Abstract

Music has been of interest in science for many years. In the past ten years, evolutionary biologists have hypothesized that music and language descended from a common predecessor in the past, a protolanguage. At this same time, developments in technology have made it possible to analyze the activation of different parts of the brain. In this literature review, we review the support for the idea of a protolanguage in evolutionary biology and the neuroscience studies that show that this relationship may exist between language and music. However, we also review a study that contradicts this relationship. The present study hopes to address this contraction by exploring the neural processing difference between spoken and sung language. This information will shed additional insight into this relationship and on this contradiction.

Background

Music is a complex subject that captivates nearly everyone in some form. Music, like language, is found in all cultures. This ubiquity piques the interest of evolutionary biologists, psychologists, musicologists, neuroscientists, and scholars in a variety of other disciplines. Each discipline has made its own theories and attempts to understand music and its origins as well as its relationship with spoken language. Each has approached this topic separately, but in combining their perspectives we can gain a greater understanding about the relationship between music and spoken language. Based on some of the theories developed by scholars who study the evolutionary origins of music, there remains an underlying connection between music and language. One of the empirical methods for researching this topic is through neuroscience, where we can record brain activity and look at the groupings of neuron activation to determine the mental similarities between music and language, allowing us to get insight into the evolution of music and language.

Music is not a new topic in the neuroscience world, but taking a more evolutionary approach to understand it is new. What is also interesting is that nearly all music studies performed by neuroscientists use piano or violin music and ignore the wide variety of other musical instruments, including the human voice. Human singing is a promising field as it combines both speech and music and blurs the line. Learning more about how the brain creates, processes, and responds to singing could provide valuable insight into the evolutionary relationship between music and language. The approach of this paper is to focus on the differences between singing and speaking.

In an effort to study this relationship, it is important to take a step back and look at other theories that support the existence of this connection between song and language and how they influence each other as tools in the greater human communicative toolkit. This groundwork is further strengthened by looking at the inherent musicality of the communication between mother and infant. Brown termed this communication between infant and mother a 'musilanguage' (Brown 2000). From this foundation of a protolanguage we can see that a study on the differences between spoken language and song can be beneficial to better understanding the relationship ship between the spoken and sung word.

In order to ascertain remnants of a protolanguage, we need to analyze how our brain processes music and language. Modern technology allows us to measure brain activity through a variety of different methods. Two common methods are positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). They use different methods to track oxygen usage in the brain, which signals brain activity. Using these tools, experiments can measure brain activity in response to different stimuli, providing us with a way to see how our brain processes music.

Angulo-Perkins and al. performed an fMRI study comparing music and language. They grouped human vocalizations and music together and compared them to non-vocal sounds. They found that the anterior portion of the superior temporal gyrus (STG) was more active in responding to music and human vocalizations than to non-vocal sounds. A similar but narrower response was found in the anterior portion of the STG when comparing human vocalizations and music to monkey vocalizations. They key analysis was the difference between music and human vocalizations. They found that the planum polare, a part of the somatosensory cortex, responded significantly more to music than to human vocalizations. There were other parts of the brain, mostly in the STG that responded more to human vocalizations than speech (Angulo-Perkins et al., 2014). The overlap in the general area of the STG seems to suggest the idea of a protolanguage because the bifurcation of the protolanguage would mean that similar parts of the brain should be active while processing music and language. Part of this defense is lost as parts of the STG (Brodmann areas 41 and 42) are responsible for auditory processing. The planum polare is an example of one area that is not tied to basic auditory processing, but that shows significant activation in response to music. This study is limited as the music stimuli were limited to piano, synthetic piano, and violin. These instruments cover a wide range of musical literature but do not shed light on the possible connections between language and song.

