

# Music and Mandarin: Differences in the Cognitive Processing of Tonality

Laura Cray

Thesis submitted for the degree of Masters of Arts

Dr. Makiko Sadakata (Primary Reader)

Dr. Kimberley Mulder (Secondary Reader)

Radboud University

August 2017

### Acknowledgements

I'd like to thank my supervisors for their inspiring interest, insight, and encouragement,  
my friends for their tireless support throughout this whole process,  
and my parents, who made it all possible.

Acknowledgements	i
Table of Contents	ii
Abstract	iii
Chapter 1. Introduction	1
1.1 Music and Language	1
1.1.1 What is Music?	1
1.1.2 Music and Memory	3
1.1.3 Musicians vs. Non-Musicians	3
1.2 Music and the Brain: Neural Representations of Music and Language	4
1.3 Why Mandarin Chinese?	5
1.3.1 What is Mandarin Tone? A Comparison of Language Features	6
1.3.2 Mandarin Tone vs. Musical Tone	8
1.4 Language and Technology	9
1.4.1 Previous Neuroimaging Music Research (fMRI, fNIRS, and PET)	9
1.4.2 Previous Neuroimaging Mandarin Tone Research (fMRI, fNIRS, and PET)	10
1.5 Why EEG?	11
1.5.1 What is EEG?	11
1.5.2 N400, P600, and P200 Background Literature	12
1.5.3 EEG studies specifically relevant to this study	15
1.5.4. Expected Outcomes	17
Chapter 2. Methodology	18
2.1 Participants	18
2.2 Materials	19
2.3 Procedure	20
2.4 EEG Recording	21
Chapter 3. Results	21
3.1 ERP Data	21
3.2 Behavioral Data	26
3.2.1 Speech vs Non-Speech identification data	26
3.2.2. Speech vs. Non-Speech correct response rate	26
3.2.3 Dutch vs. Mandarin speakers	26
3.2.4 Tone Identification test	26
Chapter 4. Discussion	27
4.1 Discussion of ERP data	27
4.2 Discussion of Behavioral Data	29
4.3 Issues with study	30
4.4 Potential applications and directions for future research	30
References	31
Appendices	37

## Abstract

This thesis questioned whether the semantic and syntactic information of Mandarin words utilise similar processing resources as musical stimuli of matching pitch. Utilising Mandarin fourth tone (tone 4) and neutral tone (tone 5) words and extracted pitch contours modified to act as musical stimuli, EEGs were performed on fourteen participants who were either native speakers of Mandarin Chinese (n=9) or Dutch (n=5), as well as behavioral tests of tonal perception. It was hypothesised that 1) the native Mandarin speaking participants would demonstrate stronger N400 for tone 4 and a stronger P600 for tone 5, that exposure to linguistic and musical stimuli as presented in this testing paradigm will result in ERP data which has similar but significantly distinguishable features, and 3) that native Mandarin speaking participants would be significantly more accurate at pitch differentiation than Dutch participants. Dutch participants were expected to show either weaker or inconsequential effects. Analysis of the ERP components N400, P600, and P200 failed to reach statistical significance due to the small sample size, but indicated potential emergent trends which may be confirmed by continuation of the study on a larger participant pool. Comparison of the grand averages did not confirm the first hypothesis, finding that tone 5 elicited a stronger N400 and tone 4 a stronger P600 effect. The difference between linguistic and musical stimuli was found to be practically indistinguishable at N400, but a much stronger P600 effect was found for musical stimuli. Behavioral tests confirmed the third hypothesis of higher pitch differentiation ability among the native Mandarin speaking participants.



## Chapter 1. Introduction

Is Mandarin musical, or is music a reflection of linguistic tonality? Both music and Mandarin are tools people use to convey meaning, both rise and fall in pitch, and both can be described in terms of tonality. But how far does this connection go? If a piece of music and a tale of woe can both elicit a description of an experience of sadness, do we arrive at that similarity by connecting disparate neural experiences, or do we arrive at this relationship through similar cognitive processing mechanisms?

This thesis will investigate the question of whether language and music are processed in a similar way by performing an EEG study on native Mandarin speakers and native Dutch speakers. By examining specific ERP components, namely N400, P600, and P200, this study aims to test three hypotheses. First, that native Mandarin speaking participants produce a stronger N400 when processing Mandarin tone 4 and a stronger P600 when processing Mandarin tone 5. Second, that exposure to linguistic and musical stimuli as presented in this testing paradigm results in ERP data which has similar but significantly distinguishable features. Third, that native Mandarin speaking participants are significantly more accurate at pitch differentiation than Dutch participants.

To begin with, a definition of music will be provided, and the impact of music on the brain will be presented. Then, having dipped a toe into the effects of musical training, the neural representations of music and language on the brain will be discussed in more depth. Having laid out the potentials of research on music to reveal further insight into cognitive processing, the benefits of studying Mandarin Chinese will be explored and Mandarin tone defined. With the impetus of the selection of music and Mandarin Chinese as the driving forces in this thesis having been explicated, a deeper look will then be taken at what unique insights particular neuroimaging techniques can offer into the relationship of music and language. Next, the benefits of using EEG will be discussed, EEG studies specifically relevant to this study will be presented, and the expected results of the EEG study will be laid out. Finally, the methodological design of this study will be described in depth, the results presented, and the potential impact of this study's findings on the field of linguistics will be discussed.

## Chapter 1.1 Music & Language

### **1.1.1 What is Music?**

Every culture known to man has a form of music (Pearce & Rohrmeir, 2012). It can communicate suffering, love, anger, and joy. Music comforts, it enthralls, it enrages, and it celebrates. But what exactly is music? How can we, as scientists, quantify this seemingly universal occurrence in the human experience?

Before diving into a discussion of the relationship of music and language, or the representation and processing of music within the brain, it is first necessary to provide a definition of what exactly is meant by "music." Pinning down a one-size-fits-all definition of what "counts" as music and operationalising that for a study is difficult, due to the subjectivity of perception and the sheer number of defined "types" of music around the globe (McDermott & Hauser, 2005), so before further discussion of how music has been operationalised in previous research it is necessary to establish a working definition of "music."

Researchers have delved into the nature and scope of human ability to create music since the mid 19th century (Nettl, 1983) with the goal of finding ways to quantify it and lock music within the scientific method. Researchers from various fields have described music with terms that recommend attention to pitch and harmony (Krumhansl, 2001) or to rhythm and intentional sound (Burke, 2015), but even today the exact elements a sound must contain in order to be quantifiably musical are not certain. Like art, the judgement of whether a sound from a violin and a sound from

a trash can lid can both be musical is highly subjective. For example the National Association for Music Education in the United States lists the elements of music as “pitch, rhythm, harmony, dynamics, timbre, texture, form, and style/articulation” (NAFME.org, 2014), but in the United Kingdom music is taught as defined through the “inter-related dimensions” of “pitch, duration, dynamics, tempo, timbre, texture, structure and appropriate musical notations” (GOV.uk, 2013). When the public education systems of two highly communicative and historically linked world powers cannot even agree on a definition, surely more research is necessary.

Across the literature an exacting definition of music seems to be largely reliant on the predilection of the researcher and their field, be that linguistics, acoustics, anthropology, psychology, or engineering. An approach to music from a psychoacoustic perspective might suggest that sound itself be analysed in terms of frequency, amplitude, and phase, with music adding additional elements of pitch, intensity, and timbre (Iakovides et al., 2004). An ethnomusicologist might approach a definition of music by describing it in terms of cultural context, artistry, and shifting personal and societal values (Nettl, 1983). From the perspective of a psychologist music can be explored with a focus on the ability of music to affect general cognitive functioning and its potential as a tool in diagnosis and treatment of cognitive impairment, with due regard given to how the research on music from a number of perspectives has led to discoveries directly relevant to psychologists (Jäncke, 2012). The role of evolution in making music definitively different from language has also been explored, with researchers such as McDermott & Hauser (2005) and Patel (2003) searching for answers as to whether and or what makes human musical ability unique, what role evolution has played in our ability to create music, and delving into what further understanding of the mechanisms that produce and process music within the human brain and physiognomy can reveal about our ability for language.

If a song is music, and a song is made up of notes, does an individual note from a song not still hold some element of musicality from the original composition it was a part of? Understanding music in this way allows for a wide array of operationalisation possibilities for music, and it was this approach that was taken to operationalise music for the EEG study conducted for this thesis. Being interested in participant reactions to both specific grammatical features of specific word stimuli and the difference between participant processing of musical and word stimuli, it was necessary to utilise some kind of musical stimuli that would not lend itself too readily to categorisation. For example, the use of stimuli from a violin was rejected because two categories (voice and instrument) would be very obvious to participants. The desire was to use music stimuli of as similar an auditory value as possible. While vocalised musical stimuli (e.g. singing) could be used, it would be difficult to make musical and linguistic stimuli match, suggesting that this be ruled out as well in any testing paradigm seeking to create comparable stimuli. Speakers of Mandarin Chinese have previously been shown to be highly sensitive to stimuli created by extracting pitch contours (Krishnan, Xu, Gandour, & Cariani, 2004), and pitch contour has been proven to be a feature highly attended by speakers of non-tonal language as well (Friedrich, Alter, & Kotz, 2001), suggesting that extracted pitch contours might contribute to a functional testing paradigm designed to test for differences in cognitive processing (through EEG) and tonal differentiation ability through behavioral tests. While previous research on the effects of exposure to short musical stimuli has been mostly focused on the perception of music after priming (Daltrozzo & Schon, 2009; Painter & Koelsch, 2011; Steinbeis & Koelsch, 2011), in studies not focused on priming effects chord-length musical stimuli of as short as one second worked effectively to elicit significant ERP data (Steinbeis & Koelsch, 2008; Daltrozzo & Schön, 2009; Hung & Lee, 2008). This suggests that if an extracted pitch contour is modified to a computerised note (e.g a sine wave) of one second in length it could still be capable of eliciting significant ERP components such as N400 or P600 in participants, particularly participants who are native speakers

of a tonal language. However as this operationalisation of tone has not been explored before this thesis the methodological design of this study broke new ground, combining aspects of previous studies into a new testing paradigm. However, discussion of priming effects and experience with music raises the question of what effects storage of experience with music might have on how musical stimuli are processed, and it is this topic which will be addressed in the next section.

### **1.1.2 Music and Memory**

This study's interest in similarities of processing begs the question of whether the relation of music and language can be seen in specific processing resource usage. The shared syntactic integration resources hypothesis, or SSIRH (Patel, 2008), proposes that they do, and in fact overlaps have been found in relation to syntactic working memory (Klajevic, 2010), executive control, and short term memory. When exposed to sentences in a syntactic complexity manipulation condition and music, participants have exhibited decreased working memory (Fiveash & Pammer, 2014), indicating that some aspect of musical stimuli input results in shared processing of syntactic working memory just as with language. The effects of short training periods can even be seen in different age groups.

Children exposed to a short period of training in music or a foreign language have been found to demonstrate no significant improvement in executive control tasks afterwards (Janus et al., 2016), however twice weekly musical therapy sessions of twenty to thirty minutes have been found to result in significant improvement in speech content and fluency of people with Alzheimer's like cognitive impairment, indicating some interaction between the regions processing music and producing language (Brotons & Kroger, 2000). This would seem to align with findings focused on the interconnectedness of music, language, and memory. Consider three common situations combining music, language, and memory: first, the infuriating inability to recall the words to a song; second, the ability to suddenly remember a song from early childhood after hearing a few bars despite have spent no conscious effort on remembering it; third, the use of music as a memory aid, such as for an advertisement for a fast food chain or professional service. These anecdotal experiences of music, language, and memory seamlessly interwoven can be seen through a scientific lens as well, and seemingly paradoxical findings such as in the studies above highlight the importance of further exploration into the effects of short training sessions in studies (such as this one) on language and music processing. Such research can even had medical implications. For example, patients with Alzheimers have been found to demonstrate more accurate recognition of sung lyrics than spoken lyrics (Simmons-Stern et al., 2010), which would seem to show that musical memory is in some way stored differently from the memory of language. It has been suggested that the improvement found by Simmons-Stern et al. (2010) might be due to a heightened state of excitation caused by participation in musical activity, however further research is necessary in order to determine the accuracy of either finding. Here the connection of music, language, and memory raises an interesting question of what applications further knowledge in this sector may have in the medical field. Debilitating ageing diseases such as Alzheimers and other similar forms of Dementia causing diseases leave millions of people every year with terrible memory issues. Clarifying whether and how cognitive processing resources are utilised by language and music not only adds to the discussion of theories such as SSIRH but also to the debate on the potential efficacy of techniques which may have strong potential as therapeutic measures debilitating diseases such as Alzheimers.

### **1.1.3 Musicians vs. Non-Musicians**

The question of music's effect on the brain leads to a natural questioning of the extent of the impact of musical training on the brain. Musicians have been shown to demonstrate faster reaction

times (measured via event related potentials) to incongruities than non-musically trained people upon exposure to musical incongruities (Besson et al., 1994). Musicians have also been found to be more accurate at differentiating the tones of Mandarin Chinese more accurately than non-musicians when presented with a three tone option judgement task (Hung & Lee, 2008). Physical differences have even been proven to exist between musicians and non-musicians (Gaser & Schlaug, 2003). Upon measuring the volume of grey matter using a magnetic resonance imaging (MRI), Gaser & Schlaug (2003) found a significant positive correlation between level of musicianship (from non-musical to professional) to greater volume in the peri-Rolandic regions of the brain. This research clearly shows that there are cognitive differences which should be accounted for between musicians and non-musicians when testing for processing differences as in this thesis. However it also raises the question of whether exposure to shorter periods of musical stimuli may have an effect on the processing linguistic stimuli. To investigate this further the neural representations of music and language will be examined in the next section.

## Chapter 1.2 Music and the Brain: Neural Representations of Music and Language

To bring this discussion of music and language to a close a more in depth look at the cognitive mechanisms and neural representations of music and language in the brain will now be taken. The chapter will close with an exploration of hemispheric specialisation.

Harmonic and dissonant chords have been shown to elicit a P600 effect that is significantly similar to that elicited by spoken stimuli in a match/mismatch (e.g. love/hate) setup (Patel et al., 1998). These findings have been supported by a joint EEG and fMRI study conducted by Steinbeis & Koelsch (2008) to investigate how semantic meaning in music is carried compared to the meaning in language. The researchers utilised a match/mismatch experimental design, using opposing words for the linguistic stimuli and consonant/dissonant chords for the musical stimuli (2). This match/mismatch setup with an elicitation task of relative harmony/dissonance is quite popular in the field, and is suitable for the research focus as proven in previous work (Patel et al., 1998). Harmonic stimuli has even been found to be able to prime target words, supported by the elicitation of the N400 ERP component (Steinbeis & Koelsch, 2008:3). For the linguistic stimuli the researchers found that the middle temporal gyrus showed heightened activity (4). N400 effects were also elicited by musical stimuli with an incongruous pattern, however interestingly the superior temporal sulcus was found to show stronger activity in response to the musical stimuli, rather than the middle temporal gyrus as for language (4). This supports the second hypothesis of this thesis by showing that music and language share some processing patterns and locations, but differ in others.

As language is a produced pattern of sound from which we draw meaning, and people draw meaning from patterns in music as well, the question arises of whether we process music as though it has syntax. This question has been confronted by Maess, Koelsch, Gunter, and Friederici (2001), who designed a study utilising magnetoencephalography (MEG) to seek out whether music has syntax like language. Maess et al. (2001) found that musical stimuli caused high processing activation in Broca's Area and its homologous right hemisphere area (543), and this, combined with their findings on the increased activation in syntax processing centers in correlation with increases in pattern violation in the musical stimuli, led to support for the conclusion that music is processed as having a syntax of some sort.

