STRESS PRODUCTION BY MANDARIN SPEAKING LEARNERS OF ENGLISH

by

George Alston Corley

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Linguistics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2019

Date of final oral examination: January 8, 2019

The dissertation is approved by the following members of the Final Oral Committee:
Eric Raimy, Professor, Language Sciences and English
Joseph C. Salmons, Professor, Language Sciences
Thomas C. Purnell, Professor, English
Rajiv G. Rao, Associate Professor, Spanish and Portuguese

To Li, Maria, and Phoebe.

Acknowledgments

I would like to thank everyone who supported me and participated in my efforts to complete my PhD. Thank you to my advisor, Eric Raimy, who has guided me through this long and arduous jouney. Thank you to my committee members: Joe Salmons, who pushed hard on me into a final revision that (I hope) improves the product and will lead me to better understanding; Tom Purnell, who introduced me to phonetics; Bei Yang, who helped me with planning for the portions of this dissertations involving Chinese; and Rajiv Rao, whose suggestions at the defense will lead me on further journeys on this topic.

I would also like to thank the attendee of the 26th Annual Conference of International Association of Chinese Linguistics who suggested Growth Curve Analysis to me (it embarrasses me that I did not get his name for proper credit), and to other attendees who gave insights into my work.

I also want to thank my family. Thank you to my parents, who raised me well and supported me, especially my mother, who took a few very concerning phone calls from me. And to my brother Burr and sisters Teresa and Laura who continually encouraged me, at least one of whom will likely be reading this page soon. But most of all my wife Li, who suffered through it all with me and still pushed me to complete this work, and my daugther Maria, who has been as understanding of Daddy's work schedule as I could expect.

All of these people have contributed to the success of my completing my dissertation and earning the title Doctor of Philosophy. I credit them for helping me make this work as good as I can make it. All flaws and errors in this paper are entirely my own.

Contents

List of Tables

List of Figures

Abstract

This dissertation explores the production of the phonetic correlates of stress in English by L2 speakers of Mandarin Chinese. English uses pitch, loudness, and vowel duration to mark stress (Fry, 1958). Chinese uses duration and loudness as well as an expansion of the pitch range (Duanmu, 2000; Chao, 1968). I examine in this paper whether Chinese speakers' production of these correlates are influence by their native stress or tonal systems.

These hypotheses are explored through an experimental phonetics task. Native English speakers and Mandarin speakers who learned English as a second language are asked to read a list of words using a carrier phrase. The wordlist has been constructed with disyllabic words with initial or final stress.

A pilot test with 19 subjects (10 Mandarin speakers, 9 English speakers) found that Mandarin speakers do use duration and loudness (measured as intensity) to mark stress. Mandarin speakers also did have similar pitch values to English speakers. A second experiment sought to confirm these initial findings with a larger subject pool (30 Mandarin speakers and 30 English speakers) and modifications to the protocol. The third and final experiment uses data from the 30 Mandarin speakers in the second experiment to compare their own English production to their Mandarin production.

The study found that Mandarin speakers use intensity and duration in a manner comparable to English speakers, with duration having a small difference in that Mandarin speakers have more lengthening of final syllables. When pitch measurements were taken, it appeared that Mandarin speakers had a tendency to have an upward curve in initial stressed syllables, where English speakers merely had a shallower fall. Comparison of curves with Mandarin tones did not, however, provide good evidence for the influence of tonal categories on stress. I concluded that Mandarin speakers likely do have some influence from their native stress system, but influence from the tonal system is not supported by current evidence.

1 Introduction

English stress is largely marked by differences in pitch, loudness, and vowel duration (Fry, 1958; Lieberman, 1960; Bolinger, 1961). Pitch is weighted particularly heavily (Fry, 1958; Chrabaszcz et al., 2014). This state of affairs raises interesting questions for second language speakers of English whose first language is a tonal language. Since tonal languages use pitch as a correlate of tone, the question arises as to how speakers would interpret and implement English stress, especially with regard to the use of pitch.

Mandarin Chinese has four tones, a high level tone (first), a rising tone (second), a low or low-rising tone (third), and a falling tone (fourth). Standard Mandarin also can be analyzed as having a stress system, with toneless or ”neutral tone” syllables, which are unstressed and have lead to proposals for a stress system (e.g. Duanmu, 2007). The correlates of stress in Mandarin include duration and intensity, but pitch primarily indexes lexical tone (Lin & Yan, 1980, 1988; Chen & Xu, 2006; Lee & Zee, 2008). There is a body of research of Mandarin speaking learners of English examining both their perception (Qin et al., 2017; Lin et al., 2013; Wang, 2010; Chrabaszcz et al., 2014; Ou, 2010; Zhang & Francis, 2010; Altmann, 2006) and production (Zhang et al., 2008; Altmann, 2006) of English stress. However, most of these previous experiments have simply reported on the there hasn’t been so much attention on precisely what second language principles are involved, or how L1 may be affecting L2 stress perception in this case.

Theories of second language acquisition generally argue that second language categories are in some way influenced by each other. The Perceptual Assimilation Model argues that L2 categories can be assimilated to L1 categories if they are sufficiently similar, or perceived as outside of L1 categories if they are not sufficiently similar (Best, 1995; Best et al., 2001). The argument of the PAM is that discrimination of sounds will depend on how this mapping takes place, with contrasting categories assimilated to the same category with an equal goodness of fit being harder to discriminate than contrasts assimilated to separate L1 categories or outside of the L1 system. The Speech Learning Model argues that native and non-native categories that are similar enough become associated with each other and even influence each other in both directions (Flege, 1995, 2007, 2011). The bottom line in each case is that the phonetic input from L2 can be assimilated to L1, resulting in a transfer of L1 phonological structure to interpret L2. Although these models have focused on segmental categories, when studying stress from the position of studying phonetics this basic insight that phonetic similarity to a native category can induce transfer is useful when studying stress systems from the perspective of phonetics.

Thus, returning to the initial questions, how will speakers of a tonal language like Mandarin interpret English stress. Mandarin speakers may detect the duration and intensity correlates that they use themselves, and index English stress to their own stress. When encountering differences in pitch, Mandarin speakers may map these onto a tonal category, allowing their tonal system to affect their perception of stress. That is, if English speakers' pitch values are pitch contours in stressed or unstressed syllables are similar enough to native tonal categories, they could be mapped onto tonal categories. By studying their production of stress I may be able to gain insights into whether this type of phonological transfer is occurring.

This dissertation outlines a series of three experiments that set out to explore this problem. The first experiment compares stress production between native English speakers and Mandarin-speaking English learners, and the second replicates the first with a larger number of subjects and improved methods. The third experiment compares the Mandarin speakers' English production to their native language tone and stress production with particular focus on pitch contours in order to probe whether a Mandarin tonal categories are being realized in or influencing Mandarin speakers' English production..

According to the results Mandarin speakers do use pitch, duration. and intensity to mark stress in English. The use of duration and intensity does indicate transfer of the native stress system, however comparisons of Mandarin speakers' English production and their Mandarin production did present some issues.

Mandarin speakers also tend to have a flat or slightly rising pitch contour on an initial stressed syllable, where English speakers are almost uniformly falling. Comparison with native tones did not definitively show this contour originating with a tone, however. The initial stressed contour most resembled a Mandarin first tone, but the similarity seems to be superficial. The difference in the slope of the curve, particularly, needs further explanation, as does the tendency to have this upward curve in initial stressed syllables and a fall in final stressed syllables. This leads to the conclusion that stress transfer is likely occurring, but I do not have good evidence to suggest tone has transferred. Further study into the source of these differences in pitch curves will be necessary.

This dissertation is laid out as follows. Chapter 2 gives background information and reviews current literature and Chapter 3 outlines the hypotheses that I chose to test. Chapter 4 outlines the methods used in all experiments, including developments that altered methods between experiments. The experiments themselves are discussed in the next three chapters. In Experiments 1 and 2 (Chapters 5, and 6, respectively), Mandarin speakers' production of English stress was compared to that of native English speakers in order to gain a baseline of how their production differed. In Experiment 3 (Chapter 7), Mandarin speakers' English production was compared to their production of Mandarin stress and tonal contrasts in order to probe what, if anything, transfers from L1 to L2. Chapter 8 is a general discussion of findings, and Chapter 9 gives the conclusion and discusses avenues for future work.

2 Background and Literature Review

2.1 Phonetic Correlates of Stress

Early work on the phonetic correlates of stress in English identified F0, vowel duration, intensity, and vowel quality (Fry, 1958; Lieberman, 1960; Bolinger, 1961). That is to say that English speakers produce stressed syllables with higher F0, longer vowels, greater intensity, and vowel quality, and their perception of stress relies on these cues.

Much of this work has identified vowel quality as an important cue for stress (Cutler, 2008; Rietveld & Koopmans-van Beinum, 1987). This is due to highly productive vowel reduction that will tend to centralize vowels. However, not all unstressed vowels are reduced (Burzio, 2007), and there may be confounds when examining vowel reduction in non-native speakers. Zhang et al. (2008) showed a variety of different results for Mandarin speakers in reducing unstressed vowels, indicating that there may be some effect from Mandarin's vowel system. My study will not examine vowel quality as a correlate of stress, instead choosing to focus on suprasegmental correlates.

Some research has pointed to spectral balance as the primary cue to stress, often using it to separate stress from accent. Sluijter & van Heuven (1996) first show this in Dutch by measuring the intensity of four frequency bands which include the fundamental frequency and the first three formants, respectively. They show that intensity in the three formant bands show significantly more variation than the intensity of the F0 band, and that this measurement was better at identifying stress than overall intensity. Sluijter et al. (1997) manipulated the intensity of the upper bands and showed that increasing their intensity affected subjects' perception of stress. Sluijter & van Heuven (1996) and (Sluijter et al., 1997) also argues that pitch is not a correlate of stress, but of accent. Accent under this theory refers to prominence in higher domains than the word, such as focus marking. Experiments in Sluijter & van Heuven (1996) show that pitch varies by focus condition, while spectral balance remains diagnostic.

Campbell & Beckman (1997) attempted to replicate the results of Sluijter & van Heuven (1996) for English and did find an effect in focus accented syllables, but did not find the same difference between stressed and unstressed syllables in in unaccented words. Okobi (2006) found another method of measuring this spectral tilt by comparing the first harmonic and third formant.

However, I have been unable to find validation of the spectral tilt phenomenon in Mandarin, and the majority of work in the area of second language stress relies on intensity. Spectral tilt also works best when focus accent is factored out, which the current experiments have not been designed to do. Vowel quality will also not be used in the current study, as I want to focus on suprasegmental effects and particularly on pitch contour. As such, this dissertation will be relying instead on mean intensity, alongside pitch, pitch contours, and duration, to examine hypotheses.

2.2 Stress in Mandarin Chinese

In Mandarin Chinese, a disyllabic word may be stressed either on the first syllable or on both syllables. As such, the strongest evidence for word-level stress in Mandarin is not in the stressed syllable, but in the presence of the unstressed syllable. Unstressed syllables, traditionally called *neutral tone* or *fifth tone* syllables, are characterized by a lack of lexical tone, shorter duration, and lower intensity than stressed syllables (Lin & Yan, 1980, 1988; Chen & Xu, 2006; Lee & Zee, 2008). Duanmu (2000, 2007) describes this in terms of syllable weight, with stressed syllables being bimoraic and heavy, while unstressed syllables are monomoraic and light, with subsequent rhyme reduction. Under this framework, the loss of tone in unstressed syllables can be seen as a type of reduction phenomenon, similar to the reduction of vowels to schwa in an unstressed syllable.

Unstressed syllables in Standard Mandarin take on a pitch contour relative to the tone of the preceding syllable. This influence of the preceding tone was noted by Chao (1968), and has been subsequently confirmed (Lee & Zee, 2008; Chen & Xu, 2006). Other studies have found that the neutral tone appears to have a weak mid or mid-low pitch target (Chen & Xu, 2006; Huang, 2012), which can still be consistent with the syllable being phonologically toneless, with said mid target merely being a default phonetic position, particularly considering that Chen & Xu (2006) found that this target is approached gradually in successive neutral tone syllables. There also exists some variation in the degree of tone reduction, as well as dialect variation in terms of how often a tone is deleted, with Taiwan Mandarin in particular deleting fewer tones than Beijing Mandarin (Huang, 2012).

Certain types of syllables are reliably unstressed. Most commonly, functional items, such as 了 *le* and 的 *de* are always unstressed and toneless, just as functional elements in English are typically unstressed and reduced (i.e. *to*, *the*). Within words, certain structures will reliably produce a stressed-unstressed combination. For instance, reduplicated nouns such as 姐姐 *jiějie*, 妈妈 *māma*,and 爸爸 *bàba* will very reliably have an unstressed final syllable (Duanmu, 2007). Also, semantically bleached suffixes, such as the -子 *-zi* in 鸭子 *yāzi* and 瓶子 *píngzi* will be reliably unstressed (Duanmu, 2007).

The distinction primarily comes in compounds. Duanmu (2000, 2007) describes compounds where the second syllable is unstressed and compounds where the both syllables are stressed, pointing to stress minimal pairs like 大意, where the variant with two full tones *dàyì* can mean ”big idea”, while a neutralized tone *dàyi* can change the meaning to ”careless” (Duanmu, 2007, p. 129)

Arguments against word level stress in Mandarin tend to be informed by a kind of theoretical intertia. Historically, linguists have presented stress languages and tone languages as entirely separate categories, with an intermediate ”pitch accent” category bridging the gap. Mandarin, with its complex contour tone system, would appear to be very far from stress languages indeed, and thus it is often categorized as a ”non-accentual” tonal language. Altmann (2006) categorizes Mandarin this way, however, her experimental data shows Mandarin behaving like a language with unpredictable stress. Hyman (2009) has criticized this framework, arguing that pitch accent itself is not a useful category, and that it is possible for a language to have both stress and tone. Purnell (1997) showed that tone systems can interact with metrical systems in ways that indicate that stress and tone are not incompatible.

Stress has been observed as a feature of Mandarin since at least Chao (1933). Although Chao (1933) states that stress is ”not important” in Chinese, he does describe neutral tone syllables as ”unstressed”, and describes vowel reduction (vowels are ”obscured”) as part of them.

While the neutral tone is clear evidence of stress, it is much more difficult to justify the primary and secondary stress posited in Duanmu (2007). Chinese speakers do not seem to perceive consistent stress differences between full tone syllables, or when they do report them they seem to be entangled with tone categories. Duanmu himself states that stress in Chinese is ”hard to hear” Duanmu (2007).

It is possible that the structure of the Chinese metrical system owes much to its morphology. Huang (2018) points out that reduplicated nouns show tone neutralization, even superceding third tone sandhi, while reduplicated verbs do not show the same pattern. She uses this as evidence for a prosodic word in Mandarin, separating reduplicated verbs into two prosodic words, while reduplicated nouns are within a single prosodic word, implying that the prosodic word is the domain for tone neutralization.

For the purposes of this dissertation, I will only consider the stress difference between full tone and neutral tone syllables, taking full tone syllables as stressed and neutral tone syllables as unstressed. As this is the clearest stress contrast in the language, and my English data focuses also on distinguishing primary stress from unstressed syllables, this should be a sufficient comparison for the purposes of the current work. Further work should be done to explore the potential reality of more levels of stress in Mandarin Chinese.

2.3 Second Language Stress

Language acquisition is a process of incorporating input and organizing it into a grammatical system. In second language acquisition, this organization is influenced by the existing language system that the learner has acquired previously as a child. Theories of second language acquisition thus have focused on how the structure of L1 influences the acquisition, perception, and production of L2 contrasts.

Major theories including the Perceptual Assimilation Model (Best, 1995), and the Speech Learning Model (Flege, 1995, 2007, 2011) have focused primarily on segmental processes, discussing how L2 segments are perceived and integrated into a speakers' linguistic system.

The Perceptual Assimilation Model Best (1995) argues that speakers perceive non-native categories differently depending on similarity to the native phonological system. (Best, 1995) argues that contrasts should be distinguished by L2 speakers when each member of that contrast can be mapped to an L1 category, but also that they can be distinguished when they are sufficiently far from L1 category (including distinctions wholly outside of the L1 space).

The Speech Learning Model Flege (1995, 2007, 2011) argues that L1 and L2 categories accommodate to each other on the basis of similarity. The phonetic realization of a segment may be affected by similar segments from the other language. This influence is bi-directional, and Flege uses evidence of this bi-directionality to argue against a critical period for L2 learning. The insights of the Speech Learning Model are of some interest to the current study, as it makes predictions about production. If L2 is stress being associated with L1 stress, we should see the L2 use of stress correlates accomodate to L1. If it is being associated with L1 tone, we should see accomodation of L2 stress categories to L1 tone categories.

Theories of L2 stress, by contrast, tend to focus on the metrical system. Peperkamp & Dupoux (2001) and Peperkamp (2004) developed a theory of stress 'deafness' which seeks to determine how a speakers' native language affects their ability to perceive stress. In order to do this, Peperkamp & Dupoux (2001) organized languages with predictable, non-morphologically conditioned stress systems into four types depending on what knowledge children must have (e.g. recognition of functional elements, word boundaries) in order to acquire the regular rules of the stress system. English falls outside of the Peperkamp & Dupoux (2001) typology, as its stress system includes lexical variations and children must memorize different stress patterns for different words. Peperkamp & Dupoux (2001) and Peperkamp (2004) predict that speakers of languages like English are certain to be able to perceive stress. We cannot be ”stress deaf” according to this theory, because we must perceive stress in order to understand our own language. If we take the account of Chinese stress from Duanmu (2000, 2007), as described in section 2.2, we should conclude that Mandarin Chinese also falls outside of the Peperkamp and Dupoux typology as a system with lexical stress.

While the theory developed in Peperkamp & Dupoux (2001) and Peperkamp (2004) is well-designed to predict stress deafness and will have implications for production of stress placement, it makes no prediction of how languages with lexical stress behave, including what other parameters may be involved. Crucially for this paper, Peperkamp & Dupoux (2001) also makes no predictions about perception or production of the phonetic correlates of stress. However, the question of phonetic correlates of stress is crucial, as L2 speakers critically must recognize the phonetic correlates of stress in order to perceive it in the physical sound wave from a native speaker, and they also must be able to reproduce those phonetic correlates to some satisfactory degree in order to be understood by native speakers.

Archibald (1993) uses a principles and parameters approach, parameterizing the metrical systems of the languages under study and studying errors made by L2 speakers and relating them either to L1 parameter settings or to default parameter settings. His argument is that L2 speakers, when building a system that differs from L1 speakers, may substitute either their own grammatical settings or construct the system from UG. Archibald (1997) investigated this with Mandarin and Cantonese speakers and concluded that they memorized stress positions in the lexicon, however his sample size was extremely small, and he operated from the assumption that Mandarin does not have stress.

2.4 Previous Experiments

Much of the work on Mandarin speakers and English stress systems has been perceptual. Generally speaking, it has been observed that they tend to use vowel reduction, duration, intensity, and pitch in ways that are similar to English speakers.

Chrabaszcz et al. (2014) used manipulated stimuli to test English, Mandarin, and Russian speakers on the phonetic cues they use to perceive English stress. They found that Mandarin speakers largely seemed to follow the same hierarchy of cues that English speakers did (vowel quality > pitch > intensity > duration). This is in contrast to Russian speakers who used intensity and duration more that pitch. Chrabaszcz et al. (2014) do note that the high ranking of vowel quality for Mandarin speakers could come from them simply seeing the reduction of /a/ to [ə] as a segmental contrast and not a correlate of stress.

There have been indications that Mandarin speaker rely more on pitch than English speakers do Wang (2010); Ou (2010). Ou (2010) changed the F0 condition by having a native speaker record nonce words in declarative and question intonation. In declarative sentences in English, stressed syllables receive a high pitch accent, while in questions they receive a marked low pitch accent. This way, it is possible to obtain naturally produced English words where the F0 is lower in the stressed syllable than in an unstressed syllable. Taiwanese students were trained to associate these words to pictures and then asked to correctly identify the words. Both high proficiency and low proficiency students had more difficulty distinguishing a stress minimal pair (*fércept* vs *fercépt*) with a low pitch accent, indicating that they rely more heavily on pitch to detect stress than English speakers. This effect was later confirmed using real English words on a similar subject pool, using an ABX task and an identification task (Ou, 2016).

There exist other differences in perception that bear mentioning. Zhang & Francis (2010) found that while Mandarin speakers do use vowel reduction as a cue to stress, but they appear to connect it with lower pitch, while English speakers use vowel reduction as an independent cue. Qin et al. (2017) found that Taiwan Mandarin speakers do not use duration as much as mainland Standard Mandarin speakers in perceiving stress. The authors attribute this to Taiwan Mandarin lacking a stress distinction, as in Taiwan Mandarin duration does not vary for neutral tone syllables, and all neutral tones tend to be a low pitch, rather than varying based on the previous tone (Huang, 2012). At the least, it can be argued that the correlates of stress being different in Taiwan Mandarin causes them to react differently than Mainland Mandarin speakers.

Mandarin speakers are relatively good at discriminating stress in English. Lin et al. (2013) found that Mandarin speakers were better than Korean speakers in perceiving and retaining English stress contrasts. Lin et al. (2013) posit that this is due to Mandarin itself having a lexical stress system, where Korean does not, although they do acknowledge that some of the correlates of stress are also used for differentiating tones in Mandarin.

On the production side, Zhang et al. (2008) examined F0, duration, intensity, and vowel quality for American English speakers and Mandarin Chinese speakers. Mandarin speakers were shown to use all four of these correlates to signal stress, but with some significantly different patterns. Notably, Zhang et al. (2008) found that, for Mandarin speakers, the peak F0 occurred later in the syllable in stressed syllables than in unstressed syllables, which suggests a difference in the pitch contour between English speakers and Mandarin speakers. This finding was my initial clue to look at pitch contours of native Mandarin speakers producing English, as having a difference in pitch contour could indicate some transfer of a tonal system.

Another mild case of evidence for tonal tranfer came from, Cheng (1968), which found that an unstressed syllable in English can trigger third tone sandhi when the word is inserted into a Chinese sentence, as in the example *hǎo professor bù duō*[1], where third tone *hǎo* (好) can become second tone *háo*. This triggering of third tone sandhi could indicate Mandarin speakers perceiving unstressed syllables as a third (low) tone, though its occurrence in Mandarin with ad-hoc loans does not necessarily reflect how speakers perceive or encode the words when speaking in English.

[1]”there are few good *professors*”, lit. ”good professor not many”, where the English word *professor*, a word with an initial unstressed syllable, was inserted. Cheng (1968) compared this with *hǎo student bù duō*, substituting *student*, with an initial stressed syllable, as the inserted word

3 Hypotheses

In order to interpret the experiments described in Chapters 5, 6, and 7, it is necessary to have well-developed and testable hypotheses to compare to the data. For my purposes, I have focused on hypotheses about L1 transfer. Although there exist potential for other kinds of interlanguage effects, I wished to focus on transfer initially, in order to give some focus to my work. In this chapter, I will list the hypotheses I wish to test in my experiments and discuss what phonetic data we should expect to see in each case. The hypotheses considered are listed in (1).

(1)
a. Hypothesis 1: Mandarin speakers associate L2 English stress categories with their L1 stress categories.
b. Hypothesis 2: Mandarin speakers associate L2 English stress categories with their L1 tonal categories.
c. Hypothesis 3: Mandarin speakers associate L2 English stress categories with L1 stress categories and also with tonal categories.

Each of these hypotheses makes different predictions about how Mandarin speakers will produce stress contrasts in English, based on what from their native language is transferred in order to encode English stress. I do not presume that these are the only possibilities. It is entirely possible that none of these hypotheses are true, and that any differences between native

English speakers and Mandarin speaking learners are attributable to other phenomena.

3.1 Hypothesis 1: Stress Transfer

Under this hypothesis, Mandarin speakers correctly perceive the stress contrast in English and relate it to their own stress system. Under this hypothesis, we expect Mandarin speakers to mark stress using duration and intensity, and their realization of these correlates in English stresses should be related to the way they realize them in Mandarin. We should also see unstressed syllables be comparable in duration and intensity to neutral tone syllables. This scenario makes no particular prediction on pitch values or pitch contours

3.2 Hypothesis 2: Tone Transfer

Under this hypothesis, Mandarin speakers perceive the higher pitch of stressed syllables in English as a tonal distinction. If this hypothesis were true, we would expect Mandarin speakers to primarily use pitch to distinguish stressed and unstressed syllables in English.

There are two possible scenarios if tone is transferred. The first is that Mandarin speakers simply hear stressed vs unstressed as a high vs low distinction and apply pitch accordingly. In this case, the pitch contours of Mandarin speakers' English production may not be particularly related the contours of their native tonal categories, but they may still differ from English speakers' contours.

The second scenario is that Mandarin speakers perceive the stress contrast according to their own tone contours. In this scenario, we expect that in Experiment 2, the contours of stressed (and possibly unstressed syllables) to to show some relationship to native tones. The tone that is selected would probably one that is similar to native speaker's pitch curve production in the same environment. It is most likely that tone 1, 2, or 4 would would influence the stressed syllable, while tone 3 is more likely influence an unstressed syllable. Such an association should be visible with a comparison of pitch contours. The observation of Cheng (1968) that English unstressed syllables can trigger third tone sandhi when inserted into Mandarin sentences might seem to support this, as it would be evidence of at least some speakers associating an unstressed syllable with Tone 3.

3.3 Hypothesis 3: Stress and Tone in Interaction

This hypothesis is a hybrid of hypotheses 1 and 2. It assumes that Mandarin speakers associate English stress categories with their own categories, but also have a secondary tonal association. This would mean that some association would be found between stressed syllables and one of the lexical tone categories.

The expected results here would combine the expectations of hypotheses

1 and 2. First, we should expect Mandarin speakers to use duration and intensity to mark stress, with the particular use of these correlates in English showing some influence from L1.

Additionally, we should expect one of the two scenarios mentioned under hypothesis 2 to be in effect, with the modification that stressed syllables would be influence by a tonal category, but unstressed syllables would resemble neutral tone contours. As discussed earlier, if contour tones are assigned, English stressed syllables should show some influence from first, second, or fourth tones, while if a simple high/low distinction is perceived, stressed syllables will receive a high tone. The important difference between expectations for hypotheses 2 and 3 is that under hypothesis 3, unstressed syllables will resemble a neutral tone, and in second position may have a neutral tone contour influenced by the contour of the stressed syllable.