Angulo-Perkins's study (2014) provides a great exploratory study but leaves several questions about the neural overlap and the processing differences along the spectrum of language and music. The neural overlap is difficult to test due to the resolution of modern imaging techniques. In Angulo-Perkins's study, the resolution of the fMRI scans (1-millimeter cubes) shows the general trends of activation but does not provide information on which neural networks were active. This is a larger issue in the common regions where it would provide insight into the protolanguage connection between music and language. The one-millimeter cube is huge compared to the size of individual neurons. Isabelle Peretz, a co-author on the paper, came up with an adaptation to fMRI studies that can show us if the same neural networks were used through the same fMRI machines. The activation of neurons has a nonlinear relationship in activity. The blood oxygen-level-dependent (BOLD) signal decreases when the same neurons are used in processing a signal whereas a larger BOLD signal will be observed when different neurons are active. By combining the stimuli in pairs and comparing the differences in activation it can be experimentally determined if the same or different neural networks are used (Peretz, 2015).

In an fMRI adaptation study by Jorge Armony, he found that the same neural networks that process acoustic signals are active in both language and music, but there are distinct 'music-preferring' and 'voice preferring' neurons in the STG (Armony et al., 2015). The acoustical overlap is intuitive as it is efficient to process auditory signals through the same networks, but the distinct music and voice neurons poses an interesting question about the continuum of brain activation between music and language.

One way to study this continuum is to start at speech and slowly work our way towards music. Carolina Mendez Orellana performed a study on the difference between speech and melodically intoned language (varying pitch in a song-like-manner). This was not an adaptation study, but as Angulo-Perkins showed, there are distinctive regions between music and language that can provide insight on the difference. In their study they found that STG and the planum polare were more active in processing the melodically intoned language than language by itself. They also noted a higher activation in the left hemisphere than the right (Orellana et al., 2014). A shift occurred to activate the musical segments of the brain just through melodically intoned language. This was further supported by a case study on two patients with bilateral lesions of the superior temporal cortex who could process language and rhythmic content, but not melodic context (Peretz et al., 1994).

Adding pitch to speech can create melodically intoned language, but other expressions contain pitch variations that are not melodically intoned. Robert Zatorre (2012) studied the pitch differences in music and speech. He came to the conclusion that we have two separate mechanisms for pitch processing. The pitch in regular language is processed in a coarse manner to catch general trends and intonations. For music, however, the brain uses a different method to process the music in a more refined manner (Zatorre and Baum, 2012). This idea is supported by several case studies that were performed on subjects that had bilateral lesions of their superior temporal cortex. They retained their ability to process rhythmic sounds and language, but they were unable to process melodic content in language or music (Peretz et al., 1994).

Besson (1998) performed a study on the electrical activity in the brain and the harmonic and semantic processing of musical phrases. Using an electroencephalogram (EEG) to measure brain activity, Besson asked the subject if the language and music were congruous. Using recordings of an opera singer singing phrases with and without harmonic and semantic incongruities, he was able to show that harmonic and semantic processing of musical phrases are processed independently, even when they are thoroughly intertwined (M. Besson et al., 1998). This study focused on western classical music as the errors were based on western music patterns. This difference in music processing appears to be culturally specific, but more research is needed to begin to ascertain the role of cultural differences in processing music and language.

The final study on the continuum explored the difference in neural networks in language, music and song perception. Daniele Schon (2010) conducted an fMRI experiment on the difference between speech, song, and vocalists. This was accomplished by having 10 non-musician subjects listen to different words that were spoken or sung and then discern if the words were the same or different, or if the melodies were the same or different. Shon also found activation in similar regions of the superior temporal gyrus bilateral, though more intense in the right hemisphere for song versus speech. The comparison of song versus vocalise (singing without words) had a stronger contrast than song versus speech. The superior temporal gyrus was more activated in song than vocalise. These networks were all active, but to different extents depending upon the stimuli (Schon et al., 2010). Schon makes the claim that the linguistic and musical portions of these stimuli were not processed independently, which contradicts what Besson, Peretz, Orellana and many other scholars have claimed through their research. While there is strong evidence for the integration of music and language in neural processing—which might suggest a protolanguage as an evolutionary antecedent to modern music and language—there are separate networks that process the different parts of the signal with overlap between the different ends of the spectrum.

