Masthead Logo
University of Tennessee, Knoxville
Trace: Tennessee Research and Creative Exchange

Masters Theses
Graduate School

5-2018

Intersensory Redundancy and Infant Selective Attention to Audiovisual Speech

Kelly C. Roth
University of Tennessee, kelcroth@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Kelly C. Roth entitled "Intersensory Redundancy and Infant Selective Attention to Audiovisual Speech." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Psychology.

Gregory D. Reynolds, Major Professor

We have read this thesis and recommend its acceptance:

Aaron T. Buss, Michael A. Olson, Caglar Tas

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Intersensory Redundancy and Infant Selective Attention to Audiovisual Speech

A Thesis Presented for the

Master of Arts

Degree

The University of Tennessee, Knoxville

Kelly C. Roth

May 2018

DEDICATION

In loving memory of Dante Nero Roth, my little master's kitty (2015-2017).
I will always love you.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Greg Reynolds, for his guidance, patience, and encouragement throughout my graduate school journey. Thank you also to my thesis committee members, Dr. Michael Olson, Dr. Aaron Buss, and Dr. Caglar Tas. Their insightful comments, expertise, and advice have helped me tremendously. A sincere thank you to Dr. David Lewkowicz and his lab for providing the stimuli used in this study. Thank you to my family for their love and reassurance. Finally, I would like to thank my cats, Xena, Jack, and Dante, for their unwavering emotional support through both this thesis and graduate school in general.

ABSTRACT

The current study utilized eye-tracking to investigate the effects of intersensory redundancy on infant visual attention and discrimination of a change in prosody in native and non-native audiovisual speech. The Intersensory Redundancy Hypothesis states synchronous and redundant presentation of bimodal stimuli selectively recruits infant attention to and facilitates processing of amodal stimulus properties (Bahrick & Lickliter, 2000). Twelve-month-old monolingual English learning infants viewed either synchronous (redundant) or asynchronous (non-redundant) video clips of a woman speaking in English (native speech) or Spanish (non-native speech). Halfway through each trial, the speaker changed prosody from adult-directed speech (ADS) to infant-directed speech (IDS) or vice versa. Participants completed four 1-min trials, counter-balanced for order. I hypothesized intersensory redundancy would direct infant attention to amodal properties of speech and facilitate discrimination of a change in prosody. Specifically, I predicted infants in the synchronous condition would demonstrate differential scanning of the face based on changes in prosody on both English and Spanish trials. I predicted infants in the asynchronous condition would only demonstrate differential scanning patterns based on a change in prosody on English trials. The analyses revealed a main effect of prosody. Infants focused their visual attention more on the mouth of the speaker on IDS trials in comparison to ADS trials regardless of language or redundancy. There was also an interaction of prosody and language on infants' selective attention. Infants focused more on the nose during English ADS speech in comparison to English IDS speech. These results indicate IDS directs infant attention to the mouth of speakers. In the analysis of detection of a change in prosody, infants in the synchronous condition showed significant differences in looking during the second block of trials depicting English ADS changing to English IDS. This effect may have been due to an interaction of the greater salience of IDS, the infants' extensive experience with their native language, and the facilitating effects of intersensory redundancy for detecting changes in prosody. Overall, these findings exemplify the complexity of development and indicate multiple factors interact to affect infants' visual attention and their ability to discriminate changes in prosody in audiovisual speech.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I INTRODUCTION

Young infants are exposed to a plethora of sensory information, from basic sounds and sights to specific caregivers and language. How do infants find a way to process and make sense of the enormous amount of perceptual information provided by the world around them? Within the context of infant development, certain properties of stimuli in the environment can facilitate perceptual processing and help to bootstrap learning and cognition, while other factors may make perceptual processing more difficult. In infancy, one such developmental process is perceptual narrowing, which refers to a decline of perceptual sensitivity to unfamiliar stimuli while sensitivity to familiar stimuli is maintained (Maurer & Werker, 2014; Pascalis, de Haan, & Nelson, 2002; Scott, Pascalis, & Nelson, 2007).

