., Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, 25, 527-542.

Lewis, J.A. (2002). *Social influences on female speakers' pitch*. Unpublished doctoral dissertation, University of California, Berkeley.

Li, A. Li, C. N., Thompson, S. A. (1977). The acquisition of tone in Mandarin speaking children. *Journal of Child Language*, 4, 185-199.

Li, C. N., Thompson, S. A. (1981). *Mandarin Chinese A Functional Reference Grammar*. Berkeley, California: University of California Press.

Lien, C. (1986). Tone merger in the dialects of northern Chinese. *Journal of Chinese Linguistics*, 14, 242-290.

Lin, H. (2001). Stress and the distribution of the neutral tone in Mandarin. *Chinese*

Phonology in Generative Grammar. Ed. Xu, D.B. San Diego: Academic Press.

Lin, H.-B., Repp, B.H. (1989). Cues to the perception of Taiwanese tones. *Language and Speech*, 32, 25-44.

Lin, Y. (2000). Vocabulary acquisition and learning Chinese as a foreign language (CFL). *Journal of the Chinese Language Teachers Association*, 35, 85-108.

Liu, H.-M. (2002). *The acoustic-phonetic characteristics of infant-directed speech in Mandarin Chinese and their relation to infant speech perception in the first year of life*. Unpublished doctoral dissertation, University of Washington.

Llisterri, J. (1995). Relationships between speech production and speech perception in a second language. *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Ed. Elenius, K., Branderud, P. Stockholm: KTH / Stockholm University

Loke, K.-K. (2002). Approaches to the teaching and learning of Chinese: a critical literature review and a proposal for a semantic, cognitive and metacognitive approach. *Journal of the Chinese Language Teachers Association*, 37, 65-112.

Lundelius, J. O. (1992). Pinyin vs. tonal spelling. *Journal of the Chinese Language Teachers' Association*, 28, 93-108.

Miller, J., Weinert, R. (1998). *Spontaneous Spoken Language*. New York, Oxford University Press.

Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of the Chinese Language Teachers Association*, 24, 49-65.

Major, R. C. (1994). *First and Second Language Phonology*. Ed. Mehmet Yavas. San Diego, California: Singular Publishing Group, Inc.

Massaro, D.W., Cohen, M.M., Tseng, C.-Y. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics*, 13, 267-289.

McGinnis, S. (1997). Tonal distinction errors by beginning Chinese language students: A comparative study of American English and Japanese native speakers. *Chinese Pedagogy: An Emerging Field, Vol. 2 CLTA Monograph #2*. Ed. McGinnis, S., Ohio State University Foreign Language Publications.

Moore, C. B. (1993). Phonetic observations on stress and tones in Mandarin Chinese. *Working Papers of the Cornell Phonetics Laboratory*, 8, 87-117.

Norman, J. (1988). *Chinese*. New York: Cambridge University Press.

Papagno, C., Valentine, T., Baddeley, A. (1991). Phonological short-term memory and foreign language vocabulary learning. *Journal of Memory and Language*, 30, 331-347.

Papousek, M., Hwang, S.-F. C. (1991). Tone and intonation in Mandarin babytalk to presyllabic infants: Comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics*, 12, 481-504.

Papousek, M., Papousek, H., Symmes, D. (1991). The meanings of melodies in Motherese in tone and stress languages. *Infant Behavior and Development*, 14, 415-440.

Richards, J.C. (2001). *Curriculum Development in Language Teaching*. Cambridge, Cambridge University Press.

Rose, P. (2002). Independent depressor and register effects in Wu dialect tonology: Evidence from Wenzhou tone sandhi. *Journal of Chinese Linguistics,* 30, 39-81.

Rummer, R., Engelkamp, J. (2001). Phonological information contributes to short-term recall of auditorily presented sentences. *Journal of Memory and Language*, 45, 451-467.

Sanders, R.M. (1992). The expression of modality in Peking and Taipei Mandarin. *Journal of Chinese Linguistics*, 20, 1-23.

Shen, J. (1985). Pitch range of tone and intonation in Beijing Dialect. *Working Papers in Experimental Phonetics*, Ed. Lin, T., & Wang, L.-J. Beijing: Beijing University Press.

Shen, X.N.S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers Association*, 24, 27-47.

Shen, X.N.S. (1990). The prosody of Mandarin Chinese. *Linguistics* (Vol.118). Berkeley, California: University of California Press.

Shen, X.N.S., Lin, M.-C. (1991). Concept of tone in Mandarin revisited: A perceptual study on tonal coarticulation. *Language Sciences*, 13, 421-432.

Shen, X.N.S., Lin, M.-C., & Yan, J.-Z. (1993). F0 turning point as an F0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *Journal of the Acoustic Society of America*, 93, 2241-2243.

Shih, C. (2000). A declination model of Mandarin Chinese. *Intonation: Analysis, Modeling & Technology*. Ed. Botinis A., Kluwer Academic Publishers, 243-268.

Snow, D. (1998). Children's imitations of intonation contours: are rising tones more difficult than falling ones? *Journal of Speech, Language & Hearing Research*, 41, 576-588.

Spinks, J.A., Liu, Y., Perfetti, C.A., & Tan, L.H. (2000). Reading Chinese characters for meaning: the role of phonological information. *Cognition*, 76, B1- B11.

Stagray, J. R., Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics*, 21, 143-168.

Sun, H.S. (1998). *The Development of a Lexical Tone Phonology in American Adult Learners of Standard Mandarin Chinese*. Second Language Teaching & Curriculum Center: University of Hawaii Press.

Swihart, D.-A.W. (2003). The two Mandarins: *Putonghua* and *Guoyu*. *Journal of the Chinese Language Teachers Association*, 38, 103-118.

Tardif, T., Shatz, M., Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, 24, 535-565.

Tseng, C.-Y. (1981). *An acoustic phonetic study on tones in Mandarin Chinese*. Unpublished doctoral dissertation, Brown University.

Tsung, C. (1987). Half-third first: on the nature of the third tone. *Journal of the Chinese Language Teachers Association*, 22, 87-101.

VanNess Simmons, R. (1999). *Chinese Dialect Classification*. Philadelphia: John Benjamins Publishing Company.

Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, 78, 332-348.

Whalen, D., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49, 25-47.

Xing, J. (2003). Toward a pedagogical grammar of Chinese: Approach, content and process. *Journal of the Chinese Language Teachers Association*, 38, 41-68.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61-83.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27, 55-105.

Xu, Y. (2001). Sources of tonal variations in connected speech. *Journal of Chinese Linguistics Monograph Series*, 17, 1-31.

Xu, Y. and Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America*, 111: 1399-1413.

Yang, L.-C. (1995). *Intonational structures of Mandarin discourse*. Unpublished doctoral dissertation, Georgetown University.

Young, R. (1989). *Language Maintenance and Language Shift Among the Chinese on Taiwan*. Taibei: The Crane Publishing Co., Ltd.

Yuasa, I. (2001). *Politeness, emotion & gender: A sociophonetic study of voice pitch modulation*. Unpublished doctoral dissertation, University of California, Berkeley.

Yue-Hashimoto, A. O. (1986). Tonal flip-flop in Chinese dialects. *Journal of Chinese Linguistics*, 14, 161-183.

Zhou, X., Marslen-Wilson, W., Taft, M., Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese compound words. *Language and Cognitive Processes*, 14, 525-565.

Appendix A

Power Point Presentation Text[8]

Part 1: Text in English

Slide 1: Welcome

Thank you for participating

in this study!

Go ahead to the next slide.

Slide 2: Test Run

The first four slides of this presentation are designed to get you comfortable with the PowerPoint presentation and to enable the researcher to adjust the volume settings.

