5-8-2019

Silent Singing: Investigating visual perceptual narrowing of rhythm from a developmental perspective

Gillian Sheehan
Bridgewater State University

Silent Singing: Investigating visual perceptual narrowing of rhythm from a developmental perspective

Gillian Sheehan

Submitted in Partial Completion of the
Requirements for Commonwealth Honors in Psychology

Bridgewater State University

May 8, 2019

Dr. Melissa Brandon, Thesis Advisor
Dr. John Calicchia, Committee Member
Dr. Joseph Schwab, Committee Member

Abstract

In the field of rhythm perception, research has focused on auditory-perceptual-narrowing with little focus on visual-perceptual-narrowing. The purpose of this study is to determine if the visual system narrows similarly to the auditory system, and if one sense is better at detecting rhythmic differences across cultural boundaries, e.g. Western vs. Non-Western rhythms. Using a within-subjects design participants watched videos of a woman singing the same or different Western and Balkan rhythms. The vocals were removed from the videos, leaving just the movement of the woman's mouth. Participants watched two videos in a row containing either Western or Balkan rhythms and determined if the videos were the same or different. The behavioral data (N=24) demonstrates perceptual narrowing to the culturally familiar rhythms with a significant difference in accuracy across Balkan (65.63%) versus Western (74.48%) visual rhythms. We examined the looking patterns across the singer's face with eye tracking during the Western versus Balkan rhythms. Eye tracking data (N=24) demonstrates a significant main effect of Dwell Time on Area of Interest. The participants looked at the mouth the majority of the time, regardless of the rhythmic condition. This indicates that visual attention alone does not lead to the perception of the visual rhythms; it is more important *how* the visual information is processed through perception. A sample of preterm adults (n=7) were tested for possible developmental delays in narrowing, but their data was not significantly different from the full-term adults, thus we collapsed their data into the full sample. This study supports the hypothesis that perceptual narrowing is occurring for visual rhythms with the behavioral data demonstrating difference in perception across the cultural rhythms despite no difference in visual attention to the rhythms.

Silent Singing: Investigating visual perceptual narrowing of rhythm from a developmental perspective

When you have a conversation with another person or go to a concert you are experiencing rhythm using both your auditory and visual senses. Rhythm can be experienced through both senses either in isolation or as a combined sensory experience. For example, when listening to music on the radio, CDs, or a streaming service, music will only be experienced through the auditory sense. However, if someone is at a live music performance or watching a music video, then the music is experienced through both the auditory and the visual senses. The two senses combine into one cohesive sensory experience of rhythm, among other shared sensory experiences.

When experiencing music many people will begin to clap along or move their bodies in time with the music. This is how many people respond to rhythm through clapping, snapping, or simply nodding their head in time with the beat. A beat is the main pulse in music; it occurs with a regular period and arises from the rhythmic structure (Honing, 2012). Music from different cultures are composed using specific beat patterns that help differentiate that culture’s music from other cultures. Musicians define the beat pattern by the time signature in musical notation, which indicates how many beats there are in a given segment or measure of music. Western music, typical of North America and Europe, uses simple beat patterns with typical time signatures of 4/4 or 3/4. Whereas, other cultures use more complex beat patterns such as Balkan music which is typically in 5/8, 7/8, or 9/8 time signatures. These more complex time signatures present uneven groupings of notes in the rhythms that make them more complex to perceive. If you consider dancing to music, Western rhythms have regular beats every two to three sounds that you emphasize with your movements. In Balkan music, the more complex time signature means that the beats to move to do not occur as regularly, making it more difficult to have fluid dance movements because the spacing of the beat keeps changing. This complex rhythm would seem very difficult to dance to for someone from a culture that listens to Western styles of music, but if you grew-up with Balkan music, the uneven pattern is what you expect to hear. This is similar to how individuals understand and become accustomed to their native language and have difficulties hearing all the sounds in a foreign language. Overall, Western rhythmic systems are less complex than other culture's rhythmic systems, making it harder for Western listeners to perceive the more complex rhythms of other cultures, such as Balkan music (Hannon & Trehub, 2005a). Regardless of its cultural origin, music is a complex sensory experience due to the fact that it combines multiple sensory experiences into one organized and integrated sensory experience across time.

Language, like music, is a complex experience that takes time for a person to learn. Humans are born able to hear every phenome in every language globally, meaning they can hear and discriminate all of the sounds in every language (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Maye, Werker, & Gerken, 2002). The same is true of human's ability to distinguish different rhythmic patterns across different cultures (Bremner, Lewkowicz, Spence, 2012). However across the first year of life, humans become selective to the sounds and rhythms that are present in their environment. This is due to the phenomenon known as perceptual narrowing. Perceptual narrowing is the process in which one's ability to discriminate differences in stimuli becomes selective to the stimuli that they encounter in their environment or culture (Bremner, Lewkowicz, Spence 2012). This also means that stimuli that are not present in their environment/culture are not as easy to distinguish. Across the first 4-months of life, infants begin to create categories based on the phonemes of their native language, and by the end of the first year of life, infants become narrowed to just the phonemes that are present in their native language (Eimas et al., 1971 Maye et al., 2002). Infants also become narrowed to different rhythmic patterns that are present in their environment and have a difficult time distinguishing differences in rhythmic patterns not experienced in their environment (Hannon & Trehub 2005a & b), especially if the rhythms are more complex than their native music systems. Thus, perceptual narrowing not only occurs in speech, but also occurs across musical rhythms.