Currently, perceptual narrowing is understood as a regressive, as opposed to progressive, development (Maurer & Werker, 2014; Lewkowicz & Ghazanfar, 2009). Over time and with experience, perceptual sensitivity moves from being broadly tuned to a wide-array of stimuli to being more narrowly focused on relevant information routinely encountered in the infant's native environment. Young infants can respond to stimulus properties characteristic of both native and non-native stimuli, as their perceptual functions are still broadly tuned and unspecialized. Thus, infancy can be considered a "sensitive window" when the child is being exposed to the foundational stimuli in their native environment that will ultimately shape their perceptual biases in later development. The ability to shift attention to relevant stimuli is critically important for the formation of perceptual expertise (Lewkowicz & Ghazanfar, 2009). For example, by 3 months of age, infants prefer human speech over non-human speech, and if raised in a monolingual home, infants will begin to show preferences for native speech over non-native speech (Maurer & Werker, 2014). At 4 months of age, infants can discriminate vowel sounds in both their native and non-native languages. However, by 10 months of age, infants are only able to discriminate vowels in their native language if raised in a monolingual environment (Polka & Werker, 1994). Perceptual narrowing is hypothesized to allow infants to focus attention on and specialize in processing their native language, instead of maintaining a broad and indiscriminate sensitivity to all languages. This heightened sensitivity results in an increased expertise in their native language at a cost of a decline in their perceptual sensitivity to non-native languages.

Intersensory redundancy is another factor that affects perceptual processing in infancy. The intersensory redundancy hypothesis states that redundant information conveyed across two or more sensory modalities facilitates infant selective attention to amodal properties of multimodal stimuli (Bahrick & Lickliter, 2000, 2002a, 2002b, 2004). Amodal properties of a stimulus include any information that can be conveyed by more than one sensory modality. Tempo can be perceived through both the visual and auditory sensory modalities, and thus provides a good example of amodal information. In contrast, modality-specific information can only be conveyed through a single modality, e.g. color can only be conveyed through a visual modality. Young infants rely on this redundancy for processing, as the temporal synchrony and shared rhythm of audiovisual stimuli recruits their selective attention and provides more information than unimodal events. With further development, older infants and children can process amodal information presented in unimodal stimuli and are less dependent on intersensory redundancy (Bahrick, Lickliter, & Flom, 2004). However, if a task is relatively difficult, older children will rely on redundancy to attend to amodal information during processing prior to processing modality-specific information of complex stimuli (Bahrick, Krogh-Jespersen, Argumosa, & Lopez, 2014).

Audiovisual intersensory redundancy is dependent on temporal synchrony between the auditory and visual amodal properties of stimuli. This form of intersensory redundancy has been shown to affect infant attention during native and non-native audiovisual speech at certain ages. A study published by de Boisferon, Tift, Minar, and Lewkowicz (2016) investigated the effects of desynchronization on infant selective attention to native and non-native audiovisual speech. Looking at the age groups 4-, 6-, 8-, 10-, and 12-months-old, infants were shown stimuli that was desynchronized by playing the auditory stream 666 milliseconds ahead of the video. This was done to control for any visual cues infants may gain from viewing the mouth movements before the audio. The offset was determined based on prior studies that show 666 ms is sufficient for infants as young as 4 months of age to discriminate (Lewkowicz, 2010; Pons & Lewkowicz, 2014). The stimuli used were full color video clips of a native English speaker and a native Spanish speaker reciting a monologue. Infants age 4 to 6 months looked at the eyes of the speaker while listening to both languages. The 8-month-old infants looked significantly more at the mouth of the speaker, also regardless of language. Infants at 12 months of age looked significantly more at the mouths of the non-native speakers and looked at both the eyes and mouth of the native speaker. However, the 10-month-olds tested responded differently than in Lewkowicz and Hansen-Tift (2012), discussed in detail later, which showed 10-month-olds look more at the mouth of native and non-native speakers during speech. Instead, in this study using desynchronization, the 10-month-olds did not look significantly more at the mouth during desynchronized audiovisual speech, regardless of language. Comparisons to Lewkowicz and Hansen-Tift (2012) show the differences in looking were due to less looking at the mouth during native speech and more to looking at the eyes during non-native speech. Lack of redundancy affected the 10-month-olds looking pattern, but the researchers used prosody, either infant-directed speech or adult-directed speech, as a between-subjects factor. IDS (infant-directed speech) has more amodal information than ADS (adult-directed speech) due to the exaggerated vocalizations and facial gestures, and may be better enhanced by intersensory redundancy.