4 General Methods

This chapter reviews the general methods used in all three experiments, as well as covering the major differences between them. Experiment 1 compared stress production of Mandarin speaking learners with that of native English speakers. Experiment 2 was an iteration on Experiment 1, improving on its testing materials and recruiting more subjects. Experiment 3 uses the Mandarin speaker subject pool from Experiment 2 to do a comparison of Mandarin speakers' English production with their own native Mandarin production.

4.1 Experimental Stimuli

In both experiments, subjects were presented with a list of of real English words to be read in a carrier phrase. Real words were chosen to provide a more natural test and avoid influences that might occur from pre-training or presentation. Experiments that use nonce words typically train participants on stress, meaning that the linguistic quality under study is known to participants (e.g Jangjamras, 2011; Altmann, 2006). Presentation of nonce words is also a challenge, given the inconsistency and lack of stress marking in English orthography, resulting in researchers needing to present stimuli auditorily (e.g. Jangjamras, 2011) or use an alternative spelling system that must be taught to participants (e.g. Altmann, 2006).

Experiment 1 Stimuli

The words for the first experiment were taken from the first 5,000 most common nouns in the Corpus of Contemporary American English (Davies, 2008). A total of 18 disyllabic nouns were chosen, 9 with initial stress and 9 with final stress. The nouns chosen for the experiment are shown in Table 4.1. Subjects were presented a list of words including these nouns and 18 monosyllabic distractors and asked to read each word in the carrier phrase *Please say the ___ again*.

Table 4.1: Stimuli used for Experiment 1.

Initial Stress	Final Stress
data	abuse
justice	approach
method	arrest
mission	attack
office	debate
purpose	device
season	disease
status	dispute
weapon	support

The experimental procedure for the first experiment presented several problems that lead to revision of the protocol for subsequent experiments. Vowel-initial and vowel-final words (such as *abuse* and *data*) combined with the carrier phrase having vowels adjacent to the target word made segmenting the words of some subjects difficult. Secondly, because the target word was near the end of the utterance, many of the English speaking subjects produced creaky phonation on the target vowels. Not controlling the particular vowels may have introduced biases in the duration data, particularly since diphthongs were present in stressed syllables. Finally, the presence of sonorant codas required me to adjust my data analysis methods (see section 4.3). These issues were addressed by using a different method of choosing stimuli in Experiments 2 and 3.

Experiment 2 and 3 Stimuli

The wordlists for the second and third experiments were taken from the SUBTLEX-US (Brysbaert & New, 2009) and SUBTLEX-CH (Cai & Brysbaert, 2010) databases for English and Chinese, respectively. Both of these databases are drawn from movie subtitles in their respective languages, meaning that they should be comparable to each other and have word frequencies that are reasonably similar to the spoken form of the language.

In order to make controlling for target vowel easier, English words were selected to match five frames which placed vowels between similar consonants.The consonants selected were all obstruents to aid in segmentation. Three frames use peripheral monophthongs; /s**ɪ**s/, /t**ɛ**k/, and /s**u**t/; while the remaining two use the central monophthong ʌ/ə, /k**ʌ**s/ and /dʒ**ʌ**s/. These frames were found in four combinations of stress (stressed and unstressed) and position (first or second syllable) in disyllabic words. The words selected are presented in Table 4.2.

Working with real words derived from corpora led to some compromises in the stimulus list. Not all vowels in English could be represented this way,

Table 4.2: English target words for experiments 2 and 3.

Frame	first stressed	second stressed	first unstressed	second unstressed
sɪs	**sis**ter	in**sist**, ass**sist**	–	ba**sis**
tɛk	**Tex**as	pro**tect**	**tech**nique	vor**tex**
sut	**suit**case	per**suit**	**Suz**anne	i**ssues**
kʌs	**cus**tom	dis**cus**	**cas**sette	fo**cus**
dʒʌs	**jus**tice	a**djust**	***Jus****tine****	gor**geous**

as it was not possible to find enough words with a particular frame in a particular position. Of the selected vowels, the frame /sɪs/ does lack an example in first position unstressed – two words with the frame in final stressed position were chosen to keep the overall number of tokens accurate. The inclusion of /sut/ required some compromise in getting precisely the same frame, accepting /suz/ and /ʃuz/ as similar enough to include in the current experiment.

Because of a typographical error, *Justine* was printed as *Justin* on one of the printed stimuli sheets. This led to *Justine* tokens being excluded from the final analysis, as upon inspection it was difficult to determine if some subjects intended *Justin* or *Justine*.

The Chinese tokens for the third experiment were selected through a similar process. In the case of Chinese, CVC frames were not feasible partly because Mandarin Chinese lacks obstruent codas entirely, but I was able to target particular syllable shapes. CV syllables were selected with obstruent onsets and monophthongs. The target syllables for Chinese were /pa/ /fu/ /tɕi/ and /tʂə/. In order to get comparisons for tones, disyllabic words were chosen with these syllables in each of the four lexical tones in both initial and final positions, as seen in Table 4.3. As the pitch curves of Mandarin tones can be affected be preceeding and following tones (Xu, 1997), the tone of the non-target syllable was controlled by selecting forth tone in non-target positions, with some exceptions noted below. This was done to prevent an overly long wordlist and potential subject fatigue, as including all tone combinations had the potential to create 64 combinations of words, and the carrier phrase would have to be modified to include different lexical tones before the first syllable, increasing the number of combinations to 192. Each of these combinations would need to be read three times, in addition to the tokens read for English and for neutral tone syllables.

Since the variation in neutral tone is much more dramatic (Chen & Xu, 2006), neutral tones were found after all four full tones where possible. The wordlist also only contains word-final neutral tones, as the neutral tone does not occur before a full tone within the same word. Neutral tone words used in the experiment are presented in Table 4.4.

Table 4.3: Full tone words for experiment 3

Tone	Position	/pa/	/fu/	/tɕi/	/tʂə/
T1	first	八卦 **bā**guà	夫妇 **fū**fù	机会 **jī**huì	遮住 **zhē**zhù
	second	第八 dì**bā**	懦夫 nuò**fū**	刺激 cì**jī**	唱歌 chàng**gē**
T2	first	拔掉 **bá**diào	服装 **fú**zhuāng	嫉妒 **jí**dù	折断 **zhé**duàn
	second	提拔 tí**bá**	幸福 xìng**fú**	立即 lì**jí**	挫折 cuò**zhé**
T3	first	把戏 **bǎ**xì	腐败 **fǔ**bài	脊柱 **jǐ**zhù	褶皱 **zhě**zhòu
	second	个把 gè**bǎ**	政府 zhèng**fǔ**	自己 zì**jǐ**	记者 jì**zhě**
T4	first	霸占 **bà**zhàn	附近 **fù**jìn	继续 **jì**xù	这样 **zhè**yàng
	second	大坝 dà**bà**	报复 bào**fù**	忘记 wàng**jì**	–

As with the English wordlist for experiments 2 and 3, some compromises

Table 4.4: Neutral tone words for experiment 3

Position	/ba/	/fu/	/tɕi/	/tʂə/
After T1	结巴 jiē**ba**	衣服 yī**fu**	书记 shū**ji**	接着 jiē**zhe**
After T2	篱笆 lí**ba**	姨夫 yí**fu**	逻辑 luó**ji**	随着 suí**zhe**
After T3	尾巴 wěi**ba**	寡妇 guǎ**fu**	–	有着 yǒu**zhe**
After T4	爸爸 bà**ba**	丈夫 zhàng**fu**	痢疾 lì**ji**	顺着 shùn**zhe**

were made in the stimulus criteria. I was unable to find a first tone or fourth tone /tʂə/ in second position, opting to replace it with /kə/ for first tone, though I was unable to find any replacement for forth tone final /tʂə/. I was also unable to find a /tɕi/ in neutral tone after a third tone syllable. There were two cases among the full tone target words where I was unable to find a fourth tone syllable for the non-target tone. For 服装 fúzhuāng, I was able to find a first tone syllable following the target second tone /fu/, which should have similar effects on the tone of the preceding syllable to a fourth tone, as both tones begin high. For 提拔 tíbá I was only able to find a second tone preceding a second tone /pa/. Finally, while I was able to separate the two syllables with obstruents in most cases, I did use 这样 zhèyàng, which has a /j/ between the two syllables, meaning this word in particular must be segmented carefully. These variations should only have small effects on the data collected.

In experiments 2 and 3, the English wordlist was presented on three sheets of paper, each with the words in a different random order. All subjects saw the same three sheets, but the order of the sheets was shuffled. To avoid problems with creaky voicing among English subjects, the carrier phrase was changed to *It was ___ that I said*, which moves the target word further toward the left of the phrase, reducing the chance of English speakers' vocal fry affecting it. The Chinese carrier phrase for experiment 3 was selected to be similar in structure to the English phrase: 你知道 ___ 是我说的 (*nǐ zhīdào ___ shì wǒ shuō de.*[1]).

4.2 Recording

All subjects were recorded in the Phonetics Laboratory on the second floor of Van Hise Hall at the University of Wisconsin-Madison. Experiment 1 subjects were recorded on a Plantronics headset mic. Subjects for experiments 2 and 3 were recorded using a Blue Snowball microphone. Due to a technical error, some of the subjects in experiments 2 and 3 were recorded on the internal microphone of the Mac used in the experiment. These tokens were kept, as I found no serious effect on the analysis. While the microphone did have an effect intensity measurements, including the effect did not appear to change the analysis in any meaningful way.

4.3 Analysis

Preparation and Measurement

TextGrids were created for English language tokens using the Penn Forced Aligner (Rosenfelder et al., 2011), and for Chinese language tokens using the

[1]Literally: "You know it was ___ that I said."

Mandarin Chinese version of the Penn Forced Aligner (Yuan et al., 2014) with alignment of relevant segments later hand-checked for accuracy. Figure 4.1 shows a demonstration of this interval correction, with the top image showing the uncorrected force-aligned token given by the Penn Forced Aligner, and the bottom showing my correction of the token. The English file for one subject in experiments 2 and 3 (18C04) had to be aligned entirely by hand due to a technical problem with the FAVE python script.

Target words were then extracted and analyzed in Praat (Boersma & Weenink, 2015). Settings for formants and pitch were hand checked per token, and pitch tracks were reviewed to correct halving and doubling errors, using Praat's Pitch object. Figure 4.2 shows an example of such pitch correction. The top image shows an unnaturally sharp change in pitch, which was corrected in the bottom image by shifting the values up one octave, verified by hand measurement of the difference between pulses. There were also sharp divides and erratic pitch measurements caused by creaky voicing, which were kept after verification using a hand-measurement. In other cases, Praat did not give a suggested pitch and the pitch curve was impossible to accurately reconstruct for the appropriate token, and the pitch curve had to be left blank for that portion, as seen in Figure 4.3, which generally would result in exclusion of that token from pitch analysis.

In experiment 1, measurements of four values were taken: duration of the vowel, duration of the sonorant portion of the rhyme, mean intensity, and maximum pitch. Pitch contours were also captured for each syllable.

Figure 4.1: Demonstration of alignment correction

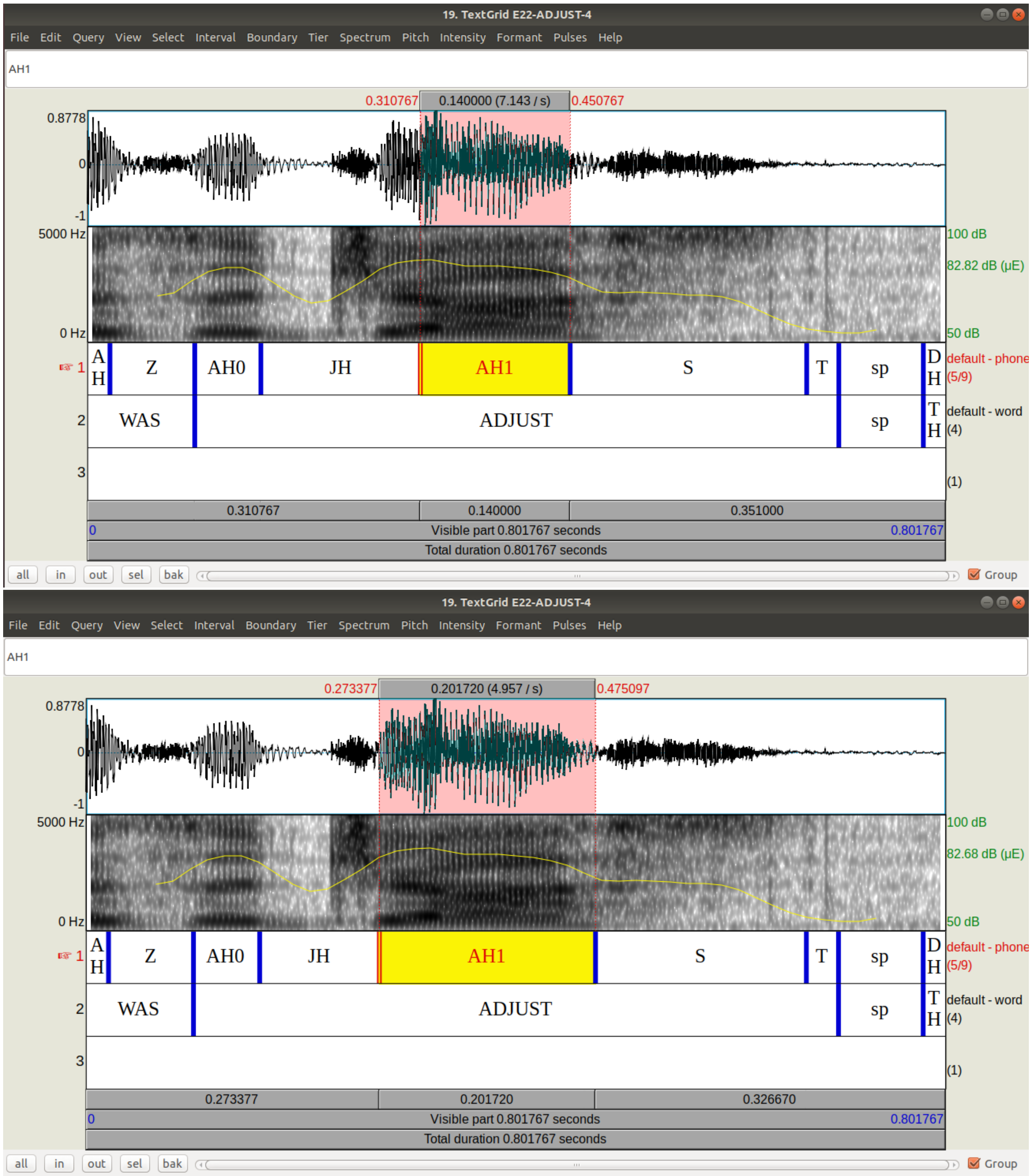


Figure 4.2: Demonstration of pitch correction.

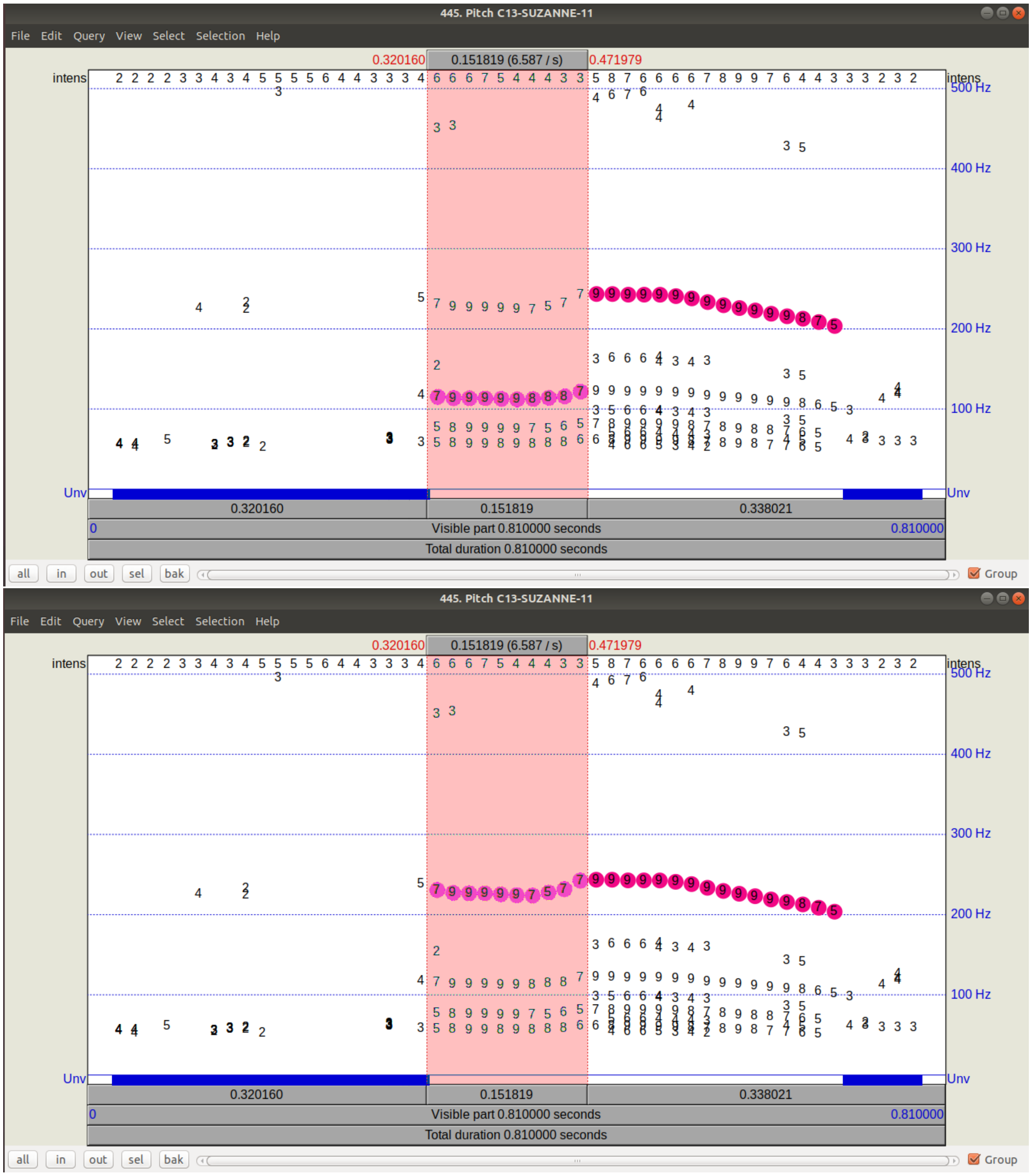
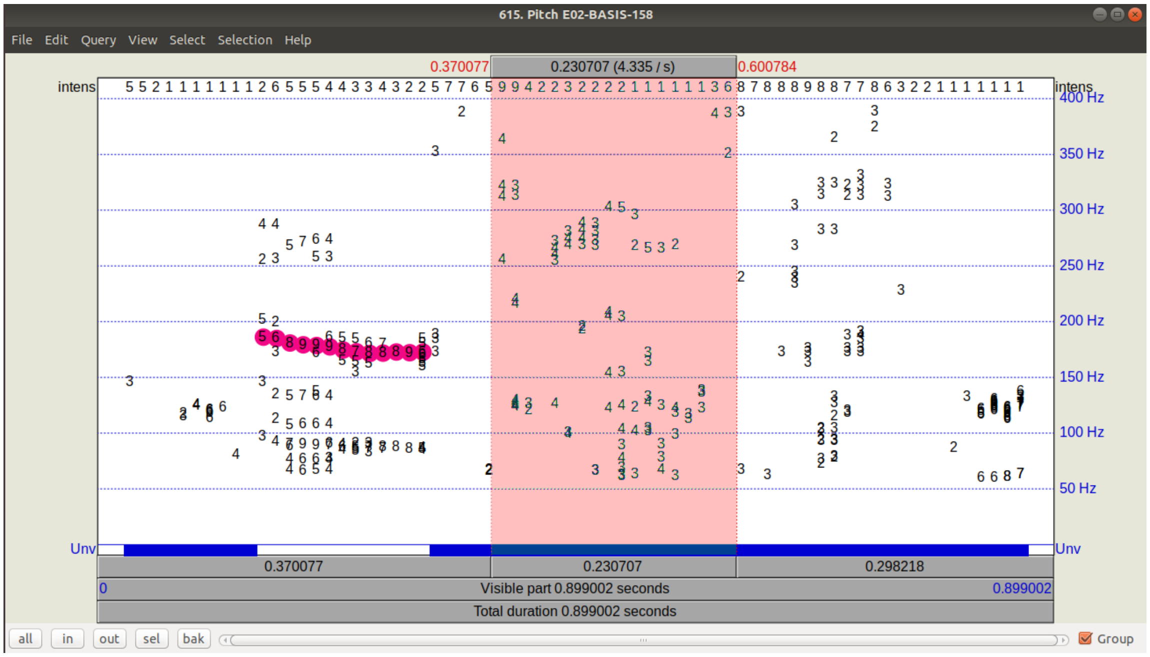


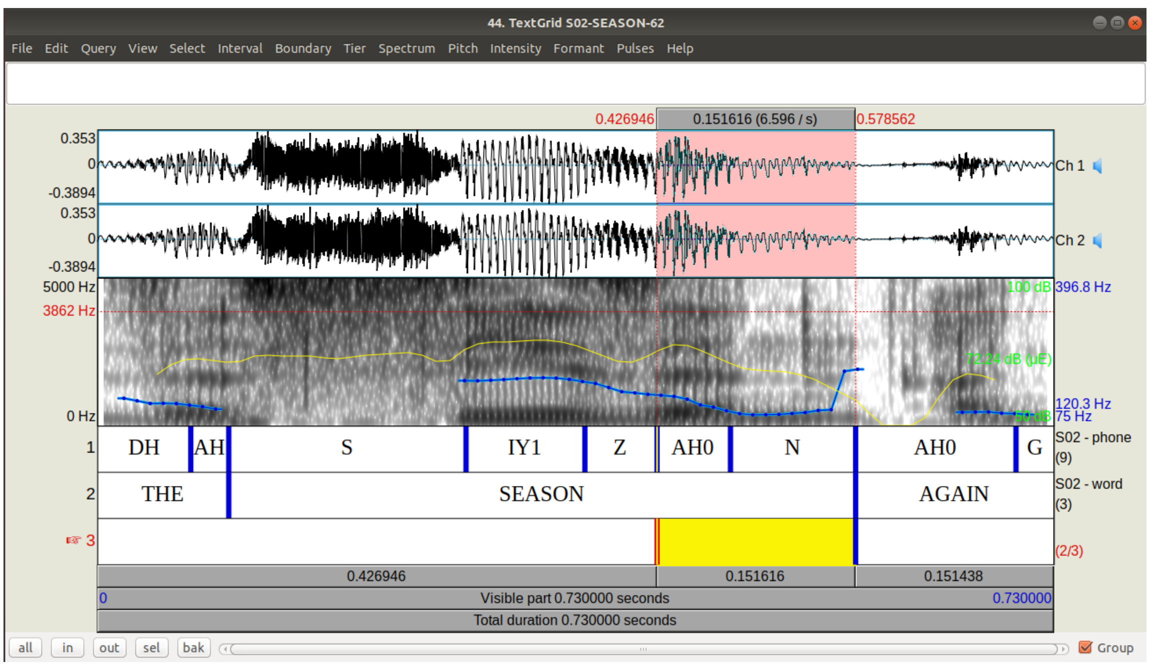
Figure 4.3: Example of unusable pitch measurements.



Measurements of mean intensity, maximum pitch, and pitch contour were done on the sonorant portion of the rhyme. The Experiment 1 wordlist (see Table 4.1) contains several words with sonorant codas or sonorant portions of codas (e.g. *support*, *mission*, *season*). This presents several problems for analysis. First, according to Duanmu (2007), in Mandarin, the coda can bear tone, with part of the tonal contour continuing onto sonorant codas. This is relevant if we wish compare L1 Mandarin speakers' English pitch curves to Mandarin tones (see Chapter 3), as we should expect that the relevant contour should extend over all sonorant portions of the rhyme. Hence, pitch measurements should be taken over those portions.

Second, the sonorant-coda syllables had a great deal of variability in realization. Many English speakers reduced the /ən/ rhymes in words such as *mission* and *season* by deleting the vowel and producing a syllabic nasal (see Figure 4.5), while several Mandarin speakers deleted the nasal coda and lengthened and nasalized the schwa (see Figure 4.6). A realization with a full schwa followed by a nasal [ən] is shown in Figure 4.4 for reference. Several Mandarin speakers also produced non-rhotic variants of *support*. A duration measurement for the sonorant portion of the rhyme was taken to observe whether the use of sonorant rhymes vs vowels would alter the duration facts for stress.

Figure 4.4: Example of an [ən] realization of /ən/.



The measurements of vowel duration and rhyme duration were calculated from the hand corrected boundaries in Praat, then converted to mil-

Figure 4.5: Example of a syllabic nasal.

