Advantages in Linguistic Tone Perception in Speakers with Tone Language Experience

Alyssa J. Yee

University of Colorado at Boulder, alye9140@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/slhs_gradetds

Part of the Bilingual, Multilingual, and Multicultural Education Commons, Linguistics Commons, and the Speech Pathology and Audiology Commons

Recommended Citation


https://scholar.colorado.edu/slhs_gradetds/51

This Thesis is brought to you for free and open access by Speech, Language and Hearing Sciences at CU Scholar. It has been accepted for inclusion in Speech, Language, and Hearing Sciences Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
ADVANTAGES IN LINGUISTIC TONE PERCEPTION IN SPEAKERS WITH TONE LANGUAGE EXPERIENCE

by

ALYSSA YEE

B.A., San Diego State University, 2015
M.A., University of Colorado-Boulder, 2017

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Master of Speech Language Pathology
Department of Speech, Language, and Hearing Sciences
2017
This thesis entitled:
Advantages in Linguistic Tone Perception in Speakers with Tone Language Experience
written by Alyssa Yee
has been approved for the Department of Speech, Language, and Hearing Sciences
University of Colorado at Boulder

Pui Fong Kan, PhD, CCC-SLP

Neeraja Sadagopan, PhD, CCC-SLP

Bhuvana Narasimhan, PhD

Christine Brennan, PhD, CCC-SLP

Date: 4/6/2017

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB protocol # 16-0152
Abstract

Yee, Alyssa (M.A., Speech, Language, and Hearing Sciences)

Advantages in Linguistic Tone Perception in Speakers with Tone Language Experiences

Thesis directed by Professor Pui-Fong Kan, Ph.D

The current study compared adult bilingual Mandarin-English speakers and monolingual English speakers’ ability to perceive linguistic (word and non-word) as well as non-linguistic (pure tone) auditory tonal contrasts to investigate the presence of both a linguistic and cross domain advantage based on tone language experience. Accuracy and reaction time were recorded for all three conditions during a tone perception discrimination task. Findings did not reveal such advantage in any context, however the bilingual Mandarin-English group was more accurate in the word condition suggesting additional semantic information contributed to the group’s decisions about tone contrasts. Additionally, the bilingual group exhibited a trend of slower reaction times across all conditions. Randomization of stimuli and both relevant and non-relevant information together may have contributed to an increased cognitive load for the bilingual group. Together, the results show that tone perception may not be a result of tone language experience and further investigation will provide insight into neural perception, processing, and access of tone.

Word count: 160

Keywords: lexical tone perception, Mandarin, language experience, bilingualism
## CONTENTS

I. INTRODUCTION AND BACKGROUND ....................................................... 1  
   a. Lexical Tone .................................................................................. 1  
   b. Learning Lexical Tones: Categorical Perception of Language (Perceptual 
      Reorganization) ......................................................................... 4  
   c. The Advantage of Tonal Language Experience on Tonal Detection ........ 6  
   d. The Bilingual Experience ............................................................. 6  
   e. Storage of Meaning and Form (the lexical access) ......................... 7  

II. THE CURRENT STUDY ......................................................................... 7  
    a. Research Questions ..................................................................... 8  

III. METHODS .......................................................................................... 9  
    a. Participants .................................................................................. 9  
    b. Stimuli ......................................................................................... 10  
    c. Word Condition .......................................................................... 10  
    d. Non-word Condition ................................................................... 11  
    e. Pure-tone Condition .................................................................... 12  
    f. Duration Condition ....................................................................... 14  
    g. Procedures ................................................................................... 16  

IV. RESULTS .......................................................................................... 19  

V. DISCUSSION ....................................................................................... 24  
    a. Experience-Dependent Neuroplasticity ........................................ 24  
    b. Language Experience Effect on Tonal Perception ......................... 25  
    c. Bilingualism ............................................................................... 26  
    d. Accuracy ..................................................................................... 27  
    e. Reaction Time ............................................................................ 28  

VI. CLINICAL IMPLICATIONS ................................................................. 29  

VII. LIMITATIONS AND FUTURE DIRECTIONS .................................. 30  
    a. Limitations ................................................................................ 30  
    b. Future Directions ....................................................................... 30  

VIII. ACKNOWLEDGMENTS ................................................................. 31  

IX. REFERENCES ................................................................................... 32
# TABLES

Table

1. The Mandarin Word “Ma” With 4 Definitions ...........................................3
2. Frequencies and Durations of the Average Female Mandarin Speaker Segmented for the Four Tones of Mandarin as Indicated by Figure 1 .........................13
3. Order of Procedures for Test Administration .............................................17
4. Stimuli Pairs for Each of the Four Conditions Created .............................18
5. Participant Performance for All Four Conditions .....................................20
6. Summary of the Repeated measures ANOVA Analysis ............................23
FIGURES

Figure

1. Stimuli components of an individual tone contrast pair ........................................11
2. Example of auditory stimuli in Praat (Tone 1-3 pair same) for word-word condition …11
3. Relative pitch changes of the four tones in Mandarin ........................................12
4. Example of auditory stimuli for the pure tone condition (Tones 1-2 4-3) ...............14
5. Example of the stimuli block presentation for the tone contrast perception task ........15
6. Example of auditory stimuli for the duration condition (0.6 seconds per tone at 275 Hz)15
7. Example of auditory stimuli for the duration condition at 330Hz (0.5-0.5 and 0.5-0.3seconds) .................................................................................................................16
8. Tone contrast perception task instructions presented on the computer screen as the auditory stimuli plays ........................................................................................................18
9. Tone perception accuracy across three conditions ..................................................21
10. Tone perception reaction time (ms) across three conditions ........................................22
Introduction and Background