Similar to the literature on narrowing to one's native language, research has investigated how humans become narrowed to rhythms across cultures. According to Hannon and Trehub (2005a & b), both infants and adults are better able to distinguish auditory rhythmic differences if the rhythms are from their native culture compared to rhythms from a non-native culture. Specifically, these studies only examined auditory perception when testing infants' and adults' abilities to distinguish differences in rhythms. Hannon and Trehub (2005a) studied adults' and 6-month-old infants' abilities to discriminate rhythms by presenting them with structure preserving or structure altering rhythmic patterns. Adults from both North America and Bulgaria or Macedonia participated and were given a 2-minute familiarization to a rhythmic pattern that either had simple (Western) or complex beat structure (Balkan). At test, participants heard music segments that had the same rhythm, a structure preserving rhythm, a structure altering rhythm, or a randomly created different rhythm. The adults indicate if the rhythms preserved or altered the structure by rating how closely the test trials matched the familiarization rhythm. To assess North American 6-month-old infants' ability to distinguish the differences in rhythms, looking times were analyzed. After the familiarization to either a Western or Balkan rhythm, infants were presented with test trials that were either structure preserving or structure altering auditory rhythms paired with a non-rhythmic moving image on a screen, and the mean looking time per trial type was calculated. For familiarization studies, infants typically look longer at unfamiliar or altered stimuli. This study found infants looked longer to the structure altering rhythms for both cultural rhythms. These results are in-line with the Bulgarian and Macedonian adults, who were better able to differentiate more complex rhythmic patterns (Balkan) than the North American adults (Western). Thus, the younger infants were not narrowed to their cultural rhythms unlike the North American adults.

A second study conducted by Hannon and Trehub (2005b) looked at older infants' and adults' abilities to differentiate Western and Non-Western rhythmic patterns, as well as looking to see how music exposure affected this ability across the age groups. Infants were studied at the age of 12-months with no prior exposure to Balkan (Non-Western) rhythms. The methods were the same as those used in the previous study for the infants and adults (Hannon and Trehub 2005a). This study found that the 12-month-olds, without previous exposure to the Balkan music, did not have the ability to discriminate structure preserving from structure altering Balkan rhythms as did the 6-month-olds in the previous study. However, if the 12-month-olds were given a 2-weeks exposer to Balkan music prior to testing, then they were able to distinguish differences in the Non-Western music, unlike the 12-month-olds who had not been exposed to Balkan music. This ability to learn Balkan music so quickly is only seen in infants and not in adults. North American adults were given either one week or two weeks exposure to Balkan music, but they still preformed at or just above chance on distinguishing the structure preserving from the structure altering rhythmic changes in Balkan music (Hannon & Trehub, 2005b).

The results from the Hannon and Trehub (2005 a & b) studies provide clear evidence that Western listeners become perceptually narrowed to their simple rhythmic systems, and this happens by the end of the first year of life. But these studies have only focused on auditory perception for distinguishing the differences in rhythm across music from different cultures. This poses the question, will perceptual narrowing to rhythm occur in vision the same way that it has been previously observed in rhythm and speech? Recall that live music is a multisensory experience. Do the senses narrow to their cultural rhythms together, or does vision narrow to rhythm at all? Our research will seek to answer these questions using a combined behavioral and eye tracking methodology with adults.

The use of eye tracking will provide information about the participants looking patterns during the task, and what visual information they use to process visual rhythm. Overall, there is very little research on visual rhythm perception and processing, so these eye tracking results will contribute to the knowledge base about how we process rhythm visually. However, several studies have examined infants' and adults' gaze patterns while watching audio-visual videos of a speaker using both their native and non-native languages. Through the use of eye tracking software, Lewkowicz and Hansen-Tift (2012) found that there are developmental shifts in attention throughout infancy (4-months – 12-months) when attending to a speaking face. The infants watched an audio-visual video of a woman speaking in either English, the infants' native language, or Spanish, the infants' non-native language, and proportion of looking time at the eyes versus the mouth on the speaker's face was measured. At 4-months of age infants who watch a video of a speaker using their native language (English), focus their gaze more on the speaker's eyes during the video, while between the ages of 4- and 8-months, infants' attention shifts to the speaker's mouth. There is a second attentional shift at the age of 12-months where the infants shift their gaze from the mouth back to the eyes if the speaker is using their native language. However, if the speaker is not using the infant's native language (Spanish), then around 12-months infants focus their gaze more on the speaker's mouth than their eyes. For both native and non-native languages an adult sample looked more at the eyes than the mouth, but the proportion on the eyes was lower for the non-native language (Lewkowicz & Hansen-Tift, 2012). Another study found similar attentional patterns in infants who were monolingual and whose native language was either Spanish or Catalan (Pons, Bosch, & Lewkowicz 2015). Infants were shown a video of a woman speaking in their native language and then were shown a video of the same woman speaking in English. The infants' eye movements were tracked using eye tracking software, with areas of interests around the speaker's eyes and mouth. The infants focused on the eyes at 4-months, and at the mouth at 8-months regardless of what language the woman was speaking. However, at the age of 12-months the infants' attention was divided equally between the eyes and the mouth if the woman was speaking in their native language, but the infants would focus on the mouth if the woman was speaking in the non-native language (Pons, Bosch, & Lewkowicz 2015). This indicates that once infants are familiar with the language that a speaker is using they are more likely to focus their attention on the eyes to gain social cues. If a speaker is using a non-native language, then the focus of the listener will be on the mouth to better understand what is being spoken. In adults, previous research has observed that attention to the mouth in a non-native language is only seen when a language processing task is required, and the listener is monolingual. In the absence of a language processing task, adults will look equally between a speaker's eyes and mouth if the speaker is using native or non-native languages (Barenholtz, Mavica, & Lewkowicz 2016).

