2016

Speech to the Beat: Infants' Processing

Alix Altow
alixaltow@gmail.com

SPEECH TO THE BEAT: INFANTS' PROCESSING

By

Alix Altow

Department of Psychology

Submitted in Fulfillment of the

requirements of the degree of

Bachelor of Arts

in

Honours Psychology

Faculty of Arts and Social Science

Huron University College

London, Canada

April 2016

HURON UNIVERSITY COLLEGE

FACSIMILE OF CERTIFICATE OF EXAMINATION
(The Original With Signatures is on file in the Department)

Advisor: Dr. Christine Tsang

Reader: Dr. Irene Cheung

The thesis by:

Alix Altow

entitled:

Speech to the Beat: Infant's Processing

is accepted in partial fulfilment of the requirements for the degree of

Bachelor of Arts

in

Honours Psychology

April 29th, 2016 — Dr. Christine Tsang

Date — Chair of Department

Abstract

Previous research suggests that infants have high sensitivity to prosodic cues when listening to continuous speech. Specifically, rhythmic patterns play a significant role for infants' language development. Kuiack (2015) conducted a head-turn preference study where infants were shown to rely on the fine durational cues to recognize novel words from familiarized words. However, Kuiack failed to control for beat regularity between words. The present study seeks to explore whether the success in detection of the contrasting vowel durations within the words in the Kuiack (2015) study was due to the regularity in timing.   The current study inserted irregular pauses between the word stimuli used in the original Kuiack (2015) study. The results of the present study show that infants do not show discrimination in absence of rhythmic regularity. These findings suggest that infants require regularity in the beat pattern to recognize fine durational differences in speech.

Keywords: Speech processing; Rhythm; Regularity; Infant language development

Acknowledgements

I am most grateful for the immense support and encouragement I have received at Huron University College. I would like to express my deep appreciation to Dr. Christine Tsang. I have been lucky to have her as my professor, thesis advisor, and mentor in helping shape my academic as well as life choices. Throughout this journey, Dr. Tsang has always been generous with her time, knowledge and fundamental interest in my studies. Her enthusiasm and patience has given me confidence in my abilities and a strong interest in this field. I am honoured to be a part of her work and cannot thank her enough for all the time and effort she has dedicated towards me.

I would also love to thank Dr. Cheung, my second reader. Dr. Cheung is beyond intuitive and enthusiastic with the assistance and feedback she has given me. I am so grateful for her willingness to assist me in my writing skills throughout my undergraduate classes and thesis. She has allowed me to understand exactly where I can improve to reach my highest academic potential. I would also like to thank the other faculty members of the Huron Psychology Department, particularly Dr. Hessels, as well as everyone else who assisted me throughout these years.

I would like to thank my family for their unconditional love and support throughout my academic journey. Mom, thank you for everything and particularly for working with me from day one to help develop the work ethic that I have today. Dad, thank you for inspiring me to work with dedication and optimism. Kara and Bari, thank you for showing me that balance in life is key to success. Bubby, thank you for the wisdom and time you have dedicated to teach me valuable academic skills. Zaidy, thank you for always being the most proud and supportive of me. I am so thankful for my support system.

Table of Contents

Page

Infants' Processing Speech to the Beat

Infants are sensitive to all types of prosodic cues, such as intonation, musical and speech stress patterns, phonotactic regularities, allophonic variations, and most significantly for the investigation of this thesis – *timing cues*. All of these features of speech ultimately assist infants in language learning and acquisition. The speed of speech can assist a non-native person in language learning (Katz-Gershon & Andruski, 2006) such that slower speeds provide most optimal assistance (Munro & Derwing, 1998). Similar to how we may speak to a non-native speaker, speech directed to infants also tends to be slower than speech directed to adults and is thought to assist infants in word segmentation and word learning. Infants are most sensitive to these cues before their first year of life and rhythmic patterns are particularly important for language learning (Nazzi, Jusczyk & Johnson, 2000). For example, as early as 4 months of age the ability to discriminate between words with different rhythmic patterns is correlated with later language development (Höhle, Pauen, Hesse, & Weissenborn, 2014).