This study sought to determine if there was an advantage of tone language experience when detecting non-linguistic tone contrasts for adult speakers of a tone language. In particular, this study examined whether there was a difference between native speakers of Mandarin and monolingual English speakers without tone language exposure in 3 auditory contexts: speech (Mandarin words), non-words, and pure tone sounds. This study may give possible additional contributions to how the human mind functions and is shaped by specific influences (linguistic tone). Previous studies in lexical and non-lexical tone perception have proven some advantage in lexical tone experience over non-tone language speakers throughout infancy and have proven differences in categorization capabilities of individuals more exposed to tone linguistically (Burnham et al., 2010; Iverson & Kuhl, 2003; Lee et al., 2010; Mattock & Burnham, 2006). Nonlinguistic tone perception stimuli have been studied before with musical tones (e.g. violin sounds), however pure tone single sine wave frequencies have not been tested in tone perception in previous studies. In what follows, I review literature in the areas of lexical tones and language experience effects on lexical tone learning.

Lexical Tone

Tone languages represent about two thirds (60-70%) of all languages used worldwide. These languages use pitch height and/or pitch contours (changes) to convey word meaning (Yip, 2002). Lexical tones refer to these distinctive changes in pitch levels that are integrated within language (i.e. at the level of a syllable or word). Tone is considered a surpasegmental part of language, meaning it is a contrastive element of speech that cannot be analyzed in segments like phonemes. The classic example of lexical tone refers to the Mandarin word “Ma.” Outlined in Table 1, the word “Ma” can mean four different words depending on the Mandarin lexical tone
contour laid atop the syllable. Although individuals of any native language with normal hearing can discriminate changes in lexical tone, speakers of a tone language have been exposed to minute changes for their entire lives that are important to indicate meaning. The perception of lexical tone therefore may be more automatic just by nature of exposure and experience. In English, pitch contours occur most commonly across the sentence level as in interrogatives and therefore have a longer period of time to denote meaning from tonal changes. The explicit attention to learn this new frequently occurring and rapidly changing auditory input when adults attempt to learn a tonal language is difficult.

Non-linguistic tones or pure tones can differ in the same way that fundamental frequency contours determine a tone in a Mandarin syllable. Gandour (1983, 1984) state that pitch in language does not contain a hierarchical (i.e., scalar) framework; there is no ‘in-’ or ‘out-of-tune’. Consequently, the perception of linguistic pitch patterns largely depends on cues related to the specific trajectory (i.e., contour) of pitch movement within a syllable rather than distances between consecutive pitches, as it does in music. Looking at tone in comparison to music, tone languages provide a unique opportunity for investigating linguistic uses of pitch as these languages exploit variations in pitch at the syllable level to contrast word meaning (Yip, 2002). In language, much like coarticulation of phonemes, pitch information that is extracted depends on the interactions between specific features of the input signal; their corresponding output representations, and the domain of pitch experience of the listener (Zatorre, 2008). Therefore, a more complex mapping of tone is taking place that has more fluid lines of perceptual categories than hearing tone in isolated syllables. It is unclear whether or not individuals whose native language is tonal present with an advantage in perceiving lexical tone that carries over to an advantage in non-lexical or non-linguistic tone.
Much of the research in auditory perception and language refers to categorization of native and non-native phoneme contrasts throughout infancy. Early in the first year of life, young infants are able to discriminate almost all phonetic contrasts they hear, whereas older infants at about 10-12 months can discriminate better between phonemes that occur in their native language rather than foreign-language phonemic contrasts (Cheour et al., 1998). Iverson & Kuhl (2003) suggest that infants over this first year develop perceptual and cognitive processes that are specialized for their native language based on linguistic experience, which are induced by learning. Non-native segments are influenced by the listener’s native phonological system and the acoustic distinctiveness of the non-native segments to be perceived (Lee, et al., 2010). The categorical perception of phonemes in speech has been studied significantly in the literature, however little is known about the perception of tone contrasts (the suprasegmental attribute to language) and has not been extensively studied. Functional MRI imaging studies show activation of both hemispheres during the perceptual processing of linguistic prosody, however processing is mediated by the right hemisphere in Chinese speakers. These studies indicate possible perceptual differences in speakers of tone languages (Gandour et al., 2003).

Table 1

The Mandarin Word “Ma” With 4 Definitions

<table>
<thead>
<tr>
<th>‘Ma’ (Tone)</th>
<th>Tone Contour [Visual]</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High-flat [-]</td>
<td>Mother</td>
</tr>
<tr>
<td>2</td>
<td>Mid-rising [/]</td>
<td>Hemp</td>
</tr>
<tr>
<td>3</td>
<td>Bi-directional (falling then rising) [√]</td>
<td>Horse</td>
</tr>
<tr>
<td>4</td>
<td>High-falling []</td>
<td>Scold</td>
</tr>
</tbody>
</table>
Learning Lexical Tones: Categorical Perception of Language (Perceptual Reorganization)