Prosody refers to patterns of stress and intonation in speech. ADS and IDS vary greatly in prosody. ADS is a more neutral or flat tone of voice often used when speaking to adults. In stark contrast, IDS is characterized by a higher pitch and amplified tone differences and is often referred to as “motherese” or “baby talk." Specifically, infants prefer the amplified pitch differences of IDS over ADS, and they show no preference for the amplitude correlated with loudness or duration patterns of IDS when compared to ADS (Fernald & Kuhl, 1987). As IDS is often accompanied by exaggerated facial and hand expressions as well, studies have shown infants’ attention and language-related learning is enhanced by IDS when compared to ADS (Adriaans & Swingley, 2012; Fernald & Mazzie, 1991; Graf Estes & Hurley, 2013; Thiessen, Hill, & Saffran, 2005). IDS may recruit attention to redundant information by drawing infants’ gaze to the exaggerated mouth movements of the speaker. When attending to the mouth, infants are able to benefit from intersensory redundancy when processing amodal information (Kubicek, de Boisferon, Dupierrix, Pascalis, Lœvenbruck, Gervain, & Schwarzer, 2014; Kubicek, Gervain, de Boisferon, Pascalis, Lœvenbruck, & Schwarzer, 2014).

The present study investigated the possibility that intersensory redundancy interacts with perceptual narrowing in language development by measuring infants’ selective attention and testing infants’ ability to detect a change in prosody, an amodal property, in both native and non-native speech. The “sensitive window” for gaining language expertise in a non-native language is thought to have occurred before 12 months of age (Cheour, Ceponiene, Lehtokoski, Luuk, Allik, Alho, & Näätänen, 1998; Mauer & Werker, 2013; Werker & Tees, 1984). However, this study examined how the presence of intersensory redundancy may be able to bootstrap monolingual 12-month-old infants' ability to discriminate a change of prosody in non-native speech. Selective attention to facial features was also measured to determine if ADS and IDS caused infants to look differently at faces during audiovisual speech. How these interactions affect where infants look at the face, either at the source of the redundancy such as the mouth, or somewhere else on the face, was also explored. The present research aimed to increase scientific knowledge regarding the effects of intersensory redundancy on attention to native and non-native audiovisual speech, and to contribute to increased understanding of the potential role of intersensory redundancy in the process of perceptual narrowing in language.

Perceptual Narrowing

Perceptual narrowing refers to a developmental process characterized by an infant's initial broad sensitivity tapering down with age and experience to more specific and relevant information encountered in the native environment. Perceptual narrowing is thought to be a domain-general developmental process, as evidence of narrowing has been found to occur across multiple forms of perception, including face recognition, language perception, musical rhythms, and even cross-species perception (Maurer & Werker, 2014; Pascalis, de Haan, & Nelson, 2002; Scott, Pascalis, & Nelson, 2007). One of the current theories on perceptual narrowing is defined by a maintenance/loss dichotomy, wherein repeated exposure to their native environment helps infants maintain early perceptual sensitivity to native stimuli while irrelevant details are ignored or no longer prioritized over time (Maurer & Werker, 2014; Lewkowicz & Ghazanfar, 2009; Kelly, Quinn, Slater, Lee, Ge, & Pascalis, 2007). This decline to non-native stimuli is not indicative of a permanent loss of sensitivity but rather a reorganization of perceptual priority

(Lewkowicz & Ghazanfar, 2009; Kelly, Quinn, Slater, Lee, Ge, & Pascalis, 2007). The reorganization is thought to be a necessary component of developing expertise in perceiving certain social signals, such as a more nuanced understanding of one's native language (Anzures, et al., 2012; Lewkowicz & Ghazanfar, 2009). Infants do not completely "lose" the ability to process non-native languages, as evidenced by adults learning a second language, but rather attribute a higher processing priority to their native language and do not process non-native languages with the same level of precision after perceptual narrowing occurs (Lewkowicz & Ghazanfar, 2009; Polley, Steinberg, & Merzenich, 2006). This hierarchy of processing aids older infants in efficiently exploring and adapting to their environment on a day-to-day basis because it allows them to focus on the specific properties of stimuli that are important for functioning in their native environment (e.g., sounds or phonemes within their native language, facial features commonly encountered in their native environment, etc.).

