THE UNIVERSITY OF CHICAGO

Does Auditory Mental Imagery Exhibit Neural Overlap With Auditory Perception?

A Paper Submitted In Partial Fulfillment of the requirements for the Master of Arts degree in the Masters of Arts Program in the Social Sciences

Rachna Ravishankar

August 2023

Faculty Advisor: Dr. Howard C. Nusbaum

Preceptor: Dr. Hannah R. Hamilton

Abstract

Imagination extends across modalities and involves numerous cognitive processes (Pelaprat & Cole, 2011). The presence of overlap in neural activity between imagination and perception (Halpern et al., 2004), prompts the questioning of whether mental representations of previously encountered auditory stimuli and the perception of said stimuli rely on the same neural mechanisms. This study was designed in order for subjects to encode the sound of pure tones after repeated exposure, and subsequently form mental images of said tones when prompted. A pitch adjustment task was utilized in order to confirm that participants were forming mental reconstructions of the tone provided. It is hypothesized that the power of the brain signals elicited during the perceived and imagined sequences will be similar. Results indicate that there is no significant difference in power between the brain signals emitted during the perception of pure tones and the mental imagery of said tones.

Keywords: mental imagery, pure tones, pitch perception

Does Auditory Mental Imagery Exhibit Neural Overlap With Auditory Perception?

How do mental reconstructions of sounds compare to the perception of sounds heard in real time? Mental imagery involves the recollection of previously perceived stimuli and in some ways is somewhat similar to perception (Brogaard & Gatzia, 2017). Mental imagery has been found to utilize memory recall that is dependent on the sensory cortices associated with the initial stimulus, such as the visual cortex being activated upon remembering an image (Pearson, 2019). However, mental imagery can also be examined in the auditory domain (Vlek et al., 2011). The present auditory perception study examines auditory mental imagery and its potential overlap with auditory perception.

Prior Research

Mental imagery is often associated with the visual domain, in part due to extensive research on visual perception (Pearson, 2019). Visual mental imagery has been considered to overlap with perception in its phenomenology (Brogaard & Gatzia, 2017) and in its neuropsychological aspects, such as activity in the visual cortex (Pearson, 2019). However, research involving individuals with brain lesions indicates that individuals can experience deficits with visual imagination while having intact visual perception (Moro et al., 2008).

While there is an extent to which visual mental imagery differs cognitively from visual perception, the auditory modality has exhibited substantial overlap between perception and imagination in regards to the cognitive representations of auditory stimuli (Halpern et al., 2004; Vlek et al., 2011). As sounds travel through the air they eventually reach the cochleas in which the basilar membranes vibrate according to the frequency of the sound, allowing for the innervation of hair cells to carry electrical impulses that eventually reach the brain (Phillips et al., 2002). However, even after sounds are heard, they can be mentally reconstructed which in turn involves activating areas of the brain associated with auditory perception (Halpern et al., 2004).

In regards to the cognitive processes essential to auditory perception, speech perception in particular relies heavily on working memory and attention (Heald & Nusbaum, 2014). Training to improve working memory has even been found to aid with speech perception in conditions in which there is excess noise pollution (Ingvalson et al., 2015). Likewise, the process of imagining speech also involves cognitive efforts such as memory retrieval and perceptual reactivation (Tian et al., 2016). This process involves increased activity in the left temporal cortex, suggesting that imagining speech involves the utilization of speech perception (McGuire et al., 1996). Therefore, it appears that the activity of imagined auditory phenomenon is in some ways cognitively comparable to the auditory perception of acoustic stimuli present within an environment.

The Current Study

The present study aims to offer further insight on the mental reconstructions of auditory stimuli by examining the potential similarities between the perception of pure tones and the subsequent mental imagery of these pure tones. Subjects were randomly provided with one of two pure tones to listen to repeatedly during a trial, which they then mentally reconstructed when prompted. A pitch adjustment task was also implemented in each trial in order to assess whether subjects were able to accurately recall the pure tone they were asked to mentally reconstruct. It was hypothesized that mental reconstructions of previously heard pure tones will emit brain signals with power that are similar to the power of the brain signals emitted when the tones are heard in real time.