If you have any problems at any time during the presentation please stop and ask the researcher.

Go ahead to the next slide.

Slide 3: Counting

Please count to ten in Chinese, slowly and in your regular speaking voice.

When you are finished you may proceed to the next slide.

Slide 4: Counting

Now, please count to ten again in Chinese, this time as quickly as you can.

When you are finished you may proceed to the next slide.

[8] Note: The entire PowerPoint presentation was presented to native speaking subjects from Taiwan and Beijing in Traditional and Simplified Chinese characters, respectively. Translation assistance came from two native speaking students of Chinese enrolled at Brigham Young University.

Slide 5: Any Questions?

If you have any questions about what you are supposed to do, please ask now.

Remember, if you need any help during the recording process just ask.

When you are ready, proceed to the next slide.

Slide 6: Interview

Please wait for a moment while I come in to join you in the recording booth.

When our conversation is over you may proceed to the next slide.

Slide 7: Picture Description

This next portion will require you to describe a series of pictures.

Each slide has four pictures that all together tell a story. Please begin describing the pictures immediately after they appear. The pictures go chronologically through the story from top to bottom and left to right.

Your response should be such that a native speaker of Chinese could visualize the pictures and understand the story you see.

When you are ready you may proceed to the next slide.

Slides 8-13 consisted of a series of pictures shown in Appendix B. The only text that appeared was a pop-up window that stated: Describe this series of pictures: Begin now!

Slide 14: Break Time!

Please take a 2-3 minute break.

Feel free to exit the booth and use the rest-room, get a drink, etc.

When you are ready, proceed to the next slide.

Slide 15: Reading Portion

This next portion will require you to read a series of sentences.

When you are ready you may proceed to the next slide.

Slides 16- 32 consisted of the carrier sentences shown in appendix H (borrowed from Xu's (1997) study).

Slide 33: Thank you for participating!

Appendix A (continued)

PowerPoint Presentation Text

Part 2: Text in Traditional Chinese

Slide 1: 歡迎！感謝您的參與！

下一頁

Slide 2: 測試開始

前四张幻灯片幫助您熟悉 PowerPoint 演示文稿。研究員借機調整音量。

演示期間如果您有任何問題，請随時向研究員提問。

下一頁

Slide 3: 数数

請用平時說話的聲音慢慢用中文從一数到十。

完成後到下一頁。

Slide 4: 数数

現在請以您最快的语速用中文從一数到十。

完成後到下一頁。

Slide 5: 有問題吗？

如果您對需要做的東西有任何問題，現在請提問。

請記住如果您在錄音過程中有任何問題，請隨時提問。

準備好後，請繼續到下一頁。

Slide 6: 面试

請稍候。我進入錄音棚後馬上開始。

在我們談話結束後,請繼續到下一頁。

Slide 7: 圖片描述

下面這個部分将要求您描述一系列的圖片。

每张幻灯片包括四张圖片，這四张圖片講述一個故事。請在看到圖片後立即開始描述。圖片由上到下，由左到右以时間顺序排列。

以中文為母語的人在聽到您的描述侯，應該能够想象出畫面并且理解您看到的故事。

準備好後，繼續到下一頁。

Slides 8-13: Pictures shown in Appendix B

Slide 14: 中間休息！

請休息 2－3 分鐘。

您尽可出入錄音棚，飲用飲料，使用洗手間等等。

準備好後，請繼續到下一頁。

Slide 15: 閱讀部分

下一部分将要求您讀出一系列句子。

準備好後，請繼續到下一頁。

Slide 16-32: Shown in Appendix H

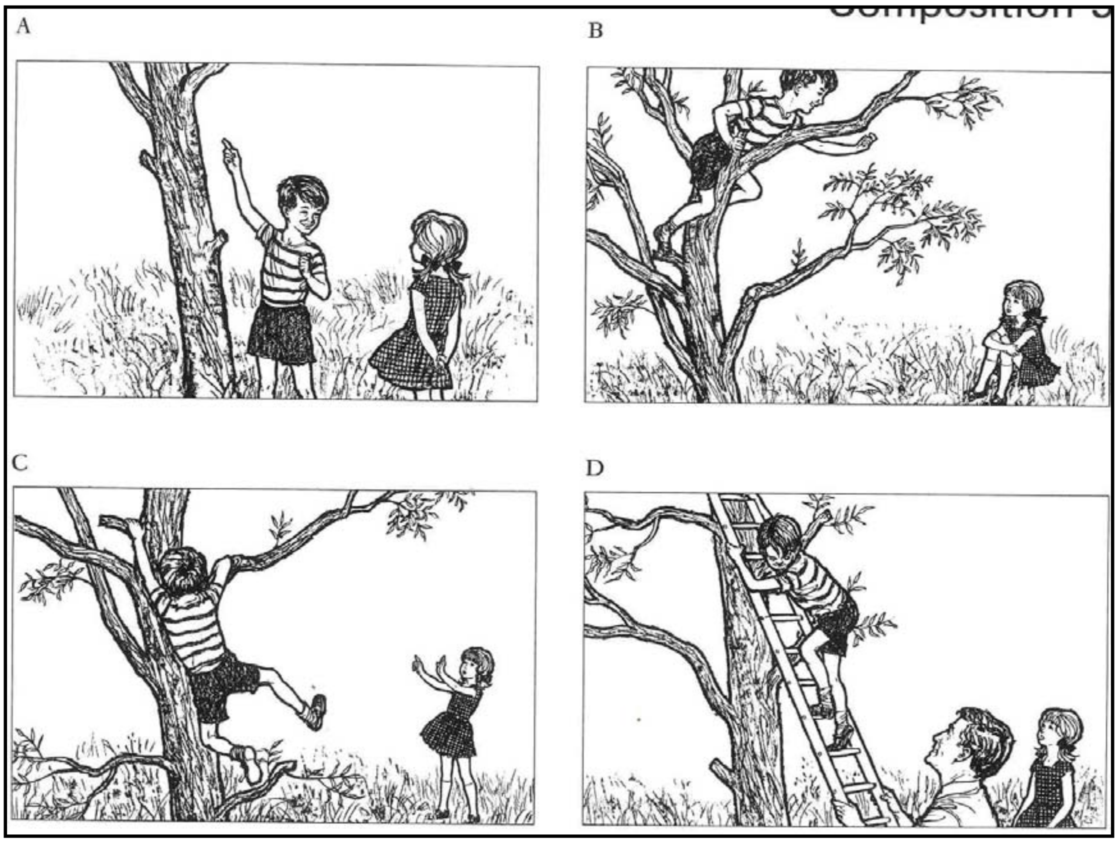
Slide 33: 謝謝您來參與!