Across these eye tracking studies on audio-visual speech perception, results suggest that when non-native speech is observed adults shift their visual attention back and forth between the eyes and the mouth, but when *processing* a non-native language, they focus on the mouth. These studies also highlight that there is a long developmental process to get to adult looking patterns.

First, auditory narrowing occurs before visual attention is split between the eyes and mouth during native speech. This split attention in vision, after the auditory sense has narrowed highlights that we do not know if visual narrowing happens and when during development it might occur. Furthermore, language contains both phonemes and prosody, the rhythm of the language, making it difficult to determine from these studies which aspects of spoken language adults have narrowed to. To understand the perceptual narrowing process in vision we need to use visual stimuli alone to see if there is evidence of visual narrowing to culturally familiar rhythms or languages. By studying the senses separately, a more complete picture of the process of perceptual narrowing of rhythm can be established. Therefore, our study will examine visual narrowing of rhythm using a video of a singer with the audio removed. This will allow us to use the same eye tracking methodology as the previous language studies but isolate the stimuli to just the visual rhythm component. Our question then becomes where will adults' eye gaze focus: on the eyes or the mouth? Also, will these gaze patterns differ across native and non-native rhythms? Given the shifts in attentional focus seen in the previous speech studies, we hypothesize that the gaze will be on the mouth for the Balkan rhythms (non-native) and on the eyes for the Western rhythms (native). Another factor that could contribute to the adults' gaze patterns could be whether or not the individual was born preterm.

Research evidence suggests that preterm infants experience delays in there perceptual development, which also impacts their language development. A longitudinal study by Jansson-Verkasalo et al. (2010) compared premature infant's abilities to distinguish native and non-native phonemes to their full-term peers. To complete this assessment both preterm and full-term infants (ages 6-months, 12-months, and 24-months) listened to clips of an individual speaking in either a native language, or a non-native language while EEG measures were recorded. Language assessment surveys were completed by parents about the infants at the ages of 12-months and 24-months. The preterm infants were still able to distinguish both native and non-native phonemes at 12-months-old, while their full-term peers lost this ability between the ages of 6- to 12-months (Jansson-Verkasalo et al., 2010). This narrowing delay is probably related to preterm infants' overall delays in development due to their immature development at birth. Given the delay seen in preterm infants' auditory sense, there will likely be a longer delay in their visual development since vision is the last sense to fully develop during infancy (Bremer, Lewkowicz, Spence, 2012). This poses an interesting developmental question: are there any lasting effects of being born preterm on adults' gaze patterns when identifying visual rhythmic differences? Specifically, do sensory developmental delays that occur early in life for preterm infants impact perceptual narrowing later in life? Therefore, we will also recruit a sample of adults born preterm for our study to help answer these questions.

The current study will examine visual perceptual narrowing, specifically in the field of rhythm perception. The existing literature on perceptual narrowing has focused on auditory narrowing or visual narrowing in other domains such as face processing, but very little is known about how the two senses might narrow together or in relation to each other. The present study will examine adults' ability to visually distinguish differences in rhythms across native (Western) and non-native (Balkan) rhythms that will be presented as videos of a female singer with the audio stream removed. We will test if perceptual narrowing to rhythm occurs in vision just as it has been previously observed for auditory rhythm. Behavioral data will be collected as participants indicate if the sung rhythms in a trial were the same or different. We predict that participants will have a higher rate of accuracy for the Western rhythms than for the Balkan rhythms. We will also look at adult's eye gaze patterns to see if they differ across the native and non-native rhythms, similar to the way that looking patterns in infants and adults were assessed for the audio-visual speech studies (Pons, Bosch, & Lewkowicz 2015; Lewkowicz & Hansen-Tift, 2012). Specifically, we will be looking to see if adults' eye gaze focuses on the mouth for the Balkan rhythm (non-native) and on the eyes for the Western rhythm (native). The final area of focus will be comparing adults born preterm, at or before 36-weeks gestational age, to their full-term peers, given that preterm infants do not develop as quickly as their full-term peers and experience developmental delays (Jansson-Verkasalo et al., 2010). We will examine if there are any lasting effects of being born preterm on adult gaze patterns when identifying visual rhythmic differences.

Method

Participants

Participants were recruited from the Psychology Department participant pool using the Sona system at Bridgewater State University, which allowed participants to receive 1 credit towards course assignments for their participation in this study. Data was collected from 24 participants with self-reported normal vision and hearing. The participants ranged in age from 18 to 24 years with the average age being 19.43 years, and 1 participant not reporting their age. The sample consisted of 4 Males, 20 Females, and 0 who identified as Other. A subset of 7 participants were born prematurely (at or before 36-weeks gestational age) and were specially recruited. These participants either received 1 course credit or $5 in BSU flex dollars for participating. The prematurely born individuals were recruited via flyers posted on campus in highly trafficked areas such as the academic buildings, dining halls, and on-campus residence halls. An additional 9 participants' data were excluded from the study: the first 3 were excluded as pilot subjects, 4 due to calibration issues with the eye tracker, 1 due to eye tracker system malfunction, and 1 due to inattention to the task. All participants signed an informed consent at the start of the study and were able to ask any questions about the procedure.

