university of groningen / faculty of science and engineering

Investigating Statistical Tone Learning by Analyzing Mismatch Negativity in ERP Results

Bachelor's Project Thesis

Jing Chen, s3336514, j.chen.32@student.rug.nl,
Supervisors: Dr J.K.Spenader & PhD M.Tang

Abstract: Statistical learning enables individuals to include regularity in their learning environment, as evidenced from the presence or absence indicated by the Mismatch Negativity (MMN) in Event Related Potential (ERP) data. In the current study, we used statistical learning to analyze 6 Mandarin tone-syllable combinations presented in two multi-featured oddball paradigms to determine the likelihood that non-tonal language speakers were able to learn tones and distinguish between tonal differences. Subsequently, a behavioral test was conducted, wherein the research findings indicated a lack of significant differences in the ERP results between early tone and syllabic learning. The participants were essentially incapable of distinguishing between tone differences during the early phase. Conversely, during the latter phase, we detected the presence of MMN in both syllabic and tone learning. In the context of late learning, participants relied on statistical learning to identify the tone differences. Moreover, the degree of accuracy in perceiving tonal and syllabic differences was above the chance level in the behavioral test. Accordingly, we concluded that non-tonal language speakers were capable of learning tones and distinguishing between tone differences.

1 Introduction

The strenuous hurdles that non-tonal language speakers face in mastering tonal languages are universally and broadly recognized (Shen, 1989). Notably, the vastly different pitch patterns, distribution, and functions between tonal and non-tonal languages are major factors (White, 1981). In this study, non-tonal language speaker participants tested two multi-featured oddball tasks through statistical learning to differentiate between six Mandarin tone-syllable combinations. Appropriately, the results can be determined by observing the presence of MMN in the ERP data to ascertain memory traces of statistical learning among the participants.

1.1 Tone and tonal language

As we mature, the innate capacity to assimilate auxiliary language features apart from our native language is greatly diminished. In general, an individual's native language substantially impacts their perception of non-native language and affects their ability to distinguish between linguistic features of non-native languages (Antoniou and Chin, 2018). In a linguistic context, tone is employed as a linguistic feature using pitch to express emotions, indicate stress, and outline contrast. In addition, it can be leveraged to ascertain the differences between lexical and grammatical meaning. In a vast majority of Western countries, tone is utilized to emphasize word stress or express emotion (Yip, 2002). However, other languages employ tone to both express emotion and distinguish between words. Furthermore, languages have these linguistic features are referred to as tonal languages, which comprise more than half of the world's languages.

Subsequently, their intonation pattern of lexical tones is a contrastive feature that discerns the differences between word meanings (Yip, 2002). From a broader perspective, different tones in conjunction with the same words could signify a diversity of meanings. Likewise, tonal languages such as Mandarin, Vietnamese, and Cantonese utilize lexical tones to distinguish between word meanings. For instance, Mandarin (the most commonly spoken tonal language) consists of four tones: 1 (flat), 2 (rising), 3 (low-dipping), and 4 (high falling). In Mandarin's linguistic context, the same word can express different meanings depending on the tone. For example, the usage of the syllable /mi/ spoken with four Mandarin tones, i.e., 'mi1 mi2 mi3 mi4' can be translated as 'to squint/narrow one's eyes', 'a riddle', 'rice', and 'honey'. Although these words have the same pronunciation, their tone differences create vastly different meanings.

In actuality, similar tone changes do not signify lexical meaning in non-tonal languages such as English. Notably, specific studies highlight the significant impact of these linguistic differences based on lexical tone processing. For example, native Mandarin speakers are regarded as proficient in recognizing tone (Gottfried and Suiter, 1997). Moreover, these speakers are viewed as having a strong categorical perception of Mandarin pitches, which British listeners who are speakers of a non-tonal language may lack (Antoniou and Chin, 2018). In addition, this innate native language advantage has also been detected in neurodisciplinary studies on tone processing. Broadly, tonal language listeners also have an advantage in cortical processing Wong, Parsons, Martinez, and Diehl (2004) owing to their devoted and robust subcortical encoding of tones (Krishnan, Xu, Gandour, and Cariani, 2005). Nevertheless, tonal languages vary significantly in the size and composition of their tonal inventories, which influences non-native tone perception.