It is known that lexical tone is an integral part of the Chinese language, Mandarin, embedded in each syllable to denote meaning. Mandarin speakers learn to listen for and use tone functionally in a linguistic context, which is much more complex than pitch contours found in English. Much like phonemic categorization early in life, tonal re-categorization based on language experience is addressed in Japanese speaker’s ability to acquire the contrasts between English /l/ and /r/ (Iverson & Kuhl, 2003). The Japanese speakers required more focused attention to the third formant frequency (F3) contours than English speakers to hear the difference between the two phonemes. Few previous studies have looked at lexical tone categorization in early childhood. Mattock & Burnham (2006) looked at infant’s abilities to perceive linguistic and non-linguistic musical/tonal contrasts (violin notes) for infants exposed to English and those exposed to Thai (also a tonal language). Results of the study indicated that English infants’ ability to discriminate lexical tone declined between 6 and 9 months of age, while their non-speech tone discrimination remained constant, which was not held true for the Thai speaking infants. This finding indicated a possible reorganization of tone perception as a function of native language environment. In a similar study, tone detection awareness was found uniquely associated with Chinese character recognition for Mandarin adult speakers even after other phonological awareness processes were taken into account (McBride-Chang et al., 2008). Thus, tone is part of the phonological and auditory processing of heard syllables in Mandarin, making the exposure, attention to, and awareness of linguistic tone highly recurring when compared to non-tonal language speakers.
In the linguistic context, phonetic implementation of Mandarin tones in conversation depend on the neighboring tones with great carryover effect from the preceding tone (Chen & Gussenhoven, 2008; Xu, 1994), however the non-linguistic contexts exhibit a more step-like contour. Therefore, if the signal being used and heard is more complex, the non-linguistic contrasts should be perceived with greater ease. This has not been proven in previous research studies.

Much of the research on linguistic pitch patterns comes from cognitive neuroscience looking at mismatch negativity of event-related potentials (ERPs) and electrical activity in the brain during auditory tasks. This research indicates that a decreased ability to perceive prosodic stress in words for bi-syllables in infancy may be a marker of risk for later language impairment (Weber et al., 2005). Possible differences in tone perception can also have implications to populations such as children with learning impairments and reading disorders, who both exhibit slower reaction times and greater difficulty recognizing auditory patterns involving discriminating frequency and durational tone patterns (Kohnert & Windsor, 2004; Walker et al., 2006). Much of this research is few and far between, giving an ambiguous indication to whether or not tone perception (lexical and non-lexical) is an imperative cognitive process for both learning and language.

The presence of a clear advantage can contribute further knowledge about brain plasticity early on in life that influences one’s ability to acquire languages and create concrete categories for language attributes beyond the critical period of language development as well as solidify an advantage based on experience-dependent neuroplasticity. Mandarin is oftentimes referred to as a complicated language and difficult to learn. Differences in auditory processing for tone may contribute to this notion. The question remains: Does a categorization and/or reorganization of
tone perception narrow if tone is not salient or present in a native language? Furthermore, if tone is salient, does this give those individuals an advantage across domains (lexical and non-lexical)?

**The Advantage of Tonal Language Experience on Tonal Detection**

Experience underlies the foundation for acquiring and remembering language. Immersion is the most effective and efficient way of learning a language and by nature an individual will be more proficient in a language given more experience. In the same light, individuals tend to pay attention to the most salient information in language to extract meaning. Since syllable-level tonal contrasts are salient in Mandarin and many tonal languages, its listeners hone more attention into that feature of language. Francis et al., (2008) stated that tone categories for native speakers of a tone language have a stronger mental representation of native tonal categories in language that can support the increased difficulty for non-tone language speakers.

**The Bilingual Experience**

In addition to tone language experience, language experience in more than one language may also influence one’s acoustic processing. Bilinguals overall may pay more detailed attention to subtle acoustic differences in language input due to the more complex language environment they encounter (Liu & Kager, 2013). Bilinguals infants also may present a cognitive advantage in heightened acoustic sensitivity compared to their monolingual peers, but may be acoustic rather than linguistic in nature (Liu, 2014). Therefore, in terms of tonal detection, the heightened acoustic sensitivity coupled with stronger categorical representations may give bilinguals with a tone language a slight advantage in detecting tone. Conversely, individuals (monolingual or bilingual) without tone language experience do not have set categories and are not attending as much to those specific acoustic cues related to lexical tone that may make it more difficult and take longer to detect tonal differences in the acoustic signal.
Storage of Meaning and Form (the lexical access)

Exposure to a tonal language as an infant may influence how one categorizes tone contrasts as a result of specific language experience, however meaning and form of language with salient tone may also be stored and accessed differently than non-tone language speakers and influenced by one’s phonological system. It has been shown that the two hemispheres play different roles in lexical tone perception; specifically the right hemisphere is involved in auditory/acoustic processing, whereas the left hemisphere is responsive to linguistic/phonological functions (Gandour et al., 2003; Wang et al., 2003). Huang & Johnson (2011) propose that perceptual salience is a function not only of raw acoustic/auditory contrast but also of the listener’s phonological systems. Influence from the phonological system may not present advantage cross-domains.

The Current Study

The purpose of this study was to determine if there was a positive effect on non-linguistic tone perception for individuals with tone language experience. The three conditions in this study were speech (words in Mandarin), non-words containing Mandarin phonemes, and sequences of pure tone frequencies. The monolingual English-speaking group had no previous exposure to Mandarin linguistic tone. For the word and non-word conditions, linguistic stimuli included auditory pairs of two 2-syllable words in Mandarin (speech) and 2-syllable non-words containing phonemes that followed the phonemic patterns of Mandarin, but did not carry meaning. Nonlinguistic stimuli included pairs of pure tone combinations of single frequency sine wave bands segmented together. The word condition contained extra semantic information that could possibly influence the Mandarin group’s performance on the tone discrimination task. The non-word condition was used to create a non-salient lexical comparison to the word condition and the
pure tones were used to determine if there was a cross-domain advantage for the tone language experienced group.