With increased exposure to their native language over the first year, infants' abilities to perceive cues from both native and foreign languages becomes more fixed, such that infants eventually discern fine distinctions that are only pertinent to their native language. The specific exposure and familiarity to particular durational patterns alter the future perception of those durational patterns. Werker and Tees (1984) showed a developmental trend of perceptual narrowing over the first year of life, such that before the age of 10 months, infants show discrimination for speech sounds that are within and outside of their native language. However, by 10 months, infants' abilities to differentiate non-native speech contrasts decline, while their ability to differentiate native language speech contrasts is maintained. This enculturation effect suggests that early experience with native language speech sounds shapes the future perception of those sounds. These findings demonstrate that familiarity and experience with native language matter.

The role of experience in perceptual narrowing impacts more than just speech contrasts. Hannon and Trehub (1987) demonstrated that enculturation is not just a linguistic phenomenon. When differentiating changes that violated the metrical structure of musical patterns from those that preserved the metrical structure (whether the first pattern was a simple or a complex meter), North American infants were similar to Bulgarian and Macedonian adults as opposed to adults from their own culture. This suggests that infants have a broad range of sounds they can hear that have not yet been fine-tuned to their native language.

Saffran, Aslin & Newport (1996) proposed that infants segment words based on statistical properties of the continuous speech stream they hear. For example, the phoneme /h/ is frequently placed at the start of a word rather than in the middle. Lexical phonotactics are formed from these regularities to assist in making generalizations of the sounds of a language (Daland, 2009). There can be regularity within a word and between words. For instance, there can be irregularity between words by having irregular pauses between the made-up words presented; or there could be irregularity within a word itself, such as different vowel durations. Non-sense words will be used in the current research because typical words may already be familiar to the infant due to their preexisting language exposure. Using non-sense words lends a fair trial to see whether or not infants can discriminate variables, such as vowel duration, in those words.

Equally frequent exposure to languages leads infants to perceive both languages equivalently well. Findings by Kuhl, Tsao, Liu (2003) proposed that American infants exposed to Mandarin were statistically equal to infants in Taiwan who were exposed to Mandarin their whole lives. This study indicates the significance of statistical learning from the impact of frequent stimuli. Depending on the context of frequent stimuli provided, it directs infants to enculturated speech processing. Thus, frequency in stimuli is dependent on one's culture.

Rhythmic regularity occurs when a sound pattern occurs at a constant and regular pace. Adults favour this regularity as it conforms to their culture-specific expectations. This regularity is a heuristic contributing to a perceptual bias (Trehub & Hannon, 2009). For example, to adults, 6-month old infants displayed a greater ability to detect subtle rhythmic changes in a conventional rhythmic sequence compared to an unconventional sequence, similar to adults. This suggests that there is little to no experience required to form this preference. A sense of expectation have formed one's perceptual capacity from this regularity and musical enculturation. Thus, it is anticipated that there is a biases for particular rhythmic patterns (Boltz, 1993).

While it remains unclear which prelexical or prosodic cue infants use in statistical learning tasks such as that described by Saffran et al. (1996), one possibility is temporal regularity. Infants are capable of using temporal cues within their native language. Five-month-old infants have demonstrated the ability to differentiate between languages from different rhythmic classes while they were not able to discriminate between languages that were derived from the same rhythmic class (Nazzi et al. 2000), Two-month-olds are able to detect a change in stressed and unstressed patterns of a multisyllabic utterance

(Jusczyk & Thompson, 1978). For example, native English-speaking 9-month-old infants have been shown to exhibit a preference for a trochaic stress pattern (strong-weak) which is common in the English language over an iambic stress patter which is uncommon in English (Jusczyk et al. 1993). Together, this body of research indicates that temporal cues such as stress and rhythm can be utilized by infant listeners.