Appendix B

Picture Description Elicitation Instruments

From Byrne (1967)

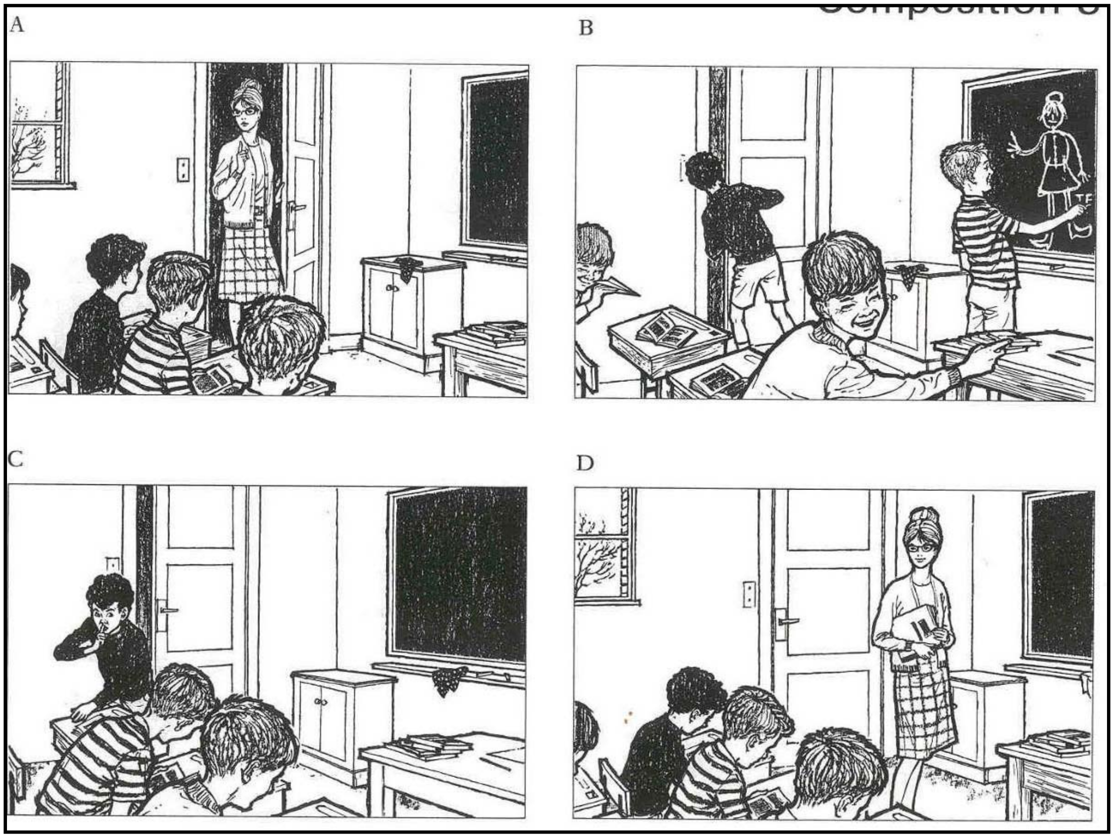
Picture 1 Slide 8



Appendix B (Continued)

From Byrne (1967)

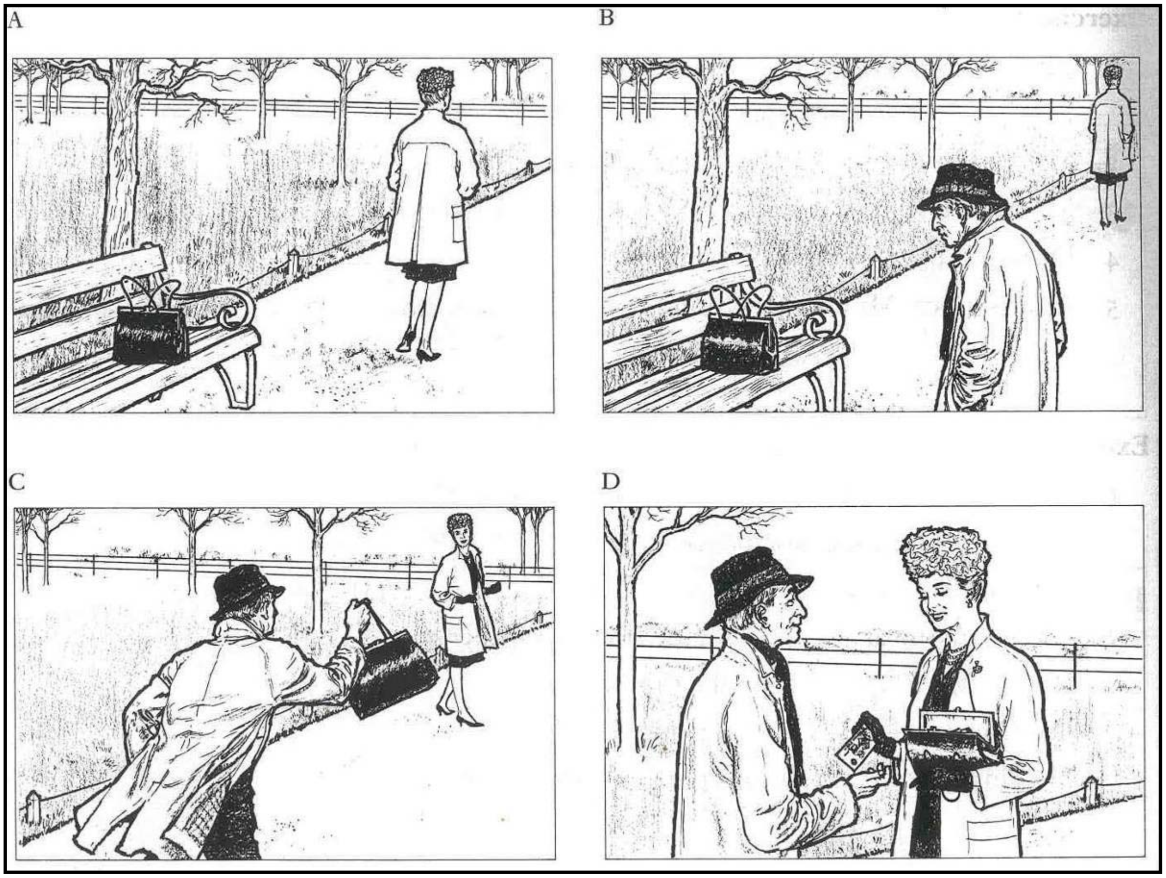
Picture 2 Slide 9



Appendix B (Continued)

From Byrne (1967)

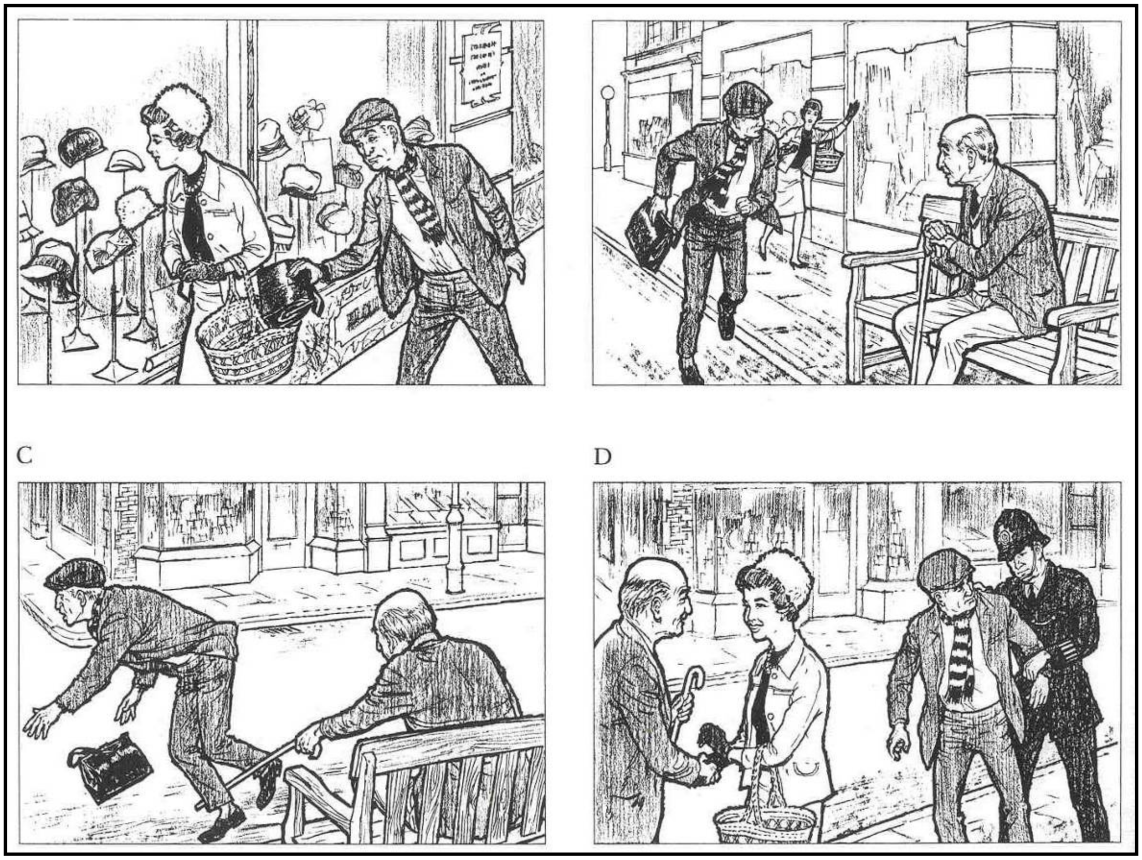
Picture 3 Slide 10



Appendix B (Continued)