To study this further, the current exploratory study will be performed to compare the brain activation in processing both spoken and sung texts. A religious group provides a setting to study this difference due to the shared familiarity with spoken and sung texts within the religious community.

Proposal

To address this conflict in the current literature, this study focuses on students from the LDS community at Brigham Young University because they share a strong musical tradition. Lyrics and tunes from the LDS hymnal are well known, and students are accustomed to hearing the lyrics sung during the weekly worship service and spoken in the context of religious talks. In addition to their shared knowledge of religious music, students at Brigham Young University also share a broader American musical tradition. By comparing the difference between reactions to sung and spoken lyrics in religious and secular tunes, we will see if there is a difference in the way religious and secular tunes are processed. Thus, the student body at Brigham Young University provides a unique opportunity to work with religious students who have a familiarity with music and spoken texts. The hymn *I Need Thee Every Hour* and the song *Oh, Shenandoah* will be used for their musical similarity. fMRI scans will be used to measure the brain activity as the subjects listen to the stimuli. A statistical analysis will compare the activation of different parts of the brain between stimuli across all subjects while accounting for random error. This is a descriptive study and qualitative comparisons will be made across subjects.

Methodology

As this is a study on humans, IRB approval was required. The required forms and following procedures were approved by the BYU IRB Office, IRB#: X17187.

Students from the advanced BYU Choirs will be paid to record the lyrics and singing of two verses of each song. Students will be recruited through the information screens in the JFSB as well as through the IHUM 201, 202 and MUSIC 201, 202 classes.

The subject population will have no professional musical training and limited non-professional music training. Angulo-Perkins found that the neurological response between music and language is drastically different between musicians and non-musicians. To ensure a consistent comparison, non-musicians will be used to prevent the differences from affecting the results of the study. We also exclude subjects who are self-defined 'tone deaf' to ensure the ability to process music. The musical experience questionnaire has specific questions to determine if the subject meets these requirements. In addition, participants will have to complete the BYU MRIF Screening Form and be eligible under the current criteria for participating in an MRI study within the BYU MRI research facility (https://mri.byu.edu/MRI_Training/BYU%20MRIRF%20Screening%20Form%20v19.pdf). Failure to meet either of these screening activities will exclude the participant from further participation in the study.

Participants will arrive at the MRI Research Facility at the McDonald building and meet with a researcher to go over the consent form as well as the screening form to make sure that they are MRI compatible. The procedures will again be explained to the participant and asked if there are any concerns or questions. The Participant will be asked to verbally state a willingness to continue participation. The participant will then be placed into the MRI for data collection. Standard MRI procedures, including safety procedures, and protocols as established and approved by the BYU MRI Research Facility will be strictly followed. All MRI personnel will be current in their certifications.

Once the participant is in the MRI machine, scans will be done to create a structural image of their brain and the fMRI protocol initiated. The participant will listen to the spoken and sung words in a random order of "I Need Thee Every Hour" and "O Shenandoah" recorded by a female singer. Subjects will be asked to press a button corresponding between verses of the pieces to verify their attention to the stimuli. Stimuli will be presented to participants in a mixed design. This process should take about one hour and 15 minutes (approximately 50 min. for task, 8 min. for structural scanning, 15 min. to situate participant in the scanner and remove them from the scanner at the end).

Project Timetable

The IRB application was submitted in April 2017. Approval was granted in December 2017. With approval, questionnaires were sent out to collect participants. The first trial fMRI scan will be March 8th, 2018. Upon a successful first data collection, the remainder of scans will be performed in March and early April. The goal is to finish data collection and submit an abstract by April 20th, for a poster at the International Network for Neuroaesthetics conference in May. The data will then be processed and the final report and paper written for the defense and publication. While we have a journal in mind, we plan to talk to the participants at the conference to see if there are journals better suited for this research.