Rhythmic cues such as accents have also been studied in the context of music. They are described as a note that protrudes from its surrounding notes and grabs the listener's attention because it may differ in its duration (Drake, Dowling, Palmer, 1991). Since accents provide a structure to music – rhythmic beats – then this should work in opposition where rhythmic beats provide a structure for accents, and more specifically vowel duration.

Kuiack (2015) examined 7-month-old infants' sensitivity to fine acoustic features in the context of speech. Infants were familiarized to a non-word with a specific pattern of fine durational cues. Infants were then tested on their ability to discriminate between two new non-words – one containing the familiar stress pattern and the other with a novel stress pattern. These novel words differed in their vowel duration and ultimately in their rhythmic patterns. The results of the study indicated that 7-month-olds have the ability to distinguish between similar novel words ('milami' and 'lamila') that have contrasting vowel durations. This difference in vowel duration thus provides a cue for the infant to discriminate between the words, and extends the findings in the literature showing that infants can segment speech based on temporal cues alone. However, one temporal aspect that Kuiak (2015) failed to control was regularity between words—due to the regularity of the inter-word intervals in the study, infants may have been using beat regularity as a basis for discriminating the two different word sets. In other words, infants' discrimination may not have been due to discrimination of vowel duration within words, but rather by the regularity created by the durational cues that occurred between words.

The purpose of the present research is to replicate the findings from Kuiack (2015) and extend them by varying the rhythmic regularity between the words. Specifically, irregular inter-stimulus intervals were introduced between word presentations. If the pattern of findings are similar to Kuiack's (2015), in that the infants are able to differentiate between the words with this added disrupted rhythm, it will suggest that the infants' abilities to discriminate are not due to the regular beat pattern but rather the matching vowel duration. If the pattern of findings is not replicated, in that the infants are unable to distinguish between the words, it is possible that infants are strictly relying on the rhythmic pattern. The deconstruction of cues will indicate the extent to which the infants use particular prosodic cues to decode speech input. From the previous literature mentioned, it was hypothesized that the irregularity between the non-sense words presented will hinder the infants' abilities to discriminate between the different vowel durations.

Method

Participants

Fifteen 6- 9-month-old infants (11 males and 4 females, mean age = 7.4 months, range = 6.3 – 9.6 months participated in this study. Participants were recruited from the developmental research participant database maintained by the Department of Psychology at Western University. At the time of testing, infants were healthy and caregivers reported no history of ear infections or history of family hearing loss.

Apparatus

The testing was conducted in a quiet and well lit testing room. The caregiver was seated in a chair in the middle of the testing room with the infant seated, facing forward, on his/her lap. The infant was positioned directly across from the experimenter who was seated behind a narrow black desk at a distance of approximately 1.5 m. A wireless computer keyboard was used by the experimenter to control the presentation of the visual and auditory stimuli and was hidden from the infant's view. There were two identical black cabinets situated at 90-degree angles, to the left and to the right of the infant and caregiver. Within each cabinet was a 13-inch standard computer monitor that was used to display the visual stimuli. On top of each cabinet was a black Bose 201-V sound speaker to present the auditory stimuli. Both the experimenter and the caregiver wore headphones and listened to masking music for the duration of the Test Phase (see Procedure) to ensure that they had no influence on the infant's head turn behaviour. See Figure 1 for the laboratory arrangement.

A Macintosh computer with customized software was used to control testing and record the looking times of the infant. This computer was located in the adjoining test room and connected to both the wireless keyboard, used during testing, and the amplifier, which controlled the stimuli presentation, through the two sound speakers.

Stimuli

The infant-directed stimuli in this study were trisyllabic non-words, consisting of three consonant vowel pairs varying in rhythmic duration. The same female speaker recorded all of the stimuli so that the voice was consistent across the presented non-words.