**Research Questions**

In this study, tone discrimination is tested within and across-language groups in three auditory contexts: words, non-words, and pure tones. Given the insufficient research of auditory linguistic perception for tone languages in adults, this study seeks to understand:

1. Is there an advantage for tone language acquisition? That is, are the speakers with tone language experience faster and/or more accurate in distinguishing auditory linguistic tone contrasts compared to speakers without tone language exposure?

2. If the bilingual Mandarin-English speaking group is faster and/or more accurate at perceiving auditory linguistic tone contrasts, does an advantage in linguistic tone perception carry over to a subsequent advantage in non-linguistic (pure tone) contrasts compared to non tone language speakers? That is, we suggest bilingual Mandarin-English speakers will have an advantage in both and monolingual English speakers should perform similarly across all contexts (linguistic and non-linguistic) or better in the non-linguistic context.

3. Are speakers of a tone language more accurate at discriminating word-word contrasts compared to non-word contrasts because of tone exposure in the Mandarin language? Are they faster or slower because of extra interfering yet meaningful semantic/phonetic information?

It has been known that speakers of a tone language tune into different acoustic features based on language-specific experience. Speakers of a tone language rely more on F0 changes and pitch direction rather than the F0 height as English speakers do (Gandour, 1983). Mattock & Burnham
(2006) found a non-linguistic tonal advantage in children with tone language experience (Thai) when compared to English-speaking children. Further, Bent et al., (2006) looked at Mandarin and English speaker’s identification in speech and non-speech discrimination tasks. Findings indicated that Mandarin speakers more often misidentified non-speech flat and falling pitch contours, suggesting that there is some influence on linguistic experience in cross-domain processing. Burnham et al (2015) state that tone language listeners have developed more sensitive brainstem mechanisms for representing pitch than non-tone language perceivers. Therefore, Mandarin speakers should be faster and more accurate at discriminating tonal contrasts in both domains given previous evidence of language experience-dependent neuroplasticity as well as a categorical perception of tone over time. The semantic information for the word condition should also be faster and more accurate for the Mandarin speaking group based on more solidified neural connections via exposure time.

Method

Participants

Thirty-eight adults, including 19 monolingual English speakers (18F, 1M; mean age 30.63 years, SD=13.07) and 19 bilingual Mandarin-English speakers (7F, 12M; mean age 21.55 years, SD=6.18), were recruited for this study. All participants presented with normal hearing and indicated no history of speech, language, hearing, or vision impairments (for vision impairments, contacts and glasses that effectively compensated for visual impairment was fine). All participants in the monolingual English-speaking group were native English speakers with no more than conversational level proficiency in another non-tone language. Participants in the bilingual Mandarin-English speaking group were native Mandarin speakers (dialects varied) with proficient English language ability. The term ‘bilingual’ meant the participants were able to
functionally communicate in both languages at the time of testing and must have used Mandarin as the primary language throughout childhood, indicating sequential bilinguals. Testing was conducted in the Speech, Language, and Hearing Center at the University of Colorado-Boulder from May-September 2016. Each participant completed a self-report language questionnaire, which specified use of one or both languages in exposure, use in various environments (i.e. social, academic, home), age of acquisition, etc. All subjects passed a hearing screen with thresholds $\leq 25$dB HL at octave frequencies from 500 to 4000 Hz. The use of human subjects in this study was reviewed and approved by the University of Colorado-Boulder Institutional Review Board.

**Stimuli**

Four different stimuli conditions were created in both linguistic and non-linguistic contexts. The linguistic contexts used the phonemes of the Mandarin language and the non-linguistic context used different combinations of single frequency bands.

**Word condition**

Two syllable words in Mandarin were recorded, uploaded to *Praat*, cut into individual two-syllable words, and placed in a tone contrast pair as represented in Figure 1. One second of silence was placed in between and the tone contrasts varied in that some words for the different pairs changed the first syllable tone and others changed the second syllable tone while preserving the same phonemes. Figure 2 shows a .wav format file of a word-word pair that is the same in tone. Individual tone contrast pairs were created for all four conditions.
Figure 1: Stimuli components of an individual tone contrast pair

Figure 2: Example of auditory stimuli in Praat (Tone 1-3 pair same) for the word-word condition.

Non-word condition

Two-syllable non-words in Mandarin were recorded, uploaded to Praat, cut into individual two-syllable non-words, and placed in a tone contrast pair as represented in Figure 1. A second of silence was placed in between and the tone contrasts varied in that some non-words for the different pairs changed the first syllable tone and others changed the second syllable tone. The average duration of the word and non-word condition syllables ranged from 0.2 to 0.8 seconds in duration. All non-words contained phonemes of the Mandarin language.
**Pure Tone condition**

Keating & Kuo (2012) state that the range of fundamental frequency measurements from isolated words for female Mandarin speakers ranged from 110-330 Hz. Figure 3 depicts the relative pitch changes of the four Mandarin tones split up into 5 ranges to indicate direction and contour of frequency across a syllable. To generate pure tone combinations that simulated the four Mandarin tones, the 110-330 Hz range was broken down into 5 frequency ranges, and the corresponding pure tone combinations are summarized in Table 2 by tone, duration, and frequency. Pure tones were generated in *Audacity* and rearranged into different combinations to parallel the linguistic stimuli. The ranges were segmented together to create pure tones that reflected a Mandarin syllable, each combined to be 1.1 seconds in duration. One second of silence was placed between the two pure tone pairs and 0.1 seconds of silence was placed between each pure tone to simulate a two-syllable word. An example of all four pure tones in *Praat* is shown in Figure 4. Each syllable that was mimicked contained a single frequency band. The fundamental frequency ranges used in the pure tone condition were consistent with the average female Mandarin speaker. The advantage of tone language experience does not necessarily apply across the board, and is mainly evident in just those sections of a fundamental frequency contour that exhibits rapid changes in pitch (Wong et al., 2007).
**Figure 3**: Relative pitch changes of the four tones in Mandarin.