A study conducted by Lewkowicz and colleagues (Lewkowicz, Leo, & Simion, 2010) provides an example of young infants' greater sensitivity to non-native stimuli. Newborns viewed two side-by-side videos on a monitor: one of a monkey producing a "coo" sound and one of the same monkey producing a grunt. Infants then heard either the synchronous audio of the coo, the grunt, or silence. Infants were tested on their ability to match the proper video with the current sound playing. Whether the infant recognized the correct visual for the sound was demonstrated by longer looking times at the correct video. One to 3-day-old infants were able to match videos of monkey faces either cooing or grunting with the correct audio sound, despite having no previous experience with such stimuli. Implications of the findings support the theory that infants are born with the ability to process bimodal events based on audiovisual temporal synchrony. Since the sound was played centrally, if the infants were unable to correctly match the monkeys' facial movements they should have shown equivalent looking to both videos. Although previous studies have shown newborns are capable of matching human faces to voices (e.g., Slater & Kirby, 1998), this study showed that newborn infants can match non-human faces and voices as well.

Within human language, infants' ability to match audiovisual speech in native and non-native speech has also been tested. Pons and colleagues (Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009) tested 6- and 11-month-olds who were learning either Spanish or English with an intersensory matching procedure. Infants viewed two side-by-side faces of a Spanish speaker silently mouthing the phonemes *ba* or *va*. After the baseline video, audio of either *ba* or *va* was played against a blank screen. The test trial was the same silent videos as before, but researchers measured if infants looked longer at the corresponding face. As these phonemes are indistinguishable in Spanish, researchers predicted the Spanish learning infants would still be "sensitive" enough to detect the difference at 6 months of age, but would be unable to discriminate the sounds at 11 months of age. As these phonemes are distinct in English, the researchers predicted both 6- and 11-month-old English learning infants would discriminate which face was mouthing which sound. Results supported these predictions. At 6 months of age, Spanish learning infants still discriminated the two sounds because they still maintained their initial broad sensitivity to perceptual differences in both native and non-native languages (Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009). The Spanish learning infants "lost" the ability to discriminate the two sounds by 11 months of age, past the onset of perceptual narrowing to their native language.

Until recently, no studies on perceptual narrowing had also examined the intersensory redundancy as a factor or utilized dynamic multimodal face stimuli. While many studies use static images and sequential audio and visual information to show experience effects, these are low in ecological validity and do not resemble "real-world" experience (e.g., Anzures, Wheeler, Quinn, Pascalis, Slater, Heron-Delaney, Tanaka, & Lee, 2012; Vogel, Monesson, & Scott, 2012; Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés 2009). A few exceptions come from recent studies conducted on infants' point of gaze when listening to native and non-native speech.

Monolingual English learning 4- to 12-month-olds, in addition to English speaking adults, were shown clips of either an English speaker or a Spanish speaker using either adult-directed or infant-directed speech and their visual scan patterns of the faces were recorded (Lewkowicz & Hansen-Tift, 2012). At 4 and 6 months of age, infants focused on the eyes of the face as the video clips played. Around 8 months, infants began to focus on the mouths of speakers regardless if the actor was talking in English or Spanish. This is theorized to coincide with infants beginning to develop language and focusing on the visual information of the mouth to aid pronunciation. At 12 months of age, infants who were listening to their native speech began to shift their attention up to the eyes of the speakers, similar to the adults, as it is hypothesized the upper area of the face has additional social information to supplement speech, such as eyes and eyebrow gestures. However, 12-month-old infants in the non-native speech trial continued to focus on the mouths of the speaker, similar to the 8- and 10-month. Since the non-native speech is less common in their monolingual English home environments, the 12-month-old infants may have experienced perceptual narrowing for their native language and become less sensitive to the Spanish language. When exposed to non-native languages, 12-month-olds may continue to focus on the mouth of the speaker to facilitate their processing of the unfamiliar language.

The results of the Lewkowicz and Hansen-Tift (2012) experiment are interesting because they show that infants may rely on the audiovisual synchrony between the mouth and voice of speakers during the early development of speech and language perception. It is hypothesized that the visual information from the mouth may aid infants in reproducing those sounds with their own mouths. At 6 months of age, infants are developing more attentional control (Colombo, 2001; Reynolds, Courage, & Richards, 2013; Ruff & Rothbart, 2001) and may be directing their attention to the mouth for the redundant information provided by audiovisual speech. At 12 months, infants have a basic level of expertise and may not need the redundant information provided by the speaker's mouth. Instead, they may begin to focus selective attention on the speaker's eyes, which provides further social cues. However, monolingual English infants would be expected to lack the same level of expertise for Spanish due to perceptual narrowing. This lack of familiarity may lead 12-month-old infants to look proportionately more at a non-native speaker's mouth for redundant information to process a more difficult, novel language (Lewkowicz & Hansen-Tift, 2012).