The syllables used were /mi/ and /la/ of a shortened vowel duration or /mee/ and /laa/ of a lengthened vowel duration. Each non-word had an ABA syllable pattern and differed in the length of the vowels and the accent of the consonants. Stimuli were either “milami”, with shortened or lengthened durations or “lamila”, with shortened or lengthened durations. At test the shortened and lengthened duration of the opposite word were presented in alternating order. The stimuli consisted of irregular interstimulus intervals between the set of words. The duration of the interstimulus interval between words varied between 0 ms to 600 ms. Each presentation, both during the Exposure Phase and the Test Phase, consisted of three slightly different versions of the same non-word presented in a quasi-random order that never repeated sequentially.

Figure 1. Head-turn preference procedure laboratory setting.

Procedure

Infants were tested individually using a head-turn preference procedure. There were two phases of this study: an Exposure (familiarization) Phase and a Test Phase. During the Exposure Phase, the caregiver and infant were brought into the test room. The caregiver was instructed to sit on a chair, between two cabinets, facing the desk of the experimenter. The caregiver was fitted with headphones to mask the sound of the stimuli during the testing procedure. The caregiver held the infant on his/her lap. The experimenter then left the room and played the familiarization stimulus for the infant on both speakers (the right and the left side) from iTunes on the computer in the adjoining test room. The infant listened passively to the stimulus for approximately 2 minutes. After the presentation of the auditory clip, the experimenter re-entered the test room, sat behind the desk directly across from the infant and began the Test Phase when the infant was facing forward and appeared to be attentive.

During the Test Phase, a standard head-turn preference procedure was used as described by Kemler-Nelson et al. (1995). The Test Phase began with the image of Mickey Mouse flashing on a computer screen to one side of the infant. When the infant turned his/her head to look at the screen, the target image stopped flashing and remained on the screen as the experimenter pressed a key from behind the desk that prompted one of the sound stimuli to begin playing from the sound speaker located directly above that computer monitor. The key press, made by the experimenter, also initiated a timer for the looking-time behaviour of the infant for this particular trial. The stimulus that played had either the familiar durational pattern or a novel durational pattern, as determined by the Exposure Phase. The familiarized stimulus had a particular rhythmic durational pattern.

For instance, the Exposure Phase may consist of the non-word “milami” in a shortened duration format with an irregular beat – creating a consistent vowel duration from the Test to Exposure Phase, with three versions of the same word repeated in random succession. The Test Phase would then consist of the sound stimulus non-word “lamila” in the same shortened duration format as the “milami” in the Exposure Phase. The difference for this trial is that the sound stimulus will be in opposite durational pattern (“lamila” presented in a lengthened duration – creating a contrasting vowel duration from the test to exposure phase). The familiar and novel stimuli were carried out for 10 trials for each stimulus (20 trials in total). The side of first presentation (left or right) was counterbalanced across participants. The first stimulus (varying in durational cues) was also counterbalanced across participants. The duration of testing was approximately 15-20 minutes per infant.

Results

A 2x2 within-subjects variable Analysis of Variance was conducted with stimulus (familiar versus novel) as the first within-subjects variable, half of the study (first half versus second half) as the second within-subjects variable. Infants’ looking-time was the dependent variable. There was a no main effect of Stimulus, F (1,14) = 0.032, p = 0.861, partial η = 0.02. There no significant main effect of Half, F (1, 14) = 2.399, p = 0.144, partial η = 0.146. There was a marginal but not significant Stimulus x Half interaction, F (1, 14) = 0.359, p = 0.060, partial η = 0.060 such that infants tended to look slightly longer to the novel stimulus in the second half (Figure 2).