Table 2

*Frequencies and Durations of the Average Female Mandarin Speaker Segmented for the Four Tones of Mandarin as Indicated by Figure 1.*

<table>
<thead>
<tr>
<th>Tone in Mandarin</th>
<th>Frequency start Hz (duration in ms)</th>
<th>Frequency middle Hz (duration in ms)</th>
<th>Frequency end Hz (duration in ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>330 (500)</td>
<td>N/A</td>
<td>330 (500)</td>
</tr>
<tr>
<td>2</td>
<td>220 (500)</td>
<td>N/A</td>
<td>330 (500)</td>
</tr>
<tr>
<td>3</td>
<td>165 (100)</td>
<td>110 (100)</td>
<td>275 (300)</td>
</tr>
<tr>
<td>4</td>
<td>330 (500)</td>
<td>N/A</td>
<td>110 (500)</td>
</tr>
</tbody>
</table>

*Note.* These measures were used to generate the pure-tone stimuli. Frequency range of average female Mandarin speaker from *Comparison of speaking fundamental frequency in English and Mandarin*, p.1057, by Keating, P., & Kuo, G., 2012, The Journal of the Acoustical Society of America, 132(2), 1050-1060.
Figure 4: Example of auditory stimuli for the pure tone condition (Tones 1-2 4-3).

Duration Condition

Xu & Pfingst (2003) state that differences in duration for Mandarin tones are rather small. However, it has been shown that syllable duration can be a strong cue for isolated syllables and syllable duration might not be a reliable cue for everyday speech (Whalen & Xu, 1992; Xu et al 2002). To decrease the influence on the variable (duration), stimuli with consistent frequency tones that vary only in duration were added to the testing blocks to focus on tone perception as the discriminating factor. The frequencies presented were taken from Table 2 and included 110, 165, 220, 275, & 330Hz. The duration stimuli resembled the pure tone condition with the first 1.1 seconds containing 0.5 seconds of pure tone, 0.1 seconds of silence, and 0.5 seconds of another pure tone. The difference in duration occurred in one of the two pure tone frequencies bands presented following the 1 second of silence, meaning that it resembled a change in duration in what is the first or second syllable of the second two-syllable word in relation to the linguistic stimuli.
Figure 5: Example of the stimuli block presentation for the tone contrast perception task.

Each segment of the duration condition ranged from 0.2 to 0.8 seconds in duration in 0.1 second increments (ex. .2, .3, .4→.8 seconds) excluding .5 as it was the same as that duration that was used for the pure tone condition.

Figure 6: Example of auditory stimuli for the duration condition (0.6 seconds per tone at 275 Hz).
Figure 7: Example of auditory stimuli for the duration condition at 330Hz (0.5-0.5 and 0.5-0.3 seconds).

Procedures

Testing administration for all subjects lasted about an hour in duration with the tone contrast discrimination task taking about 30 minutes to administer all stimuli blocks. Table 3 outlines the series of procedures for each participant. A quick hearing screen was conducted as pass/fail criteria for octave frequencies from 250-4000 Hz at 25 dB HL for both ears. A non-word repetition task from the Comprehensive Test of Phonological Processing (CTOPP) was also administered to each participant (English group mean raw score 12.53 (SD=1.93); Mandarin group mean raw score 10.95 (SD=2.50)). It should be noted that the Mandarin group scored lower than the English-speaking group overall on this task, however the CTOPP is not normed on a bilingual-speaking or non-native English-speaking population. The subtest also contains only phonemes of the English language.
### Table 3

**Order of Procedures for Test Administration**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Participants are asked to fill out a consent form and language questionnaire</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>Participants are asked to complete a quick hearing screen (500, 1,2, &amp; 4000 Hz at 25dB in both ears)</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Auditory Non-word Repetition (CTOPP non-word repetition subtest)</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Tone contrast perception task (presented in 3 blocks each 8-10 minutes in duration)</td>
<td>30</td>
</tr>
</tbody>
</table>

For the tone contrast perception task, participants were asked to make a decision on whether auditory pairs of stimuli presented through headphones were the same or different. The task was presented on a computer using the E-prime program. In each trial, (a fixation), saying “Listen Carefully” first appeared on the screen and remained there. This disappeared after the sound file played and the screen read “Decision Part (1) same or (2) different?” (Figure 8). At that time, the participants had to respond by pressing 1 on the computer keyboard if the pairs of auditory stimuli heard were identical and 2 if the pairs were non-identical, after which the next sound file played with the screen switching back to “Listen Carefully” and a new auditory pair was
Accuracy and reaction times (ms) were recorded on E-Prime. Participants were asked to perceive tone contrasts for three randomized blocks of stimuli, each containing 80 items (240 total). There was a short break between stimuli blocks to generate and run the stimuli in E-Prime, re-administer the directions, and adjust the volume if necessary. The blocks of stimuli were randomized in order of presentation by the E-Prime program for each participant. Measures of accuracy and reaction time were collected for each participant for all three blocks of stimuli, averaged for each condition, and compared within group as well as between groups.