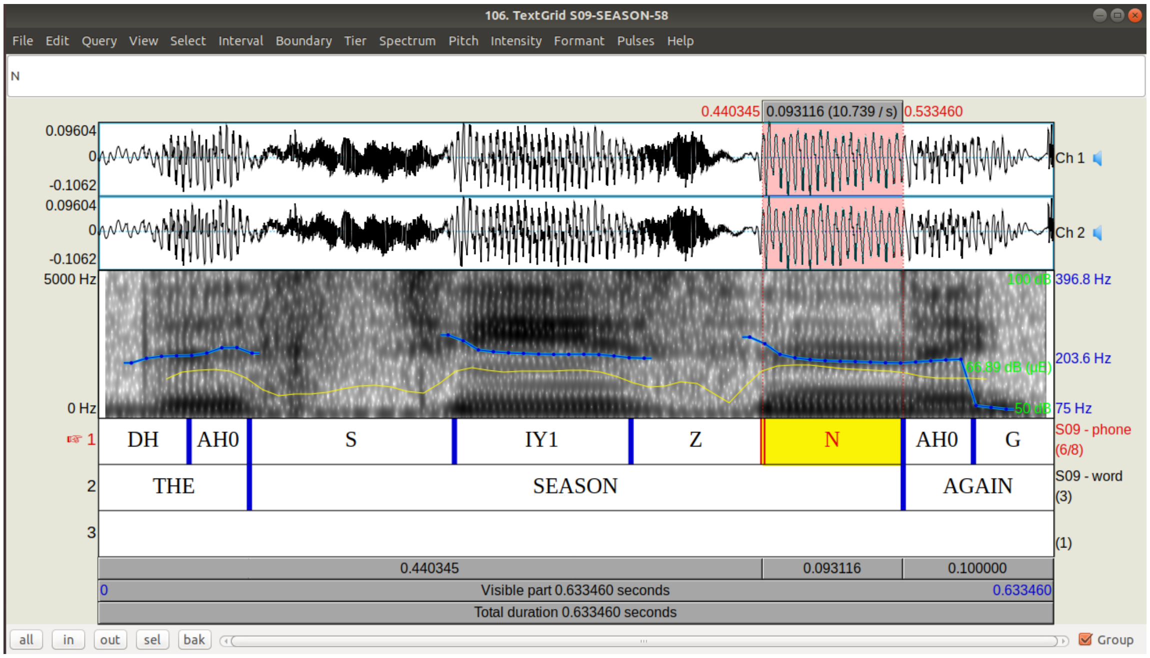
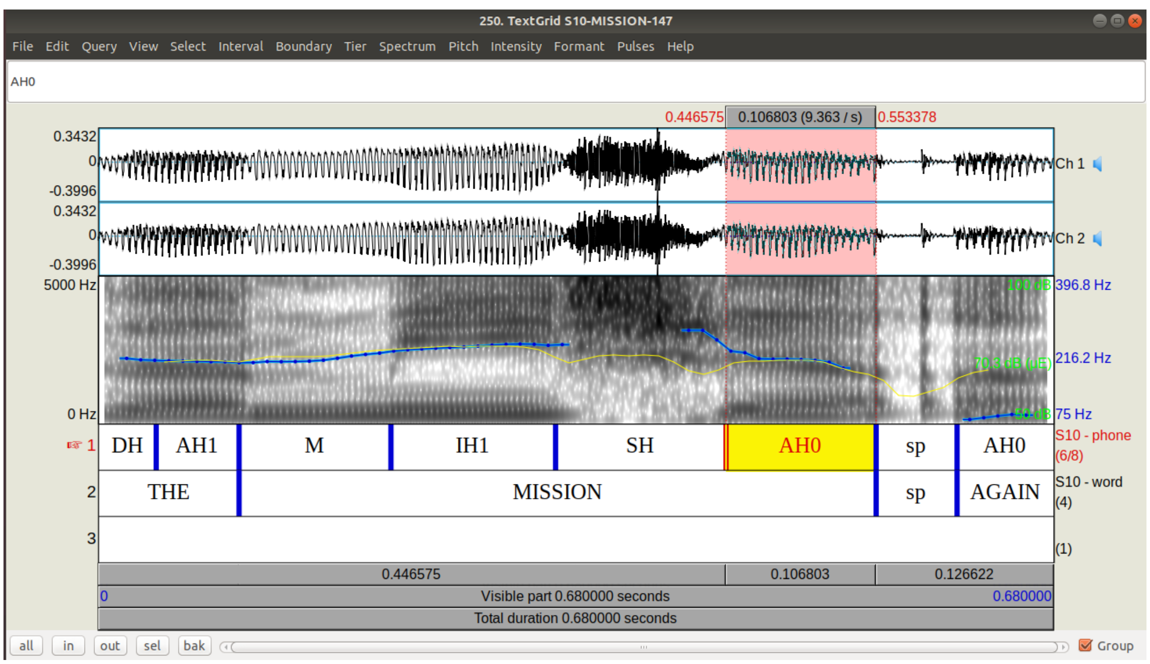


Figure 4.6: Example of a nasalized schwa.



liseconds for analysis.

Mean intensity was measured in decibels from a Praat Intensity object and was modeled as area under the curve. For analysis, the measured intensity was expressed as a ratio of the intensity of the target rhyme or vowel over the per subject mean intensity across all tokens. Calculating this ratio ensured that differences in overall intensity between subjects was minimized, as overall intensity could be affected by difficult to control factors, such as speaker distance from the microphone.

Maximum pitch was measured and analyzed in Hertz, in order to better compare it to existing research. Pitch contours were measured in Hertz and converted to semitones with a base frequency of the subject's first quartile pitch. Using the first quartile as the base provided better normalization than using the minimum frequency, as creaky voicing caused some subjects to have extremely low pitch measurements (see Shih & Lu, 2015). To normalize time in pitch contours, ten pitch measurements were taken in each sonorant rhyme. For the purpose of contour measurement, the first 10% and last 10% of each vowel were excluded, with the ten points being taken from the middle 80%. This clipping was done to prevent segmental effects from affecting the pitch curves.

The measurements for Experiments 2 and 3 are largely the same, but as all target syllables in those experiments had monophtongs with no sonorant codas, and Experiment 1 showed that the vowel duration and rhyme duration measurements followed the same pattern (see Chapter 5), the distinction between vowel duration and rhyme duration was irrelevant. As such, all measurements for Experiments 2 and 3 are taken on the vowel, which will be equivalent to the sonorant rhyme as the target syllables had no sonorant codas.

Exclusion Criteria

Tokens were excluded for dysfluencies that could alter the measurements (such as one subject who stuttered and produced three vowels in *justice*) or when errors indicated the subject selected an incorrect word or did not know the word in question. This included errors such as selecting an irregular segment ([dɪvos] rather than [dɪvaɪs]) or inserting a segment ([atɹæk] for *attack*), but not regular variants that may be the result of L2 influence of their own, such as several Mandarin speakers who consistently rhotacized all schwas, or several who deleted nasal codas while producing a nasalized vowel. In one case, for Experiment 1, a subject noticed his own segmental errors and re-read the word correctly, producing three good tokens which were kept.

I identified stress errors in the English tokens by native intuition. Leaving stress errors as noise proved to be a poor option, as particularly final stressed words in my stimuli showed a very large number of errors for non-native speakers. I did not attempt to determine the degree of reduction in Mandarin neutral tone target syllables and left any variation in reduction as noise.

Statistical Analysis

Measures of duration, intensity, and maximum pitch were submitted to mixed-effects regression analysis using the lme4 package in R (R Development Core Team, 2015; Bates et al., 2015). Pitch curves were analyzed using Growth Curve Analysis (Mirman, 2017) in order to compare curves effectively. Parameter-specific p-values for all analyses were given by the lmerTest package, using Satterthwaite’s approximation for degrees of freedom (Kuznetsova et al., 2017).

The mixed effects analyses used subject as a random effect and word as a within subjects random effect. In experiment 1, vowel was a within subjects random effect, but rhyme was a within subjects random effect for all other measurements. In experiments 2 and 3, vowel was a random effect for all cases.

When selecting models, I used three criteria: Low BIC, low AIC, and avoidance of non-significant factors. BIC was prioritized over AIC in order to choose more parsimonious models, but both were considered in choosing a model. In avoiding non-significant factors, when the inclusion of a factor or interaction lowered the information criteria, but the model output showed no significant variation, I chose a model without that factor or interaction.

Growth Curve Analysis

Growth Curve Analysis (GCA) (Mirman, 2017) was used to compare pitch curves. GCA is a technique originally developed in order to analyze longitudinal data, where multiple measurements are taken over a period of time and the characteristics of a line fitted to the data (for instance, the slope of a line indicating the rate of change). Examples of these applications of GCA can be found in Mirman (2017).

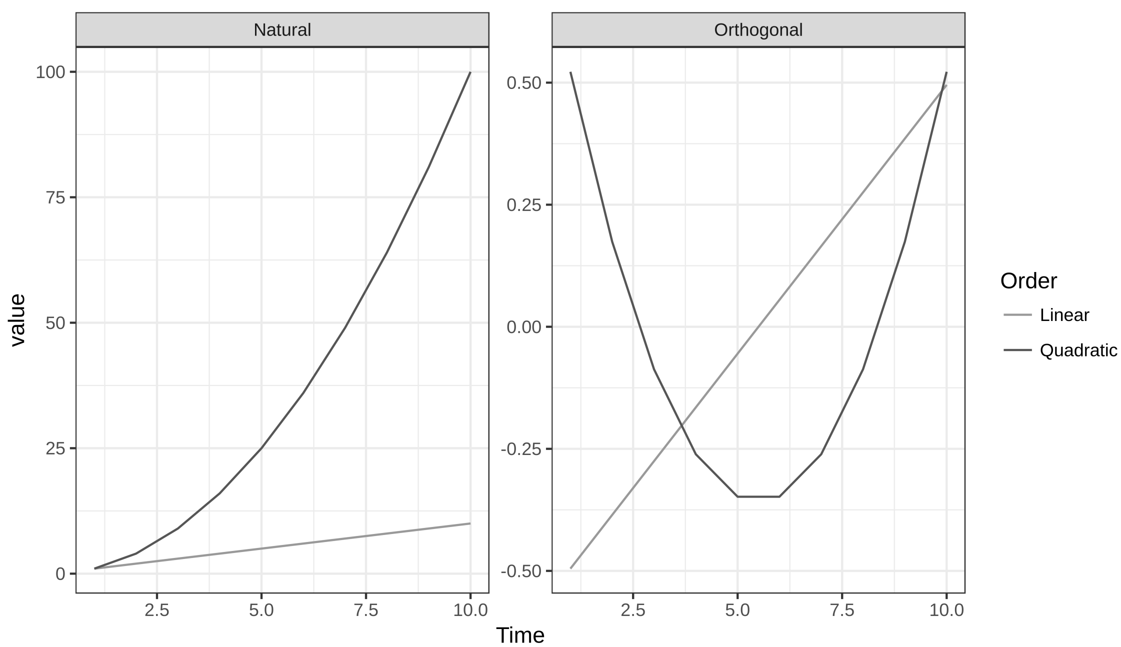
The application of Growth Curve Analysis to pitch curves is a natural extension, as a pitch curve can also be characterized with a series of measurements over time. Specifically, a pitch curve can be characterized as a series of measurements taken at different times through a vowel, rhyme, or syllable, which a curve can be fitted to. Shih & Lu (2015) used a curve fitting analysis to analyze the effects of talker-to-listener distance on tonal contours. In their analysis Shih & Lu (2015) chose to represent their pitch curves as a quadratic function, seen in (2), and compare the coefficients *a*, *b*, and *c*.

(2) $f(x) = a + bx + cx^2$

Shih & Lu (2015) reasoned that, while tones could be fit to a polynomial formula of arbitrary complexity, with increasing accuracy as more terms are added, Mandarin tones can be adequately captured by fitting a quadratic curve, as it was reasoned that the three coefficients would accurately describe the phonetically important facts about Mandarin tones. For the purpose of this dissertation, a quadratic function should also be sufficient detect any difference between Mandarin speakers' pitch curves and English speakers' pitch curves, and to compare Mandarin speakers' curves in English to their tonal curves in Mandarin.

The analysis in Shih & Lu (2015) is not the same as Growth Curve Analysis. Where Shih & Lu (2015) fit quadratic curves to each token and then did statistical tests on the resulting coefficients, Growth Curve Analysis, as described in Mirman (2017), plots a reference curve and uses a mixed-effects regression model to compare the data to that curve. GCA gives two options for the reference curve used for this, a natural curve or an orthogonal curve. Natural curves have a disadvantage in that the three coefficients are interrelated, meaning that they cannot be compared independently. Orthogonal curves are centered and scaled to the measurement area and allow the coefficients be independent of each other.

Figure 4.7: Demonstrations of natural and orthogonal curves.



Mirman (2017) advocates using orthogonal curves for most uses, with the caveat that orthogonal curves will result in the intercept term no longer representing the start point of the curve. When using orthogonal curves, the intercept instead refers to the average height of the curve. It also requires time to be invariant, meaning that using orthogonal curve requires the time normalization procedures mentioned 4.3 for pitch curves, a procedure that is also generally useful for separating pitch curves from duration.

I have chosen to use orthogonal curves for the GCA in this dissertation. In looking at the trade offs, it is more advantageous to have independent coefficients for comparison than to keep the intercept as the start of the curve. While having an accurate measure of the start point of the pitch curve may be beneficial, this dissertation benefits more from having the intercept, linear, and quadratic terms independently comparable, and these should be sufficient to compare pitch contours. While the start point of the curve may be phonologically important, as tones are characterized by sequences of high and low, for a phonetic comparison of curves it is not necessarily as valuable, especially in an analysis that does not also use the end point of the curve.

5 Experiment 1

5.1 Methods

Subjects

Subjects were recruited through the snowball method through the researcher's own social networks. 19 subjects participated, including 10 Mandarin-speaking L2 English speakers (5 male, 5 female) and 9 native English speakers (5 male, 4 female). The Mandarin speakers had a mean age of 29 (st dev: 6), a mean of 14 years studying English (st dev: 4, one subject not reporting), and had been in the US for a mean of 2.2 years (st dev: 2.3). No attempt was made to target a specific geographic area, but all subjects were from Mainland China. Of the 10 subjects, 5 reported speaking another Chinese variety in addition to Standard Mandarin. The individual data, including self reported variety, are shown on Table 5.1.

Table 5.1: Mandarin speaking subjects in first experiment.

Subject ID	gender	age	hometown	years studying English	years living in the US	Non-SM variety
15C01	male	26	Binzhou	15	1	Shandong
15C02	male	22	Beijing	10	8	none
15C03	female	23	Zibo	16	0.58	none
15C04	female	25	Huaihua	12	1.83	Hunan
15C05	male	28	Huaihua	15	4	Hunan
15C06	male	29	Jilin	10	3	none
15C07	female	23	Hailaier	–	2.5	none
15C08	female	39	Lai Wu	16	0.5	Lai Wu
15C09	male	36	Harbin	23	0.25	none
15C10	female	37	Inner Mongolia	10	0.5	Ordos

The native English speaking subjects had a mean age of 31 (st dev: 8). The majority of the English speakers (6/9) are from the Midwest, with the remaining three being from the state of New York. Demographic information for English speaking subjects is given in Table 5.2.

Table 5.2: Native English speaking subjects in first experiment.

Subject ID	gender	age	hometown
15E01	male	23	New York, NY
15E02	female	27	McGregor, IA
15E03	male	46	Mount Vernon, NY
15E04	male	45	Norwalk, WI
15E05	male	27	Appleton, WI
15E06	female	32	Armonk, NY
15E07	female	27	Shoreview, MN
15E08	female	30	Chicago, IL
15E09	male	21	Fond du Lac, WI

Experimental Stimuli and Equipment

A list of 18 disyllabic nouns, shown in Table 4.1, were selected from the first five thousand most common nouns in the Corpus of Contemporary English (Davies, 2008). Half of the words have initial stress and half have final stress. These were presented on paper to participants as a list. Participants read each word three times in the carrier phrase *Please say the ___ again.*

Analysis

The durations of the vowel and the sonorant portion of the rhyme were both taken for analysis. Measurements of mean intensity, maximum pitch, and pitch curve were all done on the sonorant portion of the rhyme. See Chapter 4 for more information.

All of these variables were submitted to mixed-effects linear regression analyses using subject's native language and the position and expected stress value of the syllable as fixed effects and subject as a random effect. All analyses had word as a within subjects random effect, and vowel duration also used vowel as a within subjects random effect, while other measures had rhyme as a within subjects random effect. For the pitch measurements (maximum pitch and pitch curves), gender was added as a fixed effects in order to factor out the tendency for men and women to have different overall pitch and pitch range.

5.2 Results

Duration

Vowel Duration

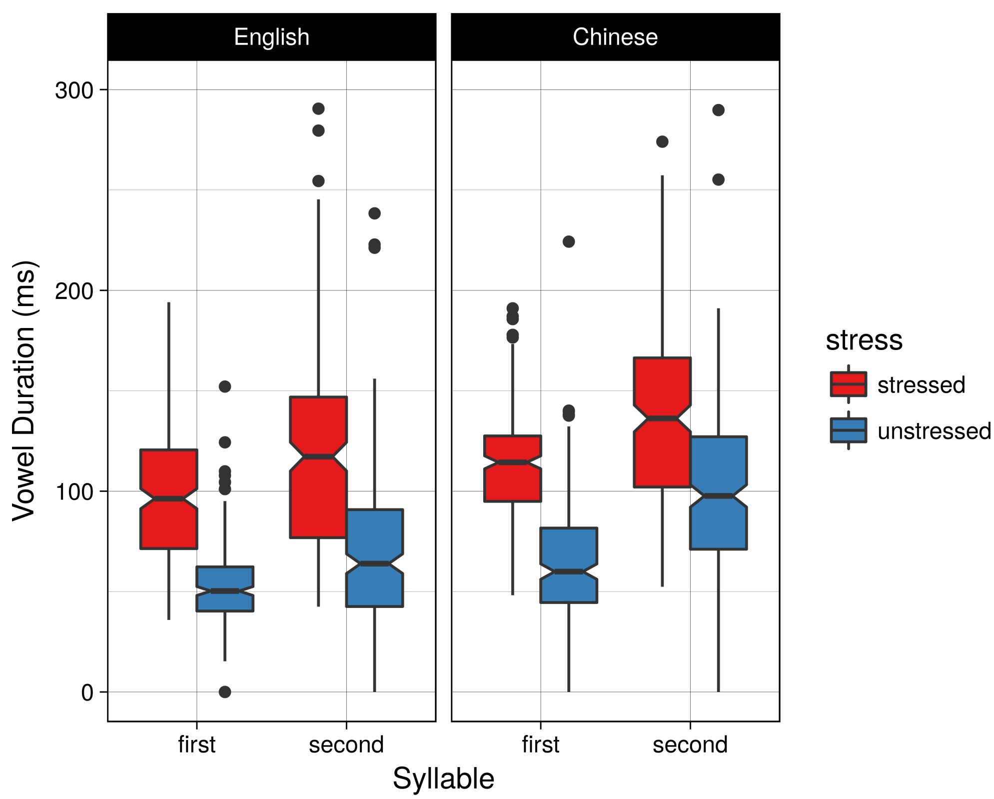
Descriptive statistics for the vowel duration model are given in Table 5.3, and a boxplot visualizing the data is seen in 5.1.

Vowel duration was submitted to a mixed-effects linear model with subject as a random effect and word and vowel as within-subjects random effects. The best model included stress, position, and the interaction of position and language. Table 5.4 shows the fixed effects of the vowel duration model, and a plot of the vowel duration models can be seen in Figure 5.1.

Table 5.3: Descriptive statistics for Experiment 1 vowel duration (milliseconds)

language	stress	position	mean	median	st. dev.	minimum	maximum
English	stressed	first	98.9	96.3	32.2	35.9	194
		second	120	119	54.4	42.5	378
	unstressed	first	52.6	50.3	21.4	0	152
		second	69.4	63.9	39.2	0	238
Chinese	stressed	first	113	114	26	48.2	191
		second	140	136	48.6	52.4	315
	unstressed	first	63.7	59.9	29.8	0	224
		second	101	97.6	41.6	0	290

Figure 5.1: Experiment 1 Vowel duration by stress, position, and language.



Stressed syllables had longer vowel duration than unstressed syllables ($t = 13$, $p < 0.001$), and the first syllable was notably shorter than the second syllable ($t = -10$, $p < 0.001$). On the interaction of position and language, the model shows Chinese speakers as having longer vowels than English speakers in both syllables, but the effect is significant in second syllables ($t = 3.37$, $p < 0.005$), while it is much smaller in the first syllable, not reaching significance ($t = 1.66$, $p > 0.1$). This indicates that while Chinese speakers and English speakers both lengthen the second syllable, Chinese speakers appear to lengthen it more.

Table 5.4: Fixed effects for vowel duration model in Experiment 1.

	Estimate	Std. Error	DF	T value	p
(Intercept)	90	5.37	19.9	16.8	3.36E-13
stress –stressed	18.1	1.4	1260	13	4.21E-36
position –first	-8.38	0.838	1810	-10	5.77E-23
position –first : language –Chinese	12.4	7.43	20.3	1.66	0.112
position –second : language –Chinese	25	7.42	20.1	3.37	0.00302

Rhyme Duration

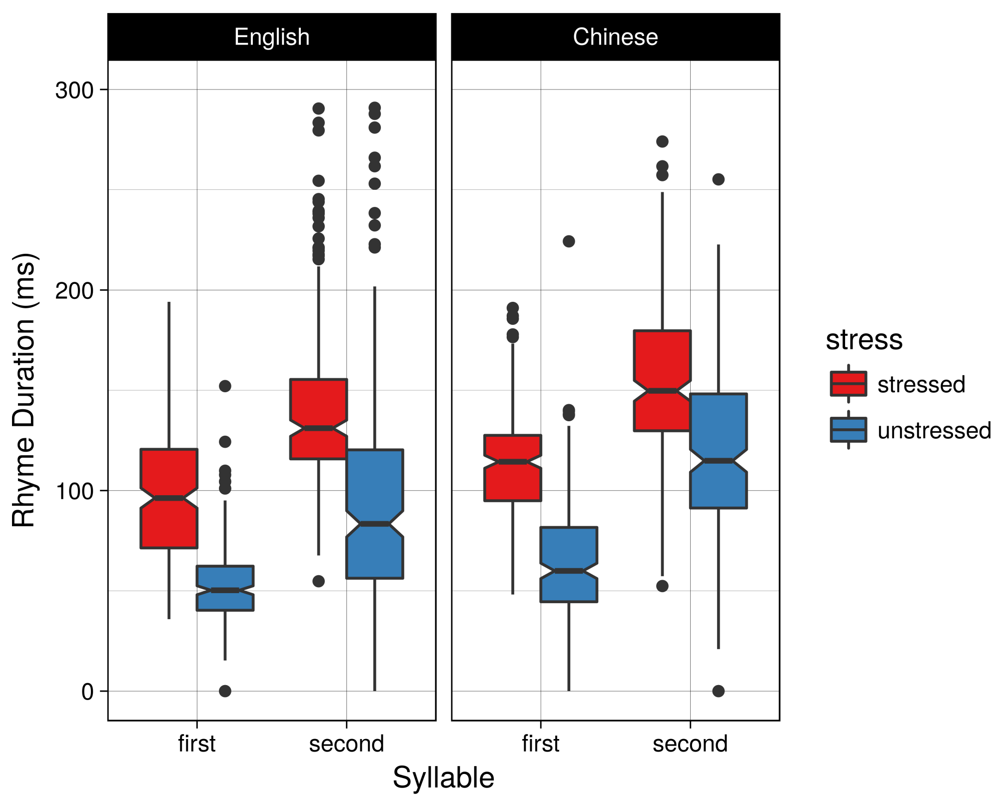
Descriptive statistics for the duration of the sonorant portion of the rhyme are given in Table 5.5, and a boxplot of the results by language, stress, and position is given in Figure 5.2

The duration of the sonorant portion of the rhyme was submitted to a mixed-effects analysis, with subject as a random effect and word and rhyme as within-subjects random effects. The best model included stress, position, and the interaction of position and language. The fixed effects for the model

Table 5.5: Descriptive statistics for Experiment 1 rhyme duration (milliseconds)

language	stress	position	mean	median	st. dev.	minimum	maximum
English	stressed	first	98.9	96.3	32.2	35.9	194
		second	143	131	47	54.8	378
	unstressed	first	52.6	50.3	21.4	0	152
		second	96	83.7	54.8	0	321
Chinese	stressed	first	113	114	26	48.2	191
		second	157	150	41.2	52.4	315
	unstressed	first	63.7	59.9	29.8	0	224
		second	121	115	45.2	0	322

Figure 5.2: Experiment 1 duration of the sonorant portion of the rhyme by stress, position, and language.



are presented in Table 5.6. The results of the rhyme duration model are similar to the vowel duration model, both showing that the first syllable is shorter ($t = -15.3$, $p < 0.001$) and stressed syllables are longer ($t = 10.1$, $p < 0.001$), and both showing Chinese speakers having longer duration in second syllables ($t = 2.42$, $p = 0.0208$) with only a slightly longer first syllable, not reaching statistical significance ($t = 1.3$, $p < 0.1$). This length difference in the second syllable somewhat smaller than in the vowel duration model shown in section 5.2 and does not reach the same level of significance, indicating that the inclusion of the sonorant coda consonants obscures the lengthening effect somewhat.

Table 5.6: Fixed effects for the rhyme duration model in Experiment 1.

	Estimate	Std. Error	DF	T value	p
(Intercept)	107	6.23	19.3	17.2	3.86E-13
stress –stressed	13.5	1.34	1150	10.1	5.59E-23
position –first	-15.2	0.998	1720	-15.3	2E-49
position –first : language –Chinese	11.3	8.7	20.3	1.3	0.208
position –second : language –Chinese	20.9	8.66	19.9	2.42	0.0254

Intensity

Table 5.7 shows descriptive statistics for mean intensity as measured in decibels. Descriptive statistics for the ratio of this measurement with the mean intensity measurement across all tokens for each subject is shown in Table 5.8. A boxplot showing this ratio by language, position, and stress is shown in Figure 5.3.