1.2 Statistical learning

Statistical learning (SL) is a crucial learning mechanism in human cognition that allow humans to understand regularity in their learning environment. SL usually occurs unconsciously within minutes of exposure. (Thiessen, Girard, and Erickson, 2016). An acquisition theory that relies on SL proposes that an infant's ability to learn a language is based on pattern perception rather than the infant's innate biological grammar (Wang and Saffran, 2014). Moreover, research has shown that the perception of non-native language features can be enhanced by implementing statistical learning language training (Erickson and Thiessen, 2015). Likewise, the intricate language learning process is aided by mechanisms that subtly track the distribution of information in the linguistic environment (Lany and Saffran, 2010). For example, when listeners are exposed to new speech patterns in a laboratory, they can easily familiarize themselves with these patterns. Further studies have argued that 9-month-old infants have demonstrated these distinct capabilities (Saffran and Thiessen, 2003; Kuhl, Ramírez, Bosseler, Lin, and Imada, 2014), while adults maintain the ability to rapidly acquire new phonemes following the complete development of their native phonemes (Goldrick, 2004).

In our study, we will adopt statistical learning to present Mandarin's tonal vocabulary to non-tonal speakers in order to determine the likelihood of non-tonal speakers being able to perceive non-native language features. We intend to apply statistical learning in order to assess these speakers' capacity to learn the tonal differences.

1.3 ERPs and MMN

Several studies have revealed the human brain's inherent capacity to detect subtle brainwave changes by capturing event-related potentials (ERPs) in an electroencephalogram (EEG), especially when participants are constantly exposed to repetitive stimuli. In short, our brain spontaneously detects statistical regularities in the environment and records 'surprising' events (Giuliano, Pfordresher, Stanley, Narayana, and Wicha, 2011). We can see evidence of statistical learning in the brain by looking for the presence of mismatch negativity (MMN) in ERP data.

In the MMN paradigm, the formulation of short-term predictive representations of environmental regularity are based on the individual's perception of recurrent events (Stefanics, Kremláček, and Czigler, 2014). Specifically, establishing a direct link between statistical learning and prediction or consolidation is a critical factor. MMN is usually evoked by using the oddball paradigm. The standard stimuli (high-frequency stimuli) in the oddball paradigm are occasionally interrupted by deviant stimuli (low-frequency stimuli) (Näätänen, Paavilainen, Rinne, and Alho, 2007). MMN is calculated by subtracting the brain's response to deviant stimuli from its response to standard stimuli. In most experiments, the employed stimuli are simple tones, while the difference between the standard and deviant stimuli in the MMN is either a frequency or amplitude Näätänen (1990). In particular, MMN is observable within the range of 200-350 ms following the stage of stimulus presentation. In this study, we complied with the multi-featured oddball paradigm published by Näätänen, Pakarinen, Rinne, and Takegata (2004) by using Mandarin tone-syllable combinations as standard stimuli (high-frequency stimuli) and deviant stimuli (low-frequency stimuli) presented in the multi-feature oddball paradigm. While research topics on exact brain regions associated with MMN remain controversial (Zhang, Yan, Wang, Wang, Wang, Wang, and Huang, 2018), we will follow Näätänen et al. (2004)'s approach and prioritize the observation of activity captured by midline frontal (FZ) cortex electrodes in the experimental results.

To summarize, this study exposed 6 Mandarin tone-syllable combinations to non-tonal language speakers to assess their ability to use statistical learning. Additionally, the presence of MMN can be detected by testing the participants' ERP results. Thus, our research question is: Do we find memory traces suggesting there is SL in non-tonal language speakers through exposure to two multi-featured oddball paradigms comprised of Mandarin tones and syllables?

2 Methods

2.1 Participants

Eighteen non-tonal language speakers were recruited for this experiment (10 males; Age range: 19-37, Mean age: 23.8 years). None of the participants reported hearing or neurological impairment. Furthermore, all participants gave informed consent and were paid 16 euros as remuneration within two weeks of the experiment. As our study involved statistical learning, we did not inform the participants in advance about the experiment content and research topic to ensure the accuracy of statistical learning.