The similarity of processing of music and language can also be seen when participants are asked to produce sentences and musical phrases. PET research by Brown, Martinez, & Parsons (2006) found that production of musical phrases resulted in notable activation in the "supplementary motor area (SMA, medial BA 6), pre-SMA, primary motor cortex (BA 41), lateral premotor cortex (BA 6), frontal operculum (BA 44/45), anterior insula, primary auditory cortex (BA 1), secondary auditory cortex (BA 22), and superior temporal pole (anterior BA 22/38, or planum polare)" (2795), whereas production of the sentences resulted in notable activation in the "pre-SMA, sensorimotor cortex (BA 4 and 3), premotor cortex (BA 6), frontal operculum (BA 44/

45), superior frontal gyrus (BA 8, 9), cingulate motor area (BA 24/32), cingulate gyrus, anterior insula, inferior parietal cortex (BA 39), primary and secondary auditory cortex, middle temporal gyrus (BA 21), hippocampus, and ventral temporal pole (BA 38)” (2795). Production of the musical phrases and sentences both resulted in activation of “bilateral SMA, left primary motor cortex (BA 4), bilateral premotor cortex (BA 6), left pars triangularis (BA 45), left primary auditory cortex (BA 41), bilateral secondary auditory cortex (BA 22), anterior insula, and left anterior cingulate cortex” (2795). Furthermore, comparative analysis of the musical and linguistic tasks revealed practically indistinguishable brain area activation in the “primary motor cortex, supplementary motor area, Broca’s area, anterior insula, primary and secondary auditory cortices, temporal pole, basal ganglia, ventral thalamus, and posterior cerebellum” (2792). These findings highlight the localisation of specific brain areas which may be activated by the production of musical versus linguistic stimuli, as well as the number of areas which show significant overlap. However, while the findings of Brown, Martinez, & Parsons (2006) reveal interesting differences and similarities in neural activity during the production of music and language which may remain relevant especially in the search to answer the second hypothesis of this thesis, their study measures neural activity during production, and so their findings can only lend support up to a certain point as the question guiding this thesis is focused on the processing of sound rather than the mechanisms involved in the production of it.

Some have argued that the differences in linguistic and musical sound processing are due to specialisation of auditory cortexes. Zatorre, Belin, & Penhune (2002) highlight these differences between music and speech such as timing, pointing out that melody is created by variation of pitch, and the distance between pitches tends to be significantly larger in speech than in music. There is also evidence that the left and right cortices may each specialise in processing the temporal data and frequency of input (Liégeois-Chauvel et al., 1999; Liégeois-Chauvel et al., 2001; Zatorre, Belin, & Penhune, 2002). This argument of specialisation is further borne out by research showing that musical training can have a physical effect on the grey matter density of the perirhinal regions of the brain (Gaser & Schlaug, 2003). A landmark longitudinal study by Schlaug et al. (2005) further supports this. In this study both participants at 9-11 years of age both with and without musical training were observed to demonstrate “strong bilateral activation of the STG,” but the musical group showed significantly more activation of the superior temporal gyrus (Schlaug et al., 2005: 226). Although only small changes were found in 5-7 year olds, significant effects of practicing musicianship on “motor and auditory” areas began to show by 9-11 years old, and further testing of adults confirmed similar patterns of cognitive processing (226). This not only highlights the important impact that experience of musical training can have on a participant, it also confirms the findings previously discussed in this section of the primary activation area differences and similarities of musical and linguistic stimuli.

With the terms of musical stimuli defined and the cognitive mechanisms responsible for processing and producing music explored, the target language of this study - Mandarin - can now be delved into. In the following chapter the benefits of operationalising Mandarin Chinese to test this thesis’ hypotheses will be presented, and relevant features of the language will be presented. The chapter will close with an exploration of the similarities and differences of Mandarin tone and musical tone.

### Chapter 1.3 Why Mandarin Chinese?

One facet of Mandarin that highly recommends its viability, even the necessity, of its inclusion in studies about the intertwined relationship of music and language in the brain, is the sheer volume of its speakers worldwide. According to an Ethnologue census there were 1,091,782,930 speakers of Mandarin Chinese worldwide as of 2013, with that number rapidly

increasing (Ethnologue.com, 2017). In the United States alone, Chinese language programs at universities have seen an increase of 115% over the last 18 years, and there is even more expansive growth predicted in the future (Walker, 2016). The Chinese speaking consumer market is also on the rise, outstripping English language based companies such as Amazon (Osawa, 2014). With this increase in speakers of Mandarin and demand for trade with Mandarin speaking countries, a firm foundation of linguistic understanding is both beneficial and necessary for harmonious relations between nations in the future.

### 1.3.1 What is Mandarin Tone? A Comparison of Language Features

What is Mandarin tone? In the previous chapter musical tone was explored, and defined in terms of a frequency, hemispheric processing centers, and ability to communicate emotion. Mandarin tone possesses many similarities to music tone, but to what extent are they truly similar? Where is that dividing line? To begin to answer these questions this chapter will delve into why Mandarin Chinese is such an increasingly common tongue worldwide, what distinguishes Mandarin tonality from other Chinese dialects, and what distinguishes Mandarin from other tonal languages,.

The meteoric rise of Mandarin Chinese's popularity over the last fifty years can be traced back to the geopolitics of China. Mandarin is far from the only language of China, but as the official language of the government in Beijing, it holds an important role in the region, and the Chinese government has enforced a strict regime of instruction for all citizens (French, 2005). Mandarin Chinese has four tones (see fig. 1): even, rising, falling-rising, and falling, referred to as first to fourth tone, respectively, and by changing the tone of a syllable the meaning of the syllable can be changed (Wang et al., 1999).

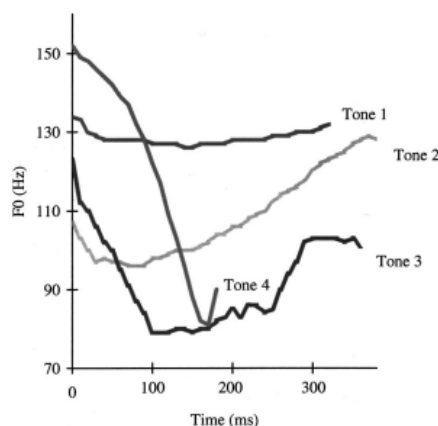


Figure 1: F0 of Mandarin Tones, from Wang et al. (1999)

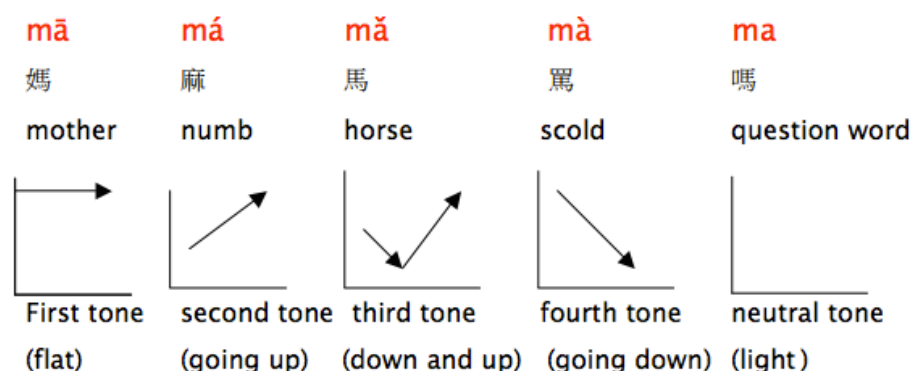


Figure 2: Example of how changing a tone changes the meaning of a syllable in Mandarin, from MissPandaChinese.com

Furthermore, there is a neutral tone (referred to as the fifth tone in this paper) which can be applied as a result of tonal sandhi or to turn a sentence into a question (Chen & Xu, 2006). For example, when the final word of the phrase ‘*Nǐ chīle ma*’ is spoken with a neutral tone (tone 5) the sentence means ‘*Have you eaten?*’, but if the final word is produced as tone 3 the sentence becomes ‘*Nǐ chīle mǎ*’ meaning ‘*You ate the horse.*’ These Mandarin tones can be differentiated based on defining characteristics such as their amplitude, duration, and F0 contour (Whalen & Xu, 1992). Much literature ignores the fifth tone -- possibly considering it relatively unimportant in comparison to the main four tones -- however when Chen & Xu (2006) examined the F0 contours of neutral fifth tone syllables they found that they had pitch targets un-reliant upon the preceding tone, indicating that it holds unique information, distinguishable from the classic four tone. This contrast of evidence highlights the need for further research on the fifth tone, a gap in research this thesis study attempts to fill.

The majority of current research on tonal languages is focused on Mandarin, Cantonese, and Thai. While all of these are tonal languages, the tonality of each must not be conflated. Thai, for example, has five tones, Mandarin has four, and Cantonese has six (Kaan, Barkley, Bao, & Wayland, 2008; Lee et al., 2015). Deeper differences exist as well. For example, native Thai speakers demonstrate higher sensitivity to late frequency contours than native speakers of the non-tonal English as well as native speakers of the tonal Mandarin Chinese (Kaan, Barkley, Bao, & Wayland, 2008).

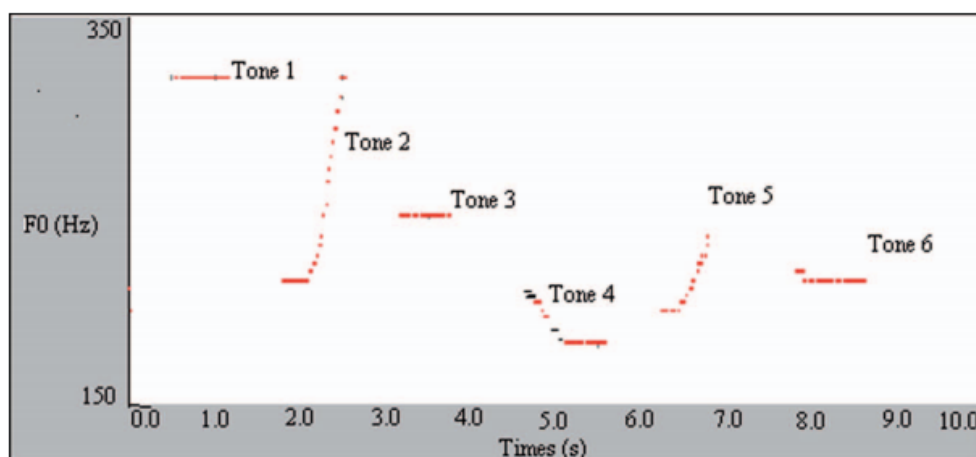


Figure 3: Cantonese Tones (Lee et al., 2015)

With tones capable of conveying lexical and musical information, the question of how - or even whether - to separate any discussion of tones must arise. This question can be further explicated as a question of how people who speak a tonal language process tone from different input sources, such as a vocalisation produced during a normal communicative act or a note produced from a musical instrument. This is what has given rise to the use of the term “tone” being utilised as a broader term, used to reference the intended or perceived pitch of a sound, and may be used in reference to musical or linguistic phenomena as it is in this study (Lee & Lee, 2009).

Lacking experience in meaningful tonal differentiation, many people from non-tonal languages such as English and Dutch struggle greatly to both perceive and produce Mandarin tones (Wang et al. 1999; Kiriloff, 1969). When American English speakers participated in a training program aimed at assisting them in the creation of new phonetic categories and then had their ability to differentiate the four tones of Mandarin Chinese tested they scored significantly better than participants who received no training (Wang et al. 1999). This highlights the influence that

experience can have participant ability to differentiate tones, as well as the importance of careful consideration being given to training setups in future research.

Difficulties with tonal differentiation are not just localised to non-native speakers. Infants who have been adopted from Mandarin speaking parents into a family from a non-tonal language culture have also been shown to lose their ability to differentiate important tonal contrasts quickly after adoption. Zhou & Broersma (2014) found from a participant pool of twenty-six adoptees that even though the children had been exposed to Mandarin for a comparatively short time (a mean of 2.4 years) and had lived in the non-tonal environment for a mean of five years, they failed to perform better than a control group of Dutch children (65). Finding this result, even after providing training (Zhou & Broersma, 2014), is very interesting, as it begins to narrow down a gap during which it can be argued that the ability to differentiate tones with native accuracy is naturally acquired. It was demonstrated by Li & Thompson (1977) that native acquisition of Mandarin tones by children was achieved quite comprehensively by age three, and beyond that they found that their participants acquired the fourth and first tones the quickest, indicating a faster integration of those neural pathways into the children's brains (187). This finding guided the selection of fourth tone as the other representative Mandarin tone in the experiment created for this thesis, as it provides a clearer contrast to the neutral tone than the first tone, and both were suitable to create musical stimuli from. But in what ways does Mandarin tone truly differ from musical tone? It is to this topic we will pivot for the rest of this chapter.

### **1.3.2 Mandarin Tone vs. Musical Tone**

With tones capable of conveying lexical and musical information, a question of separation soon arises. This can be further explicated as a question of how people who speak a tonal language process tone from different input sources, such as a vocalisation produced during a normal communicative act or a note produced from a musical instrument.

Sound, as vibration, can be visualised and discussed in a variety of terms, so it is important to define a pertinent few here. Some terms of prime importance include frequency, periodicity, pitch, and tone. As in tonal languages such as Mandarin a change in pitch changes the entire meaning of a word (Deutsch et al., 2006) but in non-tonal languages pitch functions in a significantly different fashion, reaching a definition of pitch which functions in discussions of tonal and non-tonal languages is imperative. Frequency (F0) and periodicity are measurements of sound waves: a sound wave with a longer period will be perceived as having a lower pitch, while the frequency (in hertz) gives the number of vibrations in a second and is a prime descriptor of choice of linguists researching acoustics (Langner & Ochse, 2006).

Pitch is most simply a conveyance of the height of the frequency of a sound (Patterson et al., 2002). To better visualise pitch the contour of a sound can be extracted with a program such as PRAAT (Boersma & Weekink, 2013), allowing the sound to be compared to other similar sounds, modified, et cetera. A pitch contour can be extracted from vocal sound as well as other sources such as musical instruments or even from computer generated sounds. Extracted pitch contours (EPC's) have been used in a multitude of studies to examine a variety of linguistic questions, but none have utilised all in one study the exact combination developed for this study.

Pitch perception and production can be analysed to make inferences about neurological states and developmental stage of participants (Loui et al., 2015). Pitch is used to described sound of both music and language (Jäncke, 2012), and the native language of a participant has been indicated to affect the ability of participants to differentiate pitches, with speakers of tonal languages possessing a significant advantage over speakers of non-tonal languages (Giuliano et al., 2011). For example, it has been established that native speakers of Chinese are better at differentiating pitch than native English speakers (Krishnan, Gandour, and Bidelman 2010), but the root and the extent of the differentiation still requires investigation, as provided by this thesis.