The current experiment sought to shed light on the role intersensory redundancy may play in perceptual narrowing, specifically processing of native and non-native audiovisual speech when provided temporally synchronous auditory and visual information. Currently, no studies have utilized intersensory redundancy (redundant versus non-redundant audiovisual stimuli) as a between-subjects factor to examine potential effects of redundancy on infant processing of native and non-native audiovisual speech. This research sought to understand the interactions between infant selective attention, intersensory redundancy, and perceptual narrowing during native and non-native audiovisual speech processing. Specifically, the current study tested if intersensory redundancy could be used to "overcome" perceptual narrowing and help bootstrap infants' ability to discriminate a change of prosody in non-native speech.

Intersensory Redundancy Hypothesis

Bahrick and Lickliter (2000) initially proposed the intersensory redundancy hypothesis, which states redundant amodal stimulation facilitates perceptual processing for young infants. The first iteration of the hypothesis covered three main tenets. First, intersensory redundancy facilitates infants' perceptual differentiation of amodal information. That is, redundancy across two or more modalities makes it easier for infants to differentiate and process amodal information, as it is being conveyed to them through at least two different sensory modalities. Second, temporal synchrony promotes stimulus salience. Sequential or asynchronous audio and visual amodal information do not have the same effect as when the audio and visual components are synchronized and presented simultaneously (i.e., redundant). Third, intersensory redundancy recruits infant selective attention to amodal properties of stimuli. When shown unimodal events, infants will focus on modality-specific information, i.e. focusing on the color of a hammer when shown a silent video clip of a hammer tapping a rhythm. If there is sound that is temporally synchronous with the video clip, the intersensory redundancy of the stimuli will lead infants to focus on amodal stimulus properties, i.e. infants will focus on the rhythm being conveyed by a tapping hammer as there are two redundant sources of that information: the sight of the hammer moving to the beat and the sound of the hammer tapping (Bahrick & Lickliter, 2000, 2002a, 2002b, 2004).

In an initial test of this hypothesis, 5-month-old infants viewed unimodal, synchronous bimodal, or temporally asynchronous bimodal events in the form of a hammer tapping a rhythm

(Bahrick & Lickliter, 2000). Results indicated that the infants could not differentiate changes when viewing unimodal auditory or unimodal visual presentations, e.g. when an audio track of the rhythm with a static image or when a silent video clip of a hammer tapping was shown. Infants also could not detect rhythm changes when the bimodal presentations were asynchronous, e.g. when the video of the hammer was played out of synch with the audio of the hammer tapping. Only infants exposed to synchronous bimodal presentations, where the tapping of the hammer was temporally synchronous with the video of it striking a surface, could differentiate between the familiar and novel rhythms (Bahrick & Lickliter, 2000).

As infants age, experience and more efficient perceptual processing allow them to differentiate amodal properties in both bimodal and unimodal (single-modality) stimuli (e.g., Flom & Bahrick, 2007). Thus, older infants are less likely to rely on the redundancy in bimodal stimulation to differentiate amodal properties of stimuli. Older infants can also identify modality-specific properties within a bimodal event, such as specific facial features (a visual property) of a person talking (a bimodal audiovisual event) (Bahrick, Lickliter, & Flom, 2004). However, bimodal events can overwhelm younger infants' perception of modality-specific properties. While the redundancy facilitates processing of the amodal properties, younger infants perform worse in tests of unimodal properties during bimodal events. Younger infants can differentiate between faces when the individual is silent, and differentiate between a specific voice if the person's dynamic face is not shown, but will not show this specific unimodal recognition if the event is bimodal (Bahrick, Lickliter, & Flom, 2004).