**Listen Carefully**
- Participant listens to the tone pair through headphones

**Decision Part (1) Same or (2) Different**
- Participant makes a decision using (1) and (2) on the keyboard after the pair has played

*Figure 8:* Tone contrast perception task instructions presented on the computer screen as the auditory stimuli plays

Table 4

*Stimuli Pairs for Each of the Four Conditions Created*

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Same tones: # of items</th>
<th>Different tones: # of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>60 two-syllable words in Mandarin</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>60 two-syllable non-words in Mandarin</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>60 pairs of pure tones</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>
60 pairs of pure tones that differ in duration when tones were held constant across the ‘syllable’

Each of the three blocks included 20 pairs of each (10 same and 10 different) for all 4 conditions making 80 stimuli in each block (total 240 pairs). All auditory stimuli were recorded, cut, and re-arranged into either the same or different pairs in Praat. The tone contrast pairs that were the same simply copied and pasted the first 2-syllable word, 2-syllable non-word, or pure tone combination following the second of silence. All the files were converted to wav files and uploaded to the E-prime software that randomly presented the auditory stimuli in pairs as well as recorded the reaction time and accuracy after each presentation. The stimuli created were sorted into each condition (word, non-word, pure tone, and duration) and then split up into the pairs (same & different). Each list was randomized in Microsoft Excel. The first 60 randomized items of each condition were chosen for the testing and then added to the introductions as part of the stimuli. The next three on the list were used as the trial stimuli at the beginning of each block for participants to familiarize themselves with the task. The stimuli pairs in each block were then randomized again in E-Prime. Due to combination restrictions, only 12 same pairs for the pure tone condition were created, meaning that the same pairs show up in all 3 blocks.

**Results**

The present study sought to address if there was an advantage in tone perception in speakers of a tone language that carried over to an advantage in non-linguistic tone (i.e. pure tones). Table 5 summarizes the accuracy and reaction time of the participants for all four conditions.
Table 5

*Participant Performance for All Four Conditions*

<table>
<thead>
<tr>
<th></th>
<th>Accuracy (%)</th>
<th>Reaction Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>English speakers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin speakers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real words in Mandarin (WW)</td>
<td>95.6 (SD=5.45)</td>
<td>320.13 (SD=61.1)</td>
</tr>
<tr>
<td>Non-words in Mandarin (NW)</td>
<td>95.5 (SD=5.91)</td>
<td>325.84 (SD=69.19)</td>
</tr>
<tr>
<td>Pure tones (PT)</td>
<td>97.8 (SD=2.84)</td>
<td>353.46 (SD=115.53)</td>
</tr>
<tr>
<td>Pure tones that differ in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duration when tones were held</td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant (D)</td>
<td>95.54 (SD=3.76)</td>
<td>396.21 (SD=96.67)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>494.7 (SD=182.71)</td>
</tr>
</tbody>
</table>
Figure 9: Tone perception accuracy across three conditions.
The first component of the study sought to examine whether participants relied on durational cues, rather than tones, to determine whether the pairs were the same or different. A 2 (English vs. Mandarin speakers) x 2 (PT vs. D) repeated measures ANOVA was used to examine the data of the two pure tone conditions. Results showed that, while high accuracy was found in both condition, participants were more accurate, \( F(1, 36) = 13.9, p < .01 \) and faster \( F(1,36) = 5.72, p < .05 \) in the pure-tone condition than in the duration condition. No significant group effect or group x condition interaction was found. The findings suggest that both groups of participants rely on tonal cues to determine the differences between the pairs as opposed to stimulus duration.

*Figure 10:* Tone perception reaction time (ms) across three conditions.
A 2 (English vs. Mandarin speakers) x 3 (WW, NW, & PT) repeated measures ANOVA was used to compare the performance of 2 groups in terms of the accuracy and reaction time across the linguistic and non-linguistic conditions. The results were summarized in Table 5.

Table 6

Summary of the Repeated measures ANOVA Analysis

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>df</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>Multiple Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>36</td>
<td>.279</td>
<td>&gt;.05</td>
<td>N/A</td>
</tr>
<tr>
<td>Condition</td>
<td>2</td>
<td>72</td>
<td>.53</td>
<td>&gt;.05</td>
<td>N/A</td>
</tr>
<tr>
<td>Group x</td>
<td>2</td>
<td>72</td>
<td>4.38</td>
<td>&lt;.05*</td>
<td>WW: Mandarin group &gt; English group (p &lt; .05; d = .99); no group differences in NW &amp; PT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reaction Time</th>
<th>df</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>Multiple Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>36</td>
<td>.326</td>
<td>&gt;.05</td>
<td>N/A</td>
</tr>
<tr>
<td>Condition</td>
<td>2</td>
<td>72</td>
<td>.438</td>
<td>&lt;.05*</td>
<td>Faster in NW than in PT (p &lt; .05); No difference between WW and NW; WW and PT</td>
</tr>
<tr>
<td>Group x</td>
<td>2</td>
<td>72</td>
<td>4.49</td>
<td>&lt;.05*</td>
<td>No significant group differences in WW, NW &amp; PT; However, there was a trend that the Mandarin group was slower than the English group in the WW condition (p = .05, d = .65)</td>
</tr>
</tbody>
</table>
Discussion