Stimuli

Each participant watched a total of 32 trials dived into four blocks of eight trials each. Every trial contained 2 separate videos of a female singing either the same or different Western rhythm or Balkan rhythm. Trials contained pairings of videos with Western only rhythms or Balkan only rhythms, and the audio tracks were removed to produce silent videos. These stimuli constitute the Western and Balkan variables. The order of trials was randomized across blocks such that each pairing of the videos within each culture appeared only once per block. Four different rhythmic patterns, 2 Western, and 2 Balkan, were recorded. Each video featured a woman wearing a black shirt with her hair tied back in a ponytail. The background of the video was black, and the woman had an ear piece on her right ear to hear the stimuli she was singing. The woman was centered in the frame and instructed not to move and to remain neutral in her facial expressions (See Figure 1). These instructions were given to minimize the obvious visual differences across videos to prevent that being a strategy to tell the videos apart. There were minor variations in the woman's movements across the four videos, which were accounted for with adjusted interest areas being applied to the videos for the eye tracking data analysis. The first Western rhythm was in a 4/4 time signature and known by the name "Shave & a Haircut." There were 3 repetitions of this rhythm, creating a video that was 12 seconds long. The second Western rhythm was also in a 4/4 time signature and used the same notes as the first rhythm just arranged differently. There were 3 repetitions of this rhythm, creating a video that was 12 seconds long. The first Balkan rhythm was in a 7/8 time signature and had a grouping pattern of 322. There were 4 repetitions of this rhythm, creating a video that was 8 seconds long. The second Balkan rhythm was also in a 7/8 time signature and was a rearrangement of the same notes in the primary Balkan rhythm to form a 233 grouping (See Figure 2 for notation of the 4 rhythms). There were 4 repetitions of this rhythm, creating a video that was 8 seconds long. The videos within the same time signatures were the same length and number of repetitions of the rhythm. The differences in time signature led to differences in the lengths of the rhythms because of tempo differences across the cultural conditions. That is why there were no trials pairing Western and Balkan rhythms together. Within the cultural pairings, the two videos presented on each trial were either the same or different rhythms creating our same versus different variable.

Procedure

The study began by calibrating participants' eyes with the eye tracker. To calibrate a participant the researcher first manually adjusted the infrared camera focus to best detect the pupil and the corneal reflection of the individual participant's eyes. The pupil is the darkest point in the eye and the corneal reflection is the lightest point in the eye. The eye tracker uses these two points to follow the participant's eye gaze around the screen. During the calibration sequence there were 9 white dots with a black center that appeared on a black screen. Participants were instructed to react to the appearance of each dot individually and try not to anticipate where the next dot would be. After the calibration sequence, a validation sequence was administered that had the same set-up as the calibration sequence with a series of 9 dots appearing on the screen. The validation stage reported the maximum error in the calibration defined by two different error numbers, known as the right error and the average. These numbers indicate whether a validation was successful or if there was a high error in the calibration and it needed to be repeated. The error numbers must be at or under 0.5 for the right error and at or below 1.0 for the average. If a participant could not be calibrated successfully after 5 attempts, then they were excluded at this stage because the study would not produce valid data without a successful calibration. As previously stated, 4 participants were excluded at this stage due to inability to calibrate.

Once the participants' eyes were calibrated they began the study by following the instructions on the screen to start each video with a space bar press. Participants were not told what made the videos different; they were just asked if the two videos were the same or different. No further instructions on how to determine the differences were given even if the participants asked. This was done to keep the method consistent with the past studies that had both infant and adult participants. At the end of the second video in each trial the participant answered whether the two videos were the same or different with a keypress. Stickers were placed on the keyboard to mark the response keys. An "S" sticker was placed over the F key to indicate "Same" and a "D" sticker was placed over the J key to indicate "Different." Participants were instructed to keep their index fingers on the labeled keys with their thumbs on the space bar.

Eye tracking software tracked where the participants were looking throughout the videos. For data analysis areas of interest (AOI) were applied on the singer's face and defined for each video (See Figure 1 for an example of the AOI on the singer's face). The areas were the mouth, the eyes, and the entire face. Table 1 lists the total pixel count for each AOI per video. The interest areas were defined by a rectangular interest area over the eyes and another rectangular interest area over the mouth and a contoured interest area that outlined the face. The proportion of looking to each AOI per video was calculated using the dwell time in each AOI (eyes or mouth) divided by the dwell time in the face AOI per video. The mean proportion of looking was calculated across trial types for each participant as the dependent variable.

At the end of the study the participants completed two surveys. One asked about the participants' thoughts on the study and their demographics information, and the other asked about their previous experiences with music, such as: Did you use any strategies to complete the task? Do you play any musical instruments or sing in choirs? Do you listened to music from non-western cultures? Were you born full-term or prematurely? (See Appendixes A-B for copies of the surveys). These surveys were either administered on paper (15 participants) or using a tablet (18 participants, including excluded participants) to collect electronic responses via Qualtrics. The survey method was switched to Qualtrics to save the data digitally which made analysis faster.

Results

Behavioral

An independent samples t-test was conducted to test for difference in Preterm (n = 7) vs. Full-term (n = 17) participant's overall accuracy. No significant difference was found for Preterm vs. Full-term on accuracy (t (11.87) =0.53, p = 0.609, equal variance not assumed). Therefore, we collapsed across this variable, combining Preterm and Full-term accuracy data.