From Byrne (1967)

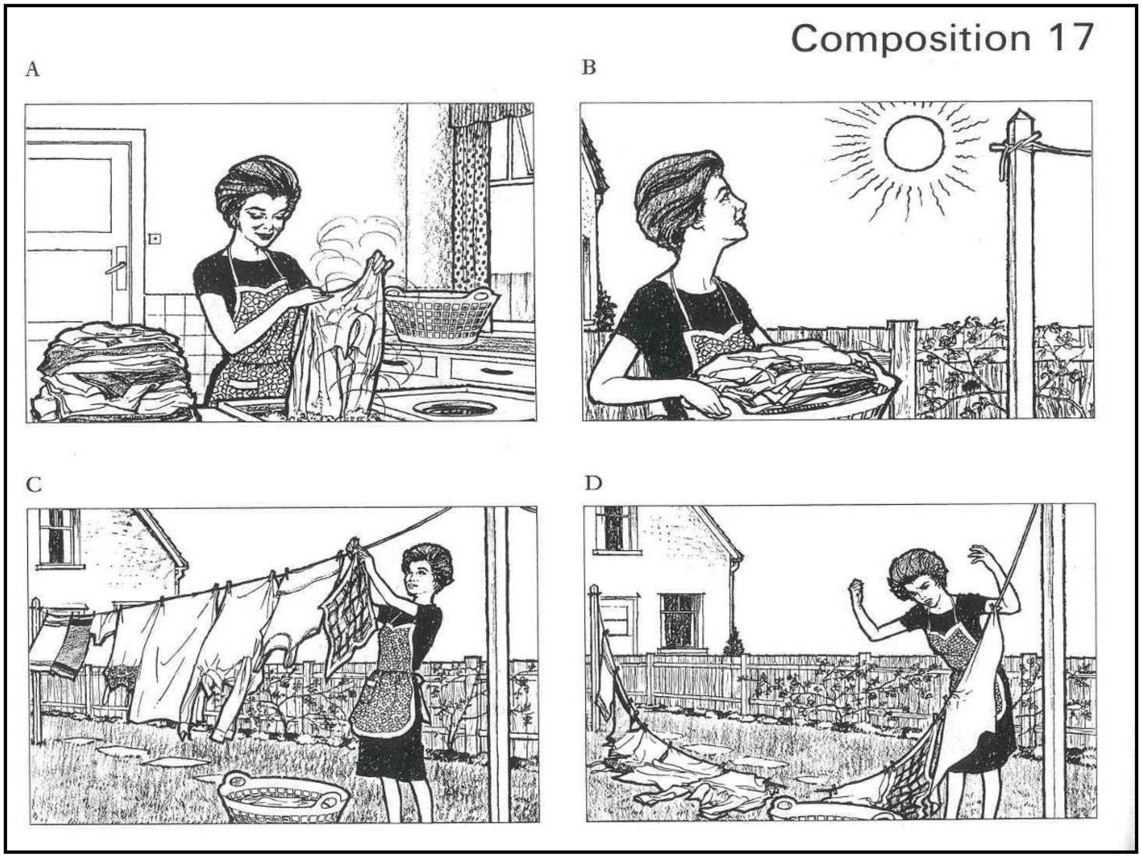
Picture 4 Slide 11



Appendix B (Continued)

From Byrne (1967)

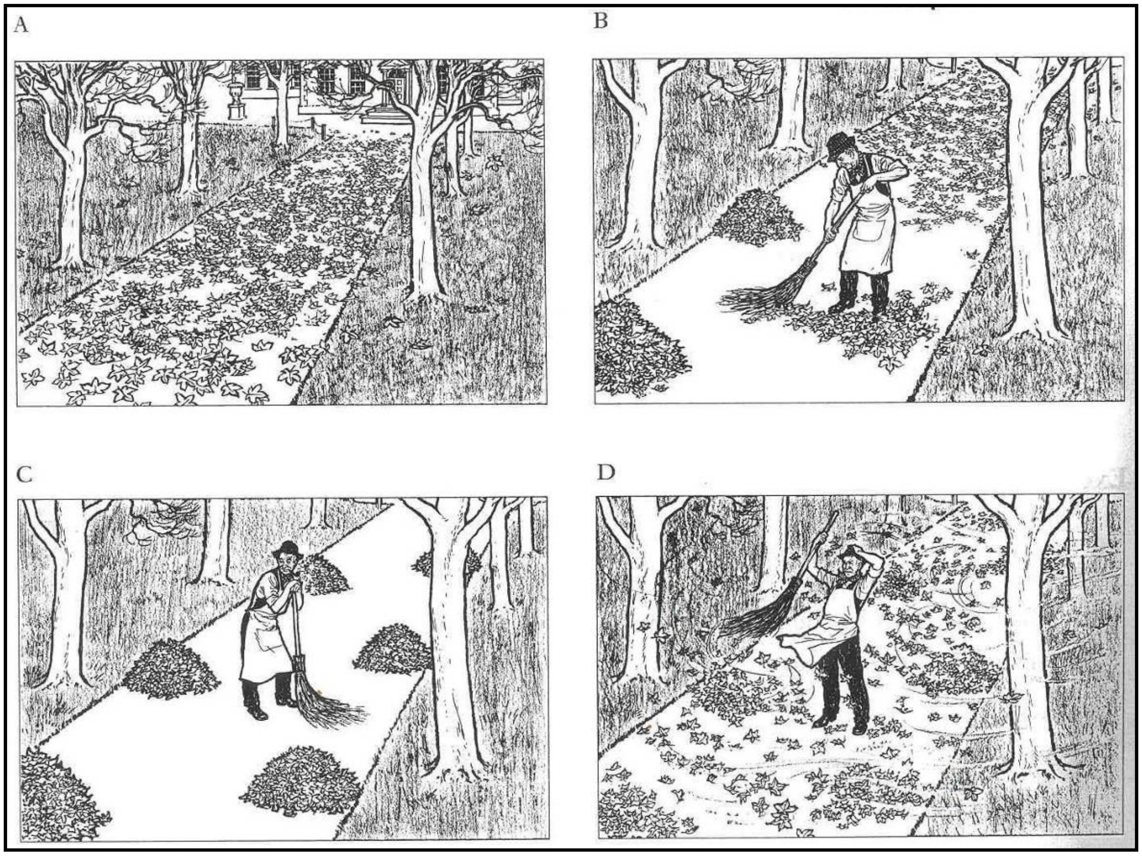
Picture 5 Slide 11



Appendix B (Continued)

From Byrne (1967)

Picture 6 Slide 12



Appendix C

Demographic Questions in English and Chinese

Note: Speech directed toward participants from Taiwan comes before the dash, speech directed towards participants from Beijing comes after the dash (/).