The present study sought to address if there was an advantage in tone perception in speakers of a tone language that carried over to an advantage in non-linguistic tone (i.e. pure tones) as well as determine if the Mandarin speaking group was faster/more accurate at the word condition over the non-word and pure tone conditions. In contrast to the proposed hypothesis, no significant group differences for the tone perception discrimination task were found for accuracy and reaction time ($p > .05$). The findings suggest that tonal language experience does not have a clear advantage in detecting linguistic or nonlinguistic tone, which are contrary to previous studies that indicated an advantage in tone perception for tone language speakers (Mattock & Burnham, 2006). Additionally for the word condition, Mandarin speakers were more accurate at discriminating those contrasts, however reaction time was insignificant. No such accuracy was seen in the non-word and pure tone conditions. This finding suggests that the knowledge and activation of word form and word meaning may play a role in distinguishing linguistically salient tonal differences. Although not significant, the ultimate finding from this study indicated that the Mandarin speakers took longer to respond to the stimuli ($p = .08$). The patterns suggest that Mandarin speaker’s tonal knowledge might interfere with their judgment about the tonal contrast (Kaushanskaya & Marian, 2007; Pavlenko, 2012). To follow are possible explanations for these findings.

Experience-Dependent Neuroplasticity

In contrast to the proposed hypothesis, no significant group differences in accuracy or reaction time were found. These findings are contrary to previous studies that indicated an advantage in tone perception in infants and young children (Mattock & Burnham, 2006) via experience-dependent neuroplasticity. Krishnan & Grandour (2009) make the argument that in
humans, experience-driven reorganization likely shapes the reorganization of brainstem mechanisms for enhanced pitch extraction at earlier stages of language development and music learning. Once this reorganization is complete, however, local mechanisms in the brainstem are sufficient to extract relevant pitch information in a robust manner without permanent corticofugal influence. This indicates that it is possible that compensatory neural networks allow non-tonal language speakers to perceive tonal contrasts to the same degree as tone language speakers. Neurophysiological evidence from cortical brain potentials suggests, for instance, that musicians exploit interval based pitch cues (Fujioka et al., 2004; Krohn et al., 2007) while tone language speakers exploit contour based cues (Chandrasekaran et al., 2007). Such “cue weighting” is consistent with each group’s unique listening experience and the relative importance of these dimensions to music (Burns & Ward, 1978) and lexical tone perception (Gandour, 1983), respectively. The non-statistically significant findings of this study in terms of a tone-language advantage to tone discrimination may be because the tonal discrimination has to do with unique listening experiences of an individual overall, not exclusive to lexical tone experience specifically. Thus, further investigation into this area of research is warranted.

**Language Experience Effect on Tonal Perception**

Contrary to the proposed hypothesis, tone language exposure and experience did not result in a significant advantage in tonal perception overall for any condition compared to the monolingual English group. The monolingual English group had no previous exposure to any tone language whereas the bilingual Mandarin-English group acquired and used Mandarin functionally at least throughout childhood. One explanation for the similarities in performance for both groups is a proposed idea by Beckman & Pierrehumbert (1986). They state that although English is a non-tonal language, its stress-based prosodic system does utilize pitch as one way to
distinguish stress accents, which may be realized as high, low, rising or falling contours and can thus be very similar to the Mandarin tones. This supports the idea that language experience in both Mandarin and English are conducive for tonal perceptive abilities linguistically and non-linguistically. This said, a tonal language does not provide any advantage over English with tone salient linguistically at the syllable level. The English language may have enough tonal components to be sufficient for more complex tonal detection in Mandarin.

Another explanation for the similarities in tonal perception among the two groups is the varying dialects of Mandarin the participants spoke. In China, Standard Mandarin has been popularized all over the country, however almost every dialect speaks Mandarin with accents to different degrees (Ma et al., 2009). Furthermore, tone shapes may be subject to influence by morphotonemic alternations between dialects of Mandarin (Chen, 2000). Huang & Johnson (2010) also propose that American English listeners may be able to detect subtle pitch differences, which may be missed by Chinese listeners’ categorical perception of tone. If tones differ between dialects, then listening to linguistic tones of Standard Mandarin may not be as strongly categorized as their native dialect, leading to slower reaction times for the Mandarin group.

**Bilingualism**

The two groups in this study differed in tone language experience as well as acquisition of a second language (L2). Language experience in more than one language, bilingualism, should be addressed as a factor contributing to this study. Wang & Saffron (2014) found that non-tonal bilinguals performed better than the monolingual Mandarin speaking group, suggesting that bilingualism alone does facilitate statistical learning when learning a novel pseudo tone language. This study proposes that tone exposure in language may not be the contributing factor,
but rather a heightened ability to track regularities in a novel linguistic environment. Therefore, it is possible that the L1-L2 experience of the participants might be a factor for further research. Further research in linguistic tone discrimination including bilingual non-tone language speakers may support bilingualism over tone language experience.

**Accuracy**

Accuracy was high for both groups across all three conditions. (>90%). This is contrary to proposed predictions that the linguistic conditions would be more challenging for the English-speaking group in terms of both accuracy and reaction time.

Results indicated that Mandarin speakers were more accurate at discriminating word-word contrasts than monolingual English speakers. There was no such accuracy advantage in the non-word or pure tone conditions. The word-word condition was the only condition that differed between the two groups in that the Mandarin words for the bilingual group contained additional semantic information (i.e. words were meaningful). This was not seen for the non-word and pure-tone conditions (i.e. containing linguistic or non-linguistic irrelevant information for both groups), indicating that higher accuracy was seen when the two words were connected to meaning.