We conducted a paired-samples t-test with Western mean percent accuracy (74.48%) versus Balkan mean percent accuracy (65.63%) and found a significant difference with participants demonstrating higher accuracy for the Western rhythms, (t (23) = 3.02, p = 0.006). This finding indicates that participants were significantly more accurate in their ability to correctly identify differences in the Western rhythms than they were in the Balkan rhythms (See Figure 3).

We also correlated the years of musical training with the accuracy scores across cultures. Mean years of musical training reported was 3.17 years (SD = 4.26 years), ranging from 0-14 years. We found that years of musical training positively correlated with accuracy for the Balkan rhythms, $r = 0.41$, $p = 0.045$. This indicates that as participants years of music training increased, their accuracy in identify the Balkan rhythms increased. This correlation suggest that musical training can reduce or mitigate perceptual narrowing's effects. None of the other self-reported measures resulted in significant correlations.

Eye tracking

For the eye tracking data, we calculated the proportion of looking time to the mouth versus the eyes for each video. To calculate the proportions, we took the total dwell time in the AOI of the mouth and the eyes and then divided each number by the total dwell time on the face per video to obtain the proportion of dwell time per AOI.

A 2x2x2x2x2 mixed factor ANOVA was conducted on the mean dwell time proportions to examine all possible outcomes and interactions. The within-subjects variables were Rhythmic Culture (Balkan vs. Western), Same versus Different trial condition, Accuracy of response (Correct vs. Incorrect), and Area of Interest (Mouth vs. Eyes). The between-subjects factor was the participants' birth timing (Full-term vs. Preterm). The within-subjects main effect of Area of Interest was significant, $F(1, 12) = 6.99$, $p = 0.021$. Overall participants spent more time looking at the mouth than the eyes in any condition. No other within-subjects main effects were significant, $p > 0.05$. The between-subjects main effect of birth timing (Full-term vs. Preterm) was not significant, $F(1, 12) = 0.64$, $p = 0.439$. There was no difference in the gaze patterns for full-term versus preterm adult participants. There was one significant interaction. The variables of Same vs. Different trials by Accuracy by AOI interacted, $F(1, 12) = 5.27$, $p= 0.041$. Given that this interaction did not have the key variable of the experiment involved, Rhythmic Culture, this interaction did not pertain to the research questions at hand. No other interactions were significant, $p > 0.05$. With minimal significant results from this overall ANOVA we proceeded to collapse across variables to examine our specific hypotheses for the eye data.

A 2x2 repeated measures ANOVA was conducted on the mean dwell time proportion data. The within-subjects factors were gaze proportions at the mouth versus the eyes and Rhythmic Culture (Balkan vs. Western). The main effect of Area of Interest was significant in this ANOVA, just as it was in the overall ANOVA, $F(1,23) = 18.04$, $p = 0.00$. The main effect of Rhythmic Culture was not significant, $F(1,23) = 1.04$, $p = 0.319$. The interaction of Area of Interest by Rhythmic Culture was not significant $F(1,23) = 0.49$, $p = 0.492$. These findings indicate that there were significant differences in dwell time per the different AOI, however this significance was not related to the different cultural conditions (See Figure 4). These findings were different than what we hypothesized would occur across the different rhythmic cultures. Even though behaviorally there is a significant difference in accuracy across rhythmic cultures, visually there is no difference across AOI dwell times by rhythmic cultures.

Discussion

This study took a novel approach by combining eye tracking data and a behavioral measure to gain a holistic view of visual perceptual narrowing. Instead of relying solely on behavioral data to record the participants experience of narrowing, we also collected eye tracking data on participants' gaze patterns to have a measure of visual attention. The participants had a specific task to perform, Same/Different judgements, while collecting their eye data to keep the participants focused on the task and using their vision to solve the task. The results of this study produced some interesting findings. First, the behavioral data indicated perceptual narrowing to the culturally familiar rhythms, Western rhythms, with significantly higher accuracy for the Western rhythms than the Balkan rhythms, as was hypothesized. We also discovered that years of music training correlated with higher accuracy for the Balkan rhythms. This could be due to exposure to music from other cultures during musical training, as well as, exposure to music in a variety of time signatures. This correlational effect of musical training further supports that the behavioral data was measuring perceptual narrowing to visual rhythms. The participants with the higher exposure to these unfamiliar time signatures, through musical training, showed less perceptual narrowing by more accurately detecting the changes in the Balkan rhythms compared to the participants with no music training.

While the behavioral data provides evidence of perceptual narrowing to visual rhythms the eye data is a bit more complicated to understand. We did not see the anticipated looking patterns of focusing on the eyes for the Western rhythm and the mouth for the Balkan rhythm. Instead, we found that the participants focused on the mouth for the majority of the videos, regardless of the cultural condition. There are a few possible reasons for the observed gaze patterns. One possible reason is that the visual only singing stimuli were atypical for the participants and they did not view them as they normally would an audio-visual face, such as the ones seen in the speech studies we drew our hypotheses from (Pons, Bosch, & Lewkowicz 2015; Lewkowicz & Hansen-Tift, 2012). Furthermore, the signer was instructed to limit other movements and facial expressions, which resulted in stimuli with the only difference being mouth movements that naturally capture attention due to motion. Also, with limited facial expressions there was not much motivation to look for expression in the eyes that would typically be seen in social encounters since no social cues were being given. Thus, the control of the singer’s movements may have inadvertently directed the participants to stay on the mouth throughout the experiment.