1. How are you today? 你今天怎么样？
2. Are you busy this semester? 这个学期忙不忙？
3. What classes are you taking? 上什么课？
4. What is your major? 你的主修/专业是什么？
5. Where is your hometown? 你老家在哪里/儿？
6. How long did you live there? 哦，你住在——多久啊？
7. Were your parents also born there? 你爸爸妈妈也在那里/儿出生的吗？
    a. If not, where were they born? 他们在哪里/儿出生的？
8. What year where you born? 你是哪一年出生的了？
9. When is your birthday? 你的生日是什么时候？
10. What other languages do you speak besides Chinese and English? 除了国语/普通话和英语以外，你还会说什么语言？
11. How about your parents? 你爸爸妈妈呢？
12. What dialects do you speak? 你会说什么方言？
13. How about your parents? 你爸爸妈妈呢？
14. What do your parents do? 你爸爸妈妈做什么？

15. What year did you start studying English in China? 你在台湾/中国的时候哪一年开始学英文？

16. Where did you go to High School? 你在哪里/儿读高中？

    a. When did you graduate? 什么时候毕业的？

17. Did you go to a Chinese University? 你有没有在台湾/大陆念过大学？

    a. What did you study there? 学什么？

18. When did you come to America? 你什么时候到美国来？

19. Do you speak more English than Chinese each day? 你每天说的英文比中文多吗？

    a. Why? 为什么？

20. Are you working now? 你现在有没有工作？

    a. What do you do? 做什么？

21. What would you like to do in the future? 你将来想做什么？

    a. Why? 为什么？

22. Do you want to go back home? 想回去台湾/大陆吗？

Appendix D

Informed Consent Form in Simplified Chinese[9]

研究对象同意书

研究目的
本研究之目的是为要更加了解中文语言，俾以改善那些针对非本地中文学生而設計的教学法和评估。进行本项研究的 Rich Torgerson，是杨百翰大学的一名研究生。在此我们邀请會讲中文的您参与此项研究。

程序
本研究需要請您到 JFSB 教学大楼的 1141 號教室做一个录音。整个录音过程需时约 30 分钟，其中包括有個 2 到 3 分钟的休息时间。 读完这份同意书后请到录音室找一台电脑，并坐下来。研究員， Rich Torgerson 会請您從 1 数到 10，幫助您熟悉如何参与本项研究，然后问您几个问题。其间若有任何问题，请随时發問。当您熟悉了整个程序之后，研究员会请您自行进行其餘的录音部分。录音过程中若您有任何疑问，研究员就在錄音室之外， 可隨時為您提供協助。

参与
任何参与本研究的情況都是建立在自愿的基础上。您在任何时候都可以选择無條件退出，或拒絕参与完整的研究計畫。

风险/不便之处
参与本研究之风险或不便之处非常小。

报酬
本研究並不提供任何實質上的 报酬。 但是本研究所带来的对中文这门语言的新的认知可能给学习中文的人带来好处。

保密性
整个研究 过程将严格保密。 在研究报告中绝不会提及您的名字。 这项研究收集到的所有数据资料都将被保存在一个安全的地方，只有参与研究的人员才能接触这些材料。

有疑问吗？
若您对这项研究有任何疑问，請直接联系 Rich Torgerson，电话：(801)294-0184，电子邮箱：rich_torgerson@byu.edu。如果您對於您参与研究計畫有何权利方面的问题，请联系机构审查委员会主席 Renea Beckstrand 教授，他的地址是：422 SWKT Brigham Young University, Provo, Utah 84602；联系电话：(801) 422-3873.

[9] Translation of consent forms was conducted by having two native speaking students correct my Chinese translation of the English document. Corrections were made, and then verified by another native speaker.

Appendix D (Continued)

Informed Consent Form in Traditional Chinese

研究對象同意書

研究目的
本研究之目的是為要更加了解中文語言，俾以改善那些針對非本地中文學生而設計的教學法和評估. 進行這項研究的 Rich Torgerson，是楊白翰大學的一名研究生. 在此我們邀請會講中文的您參與此項研究.

程序
本研究需要請您到 JFSB 教學大楼的 1141 號教室做一個錄音。整個錄音過程需時約 30 分鐘，其中包括有個 2 到 3 分鐘的休息時間。讀完這份同意書后請到錄音室找一台電腦，并坐下来。研究員，Rich Torgerson 會請您從 1 数到 10，幫助您熟悉如何參與這項研究，然后問您幾個問題。其間若有任何問題，請隨時發問。當您熟悉了整個程序之後，研究院會請您自行進行其餘的錄音部分。錄音過程中若您有任何疑問，研究就在錄音室之外，可隨時為您提供協助。

參與
任何參與本研究的情况都是建立在自愿的基础上。您在任何時候都可以选择無條件退出，或拒絕參與完整的研究計畫。

風險/不便之处
參與本研究之風險或不便之處非常小。

報酬
本研究並不提供任何實質上的報酬。但是本研究所帶來的對中文這門語言的新的認知可能給學習中文的人帶来好處。

保密性
整個研究過程將严格保密。在研究報告中絕不會提及您的名字。這項研究收集到的所有数据資料都将被保存在一個安全的地方，只有參與研究的人員才能接触這些材料。

有疑問嗎？
若您對這項研究有任何疑問，請直接聯繫 Rich Torgerson，電話：(801)294-0184，電子郵箱：rich_torgerson@byu.edu。如果您對於您參與研究計畫有何權利方面的問題，請聯繫 I.R.B.主席 Renea Beckstrand 教授，他的地址是：422 SWKT Brigham Young University, Provo, Utah 84602；(801) 422-3873.

Appendix E

Raw data from three speech styles presented in ERB, Semitones, and Hz

Raw Data Presented in ERB

Participant	Dialect	SS	10%	90%	Med	Range	S.D.
1	1	1	2.32039	3.25692	2.69873	0.9366	0.439
2	1	1	2.62625	3.70718	3.03562	1.081	0.5509
3	1	1	3.04978	3.89125	3.36365	0.8418	0.7621
6	2	1	3.17944	4.2264	3.52497	1.047	0.5372
7	2	1	2.25517	4.61669	3.83689	2.362	0.867
8	2	1	2.67725	3.70743	3.18024	1.03	0.504
1	1	2	2.38796	3.2082	2.73505	0.8203	0.3871
2	1	2	2.66959	3.72395	3.09625	1.055	0.4211
3	1	2	3.03599	3.98432	3.45595	0.9485	0.5821
6	2	2	3.21967	4.71394	3.81245	1.494	0.6471
7	2	2	3.18308	4.71874	3.95397	1.536	0.6856
8	2	2	2.8453	3.79712	3.31629	0.9519	0.4601
1	1	3	2.33955	3.07033	2.68093	0.7309	0.3359
2	1	3	2.67762	4.15643	3.29396	1.479	0.5837
3	1	3	2.45916	4.04041	3.42673	1.581	0.6189
6	2	3	3.12503	4.5054	3.74385	1.381	0.6251
7	2	3	3.21308	4.86163	3.81503	1.649	0.7537
8	2	3	2.6193	3.69924	3.03305	1.08	0.4106