Funding

Funding is coming from the Dr. Lawson's research funding and an ORCA grant. MRI scans cost $200 per hour and we are hoping to use 15-20 participants. On top of the facility cost, each participant will need to be reimbursed for their time. $1000 of Honors funding is requested to pay for an additional 5 hours of MRI time. This should allow us to expand the study by 5-10 students. This will also reduce the impact of this study on Dr. Lawson's research funding.

Conclusion

The research will contribute to the ongoing research on the theory of a protolanguage in the disciplines of musicology and evolutionary biology by showing whether there are similarities or differences in the ways song and speech are processed

REFERENCES

Angulo-Perkins A, Aube W, Peretz I, Barrios FA, Armony JL,Concha L (2014) Music listening engages specific cortical regions within the temporal lobes: Differences between musicians and non-musicians. Cortex 59:126-137.

Armony JL, Aube W, Angulo-Perkins A, Peretz I,Concha L (2015) The specificity of neural responses to music and their relation to voice processing: An fMRI-adaptation study. Neurosci Lett 593:35-39.

Aubé W, Angulo-Perkins A, Peretz I, Concha L,Armony JL (2015) Fear across the senses: brain responses to music, vocalizations and facial expressions. Social cognitive and affective neuroscience 10:399-407.

Brown S, (2000) The 'Musilanguage' model of language evolution. In: The Origins of Music, (Wallin NL ed), pp 271-300. Cambridge, Mass. [u.a.]: MIT Press.

Cross I, (2003) Music and Biocultural Evolution. In: The Cultural Study of Music, (Clayton M, Herbert T, Middleton R eds), pp 19-30. New York: Routledge.

Cross I, Morley I, (2009) The evolution of music: Theories, definitions and the nature of the evidence. In: Communicative Musicality, Vol. 1 (Malloch S, Trevarthen C eds), pp 60-81.

Darwin C, (1897) The descent of man. New York: Appleton.

Dissanayake E, (2009) Root, leaf, blossom, or bole: Concerning the origin and adaptive function of music. In: Communicative musicality : Exploring the basis of human companionship, (Malloch S, Trevarthen C eds),England:

Huron D, (2006) Sweet Anticipation. US: Mit Press.

M. Besson, F. Faïta, I. Peretz, A.-M. Bonnel,J. Requin (1998) Singing in the Brain: Independence of Lyrics and Tunes. Psychological Science 9:494-498.

Malloch S, Trevarthen C, (2009) Communicative musicality. Oxford [u.a.]: Oxford Univ. Press.

Miller GF, (2000) The mating mind : how sexual choice shaped the evolution of human nature. United States:

Mithen S, (2005) The singing Neanderthals: the origins of music, language, mind and body. London: Weidenfeld & Nicolson.

Orellana CPM, van de Sandt-Koenderman, Mieke E, Saliasi E, van der Meulen I, Klip S, van der Lugt A,Smits M (2014) Insight into the neurophysiological processes of melodically intoned language with functional MRI. Brain and Behavior 4:615-625.

Peretz I, Kolinsky R, Tramo M, Labrecque R, Hublet C, Demeurisse G,Belleville S (1994) Functional dissociations following bilateral lesions of auditory cortex. Brain 117 (Pt 6):1283-1301.

Schon D, Gordon R, Campagne A, Magne C, Astesano C, Anton J,Besson M (2010) Similar cerebral networks in language, music and song perception. Neuroimage 51:450-461.

Sidtis JJ, Strother SC, Anderson JR,Rottenberg DA (1999) Are Brain Functions Really Additive? Neuroimage 9:490-496.

Zatorre RJ,Baum SR (2012) Musical Melody and Speech Intonation: Singing a Different Tune. PLoS Biology 10:e1001372.