Another way to look at these gaze patterns is that visual perceptual narrowing might be more complex than originally thought. When considering the eye data and the behavioral data together as one combined data set, it does support the idea of perceptual narrowing. The accuracy data demonstrates the typical pattern of narrowing, but this is occurring in the context of the participants looking equally at the mouth across both rhythmic conditions. Thus, the same visual attention is resulting in different perceptual outcomes. This suggests that visual attention alone does not lead to the perception of the visual rhythms, it is actually more important *how* the participants processed the visual information. The research on perceptual narrowing in languages provides evidence that listeners have more narrow perceptual categories within one's native language and more broad categories in one's non-native languages (Eimas et al., 1971). These broader categorical boundaries mean the listener can no longer hear differences within these categories because the listener no longer perceives specific elements of the non-native language like they do in their native language. Instead they perceive broad categories across the non-native language's sounds. The same process could be occurring with the visual rhythms in this experiment. The Western rhythms, that are the native rhythms, have smaller boundaries for differences in the patterns within a 4/4 meter, and therefore, the two Western visual rhythms can easily be seen as different. But the complex Balkan rhythms may just be categorized as one complex rhythm and the participants can no longer interpret the difference in these complex visual rhythms, even when directly watching the lips moving. The fact that there is a behavioral difference with no difference in gaze patterns actually follows along with the idea of categorical perception boundaries changing within and across native and non-native stimuli as a result of perceptual narrowing.

Lastly, we did not find any significant interaction between birth timing (preterm vs. full-term), however this could be due to the small sample size of preterm participants. More research would need to be conducted to determine if there are any lasting visual perceptual effects of being born preterm.

Limitations

This study had a few limitations mainly centered on the stimuli. One confound was the different tempos across the cultural rhythms that arose from the different time signatures between the Western and Balkan rhythms. The Western rhythms were in a 4/4 time signature but the Balkan rhythms were in a 7/8 time signature which also made the tempos different between the categories. With differences in tempo the rhythms' lengths were not equal. This resulted in the stimuli having a different number of repetitions of the rhythms across the different cultures. Even though the stimuli for the Western rhythms were 12 seconds long, there were only three repetitions of the rhythm; compared to the stimuli for the Balkan rhythms which were 8 seconds long but had four repetitions due to the faster tempo of the Balkan rhythms. We decided that it was more important to use the representative rhythmic/time signatures for each culture's rhythm rather than construct artificially controlled rhythmic patterns to match the length of the stimuli, as these may not be as representative of the two cultures true rhythmic structures. Given that this is the first study looking at visual rhythm perceptual narrowing an accurate baseline was important. The only real limitation this presented in the design was that we could not have trials comparing across the cultures because of the difference in length and tempo.

Another possible limitation of this study was the actual Balkan rhythms used as the stimuli. Due to the arrangements of the notes it is possible for the different rhythms to be mistaken for the same rhythm if a participant looked away from the screen at the beginning of the pattern and started grouping the rhythms from the middle for both Balkan rhythms (See Figure 2). Although this was not likely to happen given that the participant had to hit the space bar to start each video forcing their attention to be on the screen when the stimuli began. Additionally, if a participant had several years of music training then they might couple together the end of the second note with the beginning of the third note and view the videos as the same. However, the population sampled for this study was not highly trained in music (3.17 mean years of training) so it is very unlikely that this occurred with any of the participants. In fact the correlation between musical training and Balkan accuracy actually suggested the trained musicians were better at telling the rhythms apart and were not confusing their groupings.

Finally, it is possible that participants were more focused on trying to identify differences in the videos, such as small head movements or eye blinks, rather than differences in the rhythms. One way to control for this in future studies would be to give the participants instructions at the start of the study explicitly stating that they will be asked to identify if the rhythms were the same or different. This direction could help ensure that participants are looking to identify rhythmic differences.

Future Directions

Given this was the first study to examine perceptual narrowing for visual rhythm and use this combined methodology, this study served as a baseline for the field to move forward. There are several modifications that could be made to the study proceeding forward to help refine the design and more precisely measure visual perceptual narrowing. The stimuli could be recreated to control for the tempo confound. This would involve either slowing down the Balkan rhythms or speeding up the tempo of the Western rhythms while still maintain the original difference in grouping and beat. This could possibly be done by finding a style of Western music that has a faster rhythm that might be closer to the typical Balkan tempo, such as Western rock music. This would remove the possibility that the faster tempo made the Balkan rhythms more difficult to decipher. These changes would allow the videos to be the same length and the same number of rhythm repetitions.

Another change that could be made to the stimuli would be to have the singer have more facial expressions. In the current stimuli, the woman singing was instructed to remain neutral in expression and to not move around during the video. Her faced remained neutral throughout the entire video clip which makes her mouth the only movement in the video. There is a chance that participants were drawn more to the mouth due to the movement because this was the only movement in the stimuli (as suggested before). If the singer were to exhibit facial expression more similar to the typical expressions during a vocal performance, then participants might have more variation in their looking patterns. Due to the lack of expression, participants may not have been looking for social cues at the eyes because no social cues were being given. A more natural expression could promote more real-world response to the stimuli and heighten differences in gaze patterns.