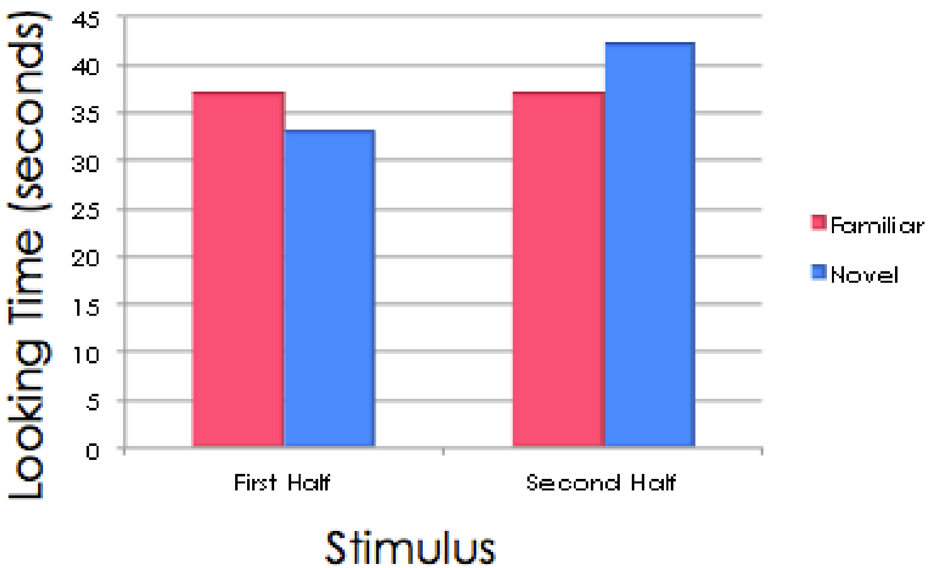

Figure 2. Mean infant looking times across familiar and novel stimuli compared across first half of testing session (10 trials) and the second half of testing session (10 trials).

Discussion

The current study did not display a significant main effect of stimulus type (novel versus familiar). This suggests that rhythmic regularity impacts infants' recognition of the stimulus. Together with the findings from Kuiack (2015), the results of the current study support the notion that infants use regularity as a cue to bootstrap their ability to detect differences in vowel duration. In the Kuiack (2015) study, the regularity in how the words were presented was congruent to the vowel duration. This coincidence of cues enabled the infants to successfully discriminate between the words. However, the infants in the current study were not able to do so when the same contrasting vowel durations were presented but the regularity between words was removed. It is evident that the irregularity makes it difficult for the infants to distinguish between the stimuli.

The findings from the current study are supported by the concept of rhythmic regularity. Previous literature indicates that at a young age infants are highly sensitive to this regularity cue. The presence of pauses has shown to have a major effect in assisting auditory stream segmentation (Carl & Gutschalk, 2013). Since there was an absence of regularity in the words presented to the infants in this study, it inhibited their ability to discriminate between the differences in vowel duration. This is a major contrast from the infants in the Kuiack (2015) study who were able to detect the difference between the words since there were no irregular pauses inserted between the words, maintaining the consistency of the rhythm.

Infants' sensitivity to these steady rhythmic patterns is also portrayed in research on languages differing in rhythmic classes. Five-month old infants have demonstrated their ability to identify a language within their native rhythmic class. Having the ability to detect these cues can be sufficient or at times necessary for speech and language acquisition. The responsiveness to rhythmic classes derives from the infants' ability to order sounds heard in their native language (Nazzi et al., 2000). Since the findings of the current study differ from Kuiack (2015) where the alteration was in the regularity between the words presented shows infants sensitivity to rhythmic regularity. This coincides with the findings of Nazzi et al (2000), demonstrating the importance of the rhythmic cue. Infants have this rhythmic sensitivity early on and there is significance in incorporating that cue in various contexts for optimal learning.