2.2 Stimuli & Experimental Setup

Block	**Stimuli**	**Syllable**	**Tone**
MO1	Standard1(S1)	bi	flat
	Deviant1(D1)	du	flat
	Deviant2(D2)	bi	rising
MO2	Standard2(S2)	kou	low-dipping
	Deviant3(D3)	pei	low-dipping
	Deviant4(D4)	kou	falling

Table 2.1: Table of syllable-tone combinations as the standard and deviant stimuli in the experiment.

To test whether participants would implicitly learn and recognize changes in tone, four Mandarin syllables (/bi/, /du/, /kou/, and /pei/) were combined with one of the four Mandarin lexical tones (flat, rising, low-dipping, and falling) to produce six tone-syllable combinations in our experiment. Namely, flat /du/, flat /bi/, rising /bi/, low-dipping /kou/, falling /kou/, and low-dipping /pei/, as 2 standards and 4 deviant stimuli in two blocks, shown in Table 2.1.

Subsequently, a voice actress used a microphone (Audio-Technica AT2020) to enunciate and record the entire set of combinations at a recording studio. Each audio file had a duration of 1000ms and was harmonized using Audacity 3.1.3 to maintain an average peak frequency for each Mandarin tone at 345, 290, 230, and 400 Hz, respectively, with a moderate intensity of 70 decibels (dB). Furthermore, the voice onset time(VOT) was 290ms, while the offset time was 750ms for each sound file.

2.2.1 Multi-feature oddball stimuli

Our experiments used six stimuli with different tones, i.e. two standard and four deviant stimuli, presented in two multi-featured oddball paradigm tasks. These experiments were created in OpenSesame 3.3.11. All stimuli were equally distributed into two blocks named Multi-featured Oddball 1

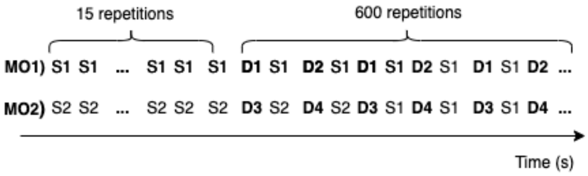

Figure 2.1: Example of multi-feature oddballs

(MO1) and Multi-featured Oddball 2 (MO2). To highlight the effects of distributed frequencies on specific stimuli, standard stimuli were also used as high-frequency stimulation (HF), appearing 50% of the time in each MO. The standard stimulus was first presented 15 times at the beginning of the task, followed by two deviant stimuli alternating with the standard stimulus. The two deviant stimuli were low-frequency stimuli (LF), and each deviant stimulus was presented 25% of the time. The two deviant stimuli were interchangeable. An example of two multi-feature oddballs is shown in Figure 2.1. Throughout the experiment, which lasted approximately 20 minutes, the standard stimuli were presented 315 times in each block, and each deviant stimulus was presented 150 times, as shown by Näätänen et al. (2004).

2.2.2 Behavioral testing stimuli

Testing trials	Which sound is more familiar to you?	
Trial 1	**Flat /du/**	**Flat /bi/**
Trial 2	**Flat /bi/**	**Rising /bi/**
Trial 3	**Flat /du/**	**Rising /bi/**

Figure 2.2: Example of behavior testing trials

To determine the participants' sensitivity to tones and syllables, we tested whether they could differentiate between different tone and syllable stimuli. All six stimuli were used to form 32 pairs of tone-syllable combination testing trials for the behavioral test. Each pair was separately presented to the participants. Pairs of stimuli varied according to tone (e.g., flat/bi/ and rising/bi/), syllable (flat/du/ and flat/bi/), or both (e.g., flat/du/ and rising/bi/). 2.2 shows an example of three pairs of the tone-syllable combination testing trial in the behavioral test.