Many researchers have theorised that a music/language connection may be found in the studying the prevalence of demonstration of absolute pitch. Deutsch et al. (2006) examined musicians from two groups - English speaking and Mandarin speaking conservatory students - and found a significant correlation between the acquisition of perfect pitch and the critical period. This seems to indicate that absolute pitch may be acquired like any feature of language, in line with the critical period hypothesis which has long been established for language acquisition (Johnson & Newport, 1991; Kuhl, 2005; Birdsong, 2001). The relationship of musical pitch and linguistic pitch has been investigated by looking specifically at native American English speaking musicians reacting to stimuli from native Mandarin speakers (Hung and Lee, 2008). However, this paper failed to pinpoint a statistically significant correlation between musical tone and lexical tone, finally unable to establish whether the results elicited by this particular operationalisation of musical tone and lexical tone could be elicited by substitution of another stimuli such as basic background noise. This highlights yet another research gap which the data from this thesis's study may provide useful progress towards filling, as well as the risks of writing where there is little previous research to duplicate and everything must be gathered from different fields. To examine the interrelationship of musical and linguistic tone further there are a variety of technologies which can be utilised to provide insight into the similar and dissimilar processing of music and language, and in the next section the correct selection of equipment to match a stated research question will be addressed and the selection of EEG for and in support of the hypotheses of this thesis will become clear.

#### Chapter 1.4 Language and Technology

The use of advanced technology has already infiltrated every aspect of everyday life, from education, to surgery, to shopping for everyday groceries. With this in mind it is not surprising then that after years of linguistic research being conducted with only behavioral experimental setups researchers would leap at the chance to both re-examine old questions in this new technological light and pose new questions, the answers to which could previously only be hypothesised. The technology now available to researchers interested in the how the human brain processes language opens up new ways to see not just how people perform language externally, but what physical processes are going on beneath the surface. In the following chapter the pros and cons of using various neuroimaging technologies in linguistic research will be presented, with a focus on fMRI, MEG, fNIRS, and PET. By explaining the different insights each technology can offer, the selection of EEG for this study will become clear. Next, previous research on music undertaken with these neuroimaging technologies will be discussed. By discussing these studies the applicability of this study's results to current research will also become evident. Finally, previous research utilising these technologies to study Mandarin tone will be explored. The absence of a discussion of EEG may be noted in this chapter. However, as the focus of the study in this thesis is EEG, it has been assigned its own full chapter in order to fully explore the potential, possibilities, and unique aspects of this technology.

##### **1.4.1 Previous Neuroimaging Music Research (fMRI, fNIRS, and PET)**

Current neuroimaging technologies have the potential to lay bare inner areas of activity of the human brain in realtime during participant completion of both active and passive tasks. The unique insight these kinds of studies can provide has been utilised by many researchers to examine a variety of effects on the brain. An fMRI study using twenty different pieces of music established that classical music is capable of affecting people's emotional states, with activity strongest in the ventral striatum, dorsal striatum, anterior cingulate, and medial temporal areas (Mitterschiffthaler et al., 2007:1150). These areas are associated with "reward experience and movement", "targeting attention", and "appraisal and processing of emotions" (1150), and validates the emotional response

people experience when listening to familiar music, and begs the further question of whether a difference would be noted if the study was re-run with unfamiliar stimuli. The role of the right hippocampus and left inferior frontal gyrus on a task involving musical recollection has also been examined using an MRI, with significant activity being found in the right hippocampus during a task of musical recollection, indicating that that region plays a significant role in musical memory (Watanabe, Yagishita, and Kikyo, 2007). With an fMRI study the difference of perception of harmonious and dissonant musical tones by musicians and non-musicians has been questioned, and the areas of highest activation located as the “inferior and middle frontal gyri, premotor cortex and inferior parietal lobule” (Minati et al., 2009: 87). Differences in the very acoustic signals of music and speech have been investigated, and a hypothesis posited that the higher activation of the right hemisphere during processing of musical stimuli might have come to exist as a complementary effect of left hemisphere specialisation in linguistic speech sound processing (Zatorre, Belin, & Penhune, 2002).

Expanding the question of hemispheric specialisation back in time developmentally, fMRI scans of newborns have allowed for the musical processing capabilities of 1-3 day old children to be analysed, revealing harmonic music as being processed primarily in the right hemisphere and dissonant music primarily in the left hemisphere (Perani et al., 2010). The accumulation of such early data from what are, in effect, world-naïve subjects, provides valuable insight into what may be the ‘factory settings’ of the human brain. For example, MEG has since been further used to prove that musical training is beneficial to the development of neural pathways in infants (Zhao & Kuhl, 2016). This interaction of nature and nurture has been reviewed in depth by Schlaug et al. (2005), with the effects of musical training on the development of cognitive functions fully examined, and an MRI study on the effects of adult musical training on grey matter volume conducted which found an increased volume in the peri-Rolandic area, which is also responsible for speech processing (Gaser & Schlaug, 2003).

Adults tested for cortical blood flow during participant exposure to musical stimuli were found to have the strongest increase during music they ‘preferred’ or found to be ‘motivational’ (Bigliassi et al., 2014), indicating an element of personal preference which could become a confound, especially in testing paradigms which utilise a full and recognisable song. The lateralisation of such blood flow in response to specific types of noise was also found to reflect the difficulty participants had in categorising the sounds due to researcher controlled noise interference (Santosa, Hong, and Hong, 2014). Furthermore, an fNIRS study conducted by Ferreri et al. (2014) demonstrated that playing music to participants while they completed a non-musically related task caused them to experience a notable strain on their episodic memory. This finding is supported by a study run by Platel et al. (2003) which showed that musical stimuli which was familiar utilised a different neural network than stimuli which were unfamiliar. Further PET research has also been conducted which shows specific areas of the brain to specific areas, for example linking Brodmann’s Area to melodic phrase structure processing (Brown, Martinez, and Parsons, 2006).

In conclusion, neuroimaging research on the cognitive processing of music and the neurological relationship of music and language appears to indicate that a significant number of the areas activated by musical input (such as Brodmann’s area, the per-Rolandic area) also happen to be heavy processors of linguistic input. A preference to process harmonic musical input in the right hemisphere is also indicated. Next, relevant neuroimaging studies of Mandarin tone will be presented.

#### **1.4.2 Previous Neuroimaging Mandarin Tone Research (fMRI, fNIRS, and PET)**

A cross-linguistic fMRI study by Gandour et al., (2003) found that the right hemisphere showed the highest activation during processing of intonation in Chinese, which is used to convey

both grammatical information (via linguistic tone) and emotional tone. Considering the high preference for right hemisphere processing of musical input discussed in the preceding section, this finding is most curious, appearing to support the second hypothesis of this thesis, that musically tone stimuli and linguistic tone stimuli may elicit results similar and yet dissimilar. Furthermore, it has been suggested that upon encountering pitch and tone, Broca's Area supports tonal processing while the superior temporal gyrus carries the weight of pitch analysis (Nan and Friederici, 2012). As this study was conducted on Mandarin speaking participants only the question arises of whether a similar activation pattern would be found in speakers of different ages and linguistic backgrounds, or whether the effects would be similar if the study were conducted with native speakers of, for example, Thai. In the meantime, however, it has further been proven with an fMRI study that participants who learned Mandarin demonstrated increased neural plasticity in their right hemispheres (Wang et al., 2003), lending further support to the theory raised in the last section that tonal languages utilise both left and right hemispheres to a significant degree, and raising the question of -- if indeed music is mainly processed in the right hemisphere -- where exactly Mandarin tone fits between language and music. In a spoken word elicitation paradigm, tones have been found to produce higher activity in the right inferior gyrus (Liu et al., 2006), and in an fMRI study of native English speaking participants who received training on how to differentiate lexical tones of Mandarin more activity was elicited in the left superior temporal gyrus and increased activity in the right inferior frontal gyrus (Wang, Sereno, Jongman, and Hirsch, 2003), findings not in discordance with the previously discussed research. In conclusion, this research indicates that perception of Mandarin tonality utilises both regions primarily associated with linguistic processing and also recruits analogous right hemisphere areas for cognitive processing of some sort. With the research of alternative neuroimaging technologies on the questions examined by this thesis the following section will highlight how EEG can produce effective support for the hypotheses of this thesis.

## Chapter 1.5 Why EEG?

In the following section the function of EEG in linguistic research will be explained, pros and cons of utilising this specific equipment weighed, and relevant background literature reviewed. The value of examining the ERP components specifically relevant to this study (N400, P600, and P200) will be discussed, and their potential contributions to the understanding of language comprehension will be discussed. The section will close with a discussion of EEG studies specifically relevant to the aspects of music and language comprehension examined in this study.

### **1.5.1 What is EEG?**

As the neural pathways dedicated to functions such as phonological, syntactic, and semantic processing are activated in response to a linguistic (or non-linguistic) stimuli a participant experiences, these areas draw electrical amperage. An electroencephalogram (EEG) may then be used to measure the alterations in the brain's electrical activity. This ability to monitor and/or record the activity in a person's brain means that EEG may be used not just in research but also in active medical situations to diagnose conditions such as epilepsy, dementia, brain injury, and sleep disorders (Mayo Clinic Staff, 2014).

When considering utilising EEG it is important to first determine the relevance of its unique capacities. The main benefits to be discussed here are the interplay of three factors: cost, accessibility, realtime monitoring, and preliminary testing.

One benefit of using EEG for linguistic research is the cost. Depending on the number of sets, channels, and batteries, but assuming a standard number of each, a fully equipped EEG setup can cost anywhere from 40,000-300,000 USD (Stemmer & Connolly, 2011). For comparison, a

fully functioning fMRI setup could easily cost upwards of one million euros (Stemmer & Connolly, 2011).

The training required for operating an EEG setup ranges from a one day seminar to a multi-week course. However many of the operations are quite basic - preparing the participant's skin for a good connection with the electrodes (known as impedance), correctly affixing a cap, monitoring the signal, caring for the equipment, et cetera. The real difficulty lies in the conception of how to operationalise the technology and in the programming of the computers, which both require knowledge of intended ERP targets and background programming knowledge, respectively. Compared to the risk of operating an fMRI machine or MEG the EEG becomes clearly apparent as an appropriate starting point for research on a specific aspect of linguistic processing, or as the foundation for a closer examination of a particular time-locked aspect of processing within EEG framework. The preparation required for participants is minimal - they might be requested to wash their hair before the experiment or fill out a background questionnaire, but aside from that they need not receive any further training or physical preparation. The setup itself lends itself to a rather comfortable situation as well; all the but those with the most extreme cases of claustrophobia should feel comfortable in the soundproof booth of an EEG setup, and while participants should restrict movement to inside the booth once connected, their physical movements are otherwise unimpeded, unlike in an fMRI or MEG.

An aspect of running an EEG study which may become problematic is the time required from participants. While an actual experiment may only take twenty minutes, the setting up process may take anywhere from thirty minutes to an hour to properly complete, depending on the experience of the researcher, the preparation done ahead of time (filling syringes and preparing applicators, etc.), and even the shape of the participant's head.

The key benefit of utilising EEG lies in how the researchers can capture not just response times or accuracy, but also a timeline of activity surrounding the brain's processing of a particular stimuli, down to the millisecond. This is done by measuring ERPs (event related potential), and then analysing the data for specific components (Drijvers et al., 2016). These ERPs are the products of averaging the results of many participants and multiple trials of a particular stimuli (Mulder et al., 2016). Depending on the question under investigation the selection may be made from a multitude of components which have been tied to specific processing tasks. Due to the complexity of the activity of linguistic processing within an environment of multiple input factors, the possibility of encountering confounds such as various types of memory or surprisal or even the effects of background processing is likely. Top components which have been tied to specific cognitive processes include N400, N100, P600, P300, and P200 (Stemmer & Connolly, 2011).

### **1.5.2 N400, P600, and P200 Background Literature**

The N400 is an ERP component which has been linked to the processing of meaning (Kutas & Hillyard, 1984; McCallum et al., 1984; Holcomb & Neville, 1990; Besson et al. 1994; Ainsworth-Darnell, Shulman, and Bolland, 1998; Koelsch et al., 2004; Kutas & Federmeier, 2011). Kutas & Hillyard (1984) linked N400 to semantic priming, utilising sentences with low, medium, and high Cloze probability words in a sentence final position to test whether a relationship could be found between the 'expectedness' of a word and the N400. Beyond simple contextual congruity, a positive correlation was found between elicitation of N400 and the semantic priming paradigm, indicating that N400 is indeed tied to semantic processing (Kutas & Hillyard, 1984). This established N400 as a significant indicator of semantic processing, making it an ideal component to be analysed to answer the research question of this thesis on the categorisation of meaning when processing music and Mandarin Chinese.

The ERP component N400 has further been established as being elicited by semantic incongruity, and it is suggested that it be treated as a late endogenous component (McCallum, Farmer, and Pocock, 1984). One aspect which must be attended to here is the difference between data elicited by visual versus audio stimuli, and this issue has, if not spawned then at least contributed, to the never-ending number of studies on priming effects. The effect of priming, be it visual or auditory, was further investigated by Holcomb and Neville (1990). Focusing on the N400 component, they found that priming effects were stronger in an auditory rather than a visual modality (Holcomb and Neville, 1990). As the component is measured in milliseconds even slight incongruities with stimuli presentation could result in skewed data, thus this finding is imperative to keep in mind during methodological design.

Previous life experience can also have a significant effect on ERP's. When musicians and non-musicians are exposed to musical incongruities (harmonic, melodic, and rhythmic) it has been demonstrated that musicians can be expected to demonstrate greater speed and accuracy than non-musicians at identifying incongruities (Besson, Faita, and Requin, 1994; Gaser & Schlaug, 2003).

If N400 is elicited by linguistic meaning, and people can draw meaning from music, the question arises of whether musical stimuli might elicit N400 as well. If indeed the N400 ERP component can be elicited by musical stimuli as well as in a strictly linguistic paradigm, this could have repercussions on how 'meaning' must be defined in future research (Steinbeis & Koelsch, 2008). In fact, musical meaning and linguistic meaning as interpreted through the ERP component N400 have been found to have slight but significant processing differences (Steinbeis & Koelsch, 2008), similar to the results of neuroimaging studies utilising technology such as fMRI and MEG. Exposure to both linguistic and musical stimuli elicited an N400 effect from participants, supporting the hypothesis that musical meaning and linguistic meaning are processed in a similar fashion (Steinbeis & Koelsch, 2008). Furthermore, the locality of this processing has been narrowed down to specific regions. By utilising an fMRI machine Steinbeis & Koelsch (2008) were able to take the vicinity measurement from their EEG study on the N400 ERP component and demonstrate that the N400 effect elicited by the linguistic stimuli was localised to the right middle temporal gyrus, whereas the N400 effect elicited by the musical stimuli (a chord) was localised to the right posterior superior temporal gyrus. As both of these areas are within a similar temporal region, it is clear that there is both a close neural connection between the processing of music and language, and that the similarity requires further studies for elucidation of localisation in order to provide a processing timeline. Furthermore, this aligns with the studies discussed in the previous chapter which aligned linguistic processing with the areas in the left hemisphere (such as Broca's area) and music, pitch, and tone with a combination of identical left and nearly homologous right hemisphere areas.