Typically, the more familiarized infants are with a stimulus the less they attend to it. Infants tend to display a habituation effect when they are presented with stimuli where a greater focus is on the initial presentations rather than the later (Klein-Radukic & Zmyj, 2015). However, in the current study, a trend towards looking longer to the novel stimulus was more apparent in the second half of the trials. This effect was also found in Kuiack's (2015) and is likely due to the complexity of the stimuli utilized in the study. The complexity of the stimuli derives from the difference between the short versus long vowel durations being so minimal that it is difficult to detect. The vowel alterations are ever so slight that adults, specifically the caregivers of the infants, were not able to perceive the difference between the stimuli. The insertion of irregular pauses between the words creating this inconsistent rhythm adds to the challenge of discriminating between the vowel durations.

While the lack of regularity is one possible reason why infants did not show significant differences in looking time to the novel and familiar stimuli, it is also possible that the lack of a secondary cue (which was present in Kuiak (2015), rather than regularity itself, accounts for the lack of effects. It is likely that infants may need supplementary cues to facilitate their ability detect the changes in vowel duration. In the infant's daily life, there are many cues that bootstrap onto one another such that there are many converging speech cues (e.g., vowel durations and regularity) present in the environment that together may aid the infant in parsing the speech signal (Saffran, Newport, Aslin, Tunick, Barrueco, 1997). The present study attempted to control the learning environment, effectively stripping the stimulus of all but one cue (vowel duration). Detecting fine patterns within spoken speech is a multimodal process that can either be all auditory (ie. beat or pitch related) or a mix between the visual and auditory modalities. Within the auditory modality, regularity is one of those many cues. Regularity may not be the only necessary cue that must be in place for infants to complete this task. There could be another cue that makes up for or provides similar enhancing benefits in replacement of that cue.

For instance, adding in a pitch emphasis cue can enhance the benefits that the regularity cue provides. On a word-learning task for adults, performance was significantly higher when pitch prominence marked the target word. This was due to the role of statistical learning for the pitch cues (Filippi, Gingras & Fitch, 2014). It would be interesting to investigate whether adding in a pitch cue instead of the regularity cue in the Kuiack (2015) study would enhance or diminish the infants ability to detect the changes in vowel duration. This would indicate if the regularity cue specifically is the cue that is necessary for the infants or rather having multiple cues is sufficient.

There could also be several prosodic cues that work in unison to enable infants to discriminate between stimuli such as speech sounds or words. For example, pitch and rhythmic regularity could simultaneously form an assistive cue for infants. A study was conducted where high school singers sang melodies that varied in pitch and rhythm combinations. The findings revealed that the rhythmic success was significantly related to pitch success. Interestingly, participants with piano experience were more efficient at performing the rhythm and pitch together than those without that instrumental practice. This may also suggests that nurture and the amount of practice one has can play a role in advancing the ability to identify how to use cues in unison to assist their learning (Henry, 2011). The overlap of these concurrent cues is how they are presented in the natural world; regularity is simply one of those numerous cues infants may be utilizing. Nevertheless, the results of the present study show that regularity facilitates n infants' ability to identify fine word patterns as demonstrated in the current study.

It is impressive how infants can overcome signal complexity to process and acquire language. These alterations in tone and pitch from speaker to speaker contribute to the complex input of speech signals the infants hear. This may suggest that multiple cues are beneficial for infants to gain an expanded range of contexts in which a phoneme or phrase can undertake (Estes & Lew-Williams, 2015). In the current study, if there were multiple cues inserted in the stimuli presented, the cues would not compete but instead assist one another. The context provided for this study may not have been rich enough in its cues for the infants to discriminate between the stimuli.

The current study is not directed to multimodal cue convergence; however, research on this topic provides further support for the importance of bootstrapping external cues to heighten the acquisition of language. Infants' high sensitivity and acquisition of this regularity cue can be further analyzed through an innate mechanism known as *statistical learning* (Saffran et al., 1997). The regularity between words may be another statistical cue that bootstraps word segmentation abilities. Having the regularity cue remain consistent in the Kuiack (2015) study may have enhanced the infants' responsiveness.