2.3 Procedure

The entire experiment had two parts - the first part was an EEG experiment with two multifeatured oddball tasks, and the second part was a behavioral test. The experiment lasted a total of one hour - half an hour for the subject's EEG setup, twenty minutes for the EEG experiment and ten minutes for the behavioural test. Participants listened to the oddball experiment while watching a silent nature documentary. We recorded the participants' EEG data using the EEG device and BioSemi. Half of the participants started with MO1 and continued with MO2. The other half started with MO2 and continued with MO1 to ensure that the experiment was counterbalanced. After the oddball experiment ended, the behavioral experiment followed. During the behavioral experiment, participants had three trial tests that were used to ensure that the participants understood the task. This was followed by a formal test with a total of thirty-two pairs of standard and deviant stimuli, where participants had five seconds to indicate and choose which stimulus sounded more familiar. We recorded the accuracy and reaction time for each testing trial. At the end of the experiment, we also recorded the language that participants used as well as their musical experience, as this may have affected their ability to discriminate between tone difference.

2.4 EEG recording and preprocessing

We used 32 Ag/AgCI active electrodes (BioSemi ActiveTwo system) to record the EEG, and digitized it at a sampling rate of 1024 Hz. In addition, we attached six additional electrodes to each participant. Two of the EOG electrodes were placed in the outer corners of each eye, while two others were placed above and below the left eye to detect eye movements. Two EOG electrodes were mounted behind the ear to detect mastoid signals. Data from each participant was re-referenced to determine the grand mean of whole-head electrodes. Scalp impedance of each electrode was maintained at 30 kΩ. We used the open-source toolbox EEGLAB (Delorme Makeig, 2004) and custom scripts from MATLAB (MathWorks, Inc.) to pre-process and analyze the EEG.

We pre-processed the raw EEG data to make it viable and analyzable. First, we divided it into two groups. The first 50% of each MO was concatenated into one set called Early Learning. The second 50% of each MO was concatenated into another group called Late Learning. This was used to analyze whether SL occurred before or after each MO was introduced for 5 minutes. Next, a 50 Hz notch filter was applied to the two sets of raw EEG data, while a 0.01 Hz high-pass filter and a 30 Hz low-pass filter were used for band-pass filtering. This removed the noise from the data to obtain a high-quality ICA decomposition during subsequent processing. The data was then re-referenced to the average reference value, and then artifacts were removed from the data using the automatic channel and a significantly different artifact. This also resulted in the loss of about 20% of the data from both sets. Hence, we re-referenced the data based on the average reference value. The two data sets were then corrected using the independent component analysis (ICA) algorithm to calculate the ICA components and return the data to the respective channels. Moreover, high currents captured by multiple electrodes were used to identify sources other than the brain, namely, eye and muscle movements, which were removed. Finally, we divided the data into epochs starting from 100 ms before and 900 ms after stimuli presentation. Each epoch was labeled in the early or late data sets as a standard stimulus, a deviant tone stimulus, or a deviant syllable stimulus.

3 Results

In this experiment, we employed 1-way repeated measure ANOVA in the EEGLab in Matlab to generate the ERPs result.

3.1 ERP results

Figure 3.1, we can see the ERP results of the syllabic learning. Specifically, Figure 3.1 a) represents the ERP result during the stage of early syllabic learning, which aligns with the participant's perceived differences between the standard stimuli in conjunction with the deviant syllable stimuli during this stage. Additionally, Figure 3.1 b) depicts the ERP result during late syllabic learning, and portrays the difference in the participant's perceived difference between the standard stimuli with the deviant syllable stimuli in this stage.

Figure 3.2 infers the ERP result in the tone learning. Notably, Figure 3.2 c) constitutes the ERP result during the early tone learning in accordance with the participant's perceived difference between the standard stimuli with the deviant tone stimuli in the early learning phase . Additionally, Figure 3.2 d) corresponds to the ERP result during the late tone learning. This highlights the difference based on the participant's perceived difference between the standard stimuli with the deviant tone stimuli during the late learning phase.

In each figure, the bold black vertical line represents the voice on set time (VOT) in our experiment, which is approximately 290ms. In addition, the red dashed rectangle portrays the typical occurrence of MMN. As a reminder, MMN generally occurs between 150-250ms following the presentation of the stimuli. (Zhang et al., 2018) In short, the event takes place after the VOT. Ideally, our MMN should occur within the time window of 400-500ms.