The examination of the overlap between auditory mental imagery and auditory perception through the usage of EEG analysis is relatively new, but has been before by Vlek et al. (2011) with results that indicate that auditory mental imagery can be decoded through the usage of EEG analysis. However, this research utilized accented and non-accented beats, thus focusing primarily on how differences in loudness influence the mental imagery of auditory stimuli. The current study, however, utilizes pure tones at a constant amplitude, and thus instead focuses on pitch perception. Pitch perception, as indicated by Perrachione et al. (2013), demonstrates shared cognitive processes between both the perception of music and the perception of language. Overlap in auditory processes and in working memory contribute to the ability to distinguish sentence prosody and determine melodies, but it is speculated that higher order cognitive functions, that have yet to be fully distinguished, are influential in pitch processing to a greater extent than musical experience alone. Therefore, the current study intends to examine pitch perception and auditory mental imagery with the intention of evaluating the cognitive functions that overlap between the two areas.

Methods

Participants

Subjects ($N = 19$) included in this study consist of undergraduate and graduate students who attend the University of Chicago. Subjects were recruited through an online platform which allowed for students to participate in ongoing research projects on the University of Chicago campus. The mean age of subjects is 19.6 years old, with 10 subjects identifying as women and 9 subjects identifying as men. Subjects either participated in order to meet a course requirement or for payment. All subjects were provided with the opportunity to earn $10 in cash if performance was well. Some subjects only completed four blocks, however, all subjects completed at least four blocks.

Materials

A questionnaire was provided to participants prior to the administration of the task. This questionnaire assessed general demographic information, as well as prior musical experience. Subjects were asked to report their initial age of musical instruction, as well as the types of instruments they have played and how many years they have played them. Subjects were also asked to report if they continue musical engagement or if they have since ceased playing instruments. Additionally, subjects were asked to report any additional languages which they are fluent in aside from English. Two pure tones were utilized in this study, 190 Hertz and 280 Hertz. The tones were played at sixty-five decibels on insert earphones with styrofoam disposable earbuds.

Procedure

During this EEG study, subjects completed a series of auditory perception tasks that spanned over five blocks, with thirty trials per block. Subjects were provided with headphones and seated at a desk in a soundproof chamber with a monitor and keyboard placed before them. Subjects were informed that during each trial they would be listening to a tone repeatedly with the expectation of being able to formulate a mental image of the given tone when prompted. In order to prompt subjects to imagine the tone they previously heard, a signifier appeared on the monitor before them without any audio. This signifier was a white plus mark located at the center of a gray screen. This signifier also appeared during the repeated playing of the pure tone itself in order to limit the influence of visual stimuli as a potential distraction. Subjects were instructed to remain as still as possible and look at this signifier when tones were played in order to limit rapid eye movements and other physical movements. A practice block occurred at the beginning of the session which included three practice trials.

Two tones were used randomly in this study. Either a tone of 190 Hertz or 280 Hertz was heard during a trail. The tones themselves lasted 500 milliseconds. Subjects initially heard one of the two tones that had been randomly selected while the signifier was presented on the monitor before them. The tone utilized during the given trial was played repeatedly on a beat for a total of four times with a gap of 1.2 seconds between each tone. On the fifth beat, the signifier reappeared on screen but no audio was played. During this brief period of silence, subjects were expected to imagine the pure tone they previously heard. After a period of 1.2 seconds, white noise played briefly. After the period of white noise, a pitch adjustment task took place in which an adjusted version of the original tone was provided. This adjusted tone ranged from ten Hertz above to ten Hertz below the original tone and lasted 500 milliseconds. The pitch difference between the adjusted tone and the pure tone used for each trial was randomized from trial to trial. Subjects were tasked with manually adjusting this pitch adjusted tone to match the original tone they first heard at the beginning of the trial. Up and down arrow keys were used in order to adjust the tone to a higher or lower pitch, respectively. Each time the subjects pressed on one of the arrow keys, subjects could hear the adjusted tone being played. This pitch adjustment task is intended to serve as a control variable which aids in indicating whether subjects recalled the tone played at the beginning of the study during the acoustic imagery period.

Analysis Plan

Trials in which there was too much noise in the data were dropped. Noise was identified as excessive eye blinks or the clenching of facial muscles during the auditory perception period.

Two electrodes on the EEG cap were placed underneath each subjects' eyes in order to account for eye blinks. These two electrodes were excluded from data analysis.