The finding that infants are not able to recognize the familiar vowel durational pattern in the context of irregularity may be due to regular patterns being easier to remember than irregular ones—i.e., higher cognitive load. The reason for this is because there are no cues to hang onto them with (Fowler, Shankweiler & Studdert-Kennedy, 2016). The inconsistent variation in the vowel durations is more difficult to detect when the timing of the words presented is also inconsistent. Having the regularity cue present in the stimuli sets the pace for the infants to attend to the contrasts in vowel duration rather than being cognitively overloaded with other unpredictable cues – the irregularity of rhythm. If given regularity, one can organize and process information in a way that files it more effectively in their memory. More favourable strategies could be used (ie., chunking) for the short-term memory and the ability to track patterns across a longer time (Holliman,Williams, Mundy, Wood, Hart, Waldron, 2014).

It should be noted that despite the apparent difficulty of recognizing familiar durational vowel patterns without regularity in beat, it is likely that some of the infants tested were still able to detect the variations, given the marginally significant interaction. Infants' age and experience may also have influenced whether or not they were able to detect the differences in vowel duration. An infant that is 8-months-old has been more exposed to their native language compared to a 6-month-old. This may increase the 8-month-olds likelihood of being able to detect the contrasts within the words and not have the irregularity between the words interfere them to the same extent it would for the 6-month-old. Even though the infants from the current study did not appear to be able to distinguish between the words, the direction of results presented itself in a similar way to Kuiack's (2015) findings.

Temporal cues (such as rhythm) may be generalized to other activities beyond speech processing and language development. The recently amalgamated team of neurologists and music therapists are spreading the value of music throughout different fields of study. For instance, rhythmic patterns or music have been used for people who have experienced a stroke to restore their speech as it is non-intrusive and therapeutic (TED, 2015, 6:50). Rhythmic regularity can help because from a production standpoint, the cue of regularity sets the pace of the speech being articulated. Not only can this cue assist an individual but it can similarly for a group as well, such as a choir. This regularity in pace can allow for a sense of social cohesion to be created permitting within-group cognitive and social synchronization (Cameron, Bentley, Grahn, 2015).

Temporal prosodic features are significant to the extent that there is likely a central mechanism in our brain for processing temporal regularity. As a speech language and music therapist, Kathleen Howland suggests how music is exercised successfully for numerous disorders (TED, 2015, 5:50). The right hemisphere of the brain is devoted mainly for melody, while the left hemisphere for processing speech (TED, 2015, 6:30). For young children with autism, a region in the brain known as the Arcuit Vasiculus, is thicker on the right hemisphere of the brain compared to the left. This leads to the finding of a greater response for sung compared to spoken cues for Autistic children (TED, 2015, 6:10). This suggests that our innate ability to process rhythm should be emphasized to override a delay in language or strengthen the networks in one's brain to have a better grasp of how to use the tools they possess for language acquisition. It could be the case that the innate tools infants have for speech processing are present; however, utilizing them to their greatest potential is unknown. Future studies should analyze cues other than regularity to investigate this further as the current study focuses on the regularity cue.

Regularity is an extremely helpful cue in our everyday lives. A sense of identity is created for a sound when a rhythm is attached to it. It enables a sound to be better distinguished amongst others. It is clear that regularity of a rhythm can be beneficial for infants, as seen from the previous literature and from children's book such as Dr. Suess where the rhyming words create a beat that allows for the prediction of what word is likely to follow. Despite the non-significant results of the current study, coupled with the findings from Kuiak (2015), it is clear that infant listeners' success in detection of the contrasting vowel durations is enhanced by temporal regularity. The current study suggests that rhythm regularity may have been the reason for their optimal success in doing so.

References

Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics*, *53*(6), 585-600.