Table 5.7: Descriptive statistics for Experiment 1 mean intensity (dB).

language	stress	position	mean	median	st. dev.	minimum	maximum
English	stressed	first	70.2	70.3	4.4	59.3	81.2
		second	68.4	68.2	4.5	56.3	81.8
	unstressed	first	66.4	66.3	4.39	50.5	75.3
		second	66.2	66.1	4.71	44.7	81.1
Chinese	stressed	first	76.4	76.9	4.9	65.5	86.1
		second	75.6	76.7	5.41	59.6	85.6
	unstressed	first	72.6	73.4	5.24	59	86.7
		second	72.6	73.6	5.27	60.6	85.3

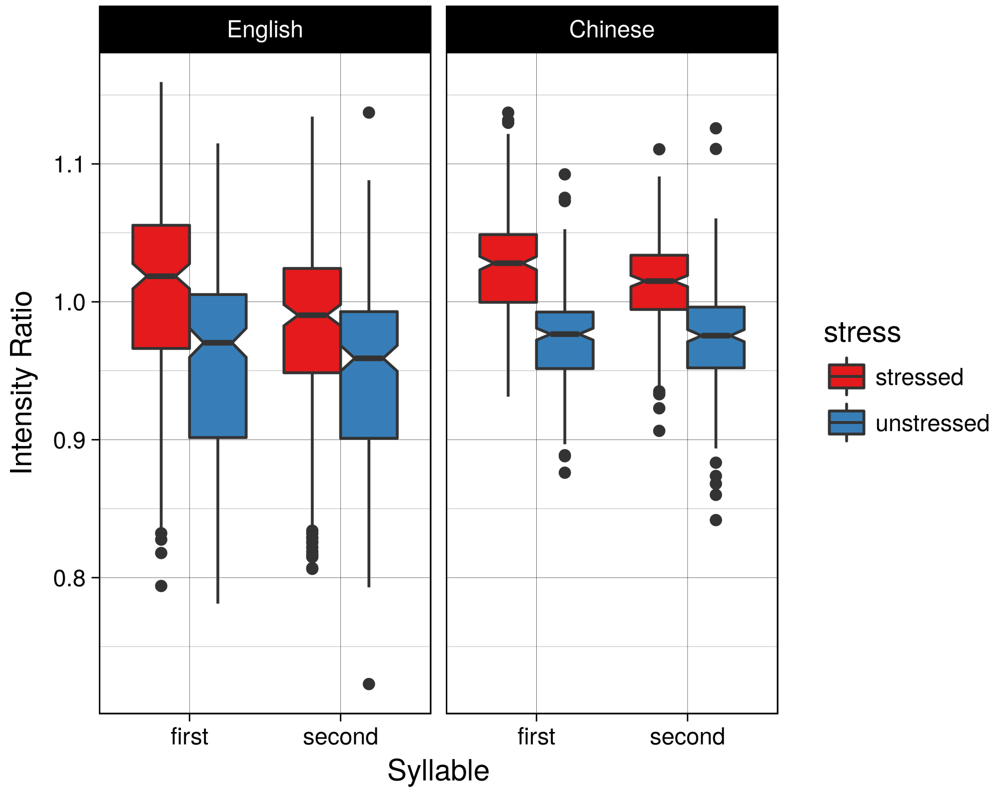
Table 5.8: Descriptive statistics for Experiment 1 ratio of rhyme mean intensity to the mean intensity for all tokens by each subject.

language	stress	position	mean	median	st. dev.	minimum	maximum
English	stressed	first	1	1.02	0.0712	0.794	1.16
		second	0.979	0.99	0.0673	0.806	1.13
	unstressed	first	0.951	0.97	0.0752	0.781	1.11
		second	0.946	0.959	0.0673	0.723	1.14
Chinese	stressed	first	1.03	1.03	0.0378	0.931	1.14
		second	1.01	1.02	0.0337	0.906	1.11
	unstressed	first	0.972	0.977	0.0349	0.876	1.09
		second	0.975	0.975	0.0393	0.842	1.13

The intensity ratio was submitted to a mixed-effects analysis with subject as a random effect and word and rhyme as within-subjects random effects. The best model showed effects for stress and the interaction of stress and position. Although there were models that may have been mathematically better in one direction or the other, this model was chosen as the most informative and closely matches the visual pattern seen in Figure 5.3. Crucially, no model that included an effect for language or its interactions performed particularly well.

In the model, stressed syllables have higher intensity than unstressed syllables ($t = 15.3$, $p < 0.001$), and stressed syllables in first position have higher

Figure 5.3: Experiment 1 ratio of mean intensity of the sonorant rhyme vs per subject mean intensity accross tokens.



intensity than those in second (t = 4.29, p < 0.001), while unstressed syllables did not have significantly higher intensity in first or second position (t = 0.413, p = 0.68). This indicates that speakers do slightly enhance intensity when producing stress on an initial syllable, but there is no appreciable difference between speakers of the two languages in normalized intensity.

Table 5.9: Fixed effects for the intensity ratio model in Experiment 1.

	Estimate	Std. Error	DF	T value	p
(Intercept)	0.985	0.00948	19.3	104	5.78E-28
stress –stressed	0.0233	0.00152	480	15.3	1.56E-43
stress –stressed : position –first	0.00734	0.00171	479	4.29	2.19E-05
stress –unstressed : position –first	0.000662	0.0016	436	0.413	0.68

Maximum Pitch

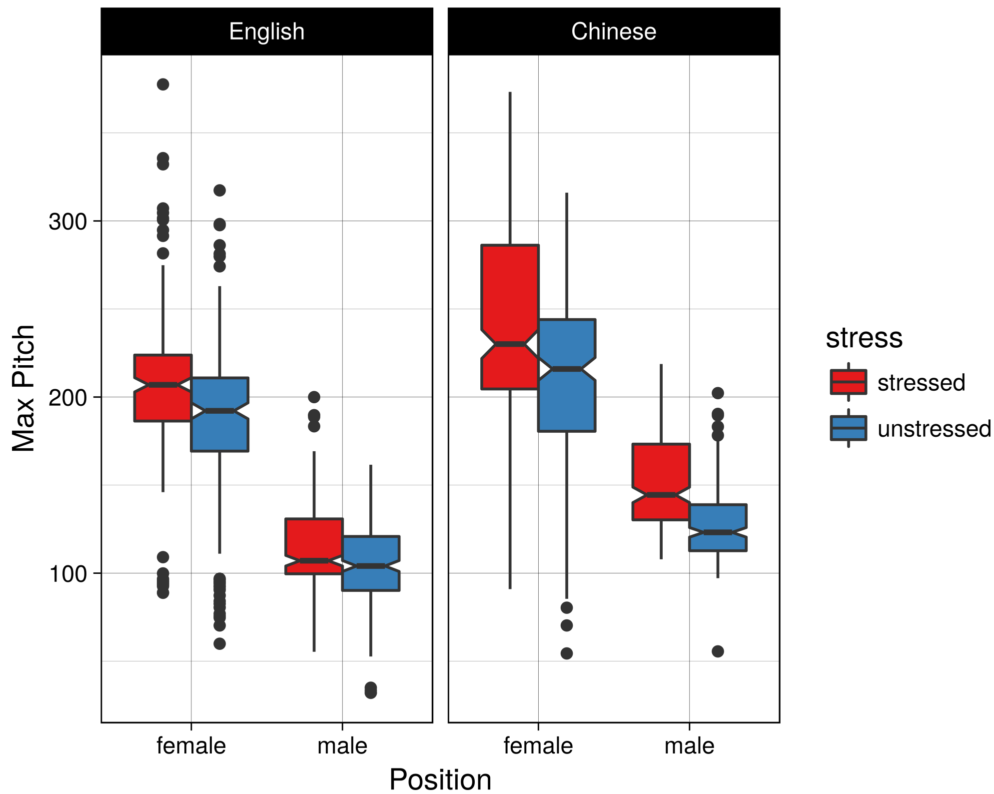
Table 5.10 shows descriptive statistics for maximum pitch as measured within the sonorant portion of the rhyme. A boxplot by language, gender, and stress is shown in 5.4.

Table 5.10: Descriptive statistics for Experiment 1 maximum F0 (Hz)

language	gender	stress	mean	median	st. dev.	minimum	maximum
English	female	stressed	205	207	42.9	88.8	378
		unstressed	186	192	45.3	59.9	317
	male	stressed	116	107	24.1	55.4	200
		unstressed	105	104	25.7	32.1	162
Chinese	female	stressed	246	230	50.5	90.9	373
		unstressed	214	216	43.8	54.4	316
	male	stressed	151	144	24.9	108	219
		unstressed	128	123	21.2	55.6	202

Maximum pitch was submitted to a mixed-effects model. The fixed effects of the model are presented in Table 5.11. The best model included effects for stress, gender, their interaction, and the interaction of stress and language. Stressed syllables had higher maximum pitch than unstressed syllables ($t = 7.78$, $p < 0.001$) and women had higher F0 than men ($t = 6.99$, $p < 0.001$). Women also had higher stressed syllables specifically ($t = 3.37$, $p < 0.001$). Chinese speakers had higher F0, but only in stressed syllables was this significant ($t = 2.96$, $p < 0.01$), while in unstressed syllables the difference was

Figure 5.4: Experiment 1 maximum pitch (F0 Hz) by language, gender, and stress.



overwhelmed by variance, though it did come close to the thresh-hold for significance (t = 1.99, p = 0.06).

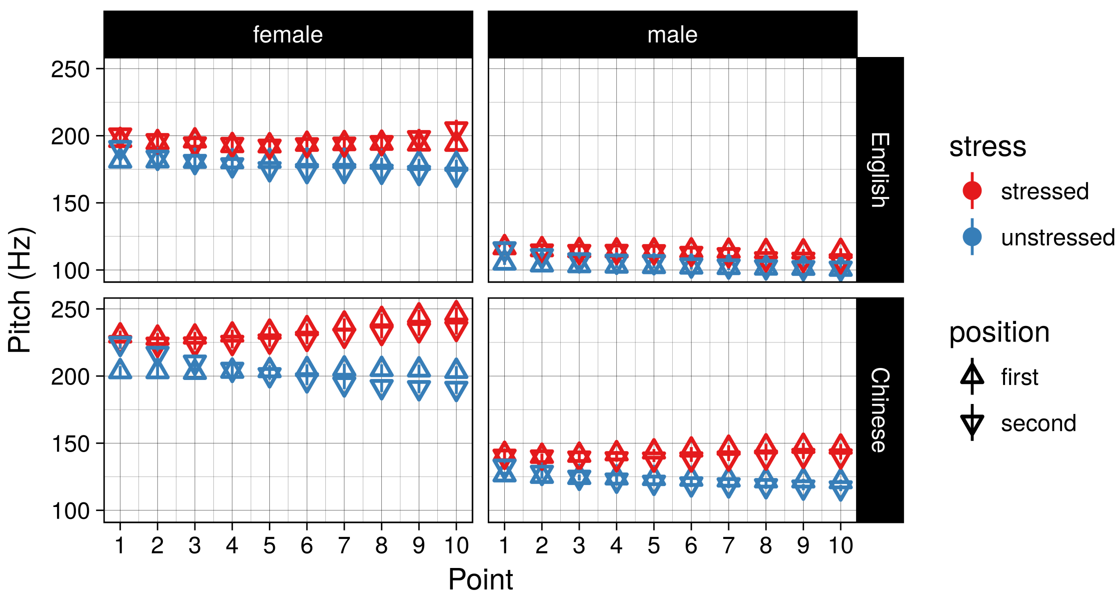
Table 5.11: Fixed effects for the max pitch model.

	Estimate	Std. Error	DF	T value	p
(Intercept)	152	9.41	19	16.2	1.4E-12
stress –stressed	8.22	1.06	229	7.78	2.5E-13
gender –female	45.3	6.48	19	6.99	1.16E-06
stress –stressed : gender –female	2.48	0.737	236	3.37	0.000876
stress –stressed : language –Chinese	38.5	13	19.2	2.96	0.00795
stress –unstressed : language –Chinese	26.1	13.1	19.9	1.99	0.06

Pitch Curves

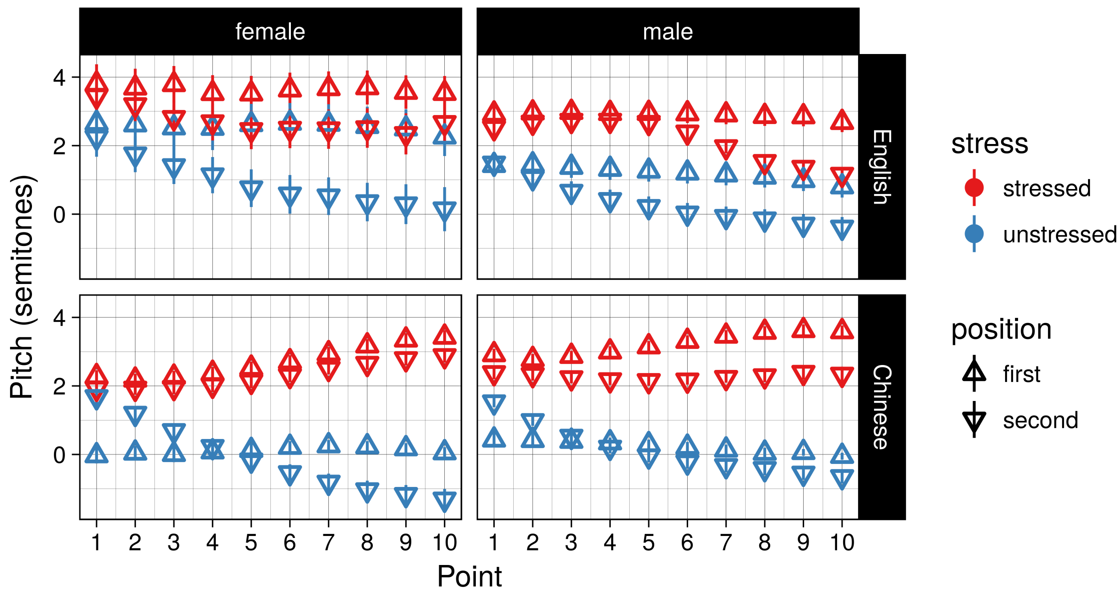
F0 was measured at 10 points in the middle 80% of the sonorant portion of each rhyme, with measurements converted into semitones. These were then submitted to Growth Curve Analysis (Mirman, 2017), using orthogonal quadratic curves. Mixed-effects models were constructed with subject as a random effect and word and rhyme as within-subjects random effects. For ease of interpretation, English speakers have been coded as the baseline for comparison in the native language dimension, and all other fixed effect variables (stress, position, gender) were recoded with sum coding.

Figure 5.5: Descriptive plot of Experiment 1 pitch curves by language, position, and stress (Hz).



The best fit model had effects for stress on the intercept and the linear

Figure 5.6: Descriptive plot of Experiment 1 pitch curves by language, position, and stress (semitones).



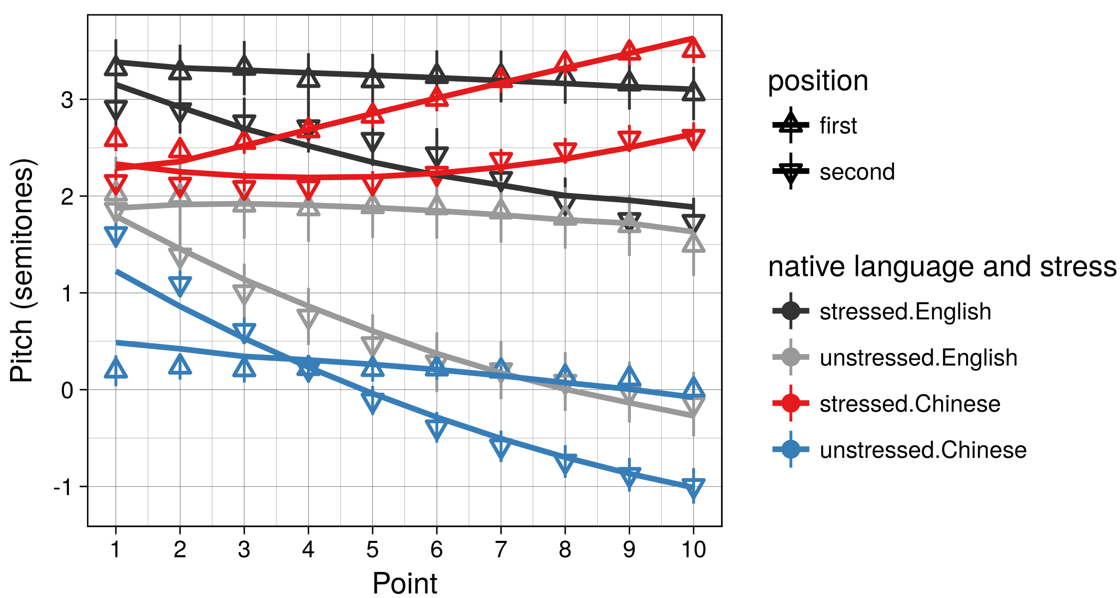
terms, position on the linear and quadratic term, and gender on the intercept only, and the interaction of stress and language. Table 5.12 lists the fixed effects of the model. Figure 5.7 shows the fitted model by stress, position, and language.

Stressed syllables had a higher intercept and a higher linear term. This translates into a high, flat mean curve for stressed syllables and a lower, descending mean curve for unstressed syllables. Position had no effect on the intercept, but did affect the linear and quadratic terms, having a higher linear coefficient, but also a slightly lower quadratic coefficient in the first syllable. This translated into a flatter, higher curve in the first syllable, vs a descending curve with a significant degree of curvature in the second sylla-

Table 5.12: Fixed effects for the pitch curves model in Experiment 1.

	Estimate	Std..Error	t.value
Intercept	14.6	1.09	13.4
Linear	-1.1	0.345	-3.17
Quadratic	0.0962	0.0745	1.29
stress - stressed	0.791	0.0584	13.5
Linear : stress - stressed	0.404	0.137	2.94
Linear : position - first	0.653	0.0601	10.9
Quadratic : position - first	-0.143	0.0471	-3.03
gender - female	3.9	0.945	4.13
Linear : stress - stressed : language - Chinese	1.29	0.444	2.9
Linear : stress - unstressed : language - Chinese	0.0351	0.496	0.0708

Figure 5.7: Experiment 1 curves model fitted to data by stress, position, and language



ble. Gender only had a significant effect on the intercept term, showing that women had overall higher curves than men. It should be noted that this is consistent with the findings for maximum pitch shown in section 5.2, where gender also increased pitch.

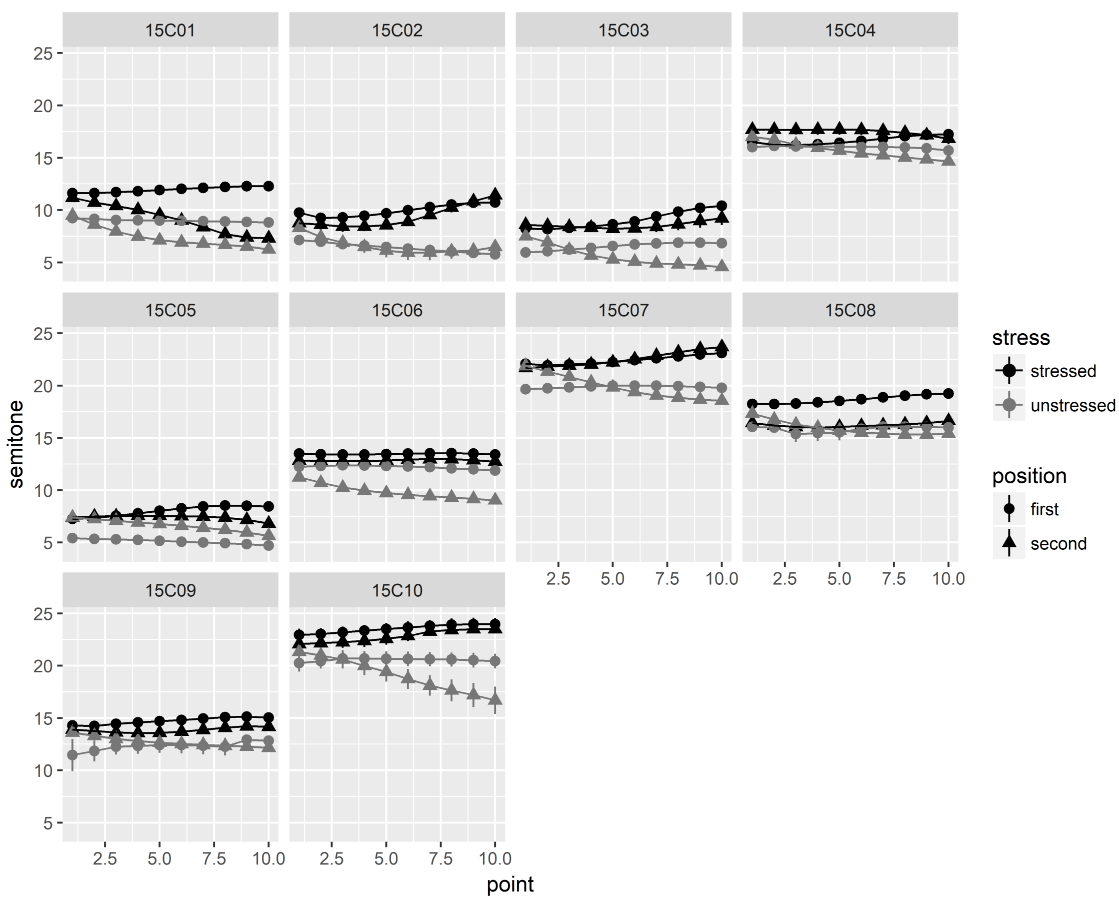
More interesting for the current study are the effects of the interactions. Chinese speakers had a higher intercept term on stressed syllables. The effect of language on unstressed syllables was not significant ($t < 1$). The higher linear term shows as a slight upward curve to the mean pitch curve, relative to English speakers, who have a downward mean pitch curve.

5.3 Discussion

The agreement of the vowel and rhyme duration measurements indicate that Mandarin speakers do use syllable duration similarly, but not identically to English speakers. This could be an indication of Mandarin speakers' native use duration affecting their English. Intensity was used the same way in both languages, with only a minor overall effect for language, possibly attributable to Chinese speakers having overall louder vowels.

The GCA results for pitch curves are interesting, but the means could potentially be misinterpreted. While the model shows an overall upward curve, this generalization does not necessarily reflect the population properly. Figure 5.8 shows the curves for each of the ten Mandarin speaking subjects by stress and syllable position. While several subjects do have clear upward curves in stressed syllables (e.g. 15C02), there is a great deal of within-subject variation. 15C01 shows an upward curve in the first syllable, but a downward curve in the second syllable, with the difference being larger in stressed syllables than unstressed syllables. Several subjects, like 15C06, show flat curves. In addition, without Mandarin tokens for comparison, I cannot attribute these differences in pitch curves to any influence from Mandarin tones.

Figure 5.8: Individual curves for Experiment 1 Chinese subjects by stress and position.



Ultimately, this experiment shows some evidence of transfer from the L1 Mandarin system of stress, most notably in the area of duration, while being inconclusive about any tonal transfer. In the next two chapters, I present the results of new experiments that build on these findings with a larger subject pool and a comparison of Mandarin speakers' L2 English and their L1 Mandarin.

6 Experiment 2

6.1 Methods

A total of 60 subjects were recorded for the experiment, 30 Mandarin speaking subjects (15 male, 15 female), and 30 native English speaking subjects (15 male, 15 female). Demographic information for the Mandarin speaking subjects is given in Table 6.1, and information for English speaking subjects is given in Table 6.2.

Subjects were recorded using the English wordlist listed in Table 4.2 using the carrier phrase *It was ___ that I said.* Further details on the experimental procedure can be found in Chapter 4.

6.2 Results

Duration

Table 6.3 shows descriptive statistics for vowel duration, and Figure 6.1 shows a boxplot of the results by stress, position, and language.

The best model had effects for stress, position, language, and the interaction of position and language. The fixed effects for this model are shown in Table 6.4. Unstressed syllables are shorter than stressed syllables ($t = -11.3$, $p < 0.001$), second syllables are longer than first syllables ($t = 18.7$, $p < 0.001$), and Chinese speakers had longer syllables overall ($t = 3.40$, $p < 0.001$). On the

Table 6.1: Mandarin speaking subjects for Experiments 2 and 3

subject	hometown	age	gender	years studying English	years living in the US	Non-SM variety
R18C01	Suizhou	36	male	10	0.00	Hubei
R18C02	Shaoyang	33	male	9	0.50	Shaoyang
R18C03	Shenzhen	20	female	15	1.00	none
R18C04	Beijing	22	female	10	7.00	none
R18C05	Beijing	19	male	13	0.08	none
R18C06	Malaysia	23	male	18	3.00	Cantonese, Hokkien
R18C07	Beijing	20	male	10	0.00	Beijing
R18C08	Hubei	29	female	17	4.00	none
R18C09	Nanchang	18	male	12	0.08	Gan
R18C10	Hangzhou	20	male	16	0.08	none
R18C11	Suzhou	20	female	14	0.08	Suzhou
R18C12	Beijing	18	female	12	0.08	none
R18C13	Taiyuan	18	female	13	0.08	none
R18C14	Weihai	19	female	16	0.08	none
R18C15	Taizhou	20	male	10	0.08	Taizhou
R18C16	Shaoyang	22	female	13	0.08	Shaoyang, Cantonese
R18C17	Nanchang	20	female	9	0.17	Gan
R18C18	Shenzhen	21	male	15	2.50	Cantonese
R18C19	Kaohsiung	18	female	15	4.00	Min Nan
R18C20	Hubei	20	female	12	0.17	Wuhan
R18C21	Kuala Lumpur, Malaysia		male	20	2.50	Cantonese
R18C22	Guangzhou	18	female	13	0.50	Cantonese
R18C23	Guangzhou	26	female	13	<1	Cantonese
R18C24	Wuhu, Anhui	32	female	20	0.17	none
R18C25	Huangmei, Hubei	55	male	30	19.00	Huangmei
R18C26	Jingzhou	38	male	20	0.13	Hubei
R18C27	Harbin	22	female	13	0.75	Dongbei
R18C28	Baishan, Jilin	39	female	25	0.67	Dongbei
R18C29	Shangrao	30	male	20	5.00	Shangrao
R18C30	Dalian	37	female	27	0.54	none

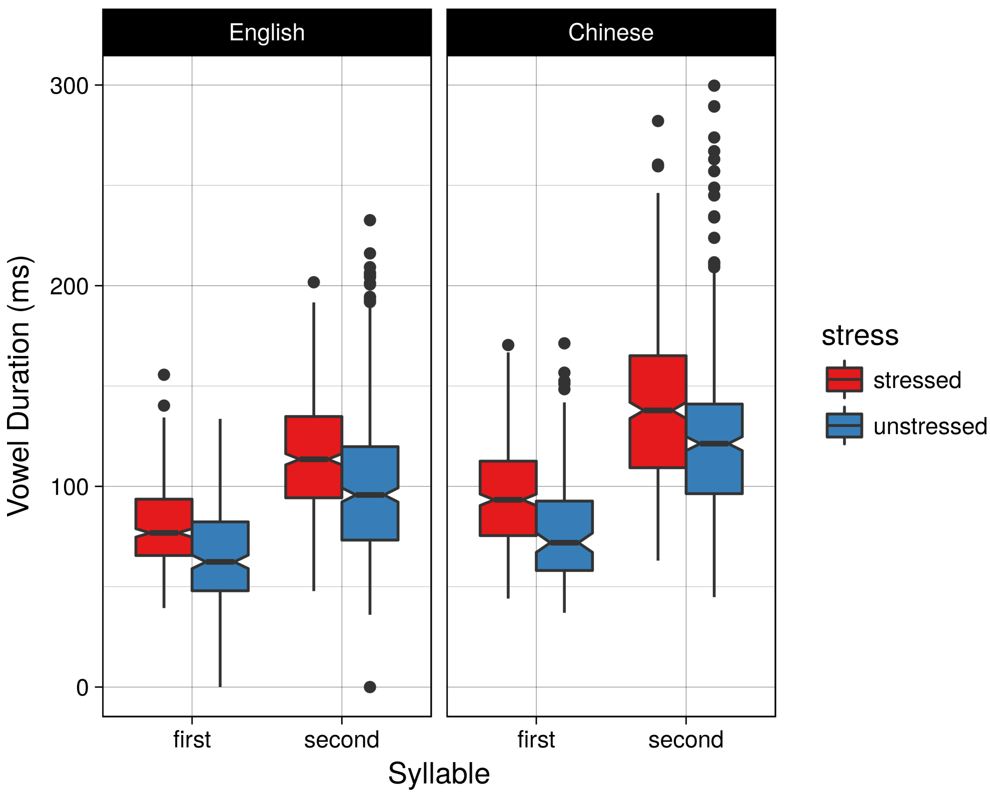
Table 6.2: Native English speaking subjects for Experiment 2

ID	hometown	age	gender
R18E01	New Glarus	18	female
R18E02	Waterford	20	female
R18E03	Beaver Dam, WI	18	female
R18E04	Milwaukee	43	female
R18E05	Pulaski, WI	19	female
R18E06	Detroit	47	female
R18E07	Appleton, WI	33	male
R18E08	Phillips,WI	21	female
R18E09	Tigerton	18	male
R18E10	Green Bay	18	female
R18E11	Green Bay	18	male
R18E12	Lake Geneva	21	male
R18E13	Green Bay	18	male
R18E14	Brookfield	19	female
R18E15	Arlington Heights, IL	57	female
R18E16	Hartland, WI	20	male
R18E17	Manitowoc,WI	22	male
R18E18	Mineral Point, WI	19	female
R18E19	Manitowoc,WI	20	male
R18E20	St Louis	19	female
R18E21	Plainview, NY	19	female
R18E22	Bayonne, NJ; Springfield, NJ	67	female
R18E23	Schaumberg		female
R18E24	Withee, WI	21	male
R18E25	Arlington Heights, IL	21	female
R18E26	LaCrosse, WI	22	female
R18E27	LaCrosse, WI	22	male
R18E28	OK, MA, CA, WI	21	female
R18E29	Green Bay	22	female
R18E30	Elgin, IL	18	female

Table 6.3: Descriptive statistics for Experiment 2 vowel duration (milliseconds)

language	stress	position	mean	median	st. dev.	minimum	maximum
English	stressed	first	80	76.8	19.7	39.4	156
		second	115	113	27.7	47.8	202
	unstressed	first	67.2	62.4	25.5	0	134
		second	101	95.3	37.8	0	233
Chinese	stressed	first	95.8	93.3	26.1	44.1	170
		second	141	138	40.7	63	319
	unstressed	first	79.4	71.9	29	37	171
		second	125	121	44	44.8	331

Figure 6.1: Vowel duration by stress, position, and language in Experiment 2.



interaction, Chinese speakers had longer second syllables (t = 3.07, p < 0.01). As with the results of Experiment 1 (see Chapter 5), Mandarin speakers seem to have more lengthening on second syllables than English speakers. The effect for language may simply indicate a small tendency for Chinese speakers to use a slower speaking rate, an effect that washed out of the model in Experiment 1 due to low sample size and less controls on vowels.