Behavioral Analysis

Correct responses for the pitch adjustment task were determined by subjects' abilities to adjust the pitch adjusted tone to exactly match the pure tone heard at the beginning of each trial. Each correct response was given a score of one. The sum of correct responses across the first four available blocks for each subject was taken in order to determine a subject's overall performance on the pitch adjustment task. Only the first four available blocks were scored for every subject because all subjects completed at least four blocks, but not every subject completed all five blocks.

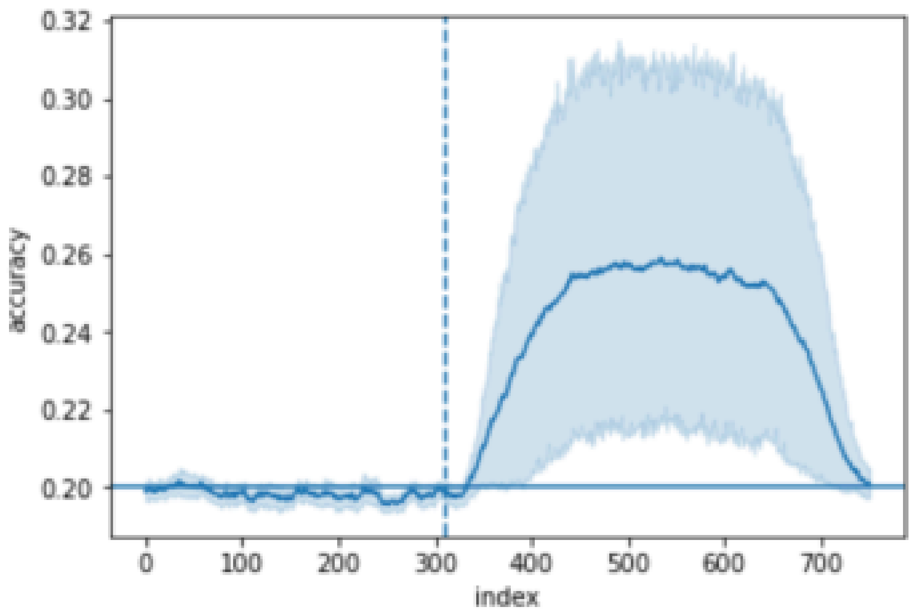
EEG Analysis

Decoding. The presence of both the tones heard in real time and the imagined tones were unable to be identified during the EEG decoding process. All electrodes, aside from the two used to filter out eye blinks, were used in this analysis. Power at each frequency was examined, and the data was examined continuously without averaging. After the shape of the raw data was established [trials x channels x time points], a fast fourier transformation was used to compute the power spectrum. After this, the transformed data [epochs x channels x frequencies x time windows] was fed through the decoder with a classifier that examines the length vector, it was found that the classifier could not identify the perceived tones and was thus not utilized to decode imagined tones.

However, this decoder was previously able to identify tones heard in real time in a previous experiment. A logistic regression classifier that utilized a stratified five-fold cross validation aided in the ability to successfully decode a tone played in real time. This experiment was then repeated with an additional sequence task in which participants had to indicate the number of times a specific tone was played in a sequence of multiple tones. In this instance, the logistic regression classifier was able to decode the identity of the target tone, although it did not produce a significant difference in accuracy for the decoding of target tones in comparison to non-target tones.