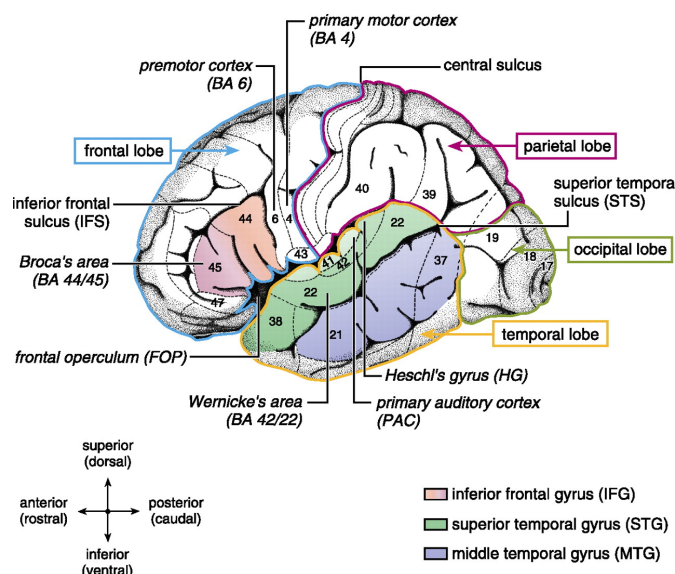


Figure 4: Brain Regions (Freiderici, 2011)

Exposure to music has also been proven to have the ability to effect language. Exposure to a musical prime has been proven to have a significant effect on an N400 ERP component elicited by exposure to a target word (Koelsch et al., 2004). N400 has also been clearly proven as beneficial to determining the difference between meaning vs. non-meaning more than linguistic vs. non-linguistic stimuli (Kutas & Federmeier, 2011). With the N400 ERP component explored the remainder of this chapter will be dedicated to literature on the ERP components N100, P200, P300, and P600.

The N100 and P200 ERP components are integral in signalling the start of perceptual processing. The components may be elicited by “clicks, speech, and abrupt changes in a continuous sound” (Hampton & Weber-Fox, 2008: 255). Furthermore, N100 has been demonstrated to reflect the processing of input sound frequency, and P200 has been suggested to reflect the processing of emotion in speech input, with emotion inferred from specific patterns in the acoustic signal (Paulmann & Kotz, 2008). The N100 and P200 ERP components may even be elicited by stimuli categorised by emotion. In a study examining vocalisations it was found that even non-verbal emotional vocalisation elicited a significant N100 and P200 depending on the emotions categorised as happy, angry, or neutral (Liu et al., 2012). These findings provide an interesting link of support for the findings of Steinbeis & Koelsch (2008), which linked the localisation of the N400 ERP component elicited by linguistic stimuli to that elicited by non-linguistic musical stimuli, similar to the verbal versus non-verbal setup of Liu et al (2012).

The P300 ERP has been argued to be tied to the “probability, quality, and duration” of a stimuli (Hampton & Weber-Fox, 2008: 256). P300 has also been suggested to be an effect of categorisation and the relatedness of words (Polich & Kok, 1995; Johnson, 1993; Hampton & Weber-Fox, 2008; Bornkessel-Schlesewsky et al., 2011), and it has been questioned whether P300 might reflect a binary system of categorisation, reflecting decisions between easily categorised accurate input and difficult to categorise inaccurate input (Bornkessel-Schlesewsky et al., 2011).

The P600 ERP component has been tied to syntactic processing (Ainsworth-Darnell et al., 1998; Patel, 2003; Hagoort, 2003; Bornkessel-Schlesewsky and Schlewsky, 2008). It can be elicited by “subject– verb agreement violations, verb inflection violations, case inflection violations, wrong pronoun inflections, and phrase structure violations” (van Herten, Kolk, and Chwilla, 2005). It has been further hypothesised that the P600 effect is not due simply to the brain re-analysing input, but that P600 is elicited by the cognitive process of general integration and reflects difficulty of syntactic processing (Kaan and Swaab, 2003). By demonstrating that P600 is elicited in sentences which are difficult but not grammatically incorrect the authors support the association of P600 with syntactic processing (Kaan and Swaab, 2003), making P600 an excellent component for analysis to answer the questions posed by this thesis.

Recently a number of studies have also noted a P600 effect that appears to have a correlation with thematic plausibility (Bornkessel, Schlewsky, and Friederici, 2003; Bornkessel-Schlesewsky and Schlewsky, 2008). This issue was further explored by van Herten, Kolk, and Chwilla (2005) who utilised Dutch sentences with semantically reversed conditions of acceptability and number and sentences providing a condition of controlled syntax (manipulated for complexity and acceptability) as well as a semantic control condition (consisting of subject relative and object relative) to explore the issue. Finding P600 in the absence of N400, the researchers found the results to indicate that their syntactic prediction hypothesis should be rejected, and instead they concluded that their results supported a definition of P600 as reflecting difficulty of syntactic integration. This not only supports the argument that P600 reflects some syntactic aspect of heuristic processing, but it also highlights the need for further research on P600, as attempted in this thesis.

### **1.5.3 EEG studies specifically relevant to this study**

Lexical tone has long been established as an imperative determiner of meaning in Mandarin Chinese. These lexical tones provide important semantic and syntactic information about a word (Ho & Bryant, 1997; Brown 2004). As language processing has been established as usually being processed mainly in the left hemisphere (Van Lancker & Fromkin; Kimura 1973, Rasumssen & Milner, 1977), and the processing of musical input has been proven to dominate the right hemisphere (Brown, 2004; Breier et al., 1999; Gootjes et al., 1999; Packard, 1986), the hemispheric processing of Mandarin as a tonal language is less certain due. This proposed left/right, language/music separation poses an interesting problem when languages which have linguistically meaningful tonality are taken into consideration. Behavioral research on Thai has indicated that exposure to linguistically meaningful tone results in left hemisphere dominant processing of the aforementioned tone with moderate right hemisphere processing as well, rather than it being treated simply as linguistic or musical input (Van Lancker & Fromkin, 1973). Exposing native Thai speaking and American native English speaking participants to three categories of tone (Thai words with normal tonality, Thai words without a defined tone, and hums that represented tone alone), Van Lancker & Fromkin (1973) found that their Thai subjects demonstrated a significant preference for the right ear - thus demonstrating dominance of left hemisphere processing - for the tonally normal words as well as the second category of Thai words with no discernible tonality, and they found no significant preference in hemispheric processing for the hums. The native English speaking participants were found to demonstrate no significant hemispheric processing dominance for the Thai words with normal tonality, a right ear (and thus left hemisphere) preference for the Thai words with no definitive tonality, and - while not reaching a point of statistical significance - they did demonstrate a trend towards a favouring of the left ear (and thus right hemisphere) for the tone only stimuli of humming (Van Lancker & Fromkin, 1973). This supports a framework of tonality being processed mainly in the left hemisphere with language, but also highlights the necessity of further research on the processing of words with normal tonality and tone only stimuli, as targeted by this thesis. the successful use of humming as non-linguistic stimuli also informed the methodological design of the study for this thesis.

A variety of ERP studies have delved into the task of providing more specific data on the processing of tonality beyond the foundational findings of Van Lancker & Fromkin (1973), investigating tone processing in Mandarin (Brown-Schmidt and Gonzalez, 2004), the interaction of tone, context, and intonation in Cantonese (Kung, Chwilla, Gussenhoven, Bögers, and Schriefers, 2010), tone and vowel characteristics in Mandarin (Li, Wang, and Yang, 2014), differences between nouns and verbs in Mandarin (Liu, Hua, and Weekes, 2007), and lexical tone categories in Mandarin (Shen and Froud, 2015).

Brown-Schmidt and Conesco-Gonzalez (2004) investigated lexical tone and segmental information in Mandarin Chinese, creating an experimental design with three conditions: correct tone and correct morpheme, wrong tone correct morpheme, wrong syllable correct tone, and wrong syllable and wrong tone. All stimuli words were placed in a sentence final position, and the correctness of each stimuli was determined with a separate experiment testing the stimuli for cloze probability, with high probability words positioned in the main experiment as correct and the low probability words positioned as the opposite. A strong N400 effect was found for each condition, and the researchers determined that the lexical tone appeared to be processed indistinguishably from linguistic stimuli as the condition with an incorrect syllable elicited the most significant negative, while the doubly wrong condition (incorrect tone and incorrect syllable) resulted in a surprising delayed N400 effect (Brown-Schmidt and Conesco-Gonzalez, 2004). Ultimately, the data from this study was found to indicate that auditorially presented stimuli that elicit N400 effects which begin sooner and last for more time than stimuli presented in a visual modality. This supports the use of

auditory stimuli presentation by future researchers interested in online processing, and highlights the importance of modality consideration during research design, as the modality (auditory versus visual) can have a significant effect on the onset time of the N400 ERP component. But what effect do context and intonation have on N400?

Tone position within a sentence has been tested in an EEG experiment focused on Cantonese by Ma et al., (2006), who upon examining Cantonese Chinese found that regardless of their written or intended form, sentence final lexical tones in a question condition demonstrated a rising F0 contour. Despite the implications for an interesting trend in tonal languages found by Ma et al., (2006), there are no completely analogous experiments in Mandarin Chinese with which to compare to their study. As Cantonese has six canonical tones as opposed to Mandarin's four tones (Ma et al., 2006; Lee et al., 2015), it would be interesting to see in future research whether the effect on F0 due to context extends to Mandarin tone as well as Cantonese tone, considering the broader categories of Mandarin tones when compared to Cantonese tones. Corpus analysis has shown that Mandarin tonality is more resistant to the type of raised fundamental frequency in question-sentence final words than Cantonese (Ho, 1997), however the ERP results from Ma et al., (2006) have not been comparatively tested in Mandarin so it is not clear whether the same frequency raising would occur in Mandarin or not. As this thesis utilises the tone 5 question word 'ma' in isolation the effects on F0 found by Ma et al. (2006) should not appear at a significant level, however their existence bears further consideration. One study which comes close to expanding on this line of enquiry is that of Kung et al., (2010), although this study as well is focused on Cantonese Chinese, not Mandarin Chinese.

With an experiment utilising a tone identification task and ERP to examine the interplay of context, tone, and intonation, Kung et al., (2010) found that increased contextual restrictions led to fewer participant errors in tone judgement, a dissipation of the P600 effect, and they found that the strongest N400 for low tones was located sentence finally in a question-sentence. This indicates that their participants had more difficulty analysing or reanalysing the semantic content of words with low tonality located word finally, and supports the argument that lexical tone is processed for semantic content. Narrowing the focus to EEG studies focused specifically on Mandarin, the perception of lexical categories has been further examined by Shen and Froud (2015).

Examining native Mandarin speakers and native English speakers with varying levels of exposure to Mandarin, Shen and Froud (2015) found that native Mandarin speakers perceived the tonality as within lexical categories (as a stronger mismatch-negativity was elicited by category defying stimuli than by stimuli which followed the normal tonal categories) significantly more than either native English speakers with no Mandarin language skills or native English speakers who acquired skills in Mandarin as adults. This is in alignment with the third hypothesis of this thesis that native Mandarin speakers will perform with more accuracy overall as well as in a behavioral test of pitch, and motivates the running of a behavioral test in the study to further examine the hypothesis that native Mandarin speakers will demonstrate greater accuracy in pitch differentiation tasks.

Furthermore, this study suggests that moderate training in Mandarin may not lead to strong enough phonetic re-categorisation in participants to have a significant impact on a study measuring their tonal perception and differentiation abilities. This result is at odds with previous research on training which did find effects (Wang et al., 1999; Kaan et al., 2007). One possible explanation for this unusual finding might be that the effect measured in the aforementioned literature was performed specifically for the study and within a short time-span of the diagnostic test; as the language experience of the participants in the study by Shen & Froud (2015) was not conducted as part of the experiment the effects seen in Wang et al. (1999) and Kaan et al. (2007) might have faded by the time of the test. With relevant EEG studies on Mandarin and pitch fully examined, the



differences between perception of nouns and verbs as investigated by Liu, Hua, and Weekes (2007) will be discussed next.

In a study by Liu, Hua, and Weekes (2007), the interaction of grammatical category and semantic category was examined by creating an EEG study with categorical priming based on semantic relation, grammatical category, and semantic category. With verbs eliciting the largest N400 effect of all the conditions, this research suggests significant processing differences for nouns and verbs due to an un-indicated aspect of their semantic relevance within Mandarin Chinese. This study highlights the necessity of further research on the topic, raising the issue of word type as an important feature which must be considered during the selection of stimuli.

The processing of vowels and tones has also been explored. Utilising patterns from Classical Chinese poetry, Li, Wang, and Yang (2014) examined the EEG data of twenty-eight native Mandarin speakers and found that their condition with an incorrect vowel elicited a stronger N400, whereas a condition with an incorrect tone elicited a P600 effect. This suggests a timeline of online processing where information from vowels is processed first around 400 milliseconds after stimuli onset, then tonal information is over-layed around 600 milliseconds, and the information of both inputs is integrated around 600 milliseconds post stimuli onset. The data from this study indicates dominant processing of incorrect tonality as similar to that of a syntactically incorrect linguistic stimuli, in support of the primary hypothesis of this thesis.

#### **1.5.4. Expected Outcomes**

With the strengths and weaknesses of various neuroimaging technologies having been laid out, the remaining chapters of this thesis detail how an EEG study was formulated and implemented to address the various problems raised in the previous chapter with the main research goal of uncovering differences in the cognitive processing of musical versus linguistic tonality. The study addresses the problems of whether specific Mandarin tones and musical tones are processed as primarily syntactic or semantic input, whether differences exist in the processing of linguistic input (from a tonal language) versus musical input, and whether there is a significant difference of pitch differentiation ability among native speakers of tonal and non-tonal languages, here Mandarin and Dutch.

In this study three main hypotheses were formed, and these can be seen summarised in Figure 5. First, it was hypothesised that native Mandarin speaking participants would demonstrate a stronger N400 effect when processing tone 4 and a stronger P600 effect when processing tone 5. Second, it was hypothesised that exposure to linguistic and musical stimuli as presented in this testing paradigm would result in ERP data which had similar but significantly distinguishable features, especially in relation to hemispheric activation where previous research has indicated that musical input should mainly show right hemispheric activation and linguistic input from a tonal language should result in both right and left hemispheric activation. Third, it was hypothesised that native Mandarin speaking participants would be significantly more accurate at pitch differentiation than Dutch participants. In general Dutch participants were expected to show either weaker or inconsequential effects.

H1	Stronger N400 effect elicited by tone 4, a stronger P600 effect by tone 5
H2	Significant difference in processing of linguistic and musical stimuli
H3	Native Mandarin speakers significantly better at pitch differentiation

Figure 5: Summary of hypotheses

The expected outcomes of the study are summarized according to effect and L1 in Figure 6. In accordance with the hypotheses stated above, speakers of Mandarin Chinese were expected to demonstrate a stronger N400 for tone 4 stimuli than Dutch participants, indicating that they were processing that input as primarily linguistic semantic data. The Mandarin speaking participants were also expected to demonstrate a stronger P600 for tone 5 stimuli than the Dutch participants, indicating that they were processing that input as data primarily relevant to syntax. In this study the Dutch participants mainly acted as a control group providing comparison for the Mandarin speakers. Their other function lay in the comparison possible by examining the third effect of interest: musical tone. For musical tone a significant difference was expected between the Mandarin speakers and Dutch speakers, with Mandarin speakers outperforming the Dutch speakers in accuracy of differentiation. This data may also be utilised in the future as comparison to native English speakers, as they could be expected to perform worse than the Dutch speakers.