Table 6.4: Fixed effects for Experiment 2 vowel duration model.

	Estimate	Std. Error	DF	T value	p
(Intercept)	79.9	2.96	88.3	27	1.99E-44
stress –unstressed	-16.9	1.5	850	-11.3	1E-27
position –second	37.6	2	833	18.7	1.2E-65
language –Chinese	14.6	4.19	88.9	3.48	0.000771
position –second : language –Chinese	9.04	2.95	838	3.07	0.00224

Intensity

Descriptive statistics for intensity are shown in Table 6.5 for mean intensity in decibels and in Table 6.6 for the ratio of mean intensity of the vowel and the mean intensity across all tokens. A boxplot of the intensity ratio by language is shown in 6.2.

The intensity ratio was submitted to a mixed-effects analysis and the best model showed effects for stress and position. Unstressed syllables had lower intensity than stressed syllables (t = -27, p < 0.001) and second syllables had lower intensity than first syllables (t = -8, p < 0.001). This result differs from the result in Experiment 1 (see Chapter 6) in that position as single factor was more important than the interaction of stress and position. Importantly,

Table 6.5: Descriptive statistics for Experiment 2 intensity (dB).

language	stress	position	mean	median	st. dev.	minimum	maximum
English	stressed	first	58.4	58.4	5.19	43.1	71.8
		second	57.9	58	4.86	45.4	70.4
	unstressed	first	55.1	55.1	4.78	37.8	66
		second	53.7	53.7	4.87	41.2	76.8
Chinese	stressed	first	59.7	59.2	6.88	43.7	79.7
		second	58.1	57.5	6.33	43.2	74.5
	unstressed	first	56	55.2	5.93	39.5	72.8
		second	54.6	53.8	6.4	40.7	76.7

Table 6.6: Descriptive statistics for Experiment 2 intensity (dB)

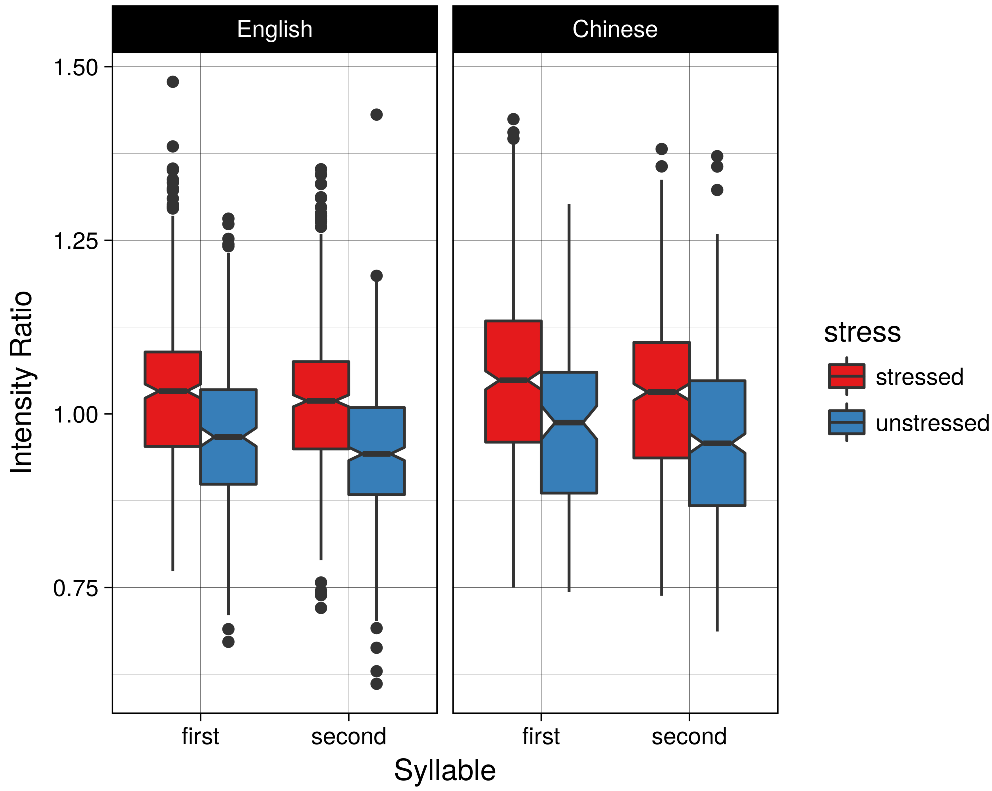
language	stress	position	mean	median	st. dev.	minimum	maximum
English	stressed	first	1.03	1.03	0.115	0.773	1.48
		second	1.02	1.02	0.109	0.721	1.35
	unstressed	first	0.973	0.967	0.111	0.672	1.28
		second	0.947	0.943	0.105	0.611	1.43
Chinese	stressed	first	1.05	1.05	0.141	0.75	1.42
		second	1.03	1.03	0.135	0.738	1.38
	unstressed	first	0.989	0.988	0.134	0.743	1.3
		second	0.965	0.958	0.133	0.687	1.37

though, models that included language were uniformly worse fits than the current model.

Table 6.7: Fixed effects for Experiment 2 intensity model.

	Estimate	Std. Error	DF	T value	p
(Intercept)	1.04	0.0145	61.5	71.8	5.42E-61
stress –unstressed	-0.0655	0.00243	869	-27	9.19E-117
position –second	-0.0194	0.00242	864	-8	3.84E-15

Figure 6.2: Experiment 2 ratio of vowel intensity over per subject mean intensity across tokens by stress, position, and language.



Maximum Pitch

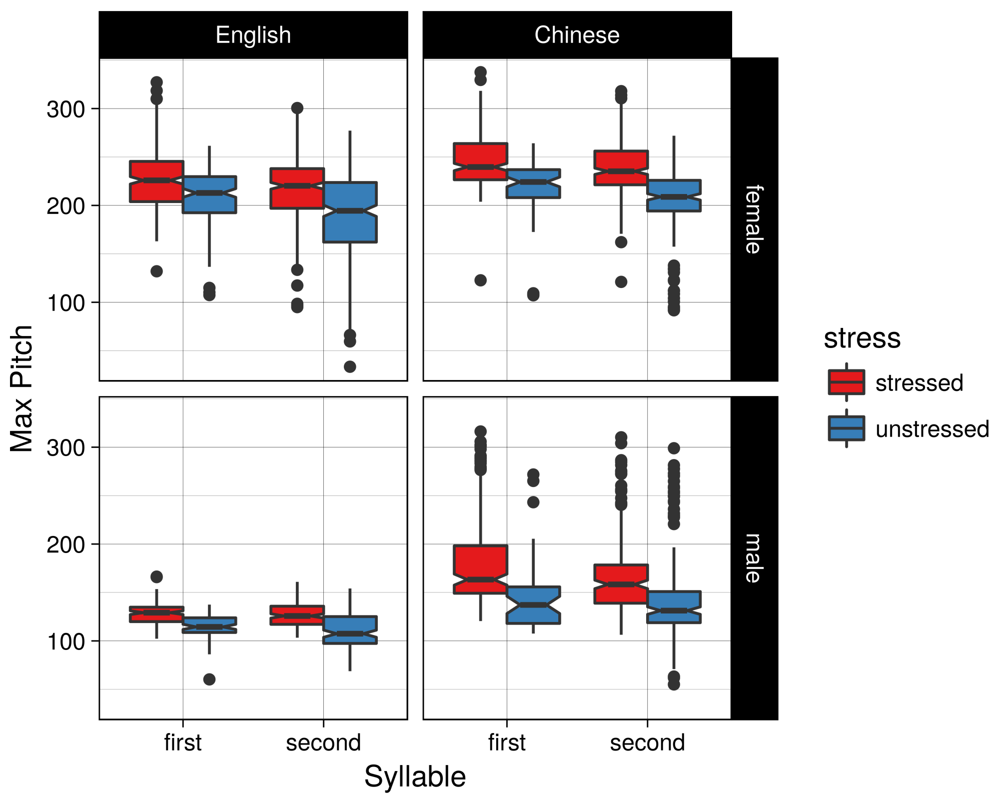
Table 6.8 shows the descriptive statistics for maximum pitch. A boxplot for the maximum pitch values is shown in Figure 6.3.

Maximum pitch was submitted to a mixed-effects analysis. The best model included effects for stress, position, gender, language, and the interaction of stress and position. Unstressed syllables were lower in pitch than stressed syllables ($t = -9.72$, $p < 0.001$), the second syllable was lower in pitch than the first syllable ($t = -4.61$, $p < 0.001$), women had higher pitch than men ($t = 11.1$,

Table 6.8: Descriptive statistics for Experiment 2 maximum pitch (Hz)

language	stress	position	mean	median	st. dev.	minimum	maximum
English	stressed	first	194	204	54	102	327
		second	188	197	50.9	95	301
	unstressed	first	180	194	50.1	60.3	262
		second	160	160	55.6	33.5	277
Chinese	stressed	first	218	228	48.7	120	338
		second	210	220	49	106	318
	unstressed	first	193	210	48.1	107	272
		second	178	188	51.5	55.1	299

Figure 6.3: Experiment 2 maximum pitch by stress, position, gender and language.



p < 0.001) and Chinese speakers had higher pitch than English speakers (t = -3.57, p < 0.001). Unstressed syllables in second position had lower pitch (t = -3.57, p < 0.001).

As the model was done on the F0 values in Herz, it is not surprising that there was an effect for gender or for subject's native language. However, with no interactions with language, this test does not show any evidence for a difference between Mandarin speakers and English speakers in terms of how pitch is used to mark stress.

Table 6.9: Fixed effects for Experiment 2 max pitch model.

	Estimate	Std. Error	DF	T value	p
(Intercept)	183	4.94	64.4	37	4.09E-45
stress –unstressed	-20.5	2.1	993	-9.72	2.08E-21
position –second	-7.28	1.58	973	-4.61	4.62E-06
gender –female	38.3	3.44	59.5	11.1	3.56E-16
language –Chinese	28.8	6.69	59.5	4.3	6.34E-05
stress –unstressed : position –second	-9.44	2.64	993	-3.57	0.000371

Pitch Curves

Figure 6.4 shows a descriptive plot of pitch curves by position, stress, and language in Hertz. Pitch curves in semitones after conversion are shown in Figure 6.5. These pitch curves were submitted to Growth Curve Analysis (Mirman, 2017), using a mixed-effects model with subject as a random effect and vowel and word as within-subjects random effects. Stress, position, language, and gender were investigated as fixed effects for the model.

The fixed effects for the chosen model are given in Table 6.10 The best

Figure 6.4: Descriptive plot of Experiment 2 pitch curves by language, position, and stress. (Hz)

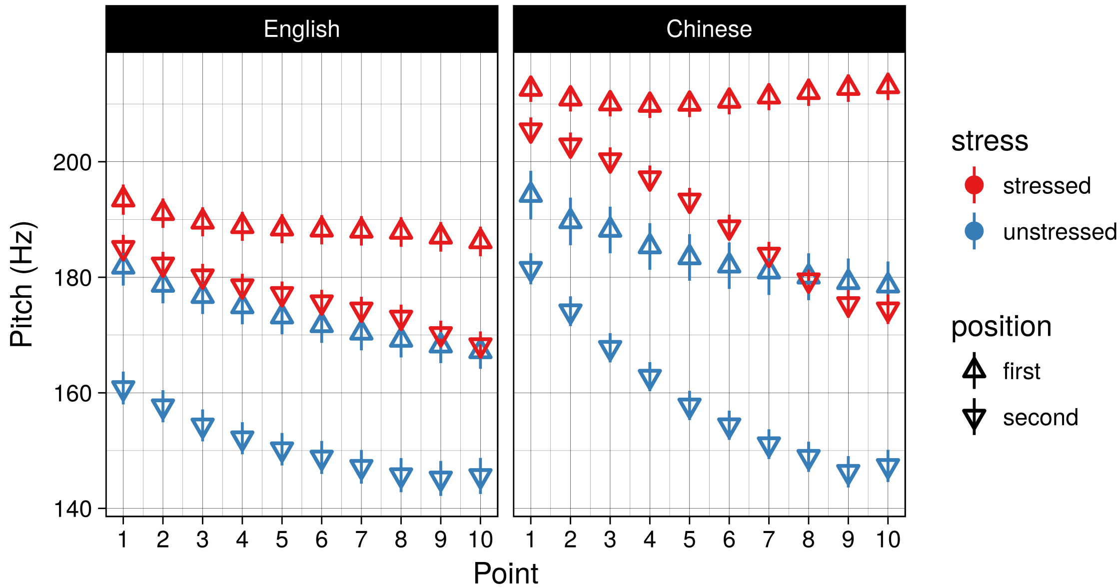
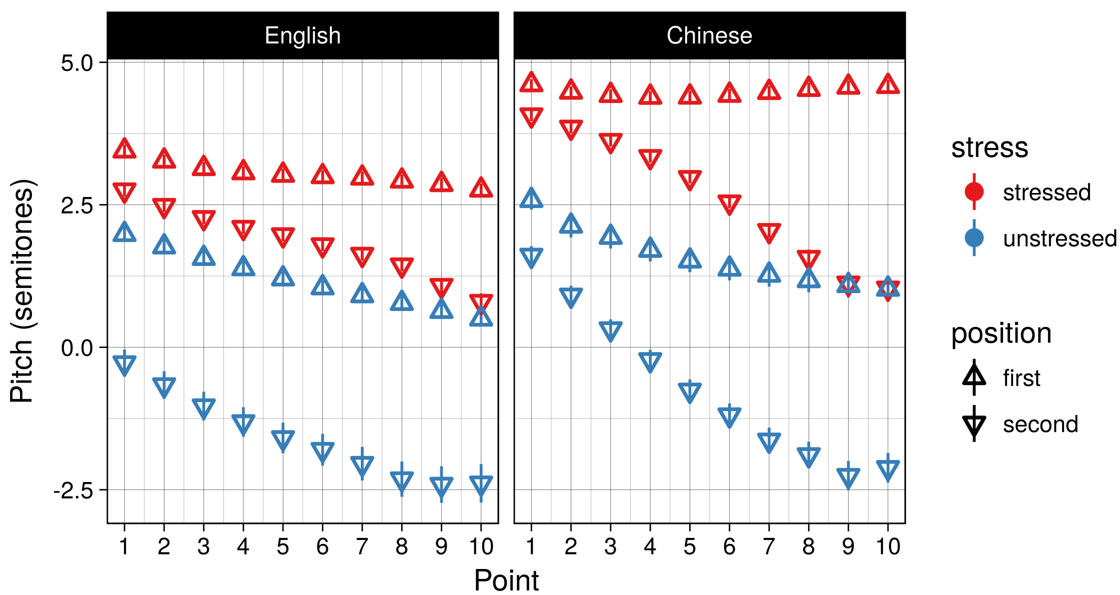


Figure 6.5: Descriptive plot of Experiment 2 pitch curves by language, position, and stress. (semitones)



model had effects for position on the intercept and the linear term, effects for stress on the intercept and quadratic term, and effects for gender and language on the intercept. It also had effects for the interaction of position and stress on the intercept, and on the linear term effects for the interaction of position and language and the interaction of stress and language.

Table 6.10: Fixed Effects for the pitch curve Growth Curve Analysis model in Experiment 2

	Estimate	Std. Error	DF	T value	p
(Intercept)	3.45	0.227	104	15.2	3.18E-28
Linear	-0.589	0.293	103	-2.01	0.0472
Quadratic	-0.0823	0.0692	92.6	-1.19	0.238
position –second	-1.59	0.18	1030	-8.82	4.72E-18
stress –unstressed	-2.29	0.234	1040	-9.76	1.31E-21
gender –male	-0.494	0.135	58.8	-3.67	0.000532
language –Chinese	0.84	0.271	58.3	3.1	0.00296
position –second : stress –unstressed	-1.31	0.288	970	-4.55	5.95E-06
Linear : position –second	-1.22	0.188	1060	-6.5	1.21E-10
Quadratic : stress –unstressed	0.44	0.0859	2160	5.13	3.22E-07
Linear : position –first : language –Chinese	0.322	0.381	92.7	0.845	0.4
Linear : position –second : language –Chinese	-1.54	0.375	86.9	-4.11	8.99E-05
Linear : stress –unstressed : language –English	-0.651	0.195	1030	-3.33	0.000902
Linear : stress –unstressed : language –Chinese	-0.954	0.216	1090	-4.41	1.13E-05

The intercept effects represent the average height of the curve, corresponding to the overall pitch of the vowel. Since the measurement here is in semitones normalized by subject, this corresponds more to the speakers' use of pitch range than to their differences in pitch, as can be seen when comparing the descriptive curves in Hertz (Figure 6.4) to the semitone values (Figure 6.5). In the Hertz values, the difference between English and Mandarin speakers looks much larger.

On the intercept second syllables had lower values than first syllables ($t = -8.82$, $p < 0.001$), and unstressed syllables had lower values than stressed syllables (t = -9.76, p < 0.001). The effect of this can be seen on the model fit graph in Figure 6.6, where we see speakers of both languages pitch stressed syllables higher and that the second syllable for the same stress value tends to be lower. The model also shows an interaction where unstressed syllables in second position have lower intercept values (t = -4.55, p < 0.001).

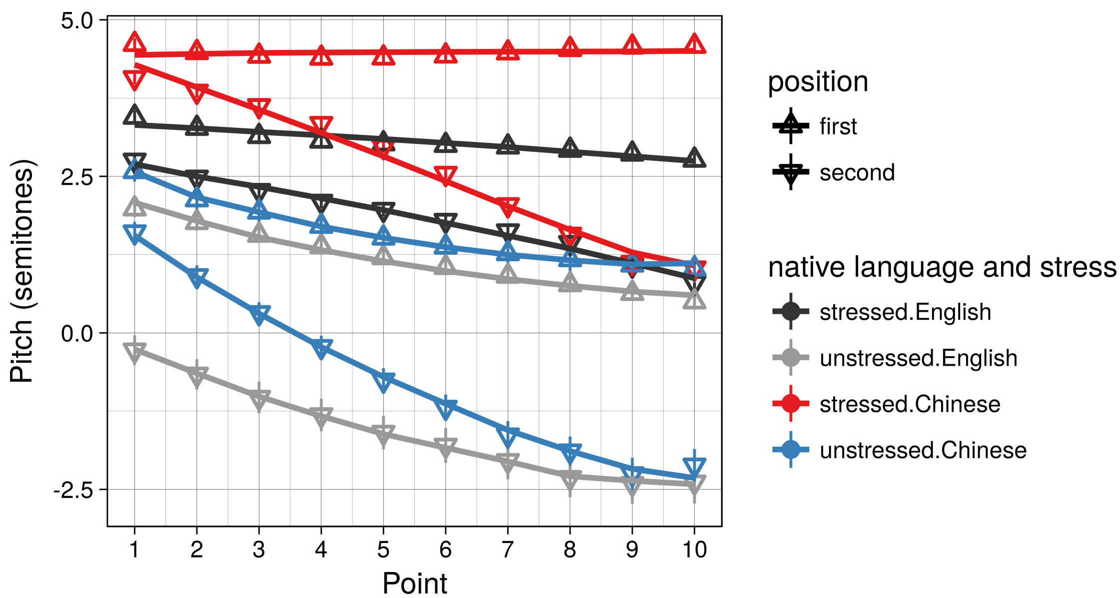


Figure 6.6: Experiment 2 curves model fitted the data by stress, position, and language

There is also a small effect for language, with Chinese speakers having higher values than English speakers (t = 3.1, p < 0.01), and small effect for gender, with male speakers having lower values than female speakers (t = -3.67, p < 0.001).

The quadratic term has only one significant factor, stress, with unstressed syllables having a higher quadratic term than stressed syllables (t = 5.13, p

< 0.001). Looking at the model fit diagram (Figure 6.6) we can see this effect as a tendency for unstressed syllables to have a greater degree of curvature.

The effects on the linear term reference the slope of the pitch curve. There is one simple factor on the linear term, position, with second position having a lower linear term than first syllables ($t = -6.5$, $p < 0.001$). This can be seen on the fit diagram as second syllables having a much steeper downward slope than first syllables.

Aside from position, two interactions were significant on the linear term, the interaction of position and language, and the interaction of stress and language. For the interaction of position and language, second syllables produced by Chinese speakers had a lower linear term ($t = -4.11$, $p < 0.001$), while there was no significant effect on the first syllable ($t < 1$, $p > 0.5$). On the interaction of stress and language, English speakers ($t = -3.33$, $p < 0.001$) and Chinese speakers ($t = -4.41$, $p < 0.001$), but the effect for Chinese speakers was much greater (estimate -0.954 st. dev. 0.216, vs English speakers' estimate -0.651, st. dev. 0.195).

Both of these interactions on the linear term can be seen in the model fit graph (Figure 6.6), and looking at the model fit aids in interpretation. First, we can see that in all cases the second syllable has a much steeper downward curve in comparison to the corresponding first syllable. This difference appears more pronounced in Chinese speakers, but is still visible in the English data. On the stressed side, we see that both languages have steeper downward curves in unstressed syllables. However, another interesting fact is that in initial stressed syllables, though the difference is small, Chinese speakers actually have a flat or slightly upward curve, where English speakers still have a downward curve.

6.3 Discussion

Experiment 2 largely confirms the Experiment 1 (Chapter 5) findings in terms of duration and intensity. Duration was used similarly to native speakers to mark stress, with a difference in the duration measures for final syllables. Intensity once again had no effect for language. Maximum pitch also followed the earlier pattern with Mandarin speakers having similar differences between stressed and unstressed syllables, but overall higher pitch values.

The pitch curve data is somewhat clearer in Experiment 2. Here we have Mandarin speakers producing a high, level pitch for stressed syllables in first position, in contrast to the slightly falling pitch of English speakers. They also have steeper fall for stressed syllables in second position. These characteristics still do not necessarily indicate influence from Mandarin tones. Experiment 3's comparisons will be needed to determine if that is a possible cause of these upward curves.

7 Experiment 3

7.1 Methods

This experiment used the Mandarin subject pool from Experiment 2, listed in Table 6.1. Their English production data from Experiment 2 was reused here, compared with Mandarin production which used the wordlist given in Tables 4.3 and 4.4. See Chapter 4 for more details on the recording procedure.

Since neutral tone syllables only occur in the second syllable, comparing all syllables on all relevant factors as done in Experiment 2 and 3 would lead to a highly unbalanced and likely difficult to interpret model. Due to this highly unbalanced nature, I decided to use several models for each variable in order to understand different parts of the data. First, I created models comparing stressed syllables in English to all full tone syllables in Mandarin. Then, I created two models for unstressed English syllables, one that compared an unstressed first syllable to the four full tones in the first syllable, and one that compared a second-position unstressed syllable to all four tones plus the neutral tone in the second syllable. For maximum pitch and the pitch curve GCA models, I broke the neutral tone into categories according to which tone it followed.

The variable of tone in these models used treatment coding, with the English syllable being tested (stressed or unstressed) as the reference value. This allows me to read the model as comparing each tone to the English syllable type (stressed or unstressed) being tested. This allowed me to see the relationships between the English stressed/unstressed syllable and each tone.

7.2 Results

Duration

Table 7.1 shows descriptive statistics for vowel duration in milliseconds for Chinese syllables by tone and position, while Table 7.2 shows the descriptive statistics for English syllables by stress and position.