3.1.1 Syllabic Learning

Conforming with our visualization of early syllabic learning from Figure 3.1 a), the lack of MMN occurrence when the p-value is greater than 0.05 in the MMN area is evident. Although the deviant stimuli elicited greater amplitudes compared to the standard stimuli, the statistics did not show difference ($t(1228.7) = 0.91704$. $p = 0.3593$). In the absence of significant difference inferred from the comparison of standard and deviant syllable stimuli, it implies that the participant failed to differentiate the tone difference between standard and deviant syllable stimuli

In the late syllabic learning from Figure 3.1 b), the deviant syllable stimuli evoked a negative component and larger amplitude ($t(1798) = 24.781$, $p < 0.05$) compared to the standard stimuli, in tandem with the appearance of MMN

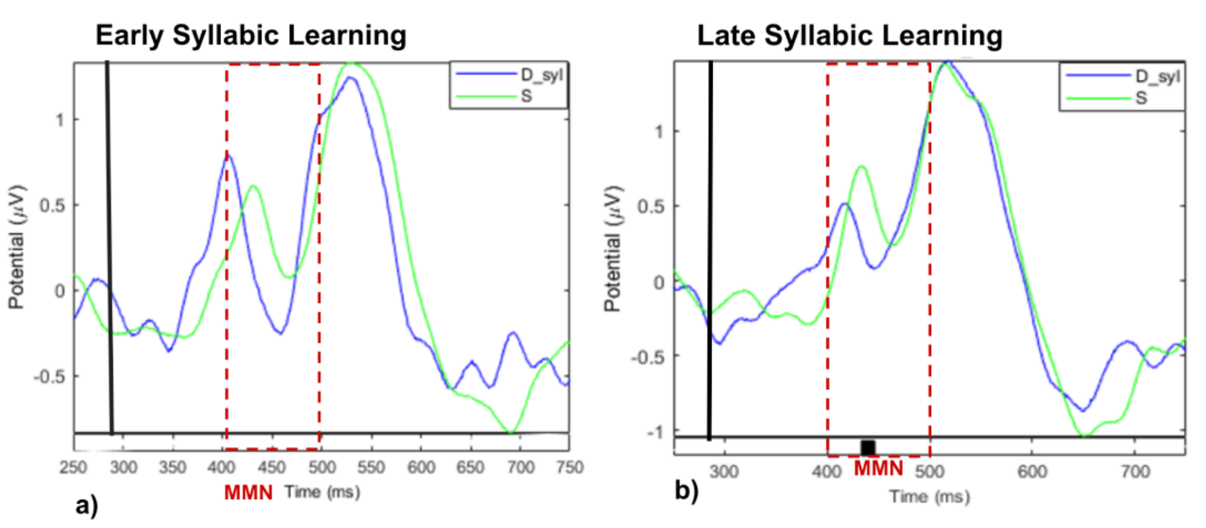

Figure 3.1: ERP results in Syllabic learning. a)Comparison of standard to deviant syllable stimuli in early phase b)Comparison of standard to deviant syllable stimuli in late phase

during its specific time range, thereby indicating the prevalence of a significant difference during late syllabic learning. Specifically, the participants failed to deduce the difference between the standard and deviant syllable stimuli.

3.1.2 Tone Learning

In Figure 3.2 c), a MMN was shown in a narrow time interval before 450 ms in the early tone learning. The deviant tone stimuli elicited a more negative component and smaller amplitude ($t(1495.1) = 21.245$, $p < 0.05$) compared to the standard stimuli. As a result of negligible disparities, we were unable to deduce this is a meaningful MMN response. Based on the time window between 670-700ms, we could observe pertinent components other than MMN. The components are termed Late Discriminative Negativity (LDN) and typically appear between the latency range of 350-450m after stimulus onset. (Rosburg, Trautner, Dietl, Korzyukov, Boutros, Schaller, Elger, and Kurthen, 2005) The LDN component exhibited positive component from the deviant tone stimuli to the standard stimuli while concurrently displaying significant differences.