<b>Effect</b>	<b>L1 Mandarin</b>	<b>L1 Dutch</b>
<b>N400</b>	Stronger for 4th tone, semantic	Similar pattern but weaker
<b>P600</b>	Stronger for 5th tone, syntactic	Similar pattern but weaker
<b>Musical Tone</b>	Significant difference between Mandarin tone & Musical tone, stronger than Dutch participants	Significant difference between Mandarin tone & Musical Tone, but not as strong as L1 Mandarin participants

Figure 6: Expected Outcomes

## Chapter 2. Methodology

### 2.1 Participants

The participants in this study consisted of fourteen L1 Mandarin Chinese speakers with high proficiency in L2 English, and fifteen L1 Dutch speakers with high proficiency in English and with no prior knowledge of Mandarin (Mean age = 25.57, SD = 3.11). Originally a 50/50 split of Mandarin/ Dutch speakers was planned, however three of the original participant pool had to be cut due to equipment malfunction, resulting in a participant pool consisting of Chinese (n=9) and Dutch (n=5). All participants were pre-screened to eliminate left-handed participants, and none reported any physical or mental disabilities during the pre-screening. Before testing all participants were

presented with the information sheet “EEG Information Document” from the CLS Lab website (Appendix 1).

## 2.2 Materials

As discussed in the previous chapter, pitch is an important feature in recognition of both linguistic material and music (Krishnan, Gandour, and Bidelman, 2010), and can even be used to recognise emotions (Dellaert, Polzin and Waibel, 1996). Thus the combination of pitch with musical and linguistic stimuli is a logical avenue of investigation when exploring the question of similarities in the cognitive processing of music and language, prompting the selection of a stimuli in this study based on pitch and conforming to a word / non-word musical note setup. To keep the musical and linguistic stimuli as similar as possible, the pitch contour of each Mandarin word, once extracted, was modified into a non-linguistic computer generated musical note.

To select which words would be used a list was drafted of Mandarin words conveying semantic and syntactic information, e.g. content words and particles (Appendix 2). The aim was to find ten of each category in fourth tone and fifth tone. The words were selected from SubtLex\_Chinese (Cai & Brysbaert, 2010), so as to account for the frequency of each word or particle’s occurrence. Fifth tone words were found to represent syntactic particles (15), bisyllabic semantic words (12), and monosyllabic content words (10). Fourth tone words were found to represent the syntactic fifth tones (15), and bisyllabic fourth tone final words (16). Upon further research all bisyllabic words were eliminated, due to the concern that preceding vocalisations might have too strong an effect on the pronunciation of the target tonal syllable. The list was then presented to a female native Mandarin speaker for input based on native speaker intuition about how “natural” each of the words sounded, in terms of likelihood of occurrence. Based on the input of the native speaker, the list was summarily adjusted (Appendix 3) and word frequencies were noted in the finalised list in terms of occurrence per million words. The fifth tone lists also had to be cut due to elimination of sandhi fifth tones. A tone was considered sandhi (and summarily eliminated) when it was only a fifth tone in combination with a preceding tone. In this way, all tones in this study exist in their final way in their own right. The final list consisted of eight fifth tone syntactic particles, nine fifth tone semantic words, and 18 semantic fourth tone words (fig. 7). Despite the slight imbalance of stimuli the list was kept at these numbers to be sure there would be enough stimuli in case of technical malfunction at any time. Having previously been proven to be effective in elicitation of ERP components, single words were selected for presentation (Drijvers, Mulder, & Ernestus, 2016).

Tone	Pitch
Tone 5 (8 syntactic, 9 semantic)	Pitch 5 (EPC tone 5)
Tone 4 (18 semantic)	Pitch 4 (EPC tone 4)

*Figure 7: Relevant characteristics of each stimuli category*

Once the words lists were finalised, the same Mandarin speaker who provided the judgement of “naturalness” recorded the stimuli. She produced three versions of each, with the highest quality version selected afterwards by the researcher. The audio was recorded in a soundproof lab at Radboud University with PRAAT. The files were then trimmed in PRAAT with the beginning and end of each file set at zero-crossing. After consultation with supervisors it was determined that five samples of each stimuli was necessary, and these were summarily recorded and edited in the same manner.

To create the musical stimuli the route of extracted pitch contours was selected. To create these sounds each Mandarin sound file was edited individually. First a word was selected to 'analyse to periodicity - pitch' (PRAAT). Since the speaker was female, the pitch range was set to a floor of 100 and ceiling of 500 Hz (Boersma & Weenink, 2013). Then under 'Sound' the 'To Sound - sine' option was selected, resulting in an electronic note classified in this paper as music, as explained in the introductory section. The music files were all then hand cut to zero crossing at each end.

Thus, the finalised stimuli list consisted of 180 Mandarin words (36 stimuli x 5 versions of each), and 180 musical sounds (extracted from the words), for a grand total of 360 sounds, each running approximately .8 seconds in length (before editing to assimilate length). During the editing process the stimuli audio files were each given unique file names indicating their word, tone, iteration, and position in the list over-all. This in turn created a master list, which was summarily run through a script in Presentation (Neurobehavioural Systems, [www.neurobs.com](http://www.neurobs.com)) to create an individual list for each participant. The stimuli were divided into five blocks, and the script ensured that each participant received a randomised list, without any stimuli playing three or more times in a row. A setup with no active task after each stimuli but rather a semi-randomised re-occurring judgement task was selected to elicit the clearest EEG signal and inspired by the work of Kutas & Federmeier (2011). Due to the intense focus and stillness required of participants in this type of study, the length of each break provided the opportunity for the participants to blink as much as necessary and stretch to relieve and anxious muscles.

### **2.3 Procedure**

Participants were seated in a sound proof booth at the Radboud University CLS Lab in room 12.21. After being welcomed into the room by the researcher, participants were seated at a setup table and presented with the consent form to sign (Appendix 4). Next, the researcher assuaged whether they had any experience with EEG, and provided a brief verbal run through of the upcoming procedure, namely a head measuring for the cap fitting, the researcher filling the cap nodes with gel, and plugging in electrodes. Once these steps were complete the participant was escorted to the soundproof room and seated while the impedance was tested. Once a clear signal was established, the researcher would show the participant their real-time brain waves, explain that they couldn't read their mind only electrical signal activity (if asked), and invite participants 'play a fun game' where they were to hold very still and then blink as hard as they could. This nearly always elicited a strong reaction from the participants (they were generally very entertained by the result of this action) and the researcher would then ask if during the experiment participants could try to just relax and blink only when they see the little cross on the screen, as movement (as they could see) created very wild data.

Next, Sennheiser HD215 headphones (fitted with hygienic covers) were fitted over the EEG cap and set, and the impedance tested one last time to make sure none of the electrodes were disrupted by the large headphones. Then, participants were handed a button box with color coded buttons and instructed to hold it in their lap and use it to register their judgements through the experiment. It was explained that after some very basic preliminary questions (handed-ness, sex, and age) they would experience instructions for the experiment and then a practice session. They were also warned that the session could only play one time, so they could take as long as needed with it. They were also informed that at the end of the practice session the researcher would return and they could ask any questions they had before starting the real experiment, wherein they would have a break every two to three minutes. The researcher then started the EEG recording, sealed the participant in the booth, and monitored from the externally linked computer monitor.

Once the practice session finished the participant was sealed in the soundproof booth and the stimuli audio began. The stimuli were divided into five block blocks of fifty trials each. Each trial started with a fixation cross of 1500 ms, followed by onset of the stimuli audio, each of which was set to 3000 ms. Within each block a question mark would appear at ten randomised intervals, after which the participant was instructed to press a button indicating which category the sound they just heard most resembled. After stimuli onset the participant had five seconds press a button indicating their judgement of the sound, and the press of any button began the next trial. At each break the researcher checked how the participant was doing, if they had any questions or needed a break, a tissue, or a sip of water.

Upon completion of the EEG test participants completed two behavioral test of pitch perception. This was based on the work of Asaridou, Hagoort, and McQueen (2015) examining the effects on tonal perception of Cantonese / Dutch bilinguals. The format of the first test was modified by replacing tone examples with Mandarin tones, and the second test remained unmodified.

## **2.4 EEG Recording**

The EEG signal was recorded with 34 active electrodes set in an elastic cap (Acticap). Electrode positions followed the international 10-20 system of 4 midline electrodes (Fz, Cz, Pz, and Oz) and 22 lateral electrodes (Fp1/Fp2, F3/F4, F7/F8, FC1/FC2, FC5/FC6, C3/C4, T7/T8, CP1/CP2, CP5/CP6, P3/P4, and P7/P8). An electrode was placed on each mastoid, and the electrooculogram (EOG) was recorded with an electrode below the right eye. Electrode impedance was held at under 5 k $\Omega$ . The EEG and EOG signals were amplified (band pass = 0.02-100 Hz) and digitised online (sampling frequency 500 Hz). The signal was also re-referenced to the average signal of the left and right mastoids and digitally filtered with a high cut-off filter (30 Hz) before the data was analysed. Next, the continuous EEG was split into stimulus-time-locked epochs, from 200 ms before the target word onset to 1000 ms after onset, and ocular artefacts were identified and removed. The output variables analysed were the N400 amplitude, P600 amplitude, accuracy scores of responses to the questions asked during the EEG test.

## **Chapter 3. Results**

In this section the grand average plots of the ERP of the Chinese and Dutch groups are presented, with descriptive attention given to specific electrodes. Each group was analysed separately to account for the difference in group sizes (Chinese: N=9, Dutch: N=5). The components analysed are P200 (mean amplitude 150-250 ms), N400 (mean amplitude 300-500 ms), and P600 (mean amplitude 600-800 ms). While this study was originally focused only on N400 and P600, the ERP data revealed some interesting activity at P200, so further analysis of this activity was conducted and description of the activity was provided as well. The possibility that this activity at P200 is due to surprisal or form-level processing will be discussed in Chapter Four.

### **3.1 ERP Data**

There was no significant interaction between the sites of the electrodes used in the analyses and the stimulus type in the N400 data [Chinese Midline:  $F = 2.744$ ,  $p < .067$ ,  $MSE = 11.051$ ,  $p < .067$ ; Dutch Midline:  $F = 2.104$ ,  $p < .0269$ ,  $MSE = 185.534$ ,  $p < .269$ ]. Potential emergent trends indicated by the results will be discussed in Chapter Four.

The ERP data on the P600 effect did not reach significance [Chinese Midline:  $F(1.198, 7.191) = 12.599$ ,  $MSE = -1.845$ ,  $p < .007$ ; Dutch Midline:  $F(1.358, 2.717) = 4.345$ ,  $MSE = 1.062$ ,  $p < .141$ ]. The emergent trends indicated by the results will be discussed in Chapter Four. The P600 data

revealed no significant interaction between electrode and stimulus type [Chinese Midline:  $F=2.541$ ,  $p<.075$ ,  $MSE = 46.566$ ; Dutch Midline:  $F= .632$ ,  $p<.555$ ,  $MSE = 49.261$ ].

The interaction between the sites of the electrodes used in the analyses and the stimulus type in the P200 data did not reach significance in the Dutch group, but did in the Chinese group [Chinese Midline:  $[F(4.035, 24.207) = 4.358$ ,  $MSE = 32.989$ ,  $p<.008$ ; Dutch Midline:  $[F(1.188, 2.376) = 4.621$ ,  $MSE = 19.077$ ,  $p<.147$ ].

No significant interaction was found between hemisphere, site, and condition except in the Dutch group for N400 [Chinese: N400:  $F=2.522$ ,  $p<.074$ ; P200:  $F= 4.401$ ,  $p<.02$ ; P600:  $F=2.162$ ,  $p<.114$ ; Dutch: N400:  $F=8.753$ ,  $p<.038$ ; P200:  $F= 4.347$ ,  $p<.157$ ; P600:  $F=.551$ ,  $p<.567$ ].

While statistical significance was not reached in most of the ERP data, it did indicate trends which may reach statistical significance when the study is conducted in a larger population. The trends can be seen in the graphs of the ERP data which will be presented next. Note that in the ERP result figures in this paper negativity is plotted upward, and the results in the figures are color-coded as follows: tone 5 (red), tone 4 (black), pitch 5 (green), pitch 4 (blue). The grand average results in Figure 8 (tone data) and Figure 9 (pitch data) alone coded differently, with red representing tone 5 and pitch 5, respectively.

A comparison of the grand averages of all native Mandarin speaking participant EEG results presented visual differences between tone 4 and tone 5 tone activation at N400 and P600, not in alignment with the prediction of the first hypothesis that the tone 4 stimuli would elicit a stronger N400 effect and tone 5 a stronger P600 effect, but rather the opposite.

Comparison of tone 4 and tone 5 revealed that within the Mandarin speaking participants the fifth tone elicited a stronger N400, and the fourth tone a stronger P600. This is not in alignment with the prediction of the first hypothesis. For the pitch stimuli the Mandarin participants had a stronger N400 for pitch 4 and a stronger P600 for pitch 5. These findings indicate an emergent trend of tone and pitch input being processed in a similar yet distinct manner.

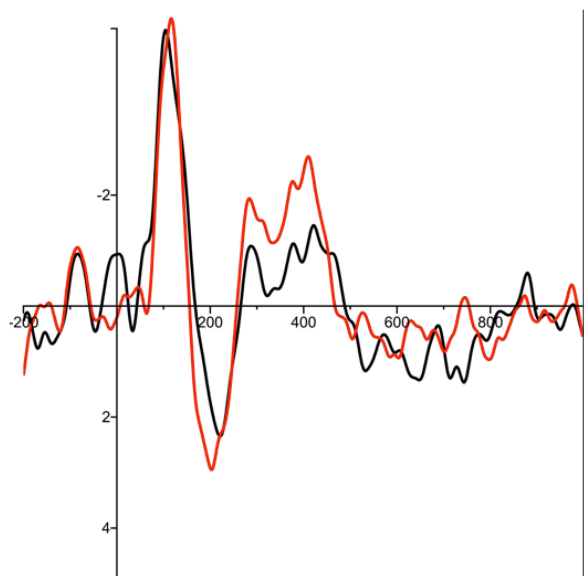


Figure 8: Grand Average ERP Data for Tone

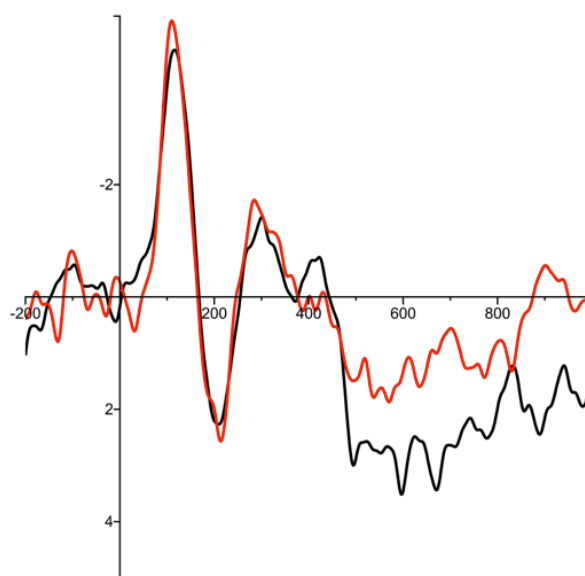


Figure 9: Grand Average ERP Data for Pitch

An examination of C2 found that tone 4 and pitch 4 were practically indistinguishable at N400, tone 5 was the strongest at N400, and pitch 4 was the strongest P600 by far with the others all still clustered in the negative. The failure here of tone 4 to elicit the strongest reaction is in opposition to the first hypothesis. A strong reaction at P200 is also notable at C2, with tone 5 showing the strongest reaction by far, followed in decreasing order by tone 4, pitch 5, then pitch 4.

No hypothesis was stated in this study about P200 effects, however this result does seem to indicate that tone 5 elicited the strongest positive reaction at the start of processing at C2.