Figure 1

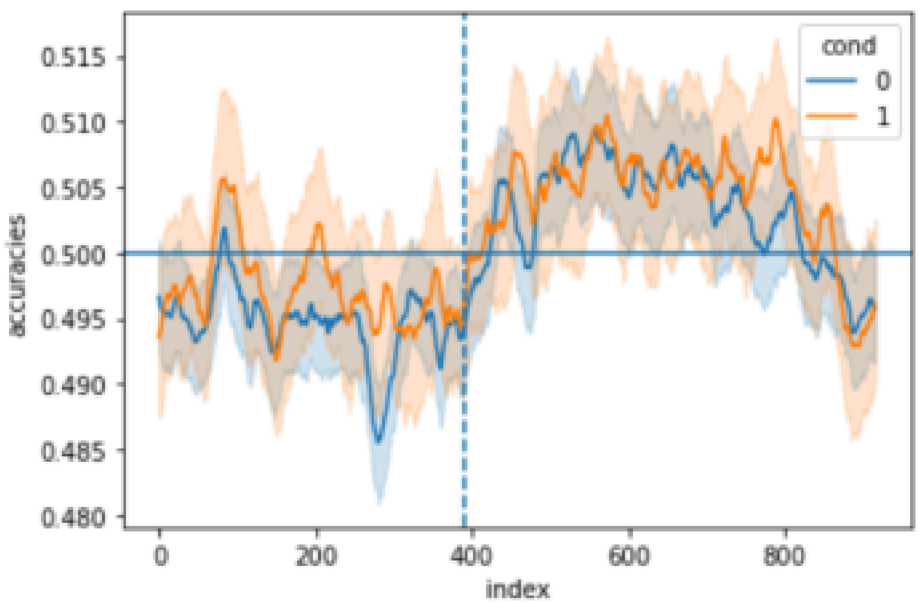
Example of Successfully Decoded Tone



Note. Tone played in real time that has been decoded with a logistic regression classifier. The vertical dashed line indicates the onset of the tone whereas the horizontal line indicates decoding at chance (20%). The shaded area surrounding the arched blue line, which is the tone, indicates the standard error of the mean.

Figure 2

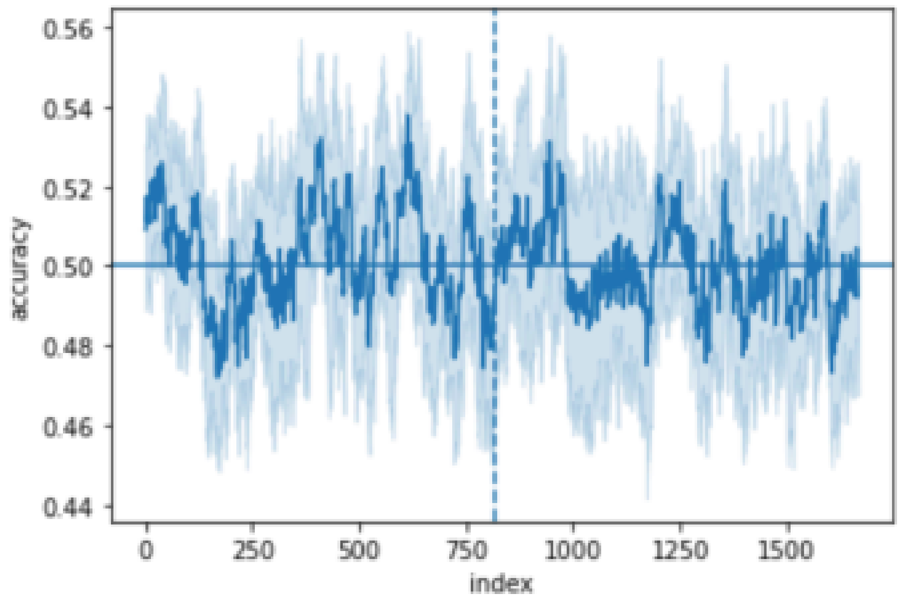
Decoding of Target Tones and Non-target Tones



Note. No difference in decoding accuracy for target tones versus non-target tones. The vertical dashed line indicates the onset of the tone whereas the horizontal line indicates decoding at chance (50%). The shaded area surrounding the conditions indicates the standard error of the mean.

Figure 3

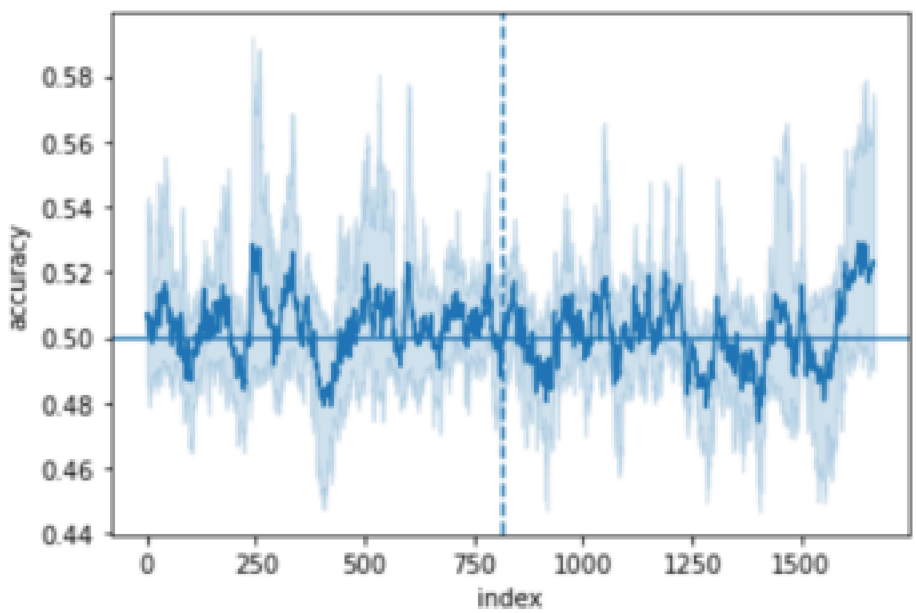
Inability to Decode Perceived Tones in Current Study