Table 7.1: Descriptive Statistics for Experiment 3 vowel duration of Chinese tones

tone	position	mean	median	st. dev.	minimum	maximum
T1	first	157	153	49.3	4.73	307
	second	172	173	66.9	0	443
T2	first	157	157	55.4	49.8	316
	second	186	183	63.3	49.5	380
T3	first	140	135	53.2	32.6	300
	second	150	132	76.8	0	419
T4	first	143	138	51.7	44.8	306
	second	146	148	51.2	42.5	285
T5	second	141	138	51.5	0	377

Figure 7.1 shows a boxplot comparing stressed syllables with each of the four full tones in Mandarin. Table 7.3 shows the fixed effects for the vowel duration model comparing stressed syllables. In the first syllable, all tones were dramatically longer than an English stressed syllable, with difference

Table 7.2: Descriptive statistics for vowel duration of English syllables

stress	position	mean	median	st. dev.	minimum	maximum
stressed	first	95.8	93.3	26.1	44.1	170
	second	141	138	40.7	63	319
unstressed	first	79.4	71.9	29	37	171
	second	125	121	44	44.8	331

estimates ranging from 46.1 ms (third tone, $p < 0.001$) to 64.2 ms (first tone, $p < 0.001$). This indicates that the English stresses in the first syllable were dramatically shorter than any of the Mandarin first syllables.

Figure 7.1: Experiment 3 vowel duration in English stressed syllables and Chinese tones

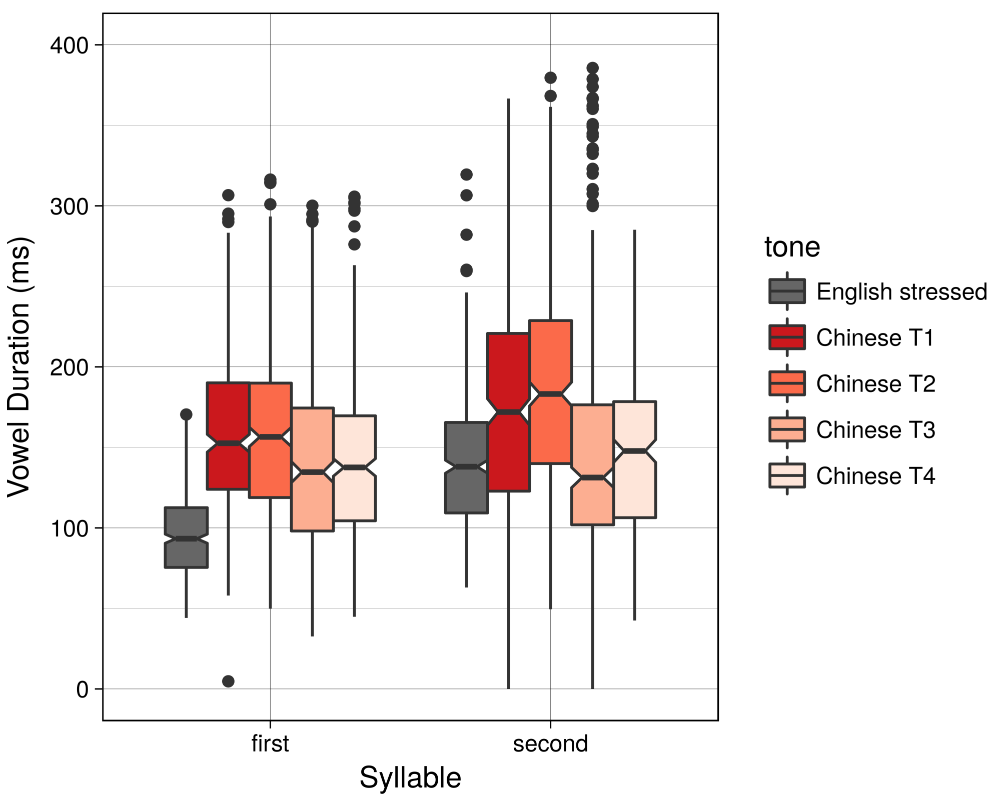


Table 7.3: Fixed effects of the Experiment 3 duration model for English stressed syllables in comparison to full tones

	Estimate	Std. Error	DF	T value	p
(Intercept)	93.9	5.74	53.3	16.4	2.01E-22
position –second	47.3	3.15	1030	15	2.35E-46
position –first : tone 1	64.2	5.14	440	12.5	7.61E-31
position –second : tone 1	30.9	5.09	422	6.06	2.98E-09
position –first : tone 2	63.4	5.15	443	12.3	3.3E-30
position –second : tone 2	44.9	5.09	422	8.82	3.09E-17
position –first : tone 3	46.1	5.16	447	8.94	1.03E-17
position –second : tone 3	9.72	5.09	422	1.91	0.0568
position –first : tone 4	49.5	5.14	440	9.62	5.02E-20
position –second : tone 4	9.03	5.35	495	1.69	0.0918

English stressed syllables in the second position were more comparable to Mandarin syllables. The differences were overall smaller, there were two tones that did not show significant differences with English stressed syllables in final position: third tone (Estimate: 9.72 ms, t = 1.91, p = 0.0568), and fourth tone (Estimate: 9.03 ms, t = 1.69, p = 0.0918).

The dramatically shorter syllables makes comparison on the basis of duration problematic. Stressed syllables in first position simply can't be compared to Mandarin syllables this way, and the similarity of final stresses to tones 3 and 4 in final position is likely to be a coincidence – with the short English syllable ending up around the same length as the shortest of the final tones. Another caveat to consider is that, due to the consonant frames described in Chapter 4, many of the English syllables had a CVC structure, while none of the Mandarin syllables do. This may have reduced the vowel length independently, leading to the overall shorter English syllables.

Figure 7.2 shows a plot of English unstressed syllables compared with all Chinese tones, including the neutral tone, separated by syllable. Separate mixed-effect models were run for first and second syllables in order to avoid issues that may be caused by neutral tone occurring only in final position. Due to the restriction of these models in terms of position, only tone was available as a fixed effect.

Figure 7.2: Experiment 3 vowel duration of English unstressed syllables vs all tones in first and second position.

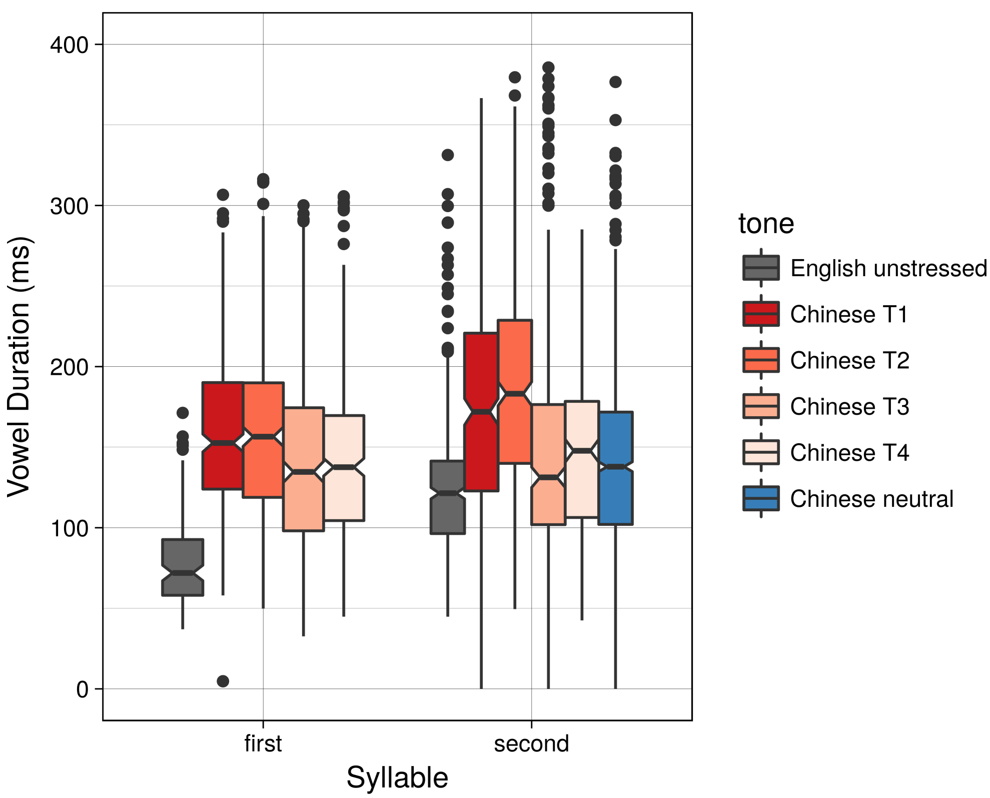


Table 7.4 shows the fixed effects for the model for the first syllable. The parameter ”tone” uses treatment coding, with the English unstressed sylla-

ble as the reference value. The model for the first syllable shows that all tones had longer durations than the unstressed syllable by a large degree, meaning that little can be drawn from this analysis.

Table 7.4: Fixed effects of the Experiment 3 duration model for unstressed syllables in comparison to full tones in the first syllable.

	Estimate	Std. Error	DF	T value	p
(Intercept)	77.2	7.66	65.8	10.1	5.46E-15
tone 1	80.8	5.87	202	13.8	6.83E-31
tone 2	80.4	5.87	203	13.7	1.22E-30
tone 3	63.1	5.88	204	10.7	1.39E-21
tone 4	66.2	5.87	202	11.3	3.42E-23

Table 7.5 shows the fixed effects for the model for the second syllable. Here, we see more interesting results. Although all tones were significantly longer than English unstressed syllable, neutral tone was still on the short side and closest to the English syllable (Est 16.2, error 4.52). However, tone 3 and tone 4 were not that much longer according to the model. These comparisons also still remain questionable due to the overall shortness of the English syllables.

Table 7.5: Fixed effects of the Experiment 3 duration model for unstressed syllables in comparison to all tones in the second syllable

	Estimate	Std. Error	DF	T value	p
(Intercept)	125	6.53	49.6	19.1	1.73E-24
tone 1	47.2	5.3	576	8.89	7.57E-18
tone 2	61.2	5.3	575	11.5	7.4E-28
tone 3	26	5.3	576	4.9	1.26E-06
tone 4	22.1	5.71	672	3.88	0.000116
tone 5	16.2	4.52	359	3.58	0.000395

Intensity

Table 7.6 shows descriptive statistics for intensity in decibels for Chinese syllables, and Table 7.7 shows the descriptive statistics for English syllables.

Table 7.6: Descriptive Statistics for Experiment 3 intensity of Chinese tones in decibels

tone	position	mean	median	st. dev.	minimum	maximum
T1	first	60.8	60.5	7.03	41.8	83.7
	second	57.6	57.3	7.86	36.5	81
T2	first	57.2	56.8	6.41	40.7	78.6
	second	56	55.4	6.72	39.6	81.8
T3	first	54.6	54.4	7.23	37.6	74.3
	second	51.8	51.7	6.86	35.1	68
T4	first	60.5	59.9	6.83	41.8	80.2
	second	56.6	55.7	8.11	38.1	82.7
T5	second	56.9	56.3	6.85	36.4	83.2

Table 7.7: Descriptive Statistics for Experiment 3 intensity of English syllables in decibels

stress	position	mean	median	st. dev.	minimum	maximum
stressed	first	59.7	59.2	6.88	43.7	79.7
	second	58.1	57.5	6.33	43.2	74.5
unstressed	first	56	55.2	5.93	39.5	72.8
	second	54.6	53.8	6.4	40.7	76.7

To normalize intensity across subjects and avoid interference of extraneous variables like distance from the microphone, I converted these base decibel values into a ratio of the measured syllables mean intensity over the mean intensity across all recorded syllables per each subject. The descriptive statistics for this ratio for Chinese syllables are given in Table 7.8, and the ratio for English syllables is given in 7.9. This ratio is what was submitted to subsequent mixed-effects analyses.

Table 7.8: Descriptive Statistics for the Experiment 3 ratio of syllable intensity to per subject mean intensity for Chinese tones

tone	position	mean	median	st. dev.	minimum	maximum
T1	first	1.07	1.06	0.149	0.745	1.56
	second	1.02	1.02	0.168	0.623	1.52
T2	first	1.01	0.999	0.144	0.723	1.52
	second	0.99	0.97	0.149	0.702	1.44
T3	first	0.965	0.949	0.155	0.641	1.47
	second	0.918	0.905	0.156	0.583	1.41
T4	first	1.07	1.06	0.146	0.747	1.51
	second	1	0.997	0.169	0.668	1.5
T5	second	1.01	0.994	0.151	0.661	1.52

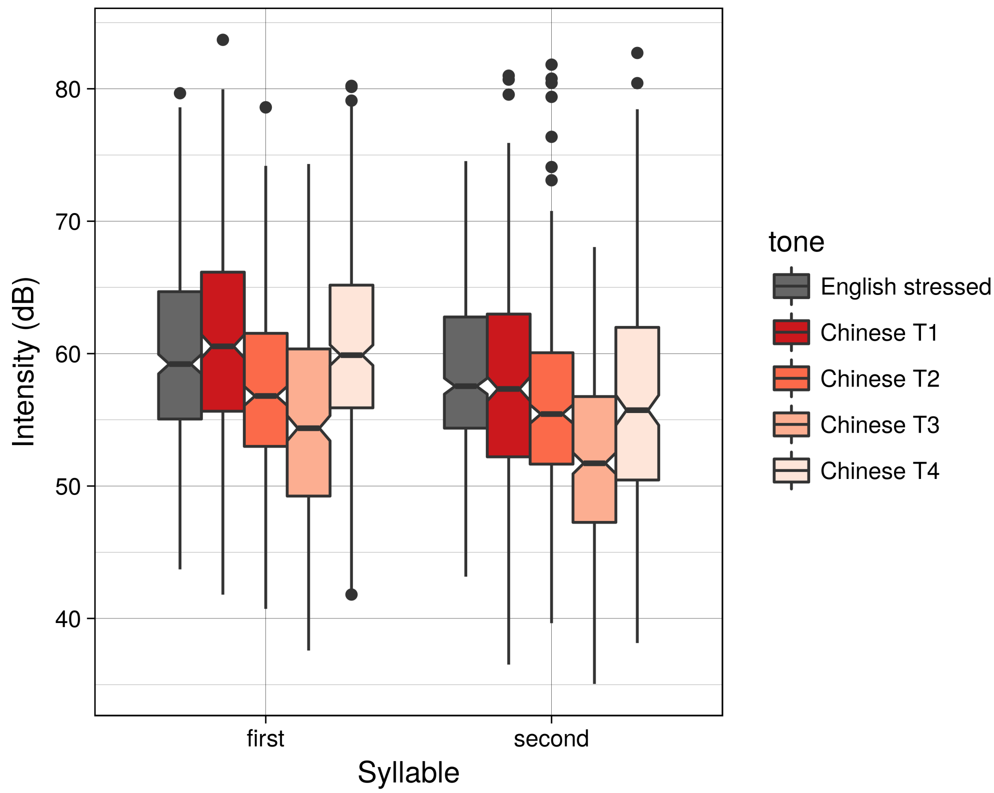
Table 7.9: Descriptive Statistics for the Experiment 3 ratio of syllable intensity to per subject mean intensity for English syllables

stress	position	mean	median	st. dev.	minimum	maximum
stressed	first	1.05	1.05	0.141	0.75	1.42
	second	1.03	1.03	0.135	0.738	1.38
unstressed	first	0.989	0.988	0.134	0.743	1.3
	second	0.965	0.958	0.133	0.687	1.37

Boxplots for English stressed syllables compared to Mandarin full tones are given in Figure 7.3 for raw decibels and 7.4 for the above described ratio.

English stressed syllables were compared with Chinese full tones in a mixed effect analysis with subject as a random effect and vowel and word as within-subjects random effects. The best model has an effect for position and an effect for the interaction of position and category. Second syllables were slightly lower in intensity ($t = -4.41$, $p < 0.001$) than first syllables. In

Figure 7.3: Experiment 3 Intensity (in dB) by position and syllable type.

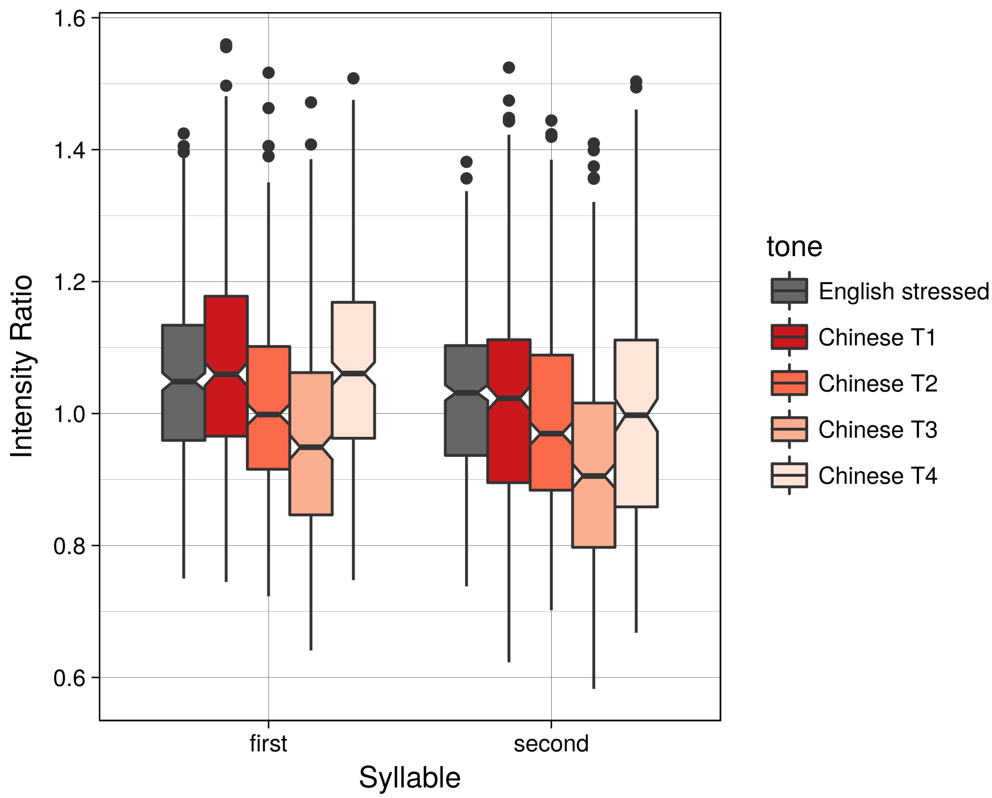


comparing English syllables to tones, in the second position, English stressed syllables were not significantly different from first tone (p = 0.299). Also, in neither position were English stressed syllables different from fourth tone syllables.

. Figure 7.5 shows a boxplot comparing the intensity (in Hz) of English syllables to Chinese syllables, including neutral tone, separated by syllable position. Figure 7.6 shows this comparison in the ratio of syllable intensity to per-subject mean intensity.

Table 7.11 shows the fixed effects of the model comparing English un-

Figure 7.4: Experiment 3 ratio of syllable intensity to per subject mean intensity by position and syllable type.



stressed syllables with tones in the first position. Tone 3 was not significantly different in intensity from the English unstressed syllable (p > 0.1).

Table 7.12 shows the fixed effects for the model comparing English unstressed syllables to all Chinese tones in second position. In this position, all of the tones had a significant difference from the English unstressed syllable.

Table 7.10: Fixed effects for the Experiment 3 mixed-effects model for the ratio of syllable intensity to per subject mean intensity comparing English stressed syllables to Chinese tones.

	Estimate	Std. Error	DF	T value	p
(Intercept)	1.05	0.0233	33	45	3.57E-31
position –second	-0.0217	0.00492	1050	-4.41	1.14E-05
position –first : tone 1	0.0248	0.00915	375	2.72	0.00693
position –second : tone 1	-0.00946	0.00909	365	-1.04	0.299
position –first : tone 2	-0.0368	0.00916	377	-4.02	7.06E-05
position –second : tone 2	-0.0379	0.00909	364	-4.17	3.81E-05
position –first : tone 3	-0.0838	0.00918	380	-9.12	4.33E-18
position –second : tone 3	-0.11	0.00909	365	-12.1	1.09E-28
position –first : tone 4	0.0192	0.00915	375	2.1	0.0364
position –second : tone 4	-0.0182	0.00944	417	-1.93	0.0545

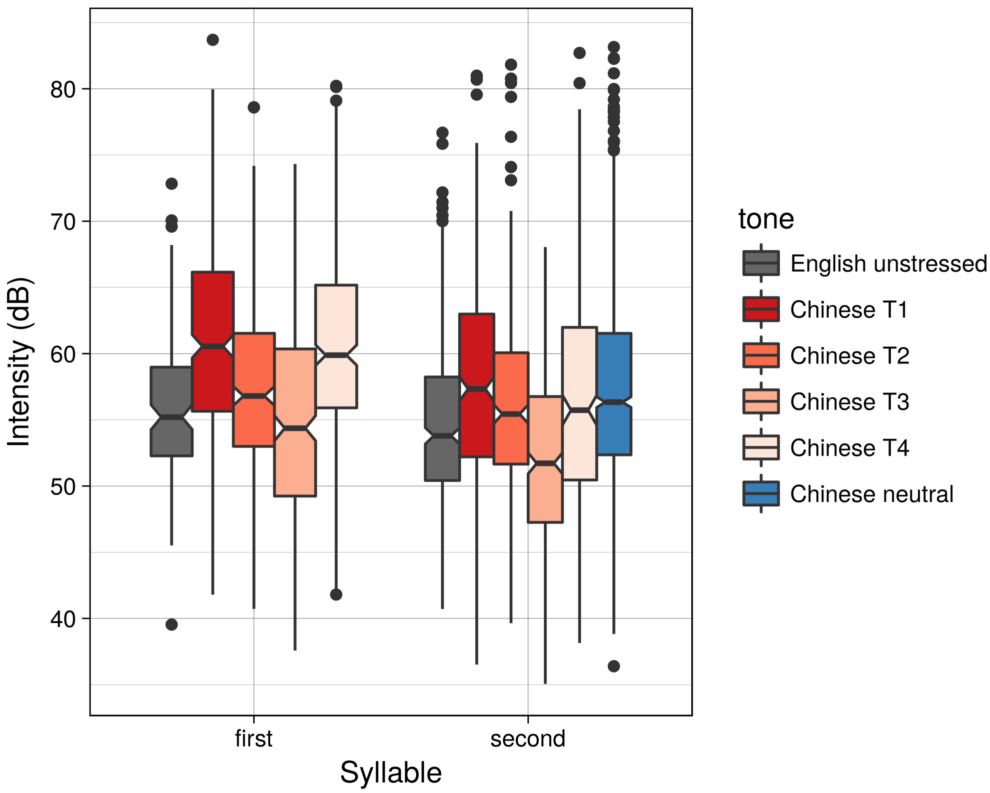
Table 7.11: Fixed effects for the Experiment 3 intensity model for comparison with English unstressed syllables in first position

	Estimate	Std. Error	DF	T value	p
(Intercept)	0.985	0.0246	40.5	40	3.5E-34
tone 1	0.0893	0.0121	225	7.37	3.18E-12
tone 2	0.0277	0.0121	226	2.29	0.0232
tone 3	-0.0196	0.0121	227	-1.61	0.108
tone 4	0.0837	0.0121	225	6.91	4.79E-11

Table 7.12: Fixed effects for the Experiment 3 intensity model for comparison with English unstressed syllables in second position

	Estimate	Std. Error	DF	T value	p
(Intercept)	0.966	0.0234	33.2	41.2	4.11E-30
tone 1	0.0516	0.0095	376	5.44	9.74E-08
tone 2	0.0232	0.00949	375	2.44	0.0151
tone 3	-0.0492	0.0095	376	-5.19	3.53E-07
tone 4	0.0396	0.00992	433	3.99	7.66E-05
tone 5	0.0379	0.00873	279	4.34	2.02E-05

Figure 7.5: Experiment 3 intensity (in dB) in English unstressed syllables and Chinese tones



Maximum Pitch

Table 7.13 shows descriptive statistics for maximum pitch, in Hertz, for Chinese syllables, and Table 7.14 shows the statistics for English syllables.

Figure 7.7 shows a boxplot comparing English stressed syllables to Mandarin full tones. This comparison was submitted to a mixed-effects analysis with subject as a random effect and word and vowel as within-subjects random effects. The fixed effects for the model are given in Table 7.15.

English stressed syllables were most similar to first tone overall (Esti-

Figure 7.6: Experiment 3 intensity ratio in English unstressed syllables and Chinese tones

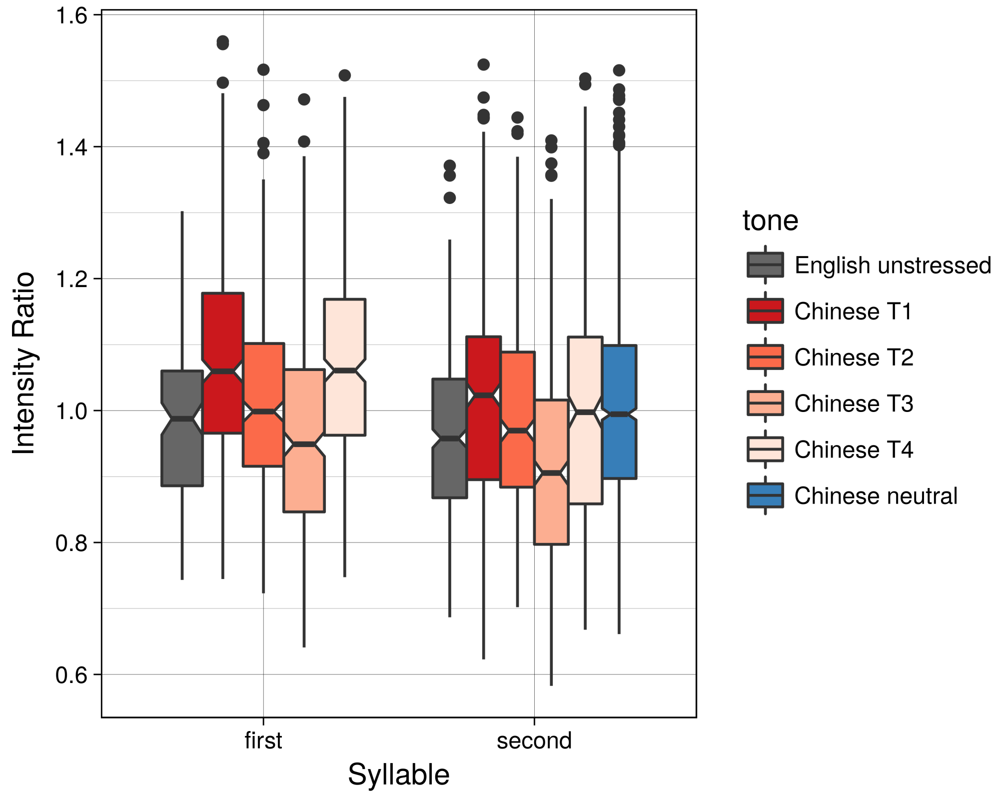


Table 7.13: Descriptive statistics for Experiment 3 maximum pitch for Chinese tones

tone	position	mean	median	st. dev.	minimum	maximum
T1	first	241	250	51.7	112	353
	second	197	212	59.7	37.7	328
T2	first	191	191	52.4	87.3	376
	second	175	187	48.5	50.2	334
T3	first	166	178	46.2	54.3	311
	second	143	138	50.5	41	288
T4	first	251	264	54.8	129	441
	second	193	197	61.2	35.7	377
T5	second	193	199	58.4	45.8	396

Table 7.14: Descriptive statistics for Experiment 3 maximum pitch for English syllables

stress	position	mean	median	st. dev.	minimum	maximum
stressed	first	218	228	48.7	120	338
	second	210	220	49	106	318
unstressed	first	193	210	48.1	107	272
	second	178	188	51.5	55.1	299

Figure 7.7: Maximum pitch by position and syllable type.