This study used a novel approach to study visual perceptual narrowing. Even though our eye data varied from our original hypothesis, this study still provides a valuable beginning to the field of visual perceptual narrowing. Our behavioral data provides the first evidence of visual perceptual narrowing to culturally familiar rhythms. The eye data indicates that perceptual narrowing is more complex than we originally thought; even though participants were looking at the mouth during both cultural conditions, they were processing the visual information very differently. This study was an important first step in understanding visual perceptual narrowing and enculturation to music across our senses, while also providing a new methodology that can be used in future studies of visual perceptual narrowing across domains.

References

Barenholtz, E., Mavica, L., & Lewkowicz, D. J. (2016). Language familiarity modulates relative attention to the eyes and mouth of a talker. *Cognition,147*, 100-105. doi:10.1016/j.cognition.2015.11.013

Bremner, A. J., Lewkowicz, D. J., & Spence, C. (2012). *Multisensory development*. Oxford, UK: Oxford University Press.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech Perception in Infants. *Science,171*(3968), 303-306. doi:10.1126/science.171.3968.303

Hannon, E. E., & Trehub, S. E. (2005a). Metrical Categories in Infancy and Adulthood. *Psychological Science,16*(1), 48-55. doi:10.1111/j.0956-7976.2005.00779.x

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences,102*(35), 12639-12643. doi:10.1073/pnas.0504254102

Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, 1252:85-91.

Jansson-Verkasalo, E., Ruusuvirta, T., Huotilainen, M., Alku, P., Kushnerenko, E., Suominen, K., ... Hallman, M. (2010). Atypical perceptual narrowing in prematurely born infants is associated with compromised language acquisition at 2 years of age. *BMC Neuroscience,11*(1), 88. doi:10.1186/1471-2202-11-88

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences,109*(5), 1431-1436. doi:10.1073/pnas.1114783109

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition,82*(3). doi:10.1016/s0010-0277(01)00157-3

Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism Modulates Infants' Selective Attention to the Mouth of a Talking Face. *Psychological Science,26*(4), 490-498. doi:10.1177/0956797614568320

Table 1

Pixel count for each Area of Interest per video stimulus

	Mouth	Eyes	Face
Primary Western Rhythm	8580	13923	64404.5
Alternate Western Rhythm	8580	13923	63326
Primary Balkan Rhythm	7920	13923	59426.25
Alternate Balkan Rhythm	7920	13923	61079.5

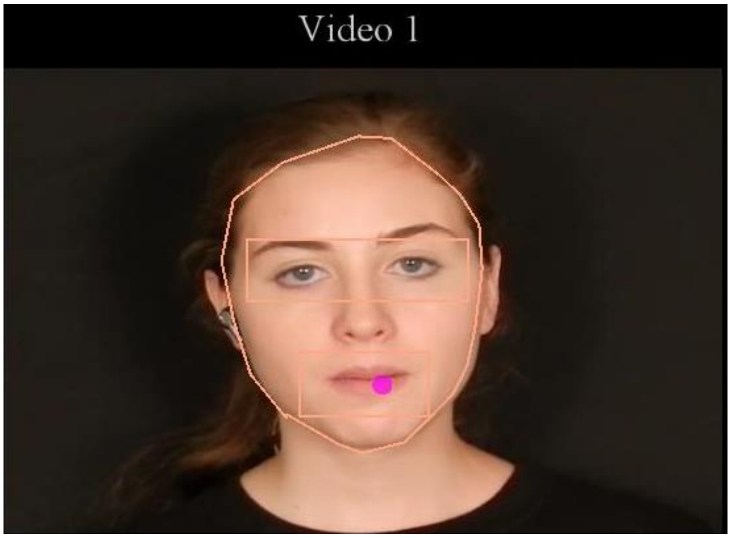

Figure 1. Areas of interest. This figure is an example of the areas of interest created per each video, overlaid on the singer's face. There is a rectangular interest area over the eyes and mouth and a contoured interest area that outlines the face.

Figure 2. Musical Notation of the Four Rhythms. This figure represents all four of the different rhythms used for this study depicted in musical notation. There were three repetitions of the Western rhythms and four repetitions of the Balkan rhythms in each respective stimulus.