Figure 3.2 d) depicts three components during late tone learning. Initially, MMN occurs within the time window of 400-500ms along with the negative ERP response and larger amplitude ($t(1797.2) = 4.1729$, $p < 0.05$) from the deviant tone stimuli to the standard stimuli. Accordingly, the results highlight the participants' ability to distinguish the tone difference between the standard and deviant tone stimuli. P300 component is also shown in the figure, and this event typically happens within the range of 250-350ms after the VOT (Dunn, Dunn, Languis, and Andrews, 1998). In addition, the LDN component is observed during the time window of 630-680ms. However, both components elicited positive ERP response in deviant tone stimuli compared to the standard stimuli, while concurrently upholding significant differences in P300 and LDN.

3.2 Behavior testing results

In the context of behavioral testing results, we investigated the likelihood of participants being able to distinguish between deviant tone stimuli and deviant syllable stimuli from their standard counterparts. Accordingly, we excluded 16 testing trials where that the stimuli differ in both tone and syllable. The remaining data were grouped into two conditions of tone learning and syllable learning, with 8 testing trails for each condition. Finally, we calculated the mean accuracy for syllabic learning as

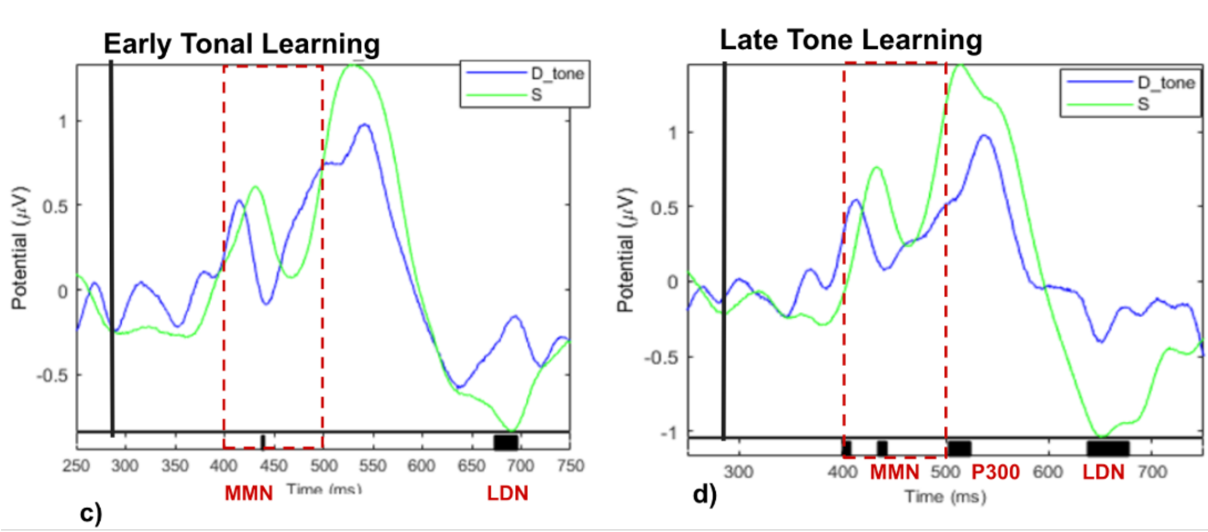

Figure 3.2: ERP results in Tone learning. c)Comparison of standard to deviant tone stimuli in early phase. d)Comparison of standard to deviant tone stimuli in late phase

71.53 and SD as 27.58. Likewise, the mean accuracy for tone learning was 69.14 and SD was 27.58. Figure 3.3 compares both tone and syllabic learning to the chance level. Afterwards, we performed the one sample T-test to compare both tonal learning and syllabic learning to the chance levels. Based on the results, there was a significant difference when comparing tone learning with the chance levels ($t(17)= 4.05$, $p < 0.05$). In addition, we saw significant differences when comparing syllabic learning with the chance levels ($t(17) = 3.22$, $p < 0.05$). Similarly, we adopted a paired t-test to compare the magnitudes of tone learning and syllabic learning. However, no differences were discovered($t(17) = -0.41$, $p = 0.685$). Taken the above results together, we can conclude that tone difference can be learned statistically. But our participants did not show more prominent learning on either tones or syllables.