Cameron, D. J., Bentley, J., & Grahn, J. A. (2015). Cross-cultural influences on rhythm processing: Reproduction, discrimination, and beat tapping. *Frontiers in Psychology*, *6*

Carl, D., & Gutschalk, A. (2013). Role of pattern, regularity, and silent intervals in auditory stream segregation based on inter-aural time differences. *Experimental Brain Research*, *224*(4), 557-570. doi:http://dx.doi.org/10.1007/s00221-012-3333-z

Daland, R. (2009). *Word segmentation, word recognition, and word learning: A computational model of first language acquisition* (Order No. AAI3355678). Available from PsycINFO. (622082764; 2009-99210-043).

Drake, C., Dowling, W. J., & Palmer, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception*, *8*(3), 315-334.

Estes, K. G., & Lew-Williams, C. (2015). Listening through voices: Infant statistical word segmentation across multiple speakers. *Developmental Psychology*, *51*(11), 1517-1528. doi:http://dx.doi.org/10.1037/a0039725

Filippi, P., Gingras, B., & Fitch, W. T. (2014). Pitch enhancement facilitates word learning across visual contexts. *Frontiers in Psychology*, *5*

Fowler, C. A., Shankweiler, D., & Studdert-Kennedy, M. (2016). "Perception of the speech code" revisited: Speech is alphabetic after all. *Psychological Review*, *123*(2), 125-150. doi:http://dx.doi.org/10.1037/rev0000013

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science, 16*(1), 48-55. doi:http://dx.doi.org/10.1111/j.0956-7976.2005.00779.x

Henry, M. L. (2011). The effect of pitch and rhythm difficulty on vocal sight-reading performance. *Journal of Research in Music Education*, *59*(1), 72-84. doi:http://dx.doi.org/10.1177/0022429410397199

Höhle, B., Pauen, S., Hesse, V., & Weissenborn, J. (2014). Discrimination of rhythmic pattern at 4 months and language performance at 5 years: A longitudinal analysis of data from German‐learning children. *Language Learning, 64,* 141-164. doi:http://dx.doi.org/10.1111/lang.12075

Howland, K. M. (2015) *Music therapy handbook: Music therapy for children with speech and language disorders.* New York: A Division of Gulfford Publications, Inc.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, *32*(3), 402-420.

Jusczyk, P.W., & Thompson, E.J. (1978). Perception of a phonetic contrast in multisyllabic utterances by two month-old infants. *Perception & Psychophysics, 23*, 105-109.

Katz-Gershon, Shiri, and Jean Andruski. "Word Extraction in Adult-directed and Infant-directed Speech in African-American-English-learning Infants." *The Journal of the Acoustical Society of America J. Acoust. Soc. Am.* 120.5 (2006): 3136. Web.

Klein-Radukic, S., & Zmyj, N. (2015). Dropout in looking time studies: The role of infants' temperament and cognitive developmental status. *Infant Behavior & Development*, *41*, 142-153. doi:http://dx.doi.org/10.1016/j.infbeh.2015.10.001

Kuhl, P. K., F-M Tsao, and H-M Liu. 2003. Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *USA* 100, (15) (Jul 22): 9096-9101

Kuiack, A. (2015). Infants' Sensitivity to Fine Durational Cues in Speech Perception. *Undergraduate Honours Theses.*

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11*(1), 122-134. doi:http://dx.doi.org/10.1111/j.1467-7687.2007.00653.x

Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning*, *48*(2), 159-182.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by english-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*(1), 1-19.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*(4), 606-621.

Saffran, J., Newport, E., Aslin, R., Tunick, R., Barrueco, S (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological*

Science 8, (2): 101-105.

TED. (2015, May 27). *How music can heal our brain and heart*. | Howland, K. M. | *TED Talks* [Video file]. Retrieved from https://www.youtube.com/watch?v=NlY4yCsGKXU

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *7*, 49–63.

Curriculum Vitae

Name:	Alix Altow
Place and Year of Birth:	Toronto, Canada, 1994
Secondary School Diploma:	Community Hebrew Academy of Toronto, Toronto, Canada