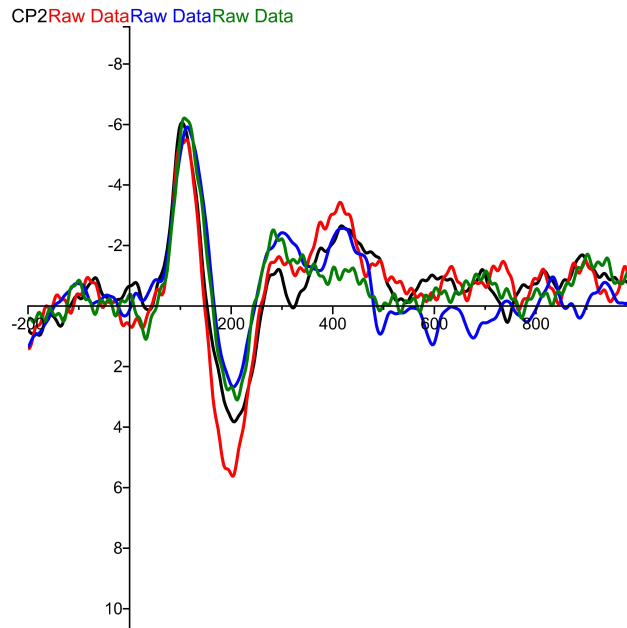


Figure 10: Mandarin ERP data at C2

At Pz Tone 5 had the strongest N400, followed by tone 4, then pitch four. Pitch five however dips into the positive. That tone 5 elicited the strongest N400 here is in opposition to the first hypothesis. The strongest elicitor by far of a P600 effect at Pz is tone 4, also in opposition to the first hypothesis which anticipated that tone 5 should elicit the strongest P600 and tone 4 the strongest N400. Interestingly, while the rest of the stimuli results are still clustered in a similar pattern, the signal strengths are (from strongest to weakest) pitch 5, tone 5, then tone 4, unlike at the C2 electrode. The Pz electrode also shows a nearly identical P200 response as found at C2, indicating that this may be a strong and common effect.

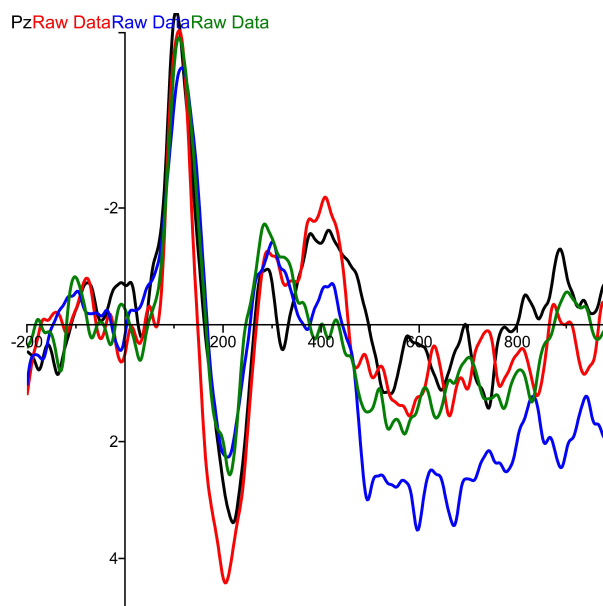


Figure 11: Mandarin ERP data at Pz



The ERP data for the Dutch participants is very visibly different from that of the native Mandarin speakers. Interestingly, a similar negative to positive dip around P200 is evident in the Dutch participant data as the data from the native Mandarin speakers, however it is much smaller. This is in accordance with the first hypothesis.

At Pz there is no significant N400, rather all results remain in the positive, with pitch 4 and tone 5 eliciting the strongest effects there, and tone 4 and pitch 5 indistinguishable. This would seem to be in partial accordance with first hypothesis: the Dutch is indeed similar but weaker at the start (through P200), however has become completely different by the 400 ms mark. At P600 the signals are similarly clustered, with pitch 4 just barely registering as the strongest, followed by pitch five, tone 5, then tone 4. This would seem to be in opposition to the first hypothesis, however as it does appear in accordance with the second hypothesis that a significant difference would be evident between the native Mandarin and native Dutch speakers in terms of the musical stimuli (pitch 4 and pitch 5).

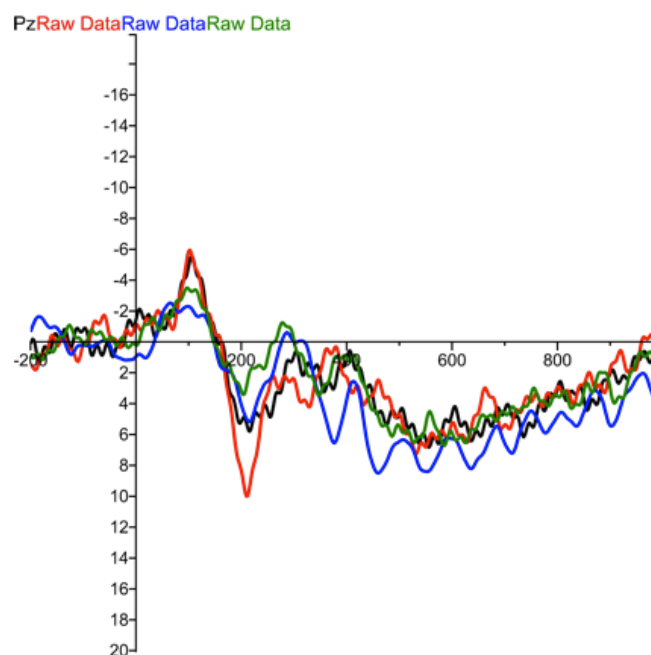


Figure 12: Dutch ERP data at Pz

At C4 a clearer N400 does appear, with pitch 5 registering the strongest, then pitch 4 and tone 4 (indistinguishable), then tone five. There is, however, no P600, as all the results are clustered in the negative.



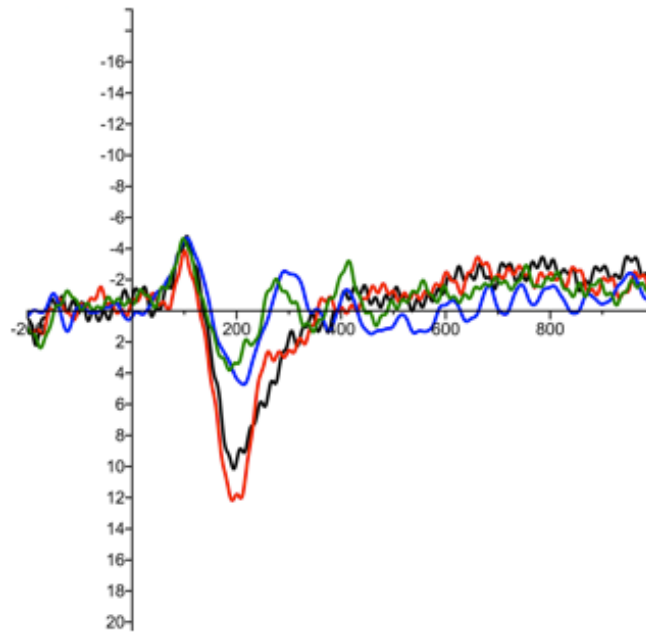


Figure 13: Dutch ERP data at C4

Finally, at Fz the pitch 5 data is shown as reaching the strongest amplitude, followed by tone 4, then pitch 4, then tone 5 at 400ms, in opposition to the first hypothesis which would predict that tone 4 should elicit the strongest N400 effect. It is interesting to note however that between the tones tone 4 does elicit a stronger N400 effect here than tone 5, which would be in accordance with the first hypothesis. Again there is no P600, however the signal reaches an even stronger level of negativity here than at C4.

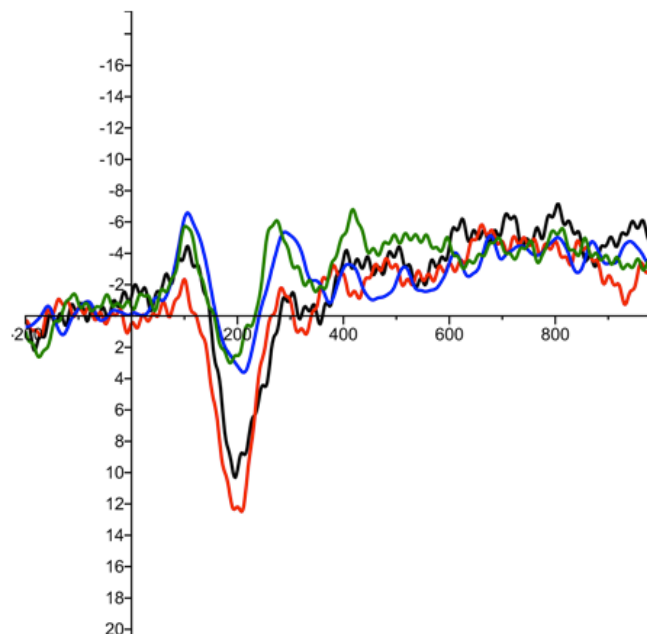


Figure 14: Dutch ERP data at Fz

### 3.2 Behavioral Data

#### 3.2.1 Speech vs. Non-Speech Identification test data

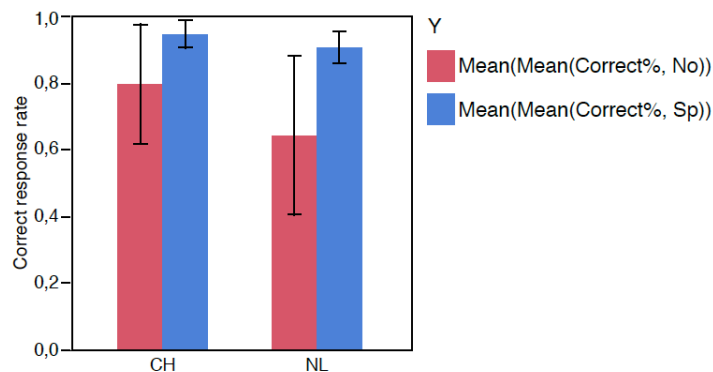


Figure 15: Mean correct response rates in Speech / Non-speech ID test

In the Speech/Music test Chinese participants (N=9) were shown to have a mean accuracy of 94.27% for Speech items and 79.44% for Non-Speech items. The Dutch participants (N=5) on the other hand demonstrated means of 90.29% accuracy with Speech items and only a 64.04% accuracy with Non-Speech items. This is represented in the graph above (fig. 15), with red representing the speech data and blue the non-speech data in the graph. These results are in accordance with the third hypothesis that the native Mandarin speakers would perform more accurately in tests of pitch differentiation ability.

#### 3.2.2 Speech vs. Non-speech correct response rate

In both groups the identification of Speech material was higher than for Non-Speech material, however a Wilcoxon Signed Rank test indicated that this was significant only for the Chinese group ( $p < .0117$ ) and not for the Dutch group ( $p < .125$ ). This is in accordance with the third hypothesis. A larger Dutch participant pool may prove the Dutch difference to be statistically significant as well.

#### 3.2.3 Dutch vs. Mandarin speakers

Neither the native Mandarin speakers nor the native Dutch speakers differed at a statistically significant level in their response rates to the Speech and Non-Speech tasks. In the Speech condition a Chi-square test indicated no significant difference between the groups ( $df=1$ , ChiSquare = 1.8015,  $p < .1795$ ; 2-Sample Test:  $Z = -1.2751$ ,  $p > .2023$ ). The Wilcoxon / Kruskal-Wallis test of rank sums resulted in a mean score of 8.61111 for Chinese participants and 5.5 for Dutch participants. In the Non-Speech condition a Chi-square test indicated no significant difference between the groups ( $df=1$ , ChiSquare = 1.9643,  $p < .1161$ ; 2-Sample Test:  $Z = -1.3348$ ,  $p > .1819$ ). The Wilcoxon / Kruskal-Wallis test of rank sums resulted in a mean score of 8.66667 for Chinese participants and 5.4 for Dutch participants. These results are in opposition to the third hypothesis which predicted that native Mandarin speakers would do better than Dutch speakers at this task. However this is not surprising due to the small group size.

#### 3.2.4 Tone Identification test

In the tone identification test native Mandarin speakers scored significantly higher in accuracy than the native Dutch speakers [ $df=1$ , ChiSquare = 6.4461,  $p > .0111$ ; 2-Sample Test:  $Z = -2.47211$ ,  $p < .0134$ ]. The Wilcoxon / Kruskal-Wallis test of rank sums resulted in a mean score of

9.61111 for Chinese participants and 3.7 for Dutch participants. This is in accordance with the third hypothesis, which predicted significantly higher accuracy in the native Mandarin speaking group than the native Dutch speaking group in tasks requiring the ability to differentiate between pitches.

#### Chapter 4. Discussion

With its ambitious design this study was able to test each of the three hypotheses (see fig. 16 below) and investigate many of the issues raised in Chapter 1. In this chapter the ERP and behavioral test results will be discussed, and the relationships between and answers to the issues raised in Chapter 1 will be explored. Finally the section will close with a summary of the a discussion of the issues with this study, and recommendations for the future applications of this research.

H1	Stronger N400 effect elicited by tone 4, a stronger P600 effect by tone 5
H2	Significant difference in processing of linguistic and musical stimuli
H3	Native Mandarin speakers significantly better at pitch differentiation

Figure 16: Summary of hypotheses

#### 4.1 Discussion of ERP data

In this study a visual examination of the grand averages of the ERP component data indicates an emergent trend which fits the prediction of the second hypothesis that the musical and linguistic stimuli would elicit similar but notably different pattern of effects. The linguistic stimuli elicited both N400 and P600 ERP components and musical stimuli elicited a very weak N400 effect but a very strong P600 effect, confirming the second hypothesis. However the prediction of the first hypothesis that a stronger N400 effect would be elicited by tone 4 stimuli, and a stronger P600 by tone 5 stimuli was not as accurate, and examination of the individual electrode sites revealed a much more complicated picture. Indeed, the linguistic tone 5 stimuli elicited a stronger N400 effect than tone 4, and in turn the linguistic tone 4 stimuli appeared to elicit a stronger P600 effect than the tone 5 stimuli. When individually examined none of the results reached statistical significance for N400 or P600 among the native Mandarin speaking participants, except for a site interaction at 200ms ( $p > .008$ ), so it is difficult to say whether this should be taken as indicative of a larger trend or whether the addition of more participant data would support the syntactic/semantic divide proposed in the first hypothesis, seeing as that that separation was hypothesised based on existing literature. Furthermore, where previous research had found that harmonic chords were capable of eliciting an N400 effect (Steinbeis & Koelsch, 2008) the musical stimuli in this study only elicited a weak N400 effect, and the most notable trend in that data is a strong P600 effect. However, an examination of the grand average ERP results and the individual site results for the native Mandarin speaking participants revealed consistent exhibition of a trend where the musical stimuli was processed in a way similar to yet still notably different from linguistic stimuli, as predicted by the second hypothesis. Building on the work of previous research which used short chords to elicit target ERP components (Steinbeis & Koelsch, 2008; Daltrozzo & Schön, 2009; Hung & Lee, 2008) this study does then seem to have demonstrated the feasibility of extracted pitch contours as part of a functional testing paradigm for examining linguistic tone in comparison to musical tone.