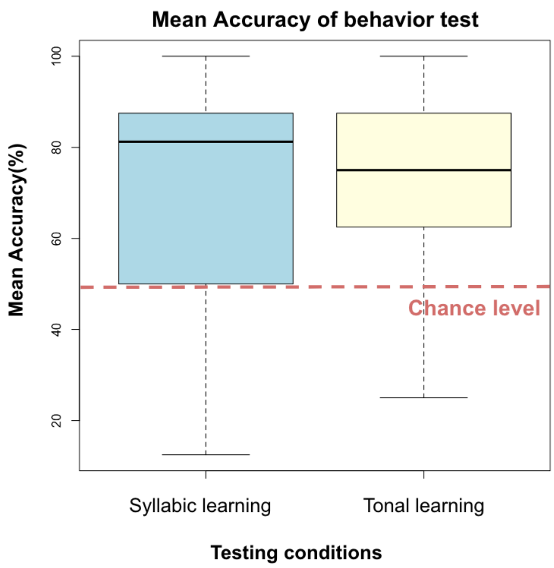

Figure 3.3: Mean accuracy of the deviant conditions for the behavior test

4 Discussion & Conclusions

The present study examined the changes in brain waves of non-tonal language speakers during the early and late learning phases by implementing statistical learning of tonal language. During the early period of statistical learning, no MMN appeared in the ERP results, which correlates with the findings of Bogaerts, Richter, Landau, and Frost (2020). Conversely, participants elicited a significant MMN in the ERP results during the second half of the experiment (the late period). During this period, participants deduced significant differences in both tone learning and syllable learning, which indicates tonal differences between standard and deviant stimuli based on the participants' stimuli exposure. This also suggests that learning to distinguish tone differences may involve the process of developing memory traces for two sounds with different tones. The behavioral test results confirmed that non-tonal language users were capable of learning tones by participating in statistical learning. Despite participants' inability to distinguish between deviant tone stimuli and deviant syllable stimuli through statistical learning, the mean accuracy of all stimuli was greater than the chance level.

4.1 Limitations

According to Craciun, Gardella, Alving, Terney, Mindruta, Zarubova, and Beniczky (2014), the upper limit of the length of the awake EEG experiment was 20 minutes. The duration of our EEG experiments was around 20 minutes. Excessively long experiments can also increase artifacts and lead to excessive noise in the experimental data, which can affect the accuracy of the results. Accordingly, shortening the duration of this experiment may enhance the overall accuracy of its results. In addition, P300 and LDN components appeared in the ERP data during the second half of the early and late learning stages. The deviant tone stimuli elicit a positive ERP response compared to the standard stimuli, whereas it should normally elicit a negative component. Notably, this strange result could be attributed to excessive noise in the data. Even though we removed a vast portion of the noise during the pre-processing phase, our laboratory's lack of fully soundproof facilities created disturbances among participants during the experiment phase, which resulted in noisy EEG data.

4.2 Future research

There are many compelling aspects that we can meticulously evaluate during future research. It is particularly evident that the participants' musical background greatly contributed to their differentiation between tones. In a study by Delogu, Lampis, and Olivetti Belardinelli (2006), a well-informed musical background can impart the ability to distinguish between tonal differences. Accordingly, the group comprised of highly melodic participants performed better than the other group in detecting tone variation. Mandarin includes only 4 different lexical tones. In future studies, we plan to focus on complex tonal languages such as Cantonese, which has 6 lexical tones (So, 1996). Moreover, we can employ different syllable combinations as deviant stimuli to assess participants' ability to detect and distinguish between tonal differences.

References

Mark Antoniou and Jessica LL Chin. What can lexical tone training studies in adults tell us about tone processing in children? *Frontiers in psychology*, 9:1, 2018.

Louisa Bogaerts, Craig G Richter, Ayelet N Landau, and Ram Frost. Beta-band activity is a signature of statistical learning. *Journal of Neuroscience*, 40(39):7523–7530, 2020.

L Craciun, Elena Gardella, Jørgen Alving, D Terney, I Mindruta, J Zarubova, and S Beniczky. How long shall we record electroencephalography? *Acta Neurologica Scandinavica*, 129(2):e9–e11, 2014.