Note. Tone identity is unidentifiable for tones heard in the present study. Tone onset is marked by the vertical dashed line. Decoding of the perceived tone did not reach above chance (50%).

Figure 4

Inability to Decode Mental Image of Tones in Current Study



Note. Tone identity is unidentifiable for mentally reconstructed tones in the present study. Tone onset is marked by the vertical dashed line. Decoding of the perceived tone did not reach above chance (50%). Due to the classifier being unable to classify perceived tones, it was not applied to the decoding process for the imagined tones.

Data Analysis. A series of paired samples t-tests were performed with the averages of the EEG signals produced upon listening to and forming mental images of both tones. The averages of the EEG signals for each of the conditions previously mentioned were recorded for each subject and placed into four columns respectively. T-tests were performed between the columns of data. A paired samples t-test was performed between the perception of 190 Hertz and the perception of 280 Hertz. Another paired samples t-test was performed between the mental reconstruction of 190 Hertz and the mental reconstruction of 280 Hertz. Another paired samples t-test was performed between the perception of 190 Hertz and the mental reconstruction of 190 Hertz, and finally a paired samples t-test was performed between the perception of 280 Hertz and the mental reconstruction of 280 Hertz.

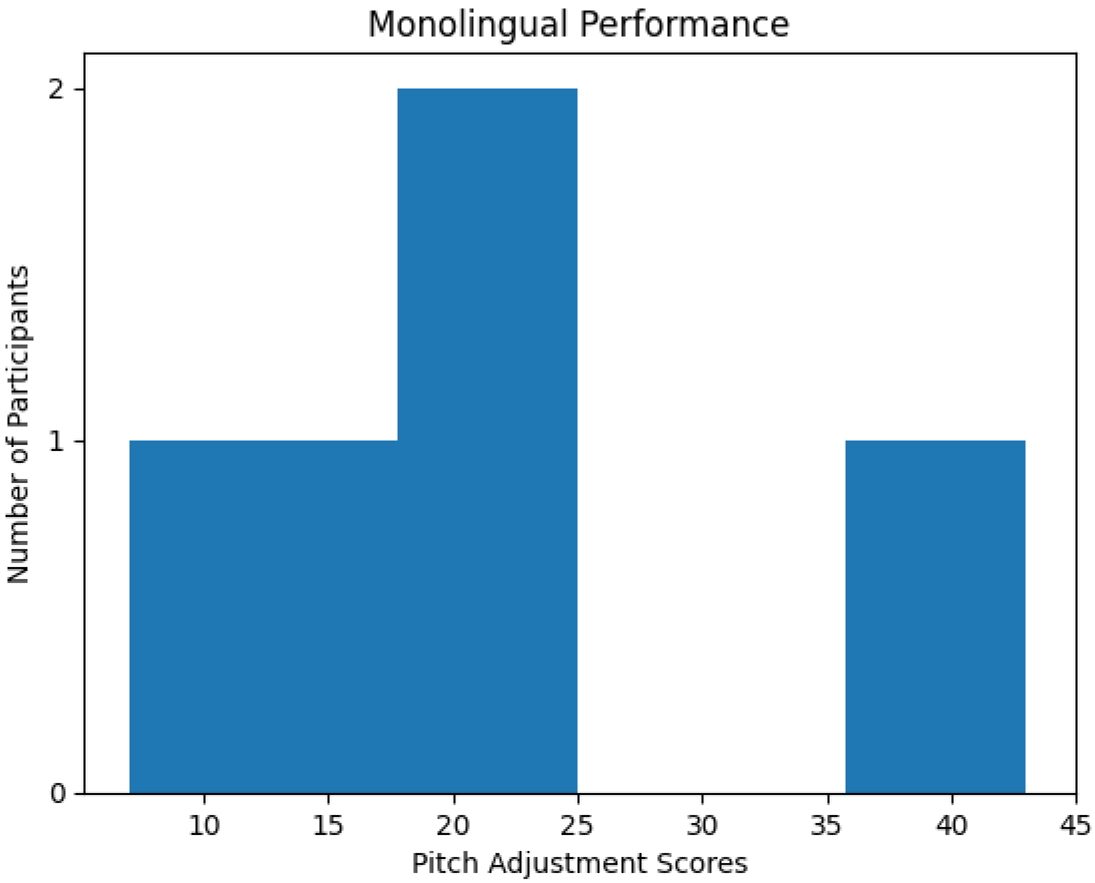