Raw Data Presented in Semitones

Participant	Dialect	SS	10%	90%	Med	Range	S.D.
1	1	1	-5.068	1.560	-2.152	6.628	2.938
2	1	1	-2.681	4.178	0.159	6.860	3.301
3	1	1	0.251	5.172	2.207	4.923	4.135
6	2	1	1.079	6.888	3.152	5.810	3.056
7	2	1	-5.612	8.753	4.883	14.370	5.219
8	2	1	-2.307	4.179	1.084	6.488	3.408
1	1	2	-4.518	1.258	-1.891	5.777	2.574
2	1	2	-2.363	4.270	0.551	6.634	2.648
3	1	2	0.161	5.660	2.752	4.887	3.840
6	2	2	1.330	9.198	4.751	7.869	3.370
7	2	2	1.102	9.220	5.502	8.120	3.905
8	2	2	-1.117	4.669	1.922	5.787	3.019
1	1	3	-4.910	0.384	-2.280	5.295	2.390
2	1	3	-2.305	6.539	1.786	8.845	3.464
3	1	3	-3.953	5.950	2.581	9.904	4.165
6	2	3	0.735	8.234	4.379	7.500	3.437
7	2	3	1.289	9.861	4.765	8.573	4.482
8	2	3	-2.733	4.134	0.142	6.867	2.627

Appendix E (continued)

Raw data from three speech styles presented in ERB, Semitones, and Hz

Raw Data Presented in Hz

Participant	Language	SS	10%	90%	Med	Range	S.D.
1	1	1	74.623	109.428	88.313	34.810	17.140
2	1	1	85.652	127.292	100.923	41.650	22.440
3	1	1	101.462	134.818	113.594	33.370	33.170
6	2	1	106.430	148.865	119.971	42.440	22.190
7	2	1	72.311	165.797	132.582	93.500	35.250
8	2	1	87.523	127.303	106.461	39.790	19.520
1	1	2	77.032	107.540	89.653	30.510	15.070
2	1	2	87.241	127.973	103.235	40.740	16.380
3	1	2	100.937	138.674	117.230	37.750	22.850
6	2	2	107.984	170.115	131.580	62.140	27.570
7	2	2	106.570	170.330	137.413	63.770	28.210
8	2	2	93.750	130.953	111.740	37.210	17.980
1	1	3	75.304	102.245	87.658	29.640	12.690
2	1	3	87.536	145.895	110.869	58.370	23.190
3	1	3	79.587	141.015	116.076	61.430	23.810
6	2	3	104.337	160.905	128.781	56.570	26.020
7	2	3	107.729	176.751	131.686	69.030	30.570
8	2	3	85.398	126.971	100.825	41.580	15.830

Appendix F

Raw Data Display of Read Tones (ERB)

Row	Language	Participant	Tone	Min	Max	Median	Range
1	1	1	1	2.71042	2.80651	2.75643	0.09608
2	1	1	1	2.86049	2.86049	2.88148	0.11120
3	1	1	1	2.71920	2.82121	2.77767	0.10200
4	1	1	2	2.38905	2.87550	2.45738	0.48650
5	1	1	2	2.59048	2.90795	2.64756	0.31750
6	1	1	2	2.54671	2.84773	2.57070	0.30100
7	1	1	3	2.39632	2.68991	2.44460	0.29360
8	1	1	3	2.22780	2.62252	2.26313	0.39470
9	1	1	3	2.41311	2.61277	2.51746	0.19970
10	1	1	4	2.90160	2.95492	2.91102	0.05333
11	1	1	4	2.54768	3.10756	2.82853	0.55990
12	1	1	4	2.39738	3.02171	2.72786	0.62430
13	1	1	5	2.77797	2.93940	2.83085	0.16140
14	1	1	5	2.65920	2.80688	2.76637	0.14770
15	1	1	5	2.56019	2.84041	2.75148	0.28020
16	1	2	1	4.07981	4.39791	4.16547	0.31810
17	1	2	1	3.84236	4.04670	3.89343	0.20430
18	1	2	1	3.69735	4.06236	4.00363	0.36500
19	1	2	2	3.02155	3.53625	3.13887	0.51470
20	1	2	2	3.29646	3.66528	3.40190	0.36880
21	1	2	2	3.11235	3.18336	3.12966	0.07102
22	1	2	3	2.46450	3.06455	2.66911	0.60010
23	1	2	3	2.69849	3.56009	2.83042	0.86160
24	1	2	3	1.63069	2.77810	2.48538	1.14700
25	1	2	4	3.34495	4.31958	3.83254	0.97460
26	1	2	4	3.35515	4.27411	4.07932	0.91900
27	1	2	4	3.85499	3.94242	3.90453	0.08743
28	1	2	5	3.87362	4.09167	4.00824	0.21810
29	1	2	5	3.69461	3.76042	3.73964	0.06580
30	1	2	5	3.58397	3.79448	3.73874	0.21050
31	1	3	1	3.37777	3.43580	3.40968	0.05803
32	1	3	1	3.45411	3.59102	3.52008	0.13690
33	1	3	1	3.35841	3.43846	3.39765	0.08005
34	1	3	2	2.99937	3.55370	3.26060	0.55430
35	1	3	2	3.33369	3.42358	3.38645	0.08989
36	1	3	2	3.16485	3.51556	3.27358	0.35070
37	1	3	3	2.39632	2.68991	2.44460	0.29360
38	1	3	3	2.22780	2.62252	2.26313	0.39470
39	1	3	3	2.12302	3.16334	2.52092	1.04000
40	1	3	4	3.30251	3.98919	3.70879	0.68670
41	1	3	4	3.08926	3.96116	3.66052	0.87190
42	1	3	4	2.60118	3.75635	3.35127	1.15500
43	1	3	5	3.53511	3.68093	3.60830	0.14580
44	1	3	5	3.58196	3.67252	3.63943	0.09056
45	1	3	5	3.30497	3.41979	3.37851	0.11480
46	2	6	1	3.98346	4.11619	4.00195	0.13270
47	2	6	1	4.14743	4.46089	4.32043	0.31350
48	2	6	1	4.07421	4.38378	4.20434	0.30960
49	2	6	2	3.64080	4.30246	3.87803	0.66170
50	2	6	2	3.57275	3.84470	3.65610	0.27200

Appendix F (Continued)