Franco Delogu, Giulia Lampis, and Marta Olivetti Belardinelli. Music-to-language transfer effect: May melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive processing*, 7(3):203–207, 2006.

Bruce R Dunn, Denise A Dunn, Marlin Languis, and David Andrews. The relation of erp components to complex memory processing. *Brain and cognition*, 36(3):355–376, 1998.

Lucy C Erickson and Erik D Thiessen. Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, 37:66–108, 2015.

Ryan J Giuliano, Peter Q Pfordresher, Emily M Stanley, Shalini Narayana, and Nicole YY Wicha. Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in psychology*, 2:146, 2011.

Matthew Goldrick. Phonological features and phonotactic constraints in speech production.

Journal of Memory and Language, 51(4):586–603, 2004.

Terry L Gottfried and Traci L Suiter. Effect of linguistic experience on the identification of mandarin chinese vowels and tones. *Journal of Phonetics*, 25(2):207–231, 1997.

Ananthanarayan Krishnan, Yisheng Xu, Jackson Gandour, and Peter Cariani. Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25(1):161–168, 2005.

Patricia K Kuhl, Rey R Ramírez, Alexis Bosseler, Jo-Fu Lotus Lin, and Toshiaki Imada. Infants' brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences*, 111(31):11238–11245, 2014.

Jill Lany and Jenny R Saffran. From statistics to meaning: Infants' acquisition of lexical categories. *Psychological science*, 21(2):284–291, 2010.

Risto Näätänen. The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and brain sciences*, 13(2):201–233, 1990.

Risto Näätänen, Satu Pakarinen, Teemu Rinne, and Rika Takegata. The mismatch negativity (mmn): towards the optimal paradigm. *Clinical neurophysiology*, 115(1):140–144, 2004.

Risto Näätänen, Petri Paavilainen, Teemu Rinne, and Kimmo Alho. The mismatch negativity (mmn) in basic research of central auditory processing: a review. *Clinical neurophysiology*, 118 (12):2544–2590, 2007.

Timm Rosburg, Peter Trautner, Thomas Dietl, Oleg A Korzyukov, Nashaat N Boutros, Carlo Schaller, Christian Erich Elger, and Martin Kurthen. Subdural recordings of the mismatch negativity (mmn) in patients with focal epilepsy. *Brain*, 128(4):819–828, 2005.

Jenny R Saffran and Erik D Thiessen. Pattern induction by infant language learners. *Developmental psychology*, 39(3):484, 2003.

Xiaonan Susan Shen. Toward a register approach in teaching mandarin tones. *Journal of the Chinese Language Teachers Association*, 24(3):27–47, 1989.

LKH So. Tonal changes in hong kong cantonese. *CURRENT ISSUES IN LANGUAGE & SOCIETY*, 3(2):186, 1996.

Gábor Stefanics, Jan Kremláček, and István Czigler. Visual mismatch negativity: a predictive coding view. *Frontiers in human neuroscience*, 8: 666, 2014.

Erik D Thiessen, Sandrine Girard, and Lucy C Erickson. Statistical learning and the critical period: how a continuous learning mechanism can give rise to discontinuous learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 7(4): 276–288, 2016.

Tianlin Wang and Jenny R Saffran. Statistical learning of a tonal language: The influence of bilingualism and previous linguistic experience. *Frontiers in psychology*, 5:953, 2014.

Carolyn M White. Tonal perception errors and interference from english intonation. *Journal of Chinese Language Teachers Association*, 16(2): 27–56, 1981.

Patrick CM Wong, Lawrence M Parsons, Michael Martinez, and Randy L Diehl. The role of the insular cortex in pitch pattern perception: the effect of linguistic contexts. *Journal of Neuroscience*, 24(41):9153–9160, 2004.

Moira Yip. *Tone*. Cambridge University Press, 2002.

Yun Zhang, Fei Yan, Liu Wang, Yubo Wang, Chunshu Wang, Qiang Wang, and Liyu Huang. Cortical areas associated with mismatch negativity: A connectivity study using propofol anesthesia. *Frontiers in Human Neuroscience*, 12:392, 2018.