One possible explanation is that the Mandarin group may have been more accurate for the word condition because the pairs may have been more predictable in addition to meaningful. They learn the contrasts in tone for ‘minimal pairs’ in that they may have learned the different tone patterns and different words when phonemes remain consistent (the anomalies in the language such as ‘ma’ in the four tones means hemp, father, etc.), much like how children and adults learning English are told not to confuse minimal pairs that differ in a single phoneme (e.g. cat, sat, mat, hat). With the word pairs that differed by one tone to change the word meaning,
there was a closed set of options that the second word could be if the tones changed, thus making it easier for them to anticipate a difference or not in the word-word condition. This may have created an advantage over the monolingual English group because they may have been relying on the closed set of words and linguistic experience in Mandarin rather than the tones themselves as the distinguishing factor.

**Reaction Time**

Although there were no significant group differences in reaction time across the three conditions, there was a pattern that the bilingual Mandarin speakers took longer to respond to the stimuli (p = .08). There are a few proposed reasons for this finding. First, the patterns suggest that Mandarin speakers’ tonal knowledge might interfere with their judgment about the tonal contrast (Kaushanskaya & Marian, 2007; Pavlenko, 2012). Second, this may suggest that the knowledge and activation of word form in addition to word meaning may play a role in distinguishing tone differences, thus slowing down the reaction time from perception to meaning to responding. Wagner et al. (2001) state that target knowledge of semantic information is not recovered through more automatic access processes. Automatic access may be insufficient or may fail either because the cue-target associative strength is not strong enough to result in automatic knowledge recovery, or because task-irrelevant representations compete with, and interfere with the automatic recovery of task-relevant knowledge. A third proposed reason for the slower reaction times for the bilingual group is the randomization of the stimuli. All four conditions were randomized across all three-presentation blocks. For the monolingual English group, all blocks contained irrelevant linguistic and non-linguistic stimuli. However, the word condition included meaningful information for the bilingual group (word-word). Chen et al., (2016) suggest that once lexical tones have been acquired, they perceive them as phonological
categories and split them from other pitch variations that do not play a phonemic role, whereas non-tone language listeners perceive lexical and non-lexical tones on a psycho-acoustical basis. Hence, this group exhibits a more unified perception of pitch across the two domains. Therefore, acoustic salience may have also required more cognitive effort to switch from irrelevant, meaningful, linguistic, and non-linguistic information throughout the thirty minutes of the tone contrast discrimination task.

The overall outcomes of the study offer insights into the perceptual abilities of tone influenced by tone language experience, broaden our understanding of language acquisition and brain plasticity, and expand on existing literature regarding bilingual advantages acquired early on in life.

**Clinical Implications**

This research has possible implications to tone language learning and results may have future relevance to developmental and acquired disorders. Kohnert & Windsor (2004) state response speed on auditory nonlinguistic detection task differences were seen between intact monolingual and bilingual language learners, compared with monolingual children with language impairment (LI). The functional daily impact of children with LI may be significant given slower reaction times for auditory processing and can possibly have implications to strategies of learning for children with LI. The same can be said about children with reading disorders (RD) who may exhibit difficulties in recognizing and processing auditory patterns involving discriminating frequency and duration tonal patterns. This may co-occur with decoding deficits in these children (Walker et al., 2006). Decreased accuracy and increased reaction time for processing auditory information in both LI and RD in children can exacerbate difficulties in learning in an academic setting and being successful as learners.
Limitations and Future Directions

The pure tone condition was a more abstract auditory signal and future studies can create non-linguistic stimuli that uses low-pass filters at 500 Hz in order to take out the segmental parts of the word and non-word condition auditory files. The low-pass filter will take out the segmental parts and preserve the harmonics, thus preserving the naturalness of human voice and making the stimuli more representational to F0 instead on unnatural tone segments. However, the duration of the syllables would be harder to control for. It is also researched that individuals with musical experience present an advantage in perceiving tonal contours in linguistic and non-linguistic auditory contexts. Gandour et al., (2011) state that musicians showed greater pitch strength than Chinese speakers due to their ability to accurately encode rapid, fine-grained changes in pitch and thus the ability to detect minute variations in pitch (e.g. in tune vs. out of tune). Future studies that compare musicians who speak both tone and non-tonal languages would determine an advantage in tonal perception based on musical non-linguistic experience.

Future Directions

Future research is needed in this area of study. First, a larger sample size and a narrower age range can increase result validity and reliability as well as isolate tone perception capabilities to a particular demographic. Second, low pass filtered speech to create the nonlinguistic condition would be more representative of the natural tonal melody of Mandarin and therefore less abstract. A nonlinguistic condition low pass filtered from the word and non-word wav files would allow for consistency throughout all conditions and further analysis may additionally determine if any tone pattern can be isolated as easier or harder to perceive. A fourth duration condition would also have to be created to rule out syllable duration as a contributing factor to tonal perception. Lastly, to truly understand what brain areas are recruited for linguistic and
nonlinguistic tonal perception, the same study using functional neural imaging (fMRI) to track brain activity throughout the tone perception discrimination task would seek to isolate the brain activity and neural pathways activated across the three conditions. This may also attempt to explain what neural networks may be involved with additional semantic meaning and determine what is attributed to the increased reaction time in the Mandarin speaking group.

Acknowledgements

I would like to thank my department advisor for this study Dr. Pui Fong Kan and Fan-Yin Cheng for their support and assistance throughout the process of this thesis project. This study was funded by the SLHS department at the University of Colorado-Boulder and presented as a poster presentation at the ASHA Convention in November 2017.
References