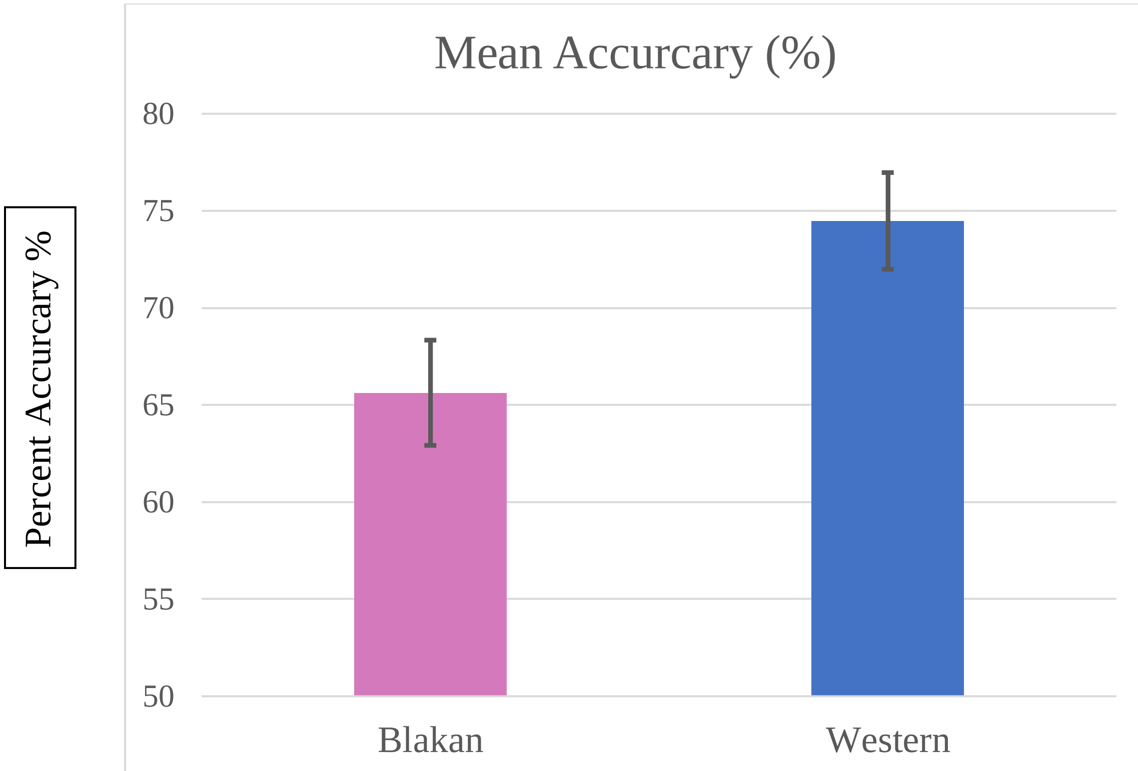

Figure 3. Mean Accuracy Graph. This graph represents the overall mean accuracy of participants across the Balkan and Western rhythm conditions on the Same/Different task. Error bars represent the standard error of the mean.

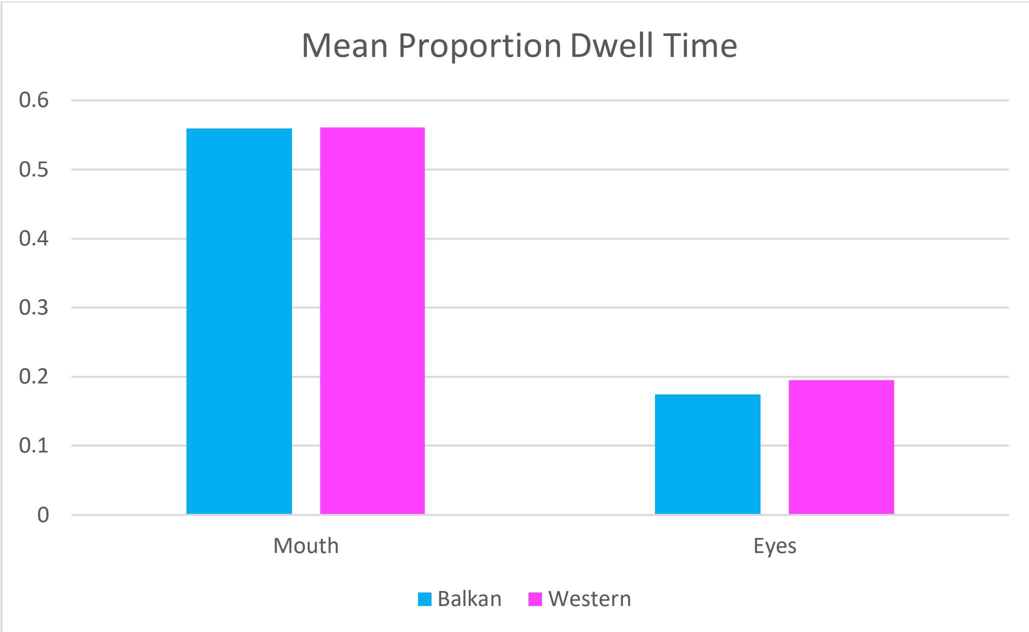

Figure 4. Mean Proportion Dwell Time Graph. This graph represents the mean proportion dwell time per each area of interest, calculated by dividing total looking time in each area of interest (mouth or eyes) by overall dwell time on the face.

Appendix A

1. Age in years: _________

2. We you born at full term or prematurely?

   Full term Prematurely (Preemie) Do not know

If preterm, do you know how preterm you were? (Days, weeks, months)

3. Gender: Male Female Other

4. Do you have any know hearing problems? If so, explain:

5. Do you have any know vision problems? If so, explain:

6. What did you think this study was about?

7. Did you use any strategies to tell if the videos were the same or different?

8. Can you speak a language other than English?

   Yes, currently Yes, previously No

If yes, please fill out the table

Language	How long have you been able to speak/could speak the language?	How proficient are you? Beginner, some experience, or fluent.
1.		
2.		

Appendix B

1. Do you CURRENTLY take lessons or are you a part of an ensemble for either a musical instrument or voice?

   Yes No Unknown

   Which instrument/vocal part? ________________

   For how long? ___________________________

2. Have you PREVIOUSLY taken lessons or have been part of an ensemble for either a musical instrument or voice?

   Yes No Unknown

   Which instrument/vocal part? ________________

   For how long? ___________________________

3. Can you read music?

   Yes No

4. How would you rank your overall music ability?

   No Knowledge　　Beginner　　Some Experience　　Competent　　Professional

   1 2 3 4 5 6 7 8 9

5. How many hours per day do you listen to music or dance videos?
6. What styles/genres of music do you most often listen to?
7. Are you exposed to music from outside the U.S.A? Yes No

   If so which cultures/countries?