Results

Behavioral Data

An independent t-test between bilingualism and performance on the pitch adjustment task produced statistically insignificant results, with bilinguals (M = 18.70, SD = 6.29) exhibiting differences in performance from monolinguals (M = 22.44, SD = 11.00); t(19) = (-0.872, p = .395). An independent t-test between prior musical experience and performance on the pitch adjustment task also produced statistically insignificant results, with those with prior musical experience (M = 21.07, SD = 7.58) exhibiting differences in performance from those without prior musical experience (M = 18.25, SD = 12.87); t(19) = (0.528, p = .604). A linear regression was conducted to determine the influence of the number of musical instruments played by an individual on overall performance on the pitch adjustment task. This resulted in a significant regression equation (F (1, 15) = 2.00, $p < .005$), with an R-squared value of 0.118. A linear regression was also conducted to determine the influence of years of musical experience on overall performance on the pitch adjustment task. This resulted in a significant regression equation (F (1, 15) = 0.39, $p < .005$), with an R-squared value of 0.026. Two histograms were created in order to convey the overall distribution of monolingual (*N* = 9) and bilingual (*N* = 10) performance on the pitch adjustment task, respectively.

Figure 5

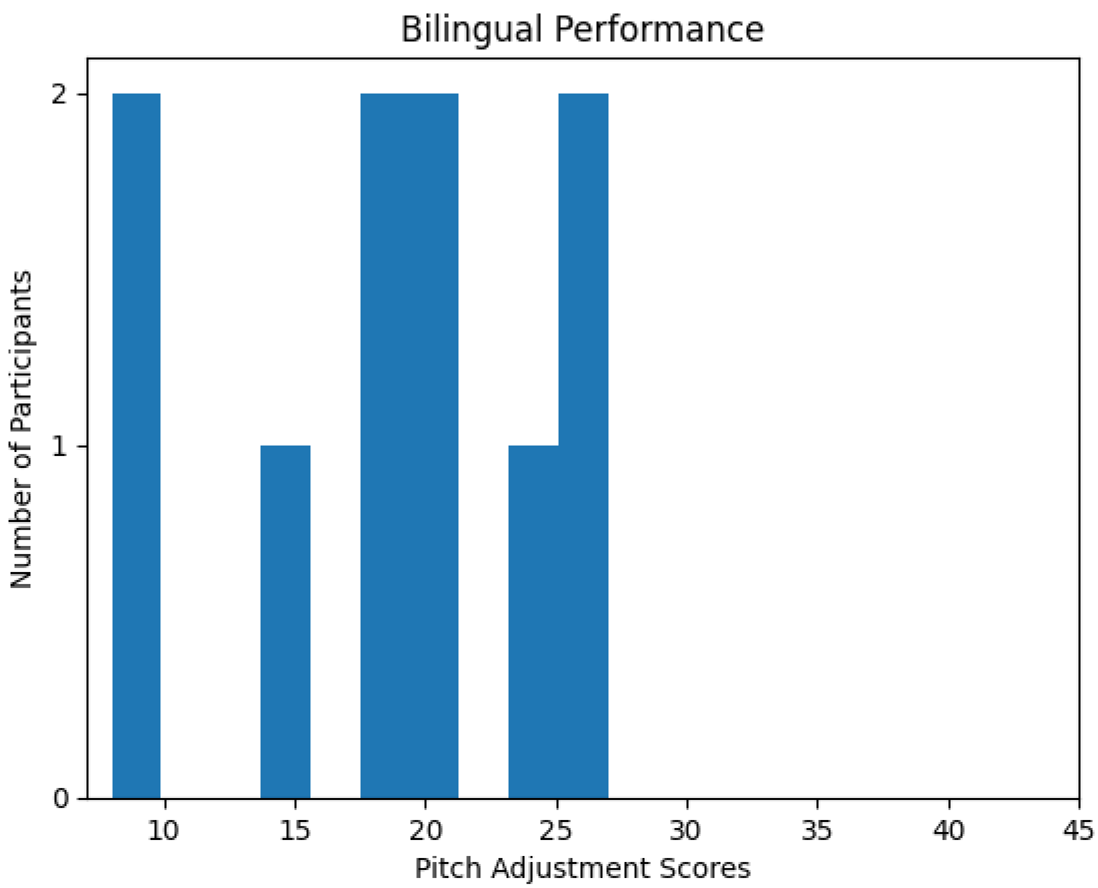
A Histogram of Pitch Adjustment Task Performance for Monolinguals