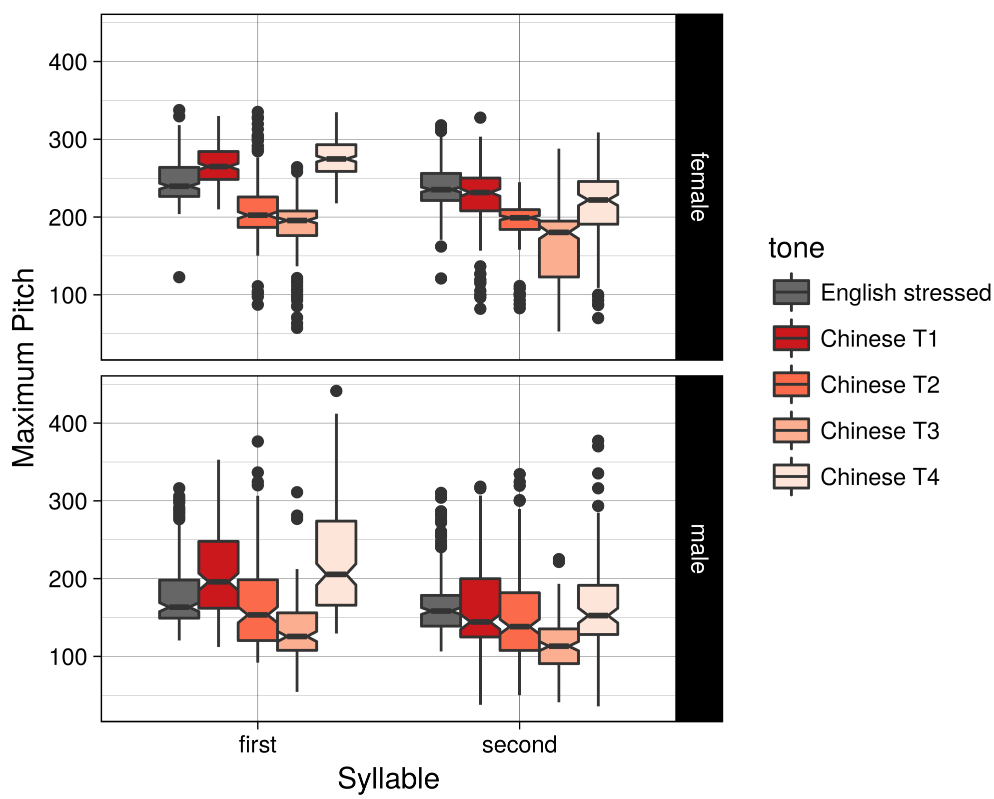


Table 7.15: Fixed effects for the Experiment 3 mixed-effects model for maximum pitch comparing English stressed syllables to full tones.

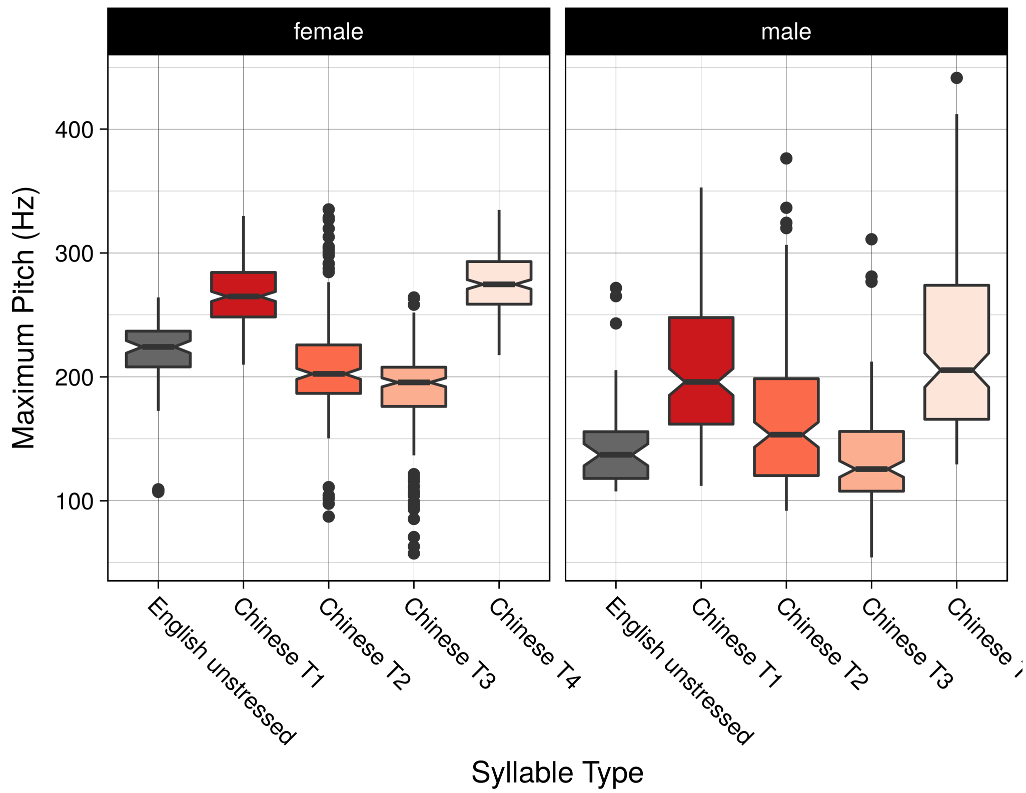
	Estimate	Std. Error	DF	T value	p
(Intercept)	213	6.21	37.9	34.4	3.7E-30
tone 1	23	3.32	1100	6.94	6.54E-12
tone 2	-24.2	3.33	1100	-7.28	6.32E-13
tone 3	-52.8	3.37	1120	-15.7	2.83E-50
tone 4	34.6	3.32	1100	10.4	2.66E-24
position –second	-7.51	2.94	1070	-2.56	0.0107
gender –female	33.1	6.01	33.2	5.5	4.15E-06
tone 1 : position –second	-36.2	4.45	1040	-8.12	1.28E-15
tone 2 : position –second	-9.11	4.46	1040	-2.04	0.0414
tone 3 : position –second	-15.7	4.5	1060	-3.49	0.000511
tone 4 : position –second	-53	4.69	1060	-11.3	5.47E-28
tone 1 : gender –female	-1.14	2.39	627	-0.477	0.633
tone 2 : gender –female	-12.5	2.4	630	-5.22	2.38E-07
tone 3 : gender –female	-7.34	2.42	650	-3.03	0.00252
tone 4 : gender –female	-6.29	2.49	683	-2.52	0.0119

mate: 23, $p < 0.001$) with no significant difference for female speakers ($p = 0.633$). In the second position, English stressed syllables were similar to second tone (Estimate: -9.11, $p = 0.0414$). Generally, all of this indicates that English stressed syllables are consistently produced with a high pitch by Mandarin speakers.

Figure 7.8 shows a plot of the maximum pitch values for unstressed syllables in first position. Figure 7.9 shows a plot for the maximum pitch values in second position.

Table 7.16 shows the fixed effects for the maximum pitch model comparing unstressed English syllables to Mandarin tones in the first position. Tone 2 is not significantly different in maximum pitch from English unstressed

Figure 7.8: Experiment 3 maximum pitch for unstressed syllables in the first position.



syllables ($p > 0.1$).

Table 7.17 shows the fixed effects for the maximum pitch model comparing unstressed English syllables to Mandarin tones in the second position. Again, tone 2 is not significantly different in maximum pitch from English unstressed syllables ($p > 0.1$).

Figure 7.9: Experiment 3 maximum pitch for unstressed syllables in the second position.

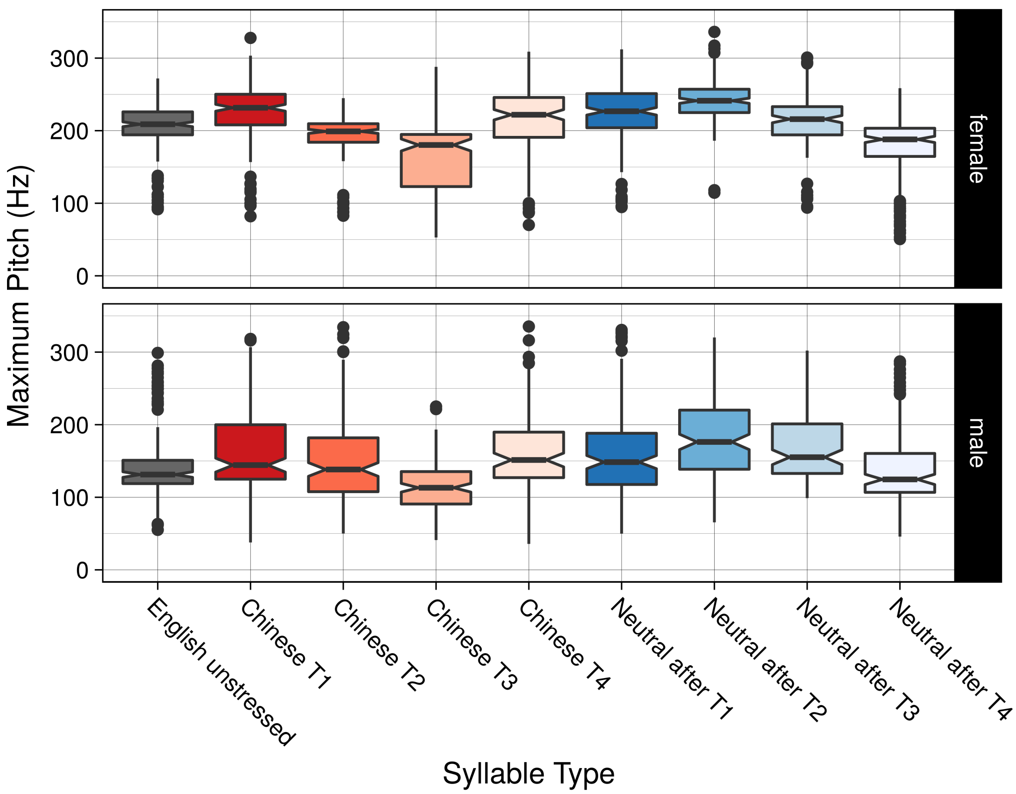


Table 7.16: Fixed effects for the Experiment 3 model for maximum pitch of unstressed syllables in the first position.

	Estimate	Std. Error	DF	T value	p
(Intercept)	211	8.94	42.2	23.6	5.75E-26
tone 1	53.1	4.52	445	11.7	6.43E-28
tone 2	4.36	4.53	446	0.964	0.336
tone 3	-23.7	4.55	450	-5.2	3E-07
tone 4	64	4.52	445	14.1	8.72E-38
gender –male	-54.2	12.4	29.9	-4.36	0.000142

Table 7.17: Fixed effects for the Experiment 3 model for maximum pitch of unstressed syllables in the second position.

	Estimate	Std. Error	DF	T value	p
(Intercept)	201	8.4	34.9	23.9	2.9E-23
tone 1	18.5	3.62	1000	5.1	4.11E-07
tone 2	-3.28	3.62	998	-0.906	0.365
tone 3	-37.8	3.65	1030	-10.4	5.47E-24
tone 4	12.6	3.97	1000	3.17	0.00159
tone 5 after T1	19.5	3.62	1000	5.39	8.65E-08
tone 5 after T2	39.3	3.63	1000	10.8	6.9E-26
tone 5 after T3	17.8	3.95	1000	4.5	7.63E-06
tone 5 after T4	-20.4	3.64	1010	-5.61	2.64E-08
gender –male	-52	12.3	30	-4.23	0.0002

Pitch Curves

Figure 7.10 shows mean curves for the four tones and for English stressed syllables in Hertz, and Figure 7.11 shows the same value for semitones. Both figures separated the curves by position and speaker gender. In all configurations, English stressed syllables in the first position show a markedly flat somewhat upward curve, while in the second position they follow a downward curve.

The semitone curves were submitted to a Growth Curve Analysis (Mirman, 2017), using a mixed-effects model with subject as a random effect and vowel and word as within-subjects random effects. When looking at this model, the most important effects to find are between English syllables and Mandarin tones, and particularly relationships between English stressed syllables in first position and the first tone, or between English stressed syllables in second position and the fourth tone, as visual inspection indicates

Figure 7.10: Descriptive curves for Experiment 3 stressed syllables in Hz.

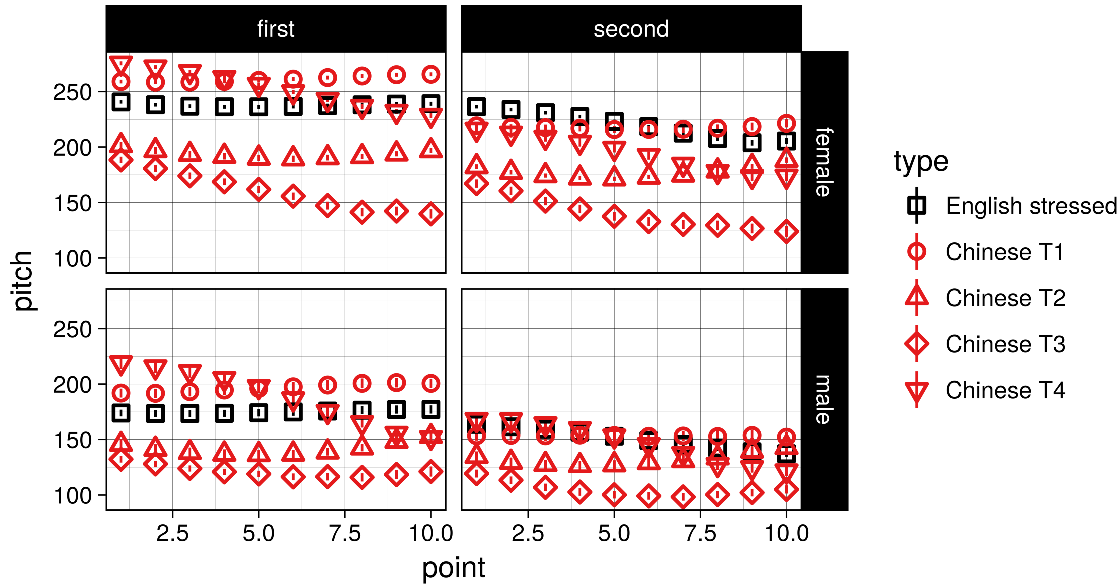
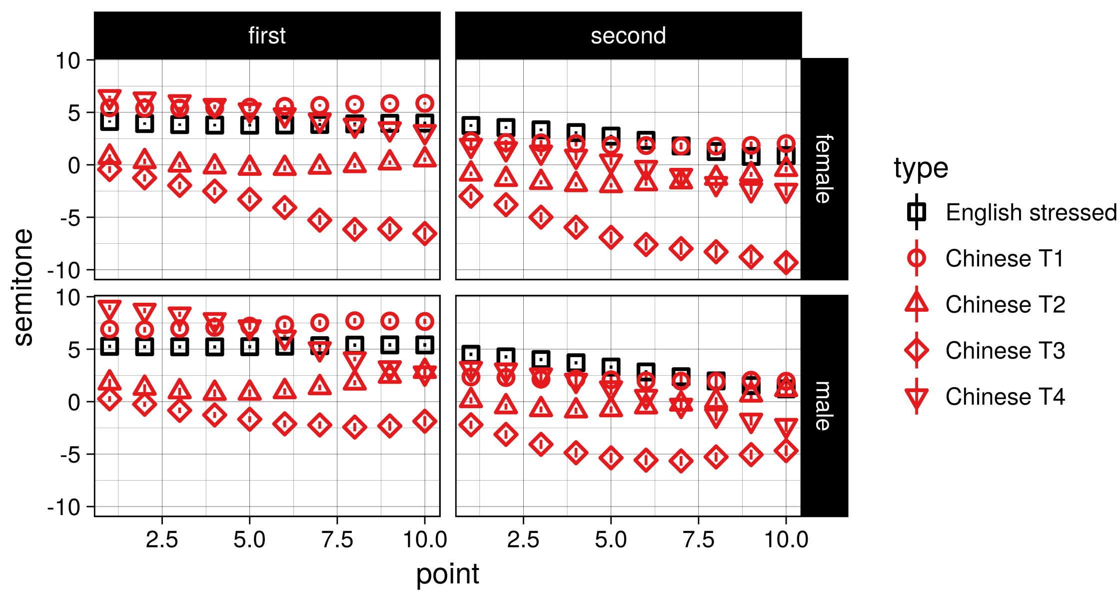


Figure 7.11: Descriptive curves for Experiment 3 stressed syllables in semitones



that these would be the closest tones in each scenario.

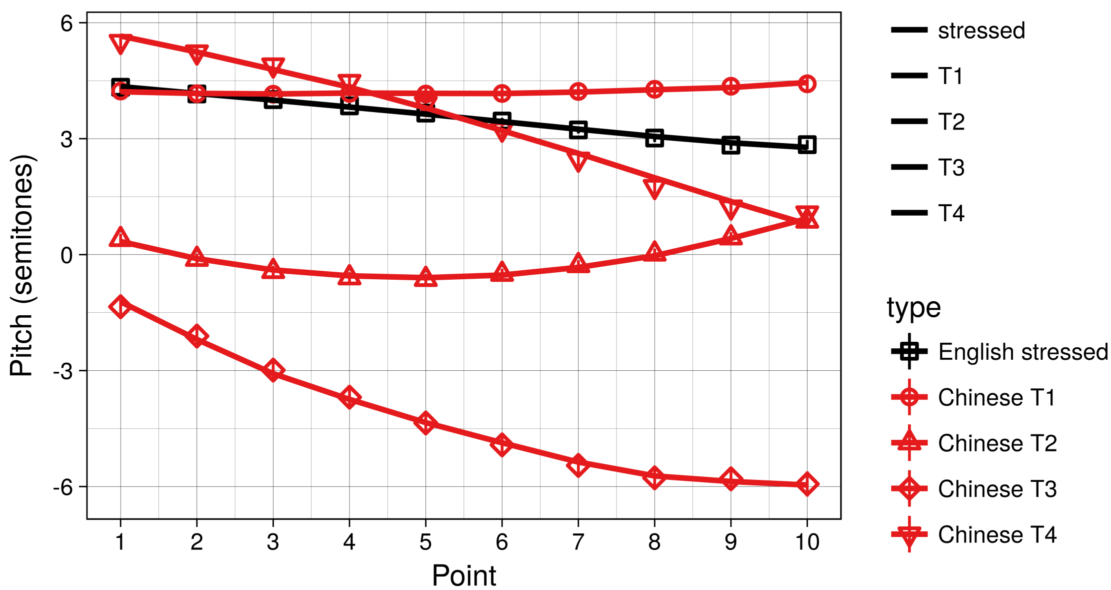
Table 7.18: Fixed effects of the Experiment 3 Growth Curve Analysis model for stressed syllables.

	Estimate	Std. Error	DF	T value	p
(Intercept)	4.58	0.31	168	14.8	3.08E-32
Linear	0.0589	0.109	32100	0.542	0.588
Quadratic	-0.0641	0.0758	32100	-0.846	0.398
tone 1	1.79	0.384	1050	4.65	3.72E-06
tone 2	-3.84	0.385	1050	-9.97	1.92E-22
tone 3	-7.6	0.392	1070	-19.4	4.72E-72
tone 4	0.949	0.384	1050	2.47	0.0137
position –second	-1.91	0.337	1100	-5.67	1.82E-08
gender –1	-0.6	0.199	30	-3.02	0.00514
tone 1 : position –second	-2.61	0.51	1090	-5.11	3.78E-07
tone 2 : position –second	0.378	0.511	1090	0.739	0.46
tone 3 : position –second	-0.95	0.523	1100	-1.82	0.0696
tone 4 : position –second	-3.57	0.541	1100	-6.6	6.44E-11
Linear : tone 1	0.671	0.161	32100	4.17	3.1E-05
Linear : tone 2	0.502	0.164	32100	3.06	0.00219
Linear : tone 3	-4.73	0.176	32100	-26.9	5.02E-158
Linear : tone 4	-5.27	0.162	32100	-32.5	1.76E-227
Linear : position –second	-3.48	0.152	32100	-22.9	1.35E-115
Quadratic : tone 1	0.165	0.114	32100	1.45	0.146
Quadratic : tone 2	1.39	0.115	32100	12	3.08E-33
Quadratic : tone 3	1.16	0.125	32100	9.31	1.38E-20
Quadratic : tone 4	-0.409	0.121	32100	-3.39	0.00071
Linear : tone 1 : position –second	2.36	0.228	32100	10.4	4.56E-25
Linear : tone 2 : position –second	3.89	0.231	32100	16.9	1.89E-63
Linear : tone 3 : position –second	3.83	0.25	32100	15.3	8.29E-53
Linear : tone 4 : position –second	3.07	0.246	32100	12.5	1.2E-35
Linear : stressed (English) gender –1	0.00308	0.0766	32100	0.0402	0.968
Linear : tone 1 : gender –1	-0.113	0.0858	32100	-1.31	0.189
Linear : tone 2 : gender –1	-0.538	0.088	32100	-6.11	1.03E-09
Linear : tone 3 : gender –1	-1.88	0.1	32100	-18.8	4.75E-78
Linear : tone 4 : gender –1	1.28	0.0956	32100	13.4	6.44E-41

Table 7.3 shows the fixed effects for the model chosen. The best model had the following effects: on the intercept, tone, position, and gender were

all factors, as was the interaction of tone and position. On the Linear term, tone and position were factors, as well as the interaction of tone and position, and the interaction of tone and gender. On the quadratic term, only tone was significant. A plot of the fit lines for this model can be seen in Figure 7.12.

Figure 7.12: Experiment 3 fitted GCA model for stressed syllables.



On the intercept all four tones differed from English syllables, with first tone (Est. 1.79, $p < 0.001$) and fourth tone (Est. 0.949, $p = 0.0137$) being the closest in overall height, though in the interaction between tone and position, second tone ($p = 0.46$) and third tone ($p = 0.0696$) in second position were not significantly different from English stressed syllables. On the linear term, we tend to see first tone being close to the English stressed syllable, with a smaller difference in the overall than when considering the second syllable. There is apparently no significant difference on the quadratic term

($p = 0.146$).

When looking at the model fit in Figure 7.12, it certainly appears that English stressed syllables in first position have a similar pitch contour to first tone syllables, while in second position they have more of a downward trend. However, these stressed syllables are still clearly somewhat lower in pitch than first tone syllables. It's also not clear what English stresses in second position would correspond to. While they do have a downward curve that most resembles a fourth tone, their relationship to the corresponding first syllable is not like that of fourth and first tones. Considering this would also raise the question of why first syllables and second syllables would be assigned different tones.

Unstressed syllables were submitted to separate mixed-effects models for initial and final syllables, with gender and tone as fixed effects, subject as a random effect, and vowel and word as within subjects random effects.

The best model for first syllables included tone and gender on the intercept, tone and the interaction of tone and gender on the linear term, and tone on the quadratic term. The fixed effects for the model are given in Table 7.19, and the fitted model is presented in Figure 7.13.

Focusing on the tone relationships, English unstressed syllables did not have a significantly different intercept from tone 2 syllables ($p = 0.0567$), but they were significantly different from all other syllables on the linear term. They were also not significantly different from first tone syllables ($p = 0.288$) or third tone syllables ($p = 0.135$) on the quadratic term.

Table 7.19: Fixed effects of the Experiment 3 Growth Curve Analysis model comparing Mandarin tones to unstressed English syllables in first postion.