Within the Dutch participant pool the only significant interaction occurred between hemisphere, site, and condition at 400ms for the Dutch participants ( $p > .038$ ). However over all, the visual representation of the Dutch data appears to reflect that the Dutch participants had great difficulty in categorising the Mandarin and musical stimuli, and instead of separating them into four fully distinct categories they perceived them as all relatively similar, as after what appears to be a standard ‘processing start’ pattern at 200ms (mirroring to a lesser extent the ERP components elicited from the native Mandarin speaking group) the contour of the data appears to significantly differ from that of the native Mandarin speakers (as predicted by the second hypothesis), with differences in amplitude very slight and always following a similar slow arching pattern, as opposed to the sharp peaks and valleys found in the ERP data of the native Mandarin speakers. At each location the tone 4 stimuli was found to elicit the strongest P200 effect, followed closely by tone 5, which could indicate that the Dutch participants managed to complete some successful categorisation of the stimuli at least initially, however this clustering effect is gone by 400ms as pitch 5 elicited the stronger amplitude there than tone 4, in opposition to what would have been predicted by the first hypothesis. Interestingly however, in this aspect the Dutch data matches the trend in the data from the native Mandarin speakers in that the results of the study were the exact opposite of those predicted by the first hypothesis. This could be due to differences in the fundamental frequency of the tone and pitch stimuli. Another explanation might lie in the influence of the training session, as previous research has indicated that the influence of short training sessions on the participants can be significant (Steinbeis & Koelsch, 2008). To test the effectiveness of the training session a future study could create a control group of non-Mandarin speaking Dutch participants and withhold the short training session. Due to the time constraints on this study further division of the participant pool to create such a control was not feasible, however such implementation in future research could establish whether the elicitation of a stronger N400 effect for Mandarin words versus musical stimuli as found here is due to innate categorisation abilities or processing primed by the training session. Finally, the similarity of stimuli results at P600 for Dutch participants is in alignment with their difficulties of categorisation, and the slight preference for musical stimuli at P600 would seem to indicate a trend in agreement with previous research which found P600 effects elicited by musical stimuli. Some research has suggested that the P600 effect may be strengthened by low context situations (Kung et al., 2010). If so, the support of the experimental design structure (e.g. training session length, number of trials, effectiveness of the attention maintaining task) could be even more important, as the design is by nature low context, forcing the participants to rely on their personal linguistic and musical processing knowledge in order to draw out an answer as to how they are processed differently.

As no significant interaction was found for hemisphere in this experiment, the question of whether participants would demonstrate similar activation of the perirolandic region could not be investigated. Future research could gain valuable insight into this issue by taking advantage of the EEG equipment’s locative abilities and analysing the data to determine hemispheric activation, as the right hemisphere has previously been found the most likely processing center for musical input and the left the most specialised for linguistic data, with tonal languages balancing somewhere in-between (Van Lanker & Fromkin, 1973). For now, this study indicates a similar split allocation of processing temporally, as the musical stimuli elicited a strong P600 effect and the linguistic stimuli generally elicited both N400 and P600 effects.

With notable N100 and P200 ERP components elicited by the testing paradigm this study also appears to support the interpretation of these components as the visualisation of the start of processing as developed by Paulmann & Kotz (2008). Future studies might benefit from focusing on a smaller number of specific sites from the beginning in order to test a more specific area of activation and thus make inferences about processing. For example this study could be refocused to

examine hemispheric activation; musical and linguistic stimuli have both been shown to elicit an N400 effect in the right gyrus, however the exact location -- right posterior superior for the musical stimuli, right middle temporal gyrus for linguistic stimuli -- differs (Steinbeis & Koelsch, 2008), so specificity could be beneficial.

H1	Stronger N400 effect elicited by tone 4, a stronger P600 effect by tone 5	✗
H2	Significant difference in processing of linguistic and musical stimuli	✓
H3	Native Mandarin speakers significantly better at pitch differentiation	✓

Figure 17: Summary hypotheses test outcome

To summarise, considering the number of participants, the results of this study are found to be indicative of a functional testing paradigm. According to the grand averages, Mandarin tone 5 elicited stronger N400 effects than tone 4, (in opposition to the first hypothesis -- although there were definitely distinct ERP data for tone 4 and tone 5), the linguistic and musical stimuli resulted in similar but distinguishable EEG data (confirming the second hypothesis), and the native Mandarin speakers performed more accurately than the native Dutch speakers in the behavioral categorisation and identification tasks (confirming the third hypothesis). Further research must next be conducted in order to declare with certainty the validity of these findings on a statistically significant scale.

## 4.2 Discussion of Behavioral Data

Of the two participant groups, the native speakers of Mandarin had significantly higher rates of accuracy. Post hoc testing confirmed that statistical significance was not reached in the other testing. However, as evident in Figure 18, a visual difference can be seen in the data. This would seem to be in alignment with previous research, which indicates that native speakers of a tonal language should be better at frequency differentiating tasks (Wang et al. 1999; Kiriloff, 1969). Considering the small size of the participant pool, this could be considered an emergent trend. Continuation of the study to build a larger participant pool of both native Mandarin and Dutch speakers could lead to the difference found in this study increasing to a level of statistical significance.

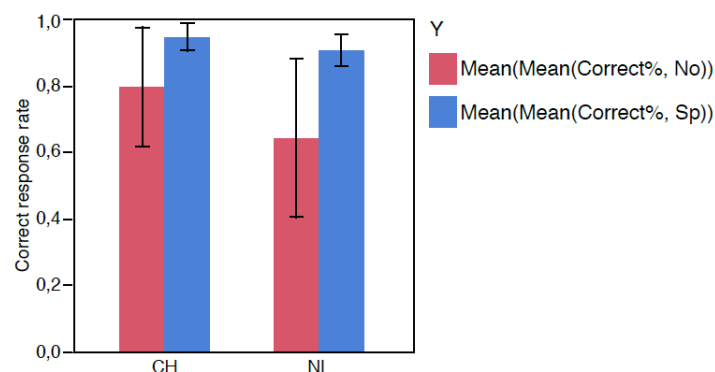


Figure 18: Mean correct response rates in Speech / Non-speech ID test

### **4.3 Issues with study**

In this study all participants were highly fluent in English (the native language of the researcher) and thus English was used as the lingua franca, with all experiment instruction given in English, not Dutch or Chinese (the target languages). While this avoids accidental priming in either of the languages under analysis (Dutch or Mandarin), it does emphasise potential confounds of bilingualism. As the field of investigation into bilingualism is complex and evolving, the bilingual state of participants in this study was not investigated. In future research it would be interesting to examine what effect the bilingualism of the participants might have had on their pitch perception abilities.

Another area that would benefit from further investigation is the training session. In this study the training session was kept quite short (only three repetitions of each sound type and an eight sound practice session) by necessity, as the goal was to create a training session just long enough to trigger the native Mandarin speakers to automatically pick up on the categories, yet with enough trials to allow the native Dutch participants to create their own mental categories for the sounds. If the Dutch participants received any more training a higher risk would have been run of simply training the Dutch speakers to create categories for the sounds beyond the speech/non-speech and semantic/syntactic cues sought within the native Mandarin speakers. Nearly all of the participants commented after the study that they wished that the training session had been longer, and just as many asked at the first post-training session break whether they could do the training again before starting the experiment again. Some participants (mainly the native Dutch speakers) also appeared concerned after a block or two about whether they were remembering the categories (or the order of the categories) correctly. When met with these concerns the researcher would attempt to reaffirm a positive attitude on their part (and avert frustration) by re-assuring them. In the future it would be interesting to examine more thoroughly the effect of the training length on the ERP data, perhaps in a shorter study with fewer blocks, as well as limit the interaction possible between researcher and participant without reducing participant comfort.

One participant reported having ADHD after the experiment had already begun. In the future the potential of this to interfere with the signal in any way should be addressed. As this was not the focus of this study, that participant's data was allowed to remain in the study after a visual examination of the EEG data revealed no obvious anomalies.

Part of the procedure for preparing participants scalps for the electrodes includes use of an alcohol based cleanser. One participant raised the question of whether there was a non-alcohol option, due to potential religious conflicts with alcoholic substances touching exposed skin. While this was not an issue for this study as no participants made a request for a non-alcoholic cleanser or preparatory gel, this experience revealed to this researcher a potential issue that should be prepared for by any lab in the future which wishes to accommodate participants of all faith backgrounds. There are rare cases of participants with skin that finds alcohol based products highly irritable as well, and in future studies it would be beneficial for researchers to be aware of this potential issue and have a non-alcoholic product available.

### **4.4 Potential applications and directions of future further research**

The findings of this study have many future applications. The findings on tonal differentiation accuracy of native Mandarin and native Dutch speakers support previous research on the improved pitch change recognition abilities of tonal language speakers. By combining many elicitation tactics and uniting them in a unique methodological design this study provides baselines for future research on the processing of Mandarin, Dutch, tonality, semantic and syntactic information, and the relation of music and language processing in the brain. Future research can be

carried out to replicate the results of this study and improve on the efficacy of the many complex aspects of its methodological design.

By daring to connect previously separated areas of study together in a unique methodological design this study has opened the door to new avenues of understanding the processing capabilities and functions of the human brain. While demonstrating an appreciation of past research this study has also contributed something new which hopefully will spark the fire of interest in future researchers, leading to a new and deeper understanding of the human brain and its unique capacities for music and language.

## References

- Ainsworth-Darnell, K., Shulman, H. G., & Boland, J. E. (1998). Dissociating brain responses to syntactic and semantic anomalies: Evidence from event-related potentials. *Journal of Memory and Language*, 38(1), 112-130.
- Asaridou, S. S., Hagoort, P., & McQueen, J. M. (2015). Effects of Early Bilingual Experience with a Tone and a Non-Tone Language on Speech-Music Integration. *PloS one*, 10(12), e0144225.
- Besson, M., Faïta, F., & Requin, J. (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neuroscience letters*, 168(1), 101-105.
- Bigliassi, M., Barreto-Silva, V., Kanthack, T. F. D., & Altimari, L. R. (2014). Music and cortical blood flow: A functional near-infrared spectroscopy (fNIRS) study. *Psychology & Neuroscience*, 7(4), 545.
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of memory and language*, 44(2), 235-249.
- Boersma, P., & Weekink, D. (2013): Praat: doing phonetics by computer [Computer program]. Version 5.3.51, retrieved 2 June 2013 from <http://www.praat.org/>
- Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., Roehm, D., & Schlewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and language*, 117(3), 133-152.
- Breier, J. I., Simos, P. G., Zouridakis, G., Wheless, J. W., Willmore, L. J., Constantinou, J. E. C., ... & Papanicolaou, A. C. (1999). Language dominance determined by magnetic source imaging A comparison with the Wada procedure. *Neurology*, 53(5), 938-938.
- Brotons, M., & Koger, S. M. (2000). The impact of music therapy on language functioning in dementia. *Journal of music therapy*, 37(3), 183-195.
- Brown, S., Martinez, M. J., & Parsons, L. M. (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *European journal of neuroscience*, 23(10), 2791-2803.
- Brown-Schmidt, S., & Canseco-Gonzalez, E. (2004). Who do you love, your mother or your horse? An event-related brain potential analysis of tone processing in Mandarin Chinese. *Journal of psycholinguistic research*, 33(2), 103-135.
- Burke, P. (2015). What is music? *Humanities*, 36(1), np.
- Chen, Y., & Xu, Y. (2006). Production of Weak Elements in Speech—Evidence from F<sub>0</sub> Patterns of Neutral Tone in Standard Chinese. *Phonetica*, 63(1), 47-75.
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS one*, 5(6), e10729.
- Daltrozzo, J., & Schön, D. (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience*, 21(10), 1882-1892.

- Dellaert, F., Polzin, T., & Waibel, A. (1996, October). Recognizing emotion in speech. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on* (Vol. 3, pp. 1970-1973). IEEE.
- Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *The Journal of the Acoustical Society of America*, 119(2), 719-722.
- Drijvers, L., Mulder, K., & Ernestus, M. (2016). Alpha and gamma band oscillations index differential processing of acoustically reduced and full forms. *Brain and language*, 153, 27-37.
- Ferreri, L., Bigand, E., Perrey, S., Muthalib, M., Bard, P., & Bugaiska, A. (2014). Less effort, better results: how does music act on prefrontal cortex in older adults during verbal encoding? An fNIRS study. *Frontiers in human neuroscience*, 8.
- Fiveash, A., & Pammer, K. (2014). Music and language: Do they draw on similar syntactic working memory resources?. *Psychology of Music*, 42(2), 190-209.
- French, H. (2005). Uniting China to speak Mandarin, the one official language: Easier said than done. *The New York Times*, 10.
- Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological reviews*, 91(4), 1357-1392.
- Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *Neuroimage*, 20, S8-S17.
- Friedrich, C. K., Alter, K., & Kotz, S. A. (2001). An electrophysiological response to different pitch contours in words. *NeuroReport*, 12(15), 3189-3191.
- Gandour, J., Wong, D., & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport*, 9(9), 2115-2119.
- Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, 23(27), 9240-9245.
- Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in psychology*, 2.
- Gootjes, L., Raij, T., Salmelin, R., & Hari, R. (1999). Left-hemisphere dominance for processing of vowels: a whole-scalp neuromagnetic study. *Neuroreport*, 10(14), 2987-2991.
- Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of cognitive neuroscience*, 15(6), 883-899.
- Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage*, 20, S18-S29.
- Hampton, A., & Weber-Fox, C. (2008). Non-linguistic auditory processing in stuttering: evidence from behavior and event-related brain potentials. *Journal of fluency disorders*, 33(4), 253-273.
- Hung, T. H., & Lee, C. Y. (2008). Processing linguistic and musical pitch by English-speaking musicians and non-musicians. In *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)* (Vol. 1).
- Ho, C. S. H., & Bryant, P. (1997). Development of phonological awareness of Chinese children in Hong Kong. *Journal of Psycholinguistic Research*, 26(1), 109-126.
- Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and cognitive processes*, 5(4), 281-312.



- Iakovides, S. A., Iliadou, V. T., Bizeli, V. T., Kaprinis, S. G., Fountoulakis, K. N., & Kaprinis, G. S. (2004). Psychophysiology and psychoacoustics of music: Perception of complex sound in normal subjects and psychiatric patients. *Annals of general hospital psychiatry*, 3(1), 6.
- Jäncke, L. (2012). The relationship between music and language. *Frontiers in psychology*, 3.
- Janus, M., Lee, Y., Moreno, S., & Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *Journal of experimental child psychology*, 144, 84-97.
- Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology*, 30(1), 90-97.
- Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39(3), 215-258.
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of cognitive neuroscience*, 15(1), 98-110.
- Kaan, E., Barkley, C. M., Bao, M., & Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *BMC neuroscience*, 9(1), 53.
- Kaan, E., Wayland, R., Bao, M., & Barkley, C. M. (2007). Effects of native language and training on lexical tone perception: An event-related potential study. *Brain research*, 1148, 113-122.
- Kimura, D. (1973). Manual activity during speaking—I. Right-handers. *Neuropsychologia*, 11(1), 45-50.
- Kimura, D. (1973). The asymmetry of the human brain. *Scientific American*, 228(3), 70-80.
- Kiriloff, C. (1969). 'On the auditory discrimination of tones in Mandarin. *Phonetica* 20, 63–67.
- Koger, S. M., Chapin, K., & Brotons, M. (1999). Is music therapy an effective intervention for dementia? A meta-analytic review of literature. *Journal of Music Therapy*, 36(1), 2-15.
- Kljajevic, V. (2010). Is syntactic working memory language specific? *Psihologija*, 43, 85–101. doi:10.2298/PSI1001085K.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nature neuroscience*, 7(3), 302.
- Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, 23(1), 81-95.
- Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequency-following response: representation of pitch contours in Chinese tones. *Hearing research*, 189(1), 1-12.
- Krumhansl, C. L. (2001). Cognitive foundations of musical pitch. Oxford University Press.
- Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period". *Language Learning and Development*, 1(3-4), 237-264.
- Kung, C., Chwilla, D. J., Gussenhoven, C., Bögels, S., & Schriefers, H. (2010). What did you say just now, bitterness or wife? An ERP study on the interaction between tone, intonation and context in Cantonese Chinese. In *Fifth International Conference on Speech Prosody, 2010 (Speech Prosody 2010)* (pp. 1-4).
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, 62, 621-647.
- Kutas, M., & Hillyard, S. A. (1984). Event-Related Brain Potentials (ERPs) Elicited by Novel Stimuli during Sentence Processing. *Annals of the New York Academy of Sciences*, 425(1), 236-241.