Note. Distribution of performance on the pitch adjustment task for monolingual subjects ($N = 9$).

Figure 6

A Histogram of Pitch Adjustment Task Performance for Bilinguals



Note. Distribution of performance on the pitch adjustment task for bilingual subjects (N = 10).

EEG Data Analysis

A paired samples t-test was performed between the perception of 190 Hertz and the perception of 280 Hertz, t(18) = (9.406, p = 3.772e-08). A paired samples t-test was also performed between the imagining of 190 Hertz and the imagining of 280 Hertz, t(18) = (9.578, p = 2.903e-08). A paired samples t-test between the perception of 190 Hertz and the imagining of

190 Hertz, $t(18) = (1.505, p = 0.151)$ resulted in no significant difference. A paired samples t-test between the perception of 280 Hertz and the imagining of 280 Hertz, $t(18) = (0.290, p = 0.775)$ resulted in no significant difference.

Discussion

The lack of significant difference in performance between bilinguals and monolinguals conflicts with prior research by Krizman et al. (2012) that suggests that bilinguals possess an advantage when perceiving auditory stimuli. This prior research pertaining to the encoding of speech syllables indicates that fundamental frequency is encoded significantly better by bilinguals than monolinguals. This finding in conjunction with bilinguals exhibiting increased selective attention during a listening task with multiple sound streams, prompts the authors to suggest that bilinguals exhibit enhanced auditory processes. However, findings in the current study suggest that bilinguals do not exhibit an advantage over monolinguals in regards to pitch perception. This difference may be due to how often bilinguals are exposed to two sets of phonology in comparison to pure tones, with the former having more opportunities to experience encoding and recall.

A lack of significant difference in performance between individuals with prior musical experience and individuals without prior musical experience is relatively unsurprising given that the number of subjects with prior musical experience greatly outnumbered the number of subjects without prior musical experience. Only four subjects lacked any prior musical experience, thus making the current sample unfit for comparison in this regard. The sample utilized for this study is not ideal for determining the long term effects of musical experience on auditory task performance. However, considering that musicians do utilize mental representations of sounds to their advantage (Zatorre & Halpern, 2005), with a more representative sample this area of research could be reexamined.

As for the influences of musical experience, the number of musical instruments an individual played was found to be somewhat influential on overall performance. Musical notes sound different on different instruments and thus having prior experience with different sounds at the same pitch may benefit individuals when performing a pitch adjustment task that involves pure tones. However, this could also depend on how long an individual takes part in learning to play two different instruments within the same span of time. It is unclear whether subjects who played more than one instrument simply ceased playing one instrument before learning the other, or if they learned to play both instruments within the same span of time. Prior studies have indicated that extensive musical experience aids in perceiving auditory stimuli in adverse conditions (Anaya et al., 2016). The years of musical experience an individual had was found to be significant but this did not account for any major discrepancies in overall performance. It is possible that some subjects may have played a musical instrument over a long span of time, but how frequently they played the musical instrument within the span of time varied. Essentially, individuals with prior musical experience may have variation in how frequently they practice playing their instruments within the same span of time. Tierney et al. (2008) have conducted research on the influence of musical experience on auditory task performance with results that suggest that trained musicians utilize working memory to their advantage when performing these tasks. These musicians outperformed other subjects on the auditory sequence task provided. It is important to note that these musicians practice the piano for an average of 24.18 hours a week. Subjects in the current study were not asked to report the average number of hours they practiced a musical instrument a week, although this factor may be influential in adjusting pitch.