	Estimate	Std. Error	DF	T value	p
(Intercept)	0.923	0.433	183	2.13	0.0344
Linear	-1.4	0.229	13900	-6.11	1.03E-09
Quadratic	0.273	0.184	13900	1.48	0.139
tone 1	4.75	0.455	410	10.4	9.23E-23
tone 2	-0.871	0.455	411	-1.91	0.0567
tone 3	-4.67	0.459	416	-10.2	7.21E-22
tone 4	3.91	0.455	411	8.59	1.83E-16
gender –male	1.45	0.392	31	3.69	0.000858
Linear : tone 1	1.92	0.271	13900	7.1	1.29E-12
Linear : tone 2	1.26	0.272	13900	4.63	3.65E-06
Linear : tone 3	-5.29	0.283	13900	-18.7	2.32E-77
Linear : tone 4	-2.22	0.271	13900	-8.19	2.77E-16
Quadratic : tone 1	-0.227	0.214	13900	-1.06	0.288
Quadratic : tone 2	0.99	0.216	13900	4.59	4.45E-06
Quadratic : tone 3	0.334	0.223	13900	1.5	0.135
Quadratic : tone 4	-0.758	0.215	13900	-3.53	0.000417
Linear : English unstressed : gender –male	0.0579	0.386	13900	0.15	0.881
Linear : tone 1 : gender –male	0.445	0.219	13900	2.03	0.0421
Linear : tone 2 : gender –male	1.46	0.226	13900	6.43	1.31E-10
Linear : tone 3 : gender –male	4.06	0.255	13900	15.9	1.92E-56
Linear : tone 4 : gender –male	-3.3	0.222	13900	-14.8	1.7E-49

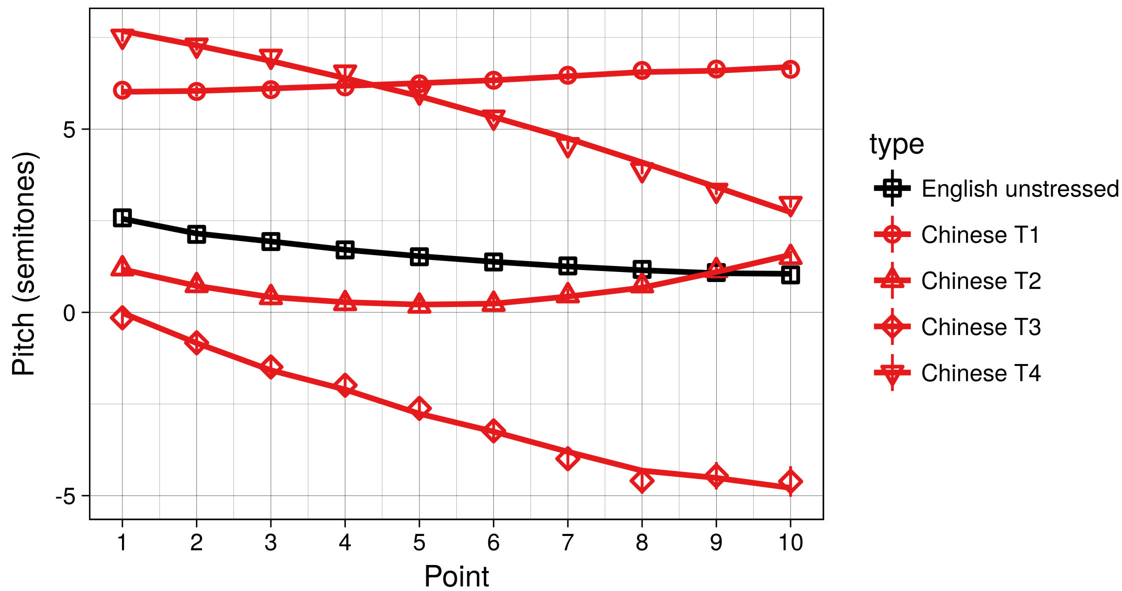
As with maximum pitch, the Growth Curve Analysis model for pitch curves in the second syllable separates neutral tone into four variants after each of the four full tones. The best model included tone on the intercept, tone, and the interaction of tone and gender on the linear term, and tone on the quadratic term. The fixed effects for the mixed-effects model are given in Table 7.20 and the fitted model is shown in Figure 7.14.

These unstressed syllables were not significantly different from second tone on the intercept ($p = 0.91$). They were not significantly different from fourth tone on the linear term ($p = 0.984$), nor were they different on the linear term from neutral tones after a first tone ($p = 0.146$). On the quadratic

Table 7.20: Fixed effects of Growth Curve Analysis model comparing Mandarin tones to unstressed English syllables in second postion.

	Estimate	Std. Error	DF	T value	p
(Intercept)	-0.941	0.417	91.9	-2.26	0.0264
Linear	-4.83	0.199	25700	-24.2	5.28E-128
Quadratic	0.629	0.149	25700	4.23	2.34E-05
tone 1	2.69	0.437	926	6.15	1.17E-09
tone 2	0.0494	0.437	926	0.113	0.91
tone 3	-5.04	0.448	942	-11.3	1.19E-27
tone 4	0.873	0.484	944	1.8	0.0714
tone 1-T5	1.17	0.44	931	2.66	0.00793
tone 2-T5	4.01	0.439	928	9.13	4.26E-19
tone 3-T5	2.5	0.478	937	5.23	2.09E-07
tone 4-T5	-2	0.443	932	-4.51	7.45E-06
Linear : tone 1	4.42	0.283	25700	15.6	1.5E-54
Linear : tone 2	5.42	0.284	25700	19.1	1.86E-80
Linear : tone 3	-1.24	0.307	25700	-4.03	5.53E-05
Linear : tone 4	0.00623	0.317	25700	0.0197	0.984
Linear : neutral after T1	-0.423	0.291	25700	-1.45	0.146
Linear : neutral after T2	2.18	0.284	25700	7.68	1.65E-14
Linear : neutral after T3	3.13	0.31	25700	10.1	5.63E-24
Linear : neutral after T4	1.43	0.295	25700	4.85	1.24E-06
Quadratic : tone 1	-0.468	0.213	25700	-2.2	0.028
Quadratic : tone 2	0.752	0.215	25700	3.5	0.000468
Quadratic : tone 3	1.01	0.233	25700	4.33	1.49E-05
Quadratic : tone 4	-1.08	0.242	25700	-4.47	8.02E-06
Quadratic : neutral after T1	-0.0504	0.219	25700	-0.23	0.818
Quadratic : neutral after T2	-0.561	0.216	25700	-2.6	0.00928
Quadratic : neutral after T3	-0.719	0.234	25700	-3.08	0.00211
Quadratic : neutral after T4	-0.2	0.223	25700	-0.9	0.368
Linear : English unstressed : gender –male	1.78	0.299	25700	5.95	2.7E-09
Linear : tone 1 : gender –male	-0.0093	0.309	25700	-0.03	0.976
Linear : tone 2 : gender –male	0.681	0.315	25700	2.16	0.031
Linear : tone 3 : gender –male	3.45	0.365	25700	9.44	4.17E-21
Linear : tone 4 : gender –male	-1.34	0.391	25700	-3.42	0.000617
Linear : neutral after T1 : gender –male	-0.803	0.327	25700	-2.46	0.0139
Linear : neutral after T2 : gender –male	-2.33	0.32	25700	-7.28	3.47E-13
Linear : neutral after T3 : gender –male	0.744	0.366	25700	2.03	0.0419
Linear : neutral after T4 : gender –male	2.04	0.336	25700	6.06	1.35E-09

Figure 7.13: Fitted curves of the Experiment 3 Growth Curve Analysis model comparing Mandarin tones to unstressed English syllables in first postion.



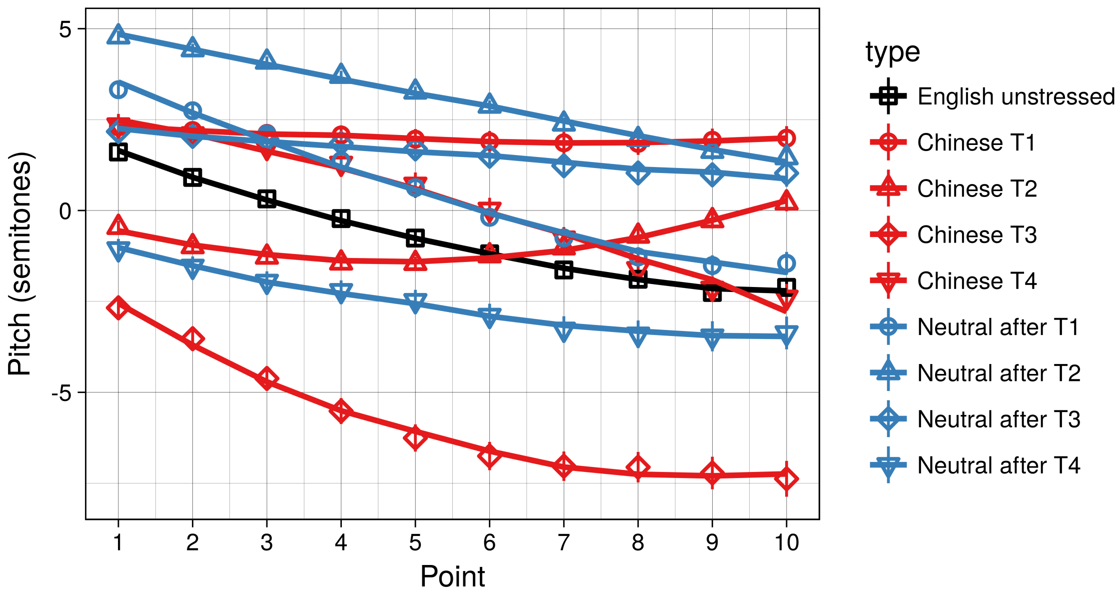
term, they were not significantly different from neutral tone after a first tone (p = 0.818) or after a fourth tone (p = 0.368).

7.3 Discussion

Some of the evidence in this chapter causes issues with interpretation. The duration data is not easy to compare, since the English syllables are so much shorter than Chinese production. Intensity also does not show a clear comparison.

Max pitch and pitch curves, however, show some interesting features. Stressed syllables are high in pitch, and have a high, flat contour in the

Figure 7.14: Fitted curves of Growth Curve Analysis model comparing Mandarin tones to unstressed English syllables in second postion.



first syllable in the first syllable and a falling contour in the second syllable. However, it's not entirely clear that the English stressed syllables are that close to tonal contours. More interestingly, the unstressed syllables do not match any particular tone, but visual examination of the contours suggest they have a target very similar to the neutral tone variations.

In the next chapter, I will discuss the evidence from all three of these experiments and examine them in the context of the hypotheses given Chapter 3. In Chapter 9 I will draw my conclusions and examine potential future directions, including further improvements to the protocol of these experiments.

8 General Discussion

8.1 Review of experimental Results

Experiments 1 and 2 indicated that in terms of intensity and duration, Mandarin speaking subjects performed very similarly to English speaking subjects. There was no difference found for intensity, and the duration measurements were mainly different in terms of the difference in length between first and second syllables. Both English speakers and Mandarin speaking learners produced louder stressed syllables than unstressed syllables, and speakers of both languages also produced longer stressed syllables than unstressed syllables. Mandarin speakers just lengthened final syllables more than English speakers.

On pitch measurements, there were interesting differences, however. While speakers of both languages had higher maximum pitch in stressed syllables than unstressed syllables, there was a significant difference in the pitch contours of those syllables. Mandarin speakers appeared to have a flatter contour for stressed syllables in initial position than English speakers did.

Experiment 3's success in sorting out these issues was limited. Unfortunately, the short length of the vowels Mandarin speakers produced in the English tokens made comparisons in duration problematic.

Intensity had interesting results. English stressed syllables were closest in intensity to first and fourth tone, which had the highest intensity values among the four regular tones. Unstressed syllables, while lower, were comparable to second or third tone, and were lower than the measurements for neutral tone syllables, which were surprisingly high. It's unclear whether this reflects a genuine difference in how intensity can indicate stress in the two languages or a sampling error in the words selected for the experiment.

In pitch, the data is more interesting. Neither stressed nor unstressed syllables precisely match any particular tonal contour, it appears that the unstressed syllables have a mid or mid-low target, which is a sort of default target for the neutral tone as well (Chen & Xu, 2006).

Although the difference between English speakers and Mandarin speakers in production of stressed syllables is interesting, It's not totally clear that it is from influence of Mandarin tones. Initial stressed syllables are similar to first tone in some ways, but differ in important ways as well.

8.2 Evaluating hypotheses

Hypothesis 1: Stress Transfer

Duration and intensity measures are somewhat unclear in terms of influence of L1 stress. The duration facts do support that Mandarin speakers' use of vowel duration is influenced by their native phonology, given that they do lengthen second syllables more than English speakers, but comparisons in Experiment 3 do not give a clear idea of how much that influence is from the stress system. It is possible that this final lengthening difference is instead just from a greater number of pauses or hesitation on the part of the Mandarin speakers. For intensity, Mandarin speakers perform very similarly to English speakers, and there isn't a clear comparison between their second language English performance and their Mandarin performance.

If we take the features of neutral tone as indicative as a toneless unstressed syllable, there is an indication that unstressed English syllables are being treated similarly. In final syllables, we see that they are somewhat comparable to neutral tones, and although no comparison was possible between English unstressed syllables and neutral tones in the first syllable, that contour does appear to have the kind of mid contour one might expect for a neutral tone syllable in that position.

Hypothesis 2: Tone Transfer

The comparisons of pitch curves do not show any clear indication of a tonal contour being used either for stressed or unstressed syllables. Although Chinese speakers differ from English speakers in their pitch contour for initial stressed syllables, it's a bit ambiguous how that is being influenced by their L1 system. The contour appears to be close to a first tone, but it is lower than a first tone. More importantly, the slope of the curve is not consistent with the first tone. If we were looking for influence of first tone on some English stressed syllables, I would expect a close match in the slope, as that is a major part of the identity of contour tones.

Although Mandarin tones themselves vary in intensity and duration, that isn't helping much here. Mandarin speakers actually seem to be producing the same intensity contrast that English speakers do, without regard to any tonal categories in Chinese. The duration measurements are also not clearly indicating anything about tone, though Chinese speakers did differ from English speakers on duration.

Hypothesis 3: Stress and Tone Interaction

If we accept that Mandarin speakers' use of duration is evidence of relating English stress to their native stress system, and associate the contour they put on initial stressed syllables with the first tone, this proposal makes some sense. However, I have not yet seen clear evidence here for the influence of tone.

8.3 Moving Toward an Answer

The complexity of the results from the experiments given in Chapters 5, 6, and 7 make a final conclusion difficult. At most, it is reasonable to say the evidence points toward stress being transferred. Mandarin speakers produce stressed syllables in English with longer duration and greater intensity. The direct comparisons between their production of English and Mandarin do cause some difficulties of interpretation, but factoring in other research discussed in Chapter 2 indicates this result falls in line with previous work, and considering the work of Qin et al. (2017) regarding differences between Beijing and Taiwan Mandarin speakers do indicate the influence of L1 on the use of duration, which would seem to point to a stress transfer hypothesis.

The implications of the pitch data warrant further study. Although the difference in pitch contours in English and Mandarin speakers observed in Experiments 1 and 2 might have indicated influence of the tone system, I don't see the evidence pointing that way definitively.

If we were to accept that the pitch contour on initial stressed syllables is influenced by Mandarin first tone, why would they have a divergent slope? Also, why would it appear that they have different tonal influences on the first and second syllable? Native English speakers do have a minor difference in slope between the two, so perhaps Mandarin speakers attend to that and exaggerate it? If we were to test words larger than two syllables, what would we find?

These questions make me hesitant to consider tonal influences on Mandarin speakers' production of English stress. It could be that there are other effects that I haven't accounted for that give a better explanation.

9 Conclusion

9.1 Findings of the Current Experiments

Data from Experiments 1 and 2 indicate that Mandarin speakers use duration and intensity similarly to English speakers. Their duration measurements also show potential L1 influence in the larger degree of final lengthening, however there is another possible expanation in that they merely had more of a tendency to pause. The data in Experiment 3, does not definitively link this with L1. The pitch contours of unstressed syllables did not match any contour tone, but they did seem to have a similar target to neutral tone syllables, indicating they were recognized and handled as unstressed and that Mandarin's . Thus, there is still tentative evidence for influence of the native stress system.

On the subject of pitch contours, Experiments 1 and 2 indicated that Mandarin speakers have an interesting difference from English speakers in that they have an upward curve in initial stressed syllables, where English speakers have a downward curve. Experiment 3 shows that this upward curve does not quite correspond to a tone, and indicating influence of tone on these realizations is questionable.

It is my conclusion that stress transfer is likely, but tonal transfer is inconclusive. More study may help to understand the pitch curves found in this experiment.

9.2 Limitations of the Current Study

There are a number of issues in the protocol of this study that have an effect on the interpretation. The presence of CVC syllables in the English wordlist for Experiments 2 and 3 may have artificially shortened those vowels, leading to my difficulty in comparing duration between languages.

Also, I did not validate my Chinese target syllables with a native listener. It's likely that there is some noise in the data from expected neutral tone syllables that were in fact not fully reduced, or expected final full-tones that were reduced toward neutral tone.

I also chose in this project not to restrict speakers to a particular geographic area. This causes a potential issue, as there is known dialectal variation in the realization of the neutral tone, particularly between Beijing and Taiwan Mandarin (Qin et al., 2017; Huang, 2012). Future studies should take geographic origin into account.

9.3 Future Work

The limitations mentioned in the last section show the importance of replication and iteration. While this dissertation has improved on the earlier stage of the research described in Experiment 1, further refinements will bring a clearer picture of the evidence.

There are a number of major changes I would like to make in future iterations. First, it is apparent that future experiments should either target

Mandarin speakers from a particular region or separate them by region. The finding in Qin et al. (2017) that Taiwan Mandarin speakers use duration differently than Beijing Mandarin speakers in perception of English stress needs further investigation.

Second, I would like to design future experiments to separate accent from stress by manipulating focus on the target word. This would help to disentangle stress and accent and allow me to investigate spectral tilt. If spectral tilt can be validated as a correlate of stress in Mandarin, it may give us a clearer picture of stress transfer than intensity.

I would also like to examine stress placement errors in more detail, in order to examine any influence of the native metrical system. In this dissertation I have limited myself to testing solely on examination of phonetic correlates of stress, but examining transfer of the metrical system is just as important. In the current study, I have noticed a pattern of Mandarin speakers having more difficulty with final-stressed words, but I have chosen not to give a formal analysis, as my experimental protocol was not designed for analysis of stress placement errors, and does not give enough variety in word shapes.

In addition to iterating this production experiment, future work should attempt to translate these findings into perception experiments. Phonology is a creature of the human mind, not of sound vibrations in the air, so perception is just as important as production when investigating it. As such, the hypotheses formed by this experiment require validation and further iteration on the perception side.

Bibliography

Altmann, Heidi. 2006. *The perception and production of second language stress: A cross-linguistic experimental study*: University of Delaware disseration. doi:10.1177/001088048102200214.

Archibald, John. 1993. *Language Learnability and L2 Phonology: The Acquisition of Metrical Parameters*. Boston: Kluwer Academic Publishers.

Archibald, John. 1997. The acquisition of English stress by speakers of nonaccentual languages: lexical storage versus computation of stress. *Linguistics* 35(1). 167–181.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. doi:10.18637/jss.v067.i01.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience: Issues in cross-language research*, 171–206.

Best, Catherine T, Gerald W McRoberts & Elizabeth Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America* 109(2). 775–794. doi:10.1121/1.1332378.

Boersma, Paul & David Weenink. 2015. Praat: Doing Phonetics by Computer. `http://www.praat.org/`.

Bolinger, Dwight L. 1961. Contrastive Accent and Contrastive Stress. *Language* 37(1). 83–96.

Brysbaert, Marc & Boris New. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41(4). 977–990. doi:10.3758/BRM.41.4.977.

Burzio, Luigi. 2007. Phonology and phonetics of English stress and vowel reduction. *Language Sciences* 29(2-3). 154–176. doi:10.1016/j.langsci.2006.12.019.

Cai, Qing & Marc Brysbaert. 2010. SUBTLEX-CH : Chinese Word and Character Frequencies Based on Film Subtitles. *PLoS One* 5(6). doi:10.1371/journal.pone.0010729.

Campbell, Nick & Mary E. Beckman. 1997. Stress, prominence, and spectral tilt. *Intonation: Theory, Models and Applications* (September). 67–70.

Chao, Yuanren. 1933. Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology* 4(2). 121–134.

Chao, Yuanren. 1968. *A grammar of spoken Chinese*. Berkely and Los Angeles: University of California Press.

Chen, Yiya & Yi Xu. 2006. Production of weak elements in speech – evidence from F(0) patterns of neutral tone in Standard Chinese. *Phonetica* 63(1). 47–75. doi:10.1159/000091406.

Cheng, Chin Chuan. 1968. English Stresses and Chinese Tones in Chinese Sentences. *Phonetica* 18(2). 77–88. doi:10.1159/000258601.

Chrabaszcz, Anna, Matthew Winn, Candise Y Lin & William J Idsardi. 2014. Acoustic Cues to Perception of Word Stress by English, Mandarin and Russian Speakers. *Journal of Speech, Language, and Hearing Research* 57. 1468–1479. doi:10.1044/2014_JSLHR-L-13-0279.

Cutler, Anne. 2008. Lexical Stress. *The Handbook of Speech Perception* 264–289. doi:10.1002/9780470757024.ch11.

Davies, Mark. 2008. The Corpus of Contemporary American English (COCA): 560 million words, 1990-present. *Literary and Linguistic Computing* `http://corpus.byu.edu/coca/`.

Duanmu, San. 2000. *The Phonology of Standard Chinese*. Oxford University Press.

Duanmu, San. 2007. *The phonology of standard chinese*. Oxford University Press 2nd edn. doi:10.1353/lan.2003.0139.

Flege, James E. 1995. Second Language Speech Learning: Theory, Findings, and Problems. In Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, chap. 8, 233–277. Baltimore: York Press. doi:10.1111/j.1600-0404.1995.tb01710.x.

Flege, James E. 2011. Assessing constraints on second-language segmental production and perception. In Hermann Reichert (ed.), *Phonology and phonetics : Phonetics and phonology in language comprehension and production : Differences and similarities*, 319–355. Walter de Gruyter.

Flege, James Emil. 2007. Language contact and bilingualism: Phonetic system interactions. In Jennifer Cole & José Ignacio Hualde (eds.), *Laboratory phonology 9*, 353–381. New York: Mouton de Gruyter.

Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and speech* 1(2). 126–152.

Huang, Karen. 2012. *A study of neutral-tone syllables in Taiwan Mandarin*: University of Hawai'i at Mānoa dissertation. `http://www.ling.hawaii.edu/graduate/Dissertations/KarenHuangFinal.pdf`.

Huang, Yaqian. 2018. The Phonological Word in Mandarin: Evidence from Tone 3 Sandhi in Reduplication. In *The 26th annual conference of international association of chinese linguistics*, .

Hyman, Larry M. 2009. How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31. 213–238. doi:10.1016/j.langsci.2008.12.007. `http://linkinghub.elsevier.com/retrieve/pii/S0388000108000466`.

Jangjamras, Jirapat. 2011. *Perception and production of English lexical stress by Thai speakers*: University of Florida dissertation.

Kuznetsova, Alexandra, Per B Brockhoff & Rune H B Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13). 1–26. doi:10.18637/jss.v082.i13.

Lee, Wai Sum & Eric Zee. 2008. Prosodic characteristics of the neutral tone in Beijing Mandarin. *Journal of Chinese Linguistics* 36(1). 1–29.

Lieberman, Philip. 1960. Some Acoustic Correlates of Word Stress in American English. *The Journal of the Acoustical Society of America* 32(4). 451–454. doi:10.1121/1.1908095.

Lin, Candise Y., Min Wang, William J. Idsardi & Yi Xu. 2013. Stress processing in Mandarin and Korean second language learners of English. *Bilingualism: Language and Cognition* 17(02). 316–346. doi: 10.1017/S1366728913000333.

Lin, M & J Yan. 1980. Beijinghua qingsheng de shengxue xingzhi [Acoustic characteristics of neutral tone in Beijing Mandarin]. *Fangyan* 3. 166–178.

Lin, M & J Yan. 1988. The characteristic features of the final reduction in the neutral-tone syllable of Beijing Mandarin. In *Phonetic laboratory annual report of phonetic research*, 37–51.

Mirman, Daniel. 2017. *Growth Curve Analysis and Visualization Using R*. CRC Press. doi:978-1-4665-8433-4.

Okobi, Anthony O. 2006. *Acoustic Correlates of Word Stress in American English*: Massachusetts Institute of Technology dissertation.

Ou, Shu-chen. 2010. Taiwanese EFL Learners' Perception of English Word Stress. *Concentric: Studies in Linguistics* 36(1). 1–23.

Ou, Shu Chen. 2016. Perception of english lexical stress with a marked pitch accent by native speakers of mandarin. *Taiwan Journal of Linguistics* 14(2). 1–31. doi:10.6519/TJL.2016.14(2).1.

Peperkamp, Sharon. 2004. Lexical Exceptions in Stress Systems: Arguments from Early Language Acquisition and Adult Speech Perception. *Language* 80(1). 98–126. doi:10.1353/lan.2004.0035.

Peperkamp, Sharon & Emmanuel Dupoux. 2001. A typological study of stress 'deafness'. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory phonology 7*, 203–240. Berlin: Mouton de Gruyter. doi:10.1016/j.specom.2010.03.005.

Purnell, Thomas C. 1997. *Principles and parameters of phonological rules: Evidence from tone languages*: University of Delaware dissera-tion.

Qin, Zhen, Yu Fu Chien & Annie Tremblay. 2017. Processing of word-level stress by Mandarin-speaking second language learners of English. *Applied Psycholinguistics* 38(3). 541–570. doi:10.1017/S0142716416000321.

R Development Core Team. 2015. R: A Language and Environment for Statistical Computing. doi:10.1007/978-3-540-74686-7. http://www.r-project.org.

Rietveld, A. C M & F. J. Koopmans-van Beinum. 1987. Vowel reduction and stress. *Speech Communication* 6(3). 217–229. doi:10.1016/0167-6393(87)90027-6.

Rosenfelder, Ingrid, Joe Fruehwald, Keelan Evanini & Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) Program Suite. http://fave.ling.upenn.edu/.

Shih, Chilin & Hsin-Yi Dora Lu. 2015. Effects of talker-to-listener distance on tone. *Journal of Phonetics* 51. 1–30. doi:10.1016/j.wocn.2015.02.002.

Sluijter, Aagath MC & Vincent J van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100(4). 2471–2485. doi:10.1121/1.417955.

Sluijter, Aagath MC, Vincent J van Heuven & Jos J A Pacilly. 1997. Spectral balance as a cue in the perception of linguistic stress. *The Journal of the Acoustical Society of America* 101(1). 503–513. doi:10.1121/1.417994.

Wang, Qian. 2010. *Perception of English stress by Mandarin Chinese learners of English: An acoustic study*: University of Victoria dissertation.

Xu, Yi. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25(1). 61–83. doi:10.1006/jpho.1996.0034. http://discovery.ucl.ac.uk/98152/.

Yuan, Jiahong, Neville Ryant & Mark Liberman. 2014. Automatic Phontic Segmentation in Mandarin Chinese: Boundary Models, Glottal Features, and Tone. In *2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, .

Zhang, Yanhong & Alexander Francis. 2010. The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics* 38(2). 260–271. doi:10.1016/j.wocn.2009.11.002.

Zhang, Yanhong, Shawn L Nissen & Alexander L Francis. 2008. Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *The Journal of the Acoustical Society of America* 123(6). 4498–4513. doi:10.1121/1.2902165.