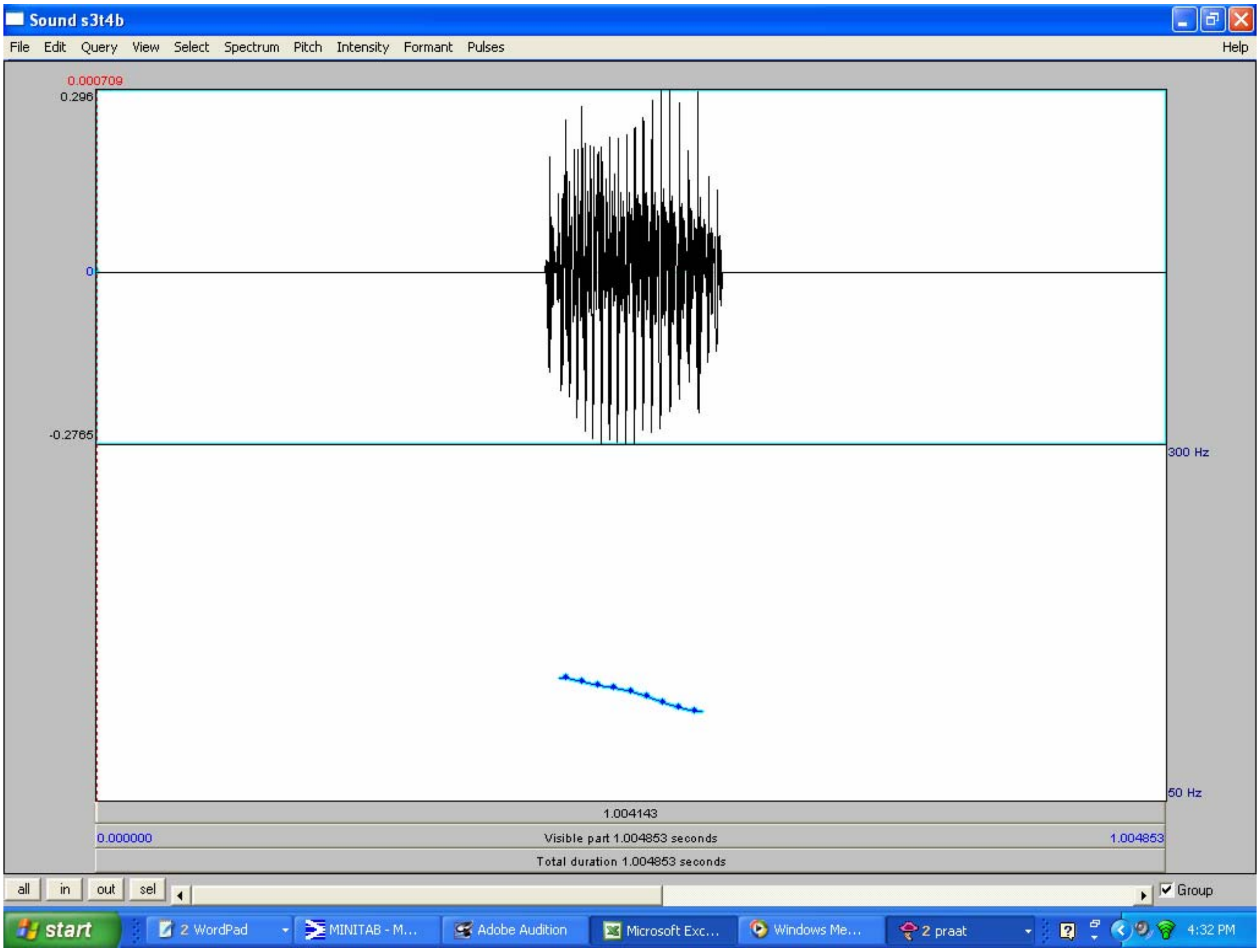
Raw Data Display of Read Tones (ERB)

51	2	6	2	3.52687	3.94616	3.64491	0.41930
52	2	6	3	2.71519	2.91891	2.75709	0.20370
53	2	6	3	2.83439	3.12756	2.94222	0.29320
54	2	6	3	2.90914	3.37804	3.01892	0.46890
55	2	6	4	3.34621	3.99635	3.62501	0.65010
56	2	6	4	3.46059	4.50045	3.95227	1.04000
57	2	6	4	3.09531	4.50680	3.97083	1.41100
58	2	6	5	4.21752	4.25188	4.23776	0.03436
59	2	6	5	4.18094	4.36349	4.24958	0.18260
60	2	6	5	4.00756	4.11181	4.06498	0.10420
61	2	7	1	4.21420	4.36448	4.25859	0.15030
62	2	7	1	4.20957	4.65405	4.42309	0.44450
63	2	7	1	3.94254	4.07522	4.03621	0.13270
64	2	7	2	3.25186	3.45255	3.29811	0.20070
65	2	7	2	3.78845	4.48861	3.92866	0.70020
66	2	7	2	3.48572	3.82472	3.58129	0.33900
67	2	7	3	3.93075	4.32553	4.07324	0.39480
68	2	7	3	2.87772	3.18758	2.93548	0.30990
69	2	7	3	1.76934	3.37589	3.16056	1.60700
70	2	7	4	3.60311	4.70842	4.08074	1.10500
71	2	7	4	3.67916	4.71542	4.16008	1.03600
72	2	5	4	3.44671	4.65943	4.01899	1.21300
73	2	7	5	3.57086	4.21850	3.74697	0.64760
74	2	7	5	3.44348	4.37651	3.83360	0.93300
75	2	7	5	3.42119	4.53570	3.78815	1.11500
76	2	8	1	3.06042	3.21640	3.12054	0.15600
77	2	8	1	3.37993	3.65854	3.52420	0.27860
78	2	8	1	3.08734	3.22222	3.17444	0.13490
79	2	8	2	2.64532	3.03826	2.70072	0.39290
80	2	8	2	2.86770	3.37802	3.15040	0.51030
81	2	8	2	2.99466	3.13082	3.02401	0.13620
82	2	8	3	1.75458	3.04893	2.59626	1.29400
83	2	8	3	1.69079	2.56000	2.02778	0.86920
84	2	8	3	2.05008	2.94081	2.39506	0.89070
85	2	8	4	2.52830	3.60777	3.13107	1.07900
86	2	8	4	2.83135	3.52031	3.05189	0.68900
87	2	8	4	2.83135	3.52031	3.05189	0.68900
88	2	8	5	1.98298	3.54341	3.06565	1.56000
89	2	8	5	2.59008	2.83061	2.67268	0.24050
90	2	8	5	2.58764	3.06715	2.81497	0.47950

Appendix G

Examples of Praat Pitch Windows for Individual Tones

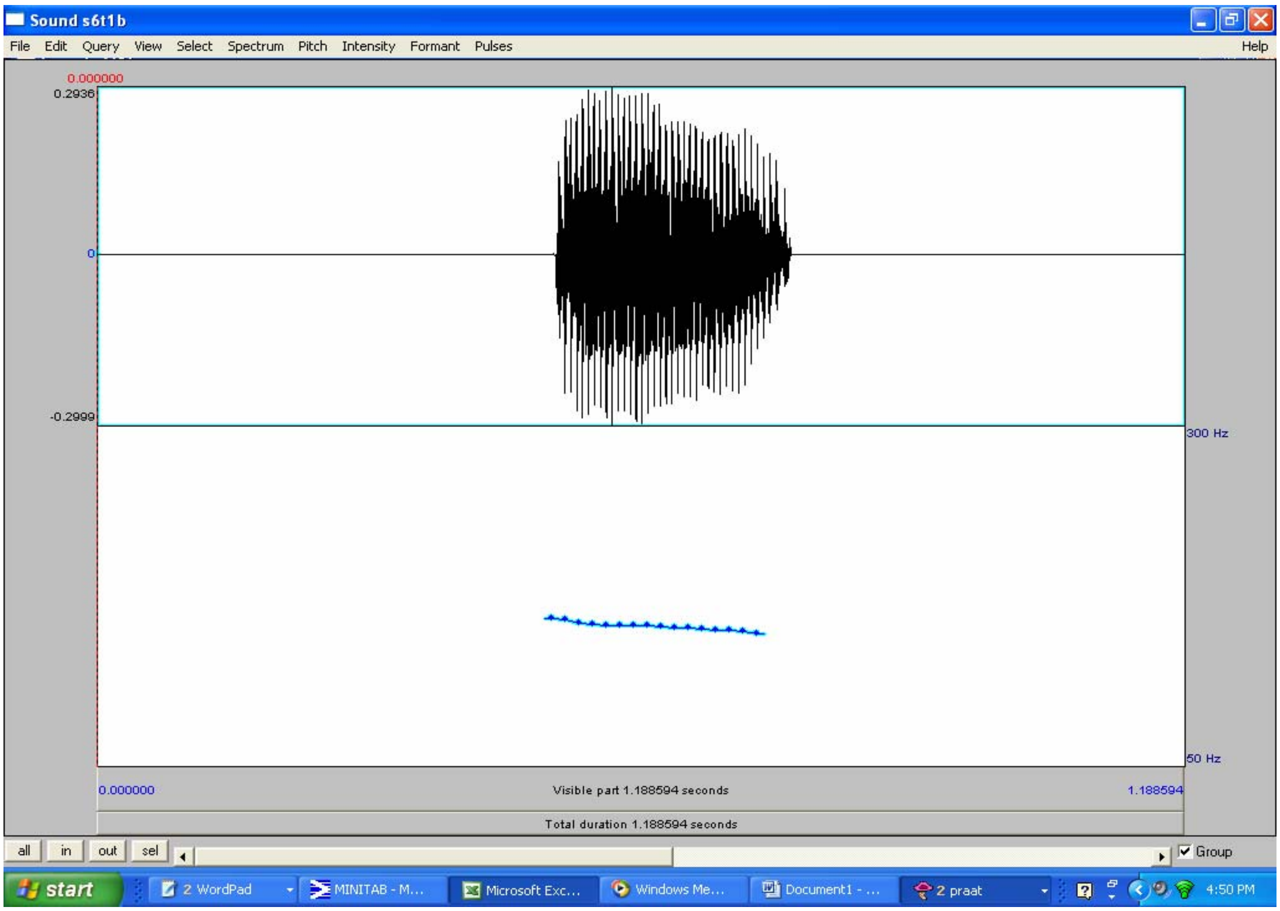
Participant 3, tone 4, sample 2 out of 3