- Kutas, M. and Hillyard, S.A., Brain potentials reflect word expectancy and semantic association during reading, *Nature*, 307 t 19841 161 163.
- Langner, G., & Ochse, M. (2006). The neural basis of pitch and harmony in the auditory system. *Musicae Scientiae*, 10(1\_suppl), 185-208.
- Lee, C. Y., & Lee, Y. F. (2010). Perception of musical pitch and lexical tones by Mandarin-speaking musicians. *The Journal of the Acoustical Society of America*, 127(1), 481-490.
- Lee, K., Chan, K., Lam, J., Van Hasselt, C., & Tong, M. (2015). Lexical tone perception in native speakers of cantonese. *International Journal of Speech-Language Pathology*, 17(1), 53-62.
- Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, 4(2), 185-199.
- Li, W., Wang, L., & Yang, Y. (2014). Chinese tone and vowel processing exhibits distinctive temporal characteristics: An electrophysiological perspective from classical Chinese poem processing. *PloS one*, 9(1), e85683.
- Liégeois-Chauvel, C., de Graaf, J. B., Laguitton, V., & Chauvel, P. (1999). Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cerebral cortex*, 9(5), 484-496.
- Liégeois-Chauvel, C., Giraud, K., Badier, J. M., Marquis, P., & Chauvel, P. (2001). Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex. *Annals of the New York Academy of Sciences*, 930(1), 117-132.
- Liu, T., Pinheiro, A. P., Deng, G., Nestor, P. G., McCarley, R. W., & Niznikiewicz, M. A. (2012). Electrophysiological insights into processing nonverbal emotional vocalizations. *NeuroReport*, 23(2), 108-112.
- Liu, Y., Hua, S., & Weekes, B. S. (2007). Differences in neural processing between nouns and verbs in Chinese: Evidence from EEG. *Brain and Language*, 103(1), 75-77.
- Liu, S., & Samuel, A. G. (2007). The role of Mandarin lexical tones in lexical access under different contextual conditions. *Language and Cognitive Processes*, 22(4), 566-594.
- Loui, P., Demorest, S. M., Pfordresher, P. Q., & Iyer, J. (2015). Neurological and developmental approaches to poor pitch perception and production. *Annals of the New York Academy of Sciences*, 1337(1), 263-271.
- Ma, J. K-Y, Ciocca, V., and Whitehill, T. (2006). Effect of intonation on Cantonese lexical tones. *JASA*, 320, 3978-3987.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature neuroscience*, 4(5), 540-545.
- Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, 23(10), 2701-2715.
- McDermott, J., & Hauser, M. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception: An Interdisciplinary Journal*, 23(1), 29-59.
- McCallum, W. C., Farmer, S. F., & Pocock, P. V. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 59(6), 477-488.
- Minati, L., Rosazza, C., D'incerti, L., Pietrocini, E., Valentini, L., Scaioli, V., ... & Bruzzone, M. G. (2009). Functional MRI/event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *Neuroreport*, 20(1), 87-92.
- Mithen, S., Morley, I., Wray, A., Tallerman, M., & Gamble, C. (2006). *The Singing Neanderthals: the Origins of Music, Language, Mind and Body*. London: Weidenfeld & Nicholson.
- Mitterschiffthaler, M. T., Fu, C. H., Dalton, J. A., Andrew, C. M., & Williams, S. C. (2007). A functional MRI study of happy and sad affective states induced by classical music. *Human brain mapping*, 28(11), 1150-1162.

- Nan, Y., & Friederici, A. D. (2013). Differential roles of right temporal cortex and Broca's area in pitch processing: evidence from music and Mandarin. *Human brain mapping*, 34(9), 2045-2054.
- National curriculum in England: music programmes of study. (2013). Retrieved from: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/239088/SECONDARY\\_national\\_curriculum\\_-\\_Music.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/239088/SECONDARY_national_curriculum_-_Music.pdf)
- Nettl, B. (1983). The study of ethnomusicology: Twenty-nine issues and concepts (No. 39). University of Illinois Press.
- Neufeld, J., Sinke, C., Dillo, W., Emrich, H. M., Szycik, G. R., Dima, D., ... & Zedler, M. (2012). The neural correlates of coloured music: A functional MRI investigation of auditory–visual synaesthesia. *Neuropsychologia*, 50(1), 85-89.
- Osawa, J. (2014). Alibaba Tackles Amazon, eBay on Home Turf. Wall Street Journal, 11.
- Packard, J. L. (1986). Tone production deficits in nonfluent aphasic Chinese speech. *Brain and Language*, 29(2), 212-223.
- Painter, J. G., & Koelsch, S. (2011). Can out-of-context musical sounds convey meaning? An ERP study on the processing of meaning in music. *Psychophysiology*, 48(5), 645-655.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, 6(7), 674-681.
- Patel, A. D. (2016). Using music to study the evolution of cognitive mechanisms relevant to language. *Psychonomic Bulletin & Review*, 1-4.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of cognitive neuroscience*, 10(6), 717-733.
- Patel, A. D., & Morgan, E. (2016). Exploring Cognitive Relations Between Prediction i Language and Music. *Cognitive Science*, 41, 303-320.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, 36(4), 767-776.
- Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport*, 19(2), 209-213.
- Pearce, M., & Rohrmeier, M. (2012). Music cognition and the cognitive sciences. *Topics in cognitive science*, 4(4), 468-484.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., ... & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, 107(10), 4758-4763.
- Platel, H., Baron, J. C., Desgranges, B., Bernard, F., & Eustache, F. (2003). Semantic and episodic memory of music are subserved by distinct neural networks. *Neuroimage*, 20(1), 244-256.
- Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological psychology*, 41(2), 103-146.
- Rasmussen, T., & Milner, B. (1977). The role of early left-brain injury in determining lateralization of cerebral speech functions. *Annals of the New York Academy of Sciences*, 299(1), 355-369.
- Santosa, H., Hong, M. J., & Hong, K. S. (2014). Lateralization of music processing with noises in the auditory cortex: an fNIRS study. *Frontiers in Behavioral Neuroscience*, 8.
- Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences*, 1060(1), 219-230.
- Shen, G., & Froud, K. (2015). Neurophysiological correlates of perceptual learning of Mandarin Chinese lexical tone categories: An event-related potential study. *The Journal of the Acoustical Society of America*, 137(4), 2384-2384.

- Simmons-Stern, N. R., Budson, A. E., & Ally, B. A. (2010). Music as a memory enhancer in patients with Alzheimer's disease. *Neuropsychologia*, 48(10), 3164-3167.
- Simons, Gary F. and Charles D. Fennig (eds.). (2017). *Ethnologue: Languages of the World*, Twentieth edition. SIL International. Retrieved from: [www.ethnologue.com/language/cmn](http://www.ethnologue.com/language/cmn)
- Stemmer, B., & Connolly, J. F. (2011). The EEG/ERP technologies in linguistic research: An essay on the advantages they offer and a survey of their purveyors. *The Mental Lexicon*, 6(1), 141-170.
- Steinbeis, N., & Koelsch, S. (2008). Comparing the processing of music and language meaning using EEG and fMRI provides evidence for similar and distinct neural representations. *PloS one*, 3(5), e2226.
- Steinbeis, N., & Koelsch, S. (2011). Affective priming effects of musical sounds on the processing of word meaning. *Journal of cognitive neuroscience*, 23(3), 604-621.
- Tones of Mandarin Chinese. (2010). Retrieved from: [www.misspandachinese.com/wp-content/uploads/2011/05/Tones-in-Mandarin-Chinese.pdf](http://www.misspandachinese.com/wp-content/uploads/2011/05/Tones-in-Mandarin-Chinese.pdf)
- Van Herten, M., Kolk, H. H., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, 22(2), 241-255.
- Van Lancker, D., & Fromkin, V. A. (1973). Hemispheric specialization for pitch and "tone": Evidence from Thai. *Journal of Phonetics*, 1(2), 101-109.
- Walker, J. (2016). A New Era for Chinese Language. *International Educator*, 25(4), 2.
- Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of cognitive neuroscience*, 15(7), 1019-1027.
- Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America*, 106(6), 3649-3658.
- Watanabe, T., Yagishita, S., & Kikyo, H. (2008). Memory of music: roles of right hippocampus and left inferior frontal gyrus. *Neuroimage*, 39(1), 483-491.
- Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49(1), 25-47.
- Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes*, 14(5-6), 609-630.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in cognitive sciences*, 6(1), 37-46.
- Zhao, T. C., & Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the National Academy of Sciences*, 113(19), 5212-5217.
- 2014 Music Standards. (2014). Retrieved from: <https://nafme.org/my-classroom/standards/core-music-standards/>

## Appendices

### Appendix 1: EEG Information document

#### **RADBOD UNIVERSITY CENTRE FOR LANGUAGE STUDIES INFORMATION ABOUT EEG RESEARCH**



#### General

Information processing in the central nervous system occurs, among other things, by electrical activity of the nerve cells. This minimal, continuous electrical activity of the brain, which is produced by the brain itself, can be measured and recorded by using electrodes. The result of such a measurement is called an ElectroEncephaloGram (EEG).

Depending on the research goal, the duration of an EEG experiment varies from 1 hour up to 2,5 hours.

The research data collected in the CLS Lab can't be viewed from a clinical perspective. Your participation in the research is therefore not a clinical test.

#### Preparation at home

To make the EEG measurement run smoothly, we ask you to prepare the following at home:

- Wash and dry your hair beforehand;
- Do not use gel, hairspray, etc.;
- Do not use face cream or make-up;
- If needed, bring a comb or hair brush;
- Bring your (reading) glasses. Also if you are wearing contact lenses.

#### Preparation at the CLS Lab

A cap (sort of bathing cap) will be put onto your head. In this cap a large amount of measuring electrodes will be attached. In addition, a few single electrodes will be attached around your eyes and behind your ears using small stickers. Your eyes, nose, mouth and the bottom part of your face will remain free.

To obtain good signals it is important that the resistance of the skin is not too high. If necessary, the experimenter will make sure the resistance between your skin and the electrodes drops to the desired value by using some alcohol and conducting gel.

#### The experiment

The experimenter will instruct you on what you have to do during the experiment. It may be that you have to look at a computer screen, listen to sounds, carry out a reaction-time task, or just sit and be relaxed. The EEG experiment takes place in a sound attenuated booth. During the measurement the door of the booth is closed, but not locked. The experimenter can see you by use of a video camera and talk to you by means of an intercom. The measurement itself will not be noticeable for you.

When the experiment is finished, the experimenter will remove the cap with the electrodes. If you want you can rinse your hair, wash and dry it. For this purpose shampoo and towels are available. For hygienic reasons it is practical if you bring your own comb.

### Additional information

The risk associated with participation can be considered as negligible risk and minimal burden. No invasive procedures are involved.

You can NOT participate in a EEG experiment if one of the following applies:

- 1) You had head/brain surgery.
- 2) You suffer from epilepsy.
- 3) You suffer from claustrophobia.
- 4) You are younger than 18 years of age.

If one of the above is applicable, please contact the researcher before the day of the experiment.

### Appendix 2: Draft list of Mandarin Words

#### **Neutral Fifth Tone:**

##### **1a**

1. ma5 吗
2. de5 的
3. le5 了
4. men5 们
5. me5 么
6. de5 得
7. zhe5 着
8. guo5 过
9. bei5 呗

##### **1b**

1. ba4ba5 爸爸
2. chi2zi5 池子
3. di4dao5 地道
4. sheng1yi5 商业
5. gen1tou5 跟头
6. huang3zi5 幌子
7. jiao1qing5 交情
8. kuang4zi5 框子
9. liang2kuai5 凉快
10. mu4tou5 木头

##### **1c**

1. ba 5 吧
2. a5 啊
3. ma 5 嘛
4. ne 5 呢
5. na 5 呐
6. na 5 哪
7. la 5 啦

8. ya 5 呀

9. wa 5 哇

#### **Fourth Tone**

##### **2a**

##### TOP 10 matched 4th tones (5x5)

1. ma4 骂

2. le4 乐

3. men4 闷

4. zhe4 (“this”)

5. guo4 过

6. ba4 爸

7. a4 啊

8. ne4 (“to speak”)

9. na4 捺

10. la4 辣

##### Bisyllabic: 4th tone final

##### **2b**

1. xian4zai4 现在

2. zhe4yang4 这样 / yi2yang4 一样

3. yin1wei4 因为

4. dan4shi4 但是 / jiu shi 就是

5. xu1yao4 需要

6. gong1zuo4 工作

7. ran2hou4 然后

8. yi2xia4 一下

9. kan4dao4 看到

10. xi1wang4 希望

#### Appendix 3: Final list of Mandarin Words with word frequency per million

##### **Tone 5: Syntactic**

1. ma5 吗 7428

2. de5 的 50155

3. le5 了 28211

4. men5 们 1191.12

5. me5 么 1166.14

6. de5 得 3135.62

7. zhe5 着 2983.65

8. bei5 呗 9.3

##### **Tone 5: Semantic**

1. ba 5 吧 5785.52

2. a5 啊 3903.92

3. ma 5 嘛 491.02
4. ne 5 呢 2287.93
5. na 5 呐 47.78
6. na 5 哪 786.52
7. la 5 啦 398.52
8. ya 5 呀 318.96
9. wa 5 哇 296.63

**Tone 4:**

1. ma4 骂 21.88
2. le4 乐 52.08
3. men4 闷 10.49
4. zhe4 这 8554.9
5. guo4 过 3350.21
6. ba4 爸 272.87
7. a4 啊 3903.92
8. ne4 讷 0.54
9. na4 那 5666.37
10. la4 辣 50.35
11. di 地 536.84
12. bei 备 11.24
13. ya 轧 1.7
14. wa 袜 7.96
15. ni 腻 4.95
16. lu 鹿 15.47
17. shu 树 64.09
18. da大 1334.24

Appendix 4: Consent Form (English)

**CONSENT FORM**

**Title of the research study:**

.....

**Researcher responsible:**

.....

**Statement of Participant**

The aim of the research study has been outlined to me. I was given the opportunity to ask questions regarding the research study. I participate voluntarily in the research study. I



understand that I can stop at any point during the research study, should I wish to do so. I understand how the data of the research study will be stored and how they will be used. I consent to participating in the research study.

Name: ..... Date of birth: .....

Signature: ..... Date: .....

***Statement of executive researcher***

I declare that I have informed the above-mentioned person correctly about the research study and that I abide by the guidelines for research as stated in the protocol of the Ethics Assessment Committee Humanities.

Name: .....

Signature: ..... Date: .....