It should also be said that issues with the decoding process do not negate the significance of prior research involving the overlap between auditory mental images and auditory perception. Halpern et al. (2004) utilized neuroimaging in order to explore areas of potential overlap between imagining musical timbres and perceiving musical timbres. Their results indicate that there is sufficient activity in the auditory cortices during both the perception of these auditory stimuli and the imagining of them. Vlek et al. (2011) investigated the overlap between auditory perception and imagination further and uncovered similarities between the brain signatures of imagined beats and perceived beats through the usage of EEG analysis.

As for the current study, the significant difference between the power of each signal during the perception of each tone suggests that different frequencies emit different electrophysiological responses. The significant difference between the power of each signal during the mental reconstruction of each tone suggests that mental images of auditory stimuli emit different electrophysiological responses depending on the stimuli which is being recalled. The lack of a significant difference between the power of each signal during the mental reconstruction and perception of the same pure tone in both conditions suggests that there is a failure to reject the null hypothesis.

Future Directions

The utilization of a larger sample that consists of equatable sections of individuals who lack musical experience, those with intermediate experience, and those who have extensive musical experience would offer further insight on the potential connections between musical experience and auditory mental images. Alongside this, having subjects report their perceived performance may aid in assessing the influence of personal motivation on performance of the pitch adjustment task.

In regards to considering the cognitive functions involved in the experiment, including more pure tones while maintaining the current number of trials would account for selective adaptability. In a study with two groups of subjects, implementing a series of auditory working memory training intervals in one randomized group of subjects while presenting both the training group and the control group with the same mental imagery and pitch perception tasks could offer further insight on how auditory working memory can be improved. Doing so would also assist in determining whether or not training would strengthen the similarities in power between the signals emitted when hearing and mentally reconstructing a tone.

References

Anaya, E. M., et al. (2016). Long-term musical experience and auditory and visual perceptual abilities under adverse conditions. *The Journal of the Acoustical Society of America, 140*(3), 2074–2081. https://doi.org/10.1121/1.4962628

Brogaard, B., & Gatzia, D. E. (2017). Unconscious imagination and the mental imagery debate. *Frontiers in Psychology, 8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00799

Halpern, A. R., et al. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia, 42*(9), 1281–1292. https://doi.org/10.1016/j.neuropsychologia.2003.12.017

Heald, S. L., & Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Frontiers in systems neuroscience, 8*, 35. https://doi.org/10.3389/fnsys.2014.00035

Ingvalson, E. M., et al. (2015). Working memory training to improve speech perception in noise across languages. *The Journal of the Acoustical Society of America, 137*(6), 3477–3486. https://doi.org/10.1121/1.4921601

Krizman, J., et al. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences, 109*(20), 7877–7881. https://doi.org/10.1073/pnas.1201575109

McGuire, P. K., et al. (1996). Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine, 26*(1), 29–38. https://doi.org/10.1017/S0033291700033699

Moro, V., et al. (2008). Selective deficit of mental visual imagery with intact primary visual cortex and visual perception. *Cortex, 44*(2), 109–118. https://doi.org/10.1016/j.cortex.2006.06.004

Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience, 20*(10), 624–634. https://doi.org/10.1038/s41583-019-0202-9

Pelaprat, E., & Cole, M. (2011). "Minding the gap": Imagination, creativity and human cognition. *Integrative Psychological and Behavioral Science, 45*(4), 397–418. https://doi.org/10.1007/s12124-011-9176-5

Perrachione, T. K., et al. (2013). Evidence for shared cognitive processing of pitch in music and language. *PLOS ONE, 8*(8), e73372. https://doi.org/10.1371/journal.pone.0073372

Phillips, D. P., et al. (2002). Central auditory onset responses, and temporal asymmetries in auditory perception. *Hearing Research, 167*(1), 192–205. https://doi.org/10.1016/S0378-5955(02)00393-3

Tian, X., et al. (2016). Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex; a journal devoted to the study of the nervous system and behavior, 77*, 1–12. https://doi.org/10.1016/j.cortex.2016.01.002

Tierney, A. T., et al. (2008). Effects of early musical experience on auditory sequence memory. *Empirical Musicology Review : EMR, 3*(4), 178–186. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3050509/

Vlek, R. J., et al. (2011). Shared mechanisms in perception and imagery of auditory accents. *Clinical Neurophysiology, 122*(8), 1526–1532. https://doi.org/10.1016/j.clinph.2011.01.042

Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron, 47*(1), 9–12. https://doi.org/10.1016/j.neuron.2005.06.013