Appendix G

Examples of Praat Pitch Windows for Individual Tones

Participant 6, tone 1, sample 2 out of 3



Appendix H

Sentences from Reading Portion of Presentation

1
我叫麻馬練習
我教麻馬練習
我叫麻馬聯絡
我教麻馬聯絡

我叫媽媽練習
我教媽媽練習
我叫媽媽聯絡
我教媽媽聯絡

我叫馬馬練習
我教馬馬練習
我叫馬馬聯絡
我教馬馬聯絡

2
我叫媽媽練習
我教媽媽練習
我叫媽媽聯絡
我教媽媽聯絡

我叫罵麻練習
我教罵麻練習
我叫罵麻聯絡
我教罵麻聯絡

我叫馬罵練習
我教馬罵練習
我叫馬罵聯絡
我教馬罵聯絡

3
我叫媽媽練習
我教媽媽練習
我叫媽媽聯絡
我教媽媽聯絡

我叫罵媽練習
我教罵媽練習
我叫罵媽聯絡
我教罵媽聯絡

我叫麻麻練習
我教麻麻練習
我叫麻麻聯絡
我教麻麻聯絡

4
我叫麻麻練習
我教麻麻練習
我叫麻麻聯絡
我教麻麻聯絡

我叫麻罵練習
我教麻罵練習
我叫麻罵聯絡
我教麻罵聯絡

我叫麻罵練習
我教麻罵練習
我叫麻罵聯絡
我教麻罵聯絡

5
我叫麻媽練習
我教麻媽練習
我叫麻媽聯絡
我教麻媽聯絡

我叫罵罵練習
我教罵罵練習
我叫罵罵聯絡
我教罵罵聯絡

我叫馬媽練習
我教馬媽練習
我叫馬媽聯絡
我教馬媽聯絡

6
我叫馬罵練習
我教馬罵練習
我叫馬罵聯絡
我教馬罵聯絡

我叫麻麻練習
我教麻麻練習
我叫麻麻聯絡
我教麻麻聯絡

我叫罵馬練習
我教罵馬練習
我叫罵馬聯絡
我教罵馬聯絡

7
我叫麻媽練習
我教麻媽練習
我叫麻媽聯絡
我教麻媽聯絡

我叫罵罵練習
我教罵罵練習
我叫罵罵聯絡
我教罵罵聯絡

我叫馬媽練習
我教馬媽練習
我叫馬媽聯絡
我教馬媽聯絡

8
我叫馬罵練習
我教馬罵練習
我叫馬罵聯絡
我教馬罵聯絡

我叫麻麻練習
我教麻麻練習
我叫麻麻聯絡
我教麻麻聯絡

我叫罵馬練習
我教罵馬練習
我叫罵馬聯絡
我教罵馬聯絡

9
我叫罵麻練習
我教罵麻練習
我叫罵麻聯絡
我教罵麻聯絡

我叫媽麻練習
我教媽麻練習
我叫媽麻聯絡
我教媽麻聯絡

我叫罵罵練習
我教罵罵練習
我叫罵罵聯絡
我教罵罵聯絡

10
我叫媽罵練習
我教媽罵練習
我叫媽罵聯絡
我教媽罵聯絡

我叫媽罵練習
我教媽罵練習
我叫媽罵聯絡
我教媽罵聯絡

我叫麻馬練習
我教麻馬練習
我叫麻馬聯絡
我教麻馬聯絡

11
我叫馬媽練習
我教馬媽練習
我叫馬媽聯絡
我教馬媽聯絡

我叫麻媽練習
我教麻媽練習
我叫麻媽聯絡
我教麻媽聯絡

我叫罵媽練習
我教罵媽練習
我叫罵媽聯絡
我教罵媽聯絡

12
我叫馬麻練習
我教馬麻練習
我叫馬麻聯絡
我教馬麻聯絡

我叫麻罵練習
我教麻罵練習
我叫麻罵聯絡
我教麻罵聯絡

我叫馬馬練習
我教馬馬練習
我叫馬馬聯絡
我教馬馬聯絡

13
我叫麻馬練習
我教麻馬練習
我叫麻馬聯絡
我教麻馬聯絡

我叫罵罵練習
我教罵罵練習
我叫罵罵聯絡
我教罵罵聯絡

我叫媽罵練習
我教媽罵練習
我叫媽罵聯絡
我教媽罵聯絡

14
我叫馬馬練習
我教馬馬練習
我叫馬馬聯絡
我教馬馬聯絡

我叫麻媽練習
我教麻媽練習
我叫麻媽聯絡
我教麻媽聯絡

我叫馬媽練習
我教馬媽練習
我叫馬媽聯絡
我教馬媽聯絡

15
我叫罵麻練習
我教罵麻練習
我叫罵麻聯絡
我教罵麻聯絡

我叫馬罵練習
我教馬罵練習
我叫馬罵聯絡
我教馬罵聯絡

我叫媽馬練習
我教媽馬練習
我叫媽馬聯絡
我教媽馬聯絡

16
我叫罵媽練習
我教罵媽練習
我叫罵媽聯絡
我教罵媽聯絡

我叫媽麻練習
我教媽麻練習
我叫媽麻聯絡
我教媽麻聯絡

我叫馬麻練習
我教馬麻練習
我叫馬麻聯絡
我教馬麻聯絡

17
我叫媽麻練習
我教媽麻練習
我叫媽麻聯絡
我教媽麻聯絡

我叫馬麻練習
我教馬麻練習
我叫馬麻聯絡
我教馬麻聯絡

我叫媽馬練習
我教媽馬練習
我叫媽馬聯絡
我教媽馬聯絡

Appendix I

Participant Demographic Information Table: Parent Languages

Participant	Father's Birthplace	Mother's Birthplace	Languages or dialects spoken by father	Languages or dialects spoken by mother
1	Taiwan Jidong	Taiwan Taibei	*Guoyu* *Taiyu*	*Guoyu* *Taiyu* Japanese
2	Taiwan Jiayi	Taiwan Jiayi	*Guoyu* *Taiyu*	*Guoyu* *Taiyu*
3	Taiwan Jiayi	Taiwan Jiayi	*Guoyu* *Taiyu*	*Guoyu* *Taiyu*
4	China Linyi	China Shandong Province	*Putonghua* *Shandonghua*	*Putonghua* *Shandonghua*
5	China Tianjin	China Wuhan	*Putonghua* *Tianjinhua*	*Putonghua*
6	China Tianjin	China Beijing	*Putonghua*	*Putonghua*
7	China Sichuan Province	China Liaoning	*Putonghua* *Sichuanhua*	*Putonghua* *Sichuanhua*
8	China Beijing	China Beijing	*Putonghua* *Chaozhouhua*	*Putonghua*