Flom and Bahrick (2007) investigated the developmental course of infants' ability to discriminate changes in affect in unimodal and bimodal events. Affect is an amodal stimulus property, as emotions are conveyed through speech in the auditory modality and facial expressions in the visual modality. When shown video clips of actors displaying one affect, either happy, angry, or sad, then shown the same actor displaying a different affect, infants from 4 months of age on could discriminate the change in affect in bimodal synchronous conditions. However, if the affect was conveyed unimodally, sensitivity to the change appeared gradually. By 5 months of age, infants detected a change in a unimodal auditory event, but were not able to detect a change in affect in unimodal visual events until 7 months of age. Thus, temporal synchrony was also found to be necessary for younger infants to be able to detect changes in affect. Only with age and experience can infants recognize changes in amodal properties during unimodal events (Flom & Bahrick, 2007).

In a study by Bahrick, Lickliter, and Castellanos (2013), 2-month-old infants also showed recognition of amodal properties at the expense of modality-specific information during synchronous bimodal events. Those exposed to synchronous audiovisual speech could not discriminate a novel face, but were able to do so when exposed to asynchronous audiovisual speech or unimodal visual speech. When there was a lack of synchrony, unimodal facilitation occurred wherein the modality-specific information, the faces, were prioritized for processing. This processing priority allowed the 2-month-olds to discriminate novel from familiar faces. These findings indicate that intersensory redundancy's recruitment of selective attention to amodal properties of stimuli can hinder facial perception early on, while unimodal or asynchronous stimulations can facilitate discrimination between novel and familiar faces, i.e. modality-specific visual information, a process referred to as unimodal facilitation (Bahrick, 2010; Bahrick & Lickliter, 2000; Bahrick & Lickliter, 2002b; Bahrick & Lickliter, 2012; Bahrick, Lickliter, & Flom, 2004). Taken together, these results show that infants begin by processing amodal properties in bimodal events and modality-specific properties in unimodal events.

Further studies have shown that increasing task difficulty can cause children to revert to relying on intersensory redundancy for perceptual processing and learning. In a study by Bahrick, Lickliter, and Castellanos (2013), 3.5 to 4-year-old children were asked to discriminate between novel and familiar faces of women in a paired-comparison task. When asked to identify which face of a pair was the familiar one ("Which one was at my [birthday] party?"), the children who were exposed to the audiovisual redundant stimuli performed worse than those who encountered either unimodal visual or asynchronous audiovisual stimuli (Bahrick, Krogh-Jespersen, Argumosa, & Lopez, 2014). This led the authors to add a fourth tenet to the intersensory redundancy hypothesis related to task difficulty: older infants or children, when presented with a relatively difficult task, resort to relying on intersensory redundancy similar to younger infants, thus focusing on initially processing amodal properties at the expense of modality-specific properties of bimodal stimuli.

These studies on the intersensory redundancy hypothesis show the importance of temporal synchrony in recruiting attention to amodal properties of stimuli (Bahrick & Lickliter, 2000). Intersensory redundancy can be used to bootstrap infants' processing of amodal information, but comes at a cost of not prioritizing processing modality-specific information (Bahrick & Lickliter, 2002a, 2002b). Alternatively, when shown unimodal or temporally asynchronous bimodal events, infants do not preferentially process amodal but rather modality-specific information first (Bahrick, 2010; Bahrick, Krogh-Jespersen, Argumosa, & Lopez, 2014; Bahrick & Lickliter, 2000; Bahrick & Lickliter, 2002b; Bahrick & Lickliter, 2012; Bahrick, Lickliter, & Castellanos, 2013; Bahrick, Lickliter, & Flom, 2004). The presence or absence of intersensory redundancy influences what properties of stimuli infants attend to, subsequently affecting whether they prioritize processing amodal or modality-specific information. The current study measured infant selective attention to facial features to investigate if attention would be drawn to the mouth, the source of redundancy, depending on if the speech was native or non-native and if the audio and video were temporally synchronous or not. Many of the previous studies mentioned utilized affect as an amodal variable. Similar to affect, prosody is an amodal property that can be conveyed unimodally depending on modality of presentation.

Prosody

Prosody is a characteristic of speech that facilitates infant attention and learning. Prosody refers to patterns of stress and intonation used in linguistics to convey emotional information or to accent certain words or syllables. IDS is a form of speech characterized by exaggerated prosody with shorter sounds and longer pauses in between, higher pitch, an exaggerated rhythm and difference in pitch contours, and a slower rate of speech used when talking to infants or young children. Research has shown IDS facilitates infants' learning and selective attention when compared to ADS (Fernald & Mazzie, 1991; Graf Estes & Hurley, 2013). Prosody is an amodal property as the inflections and emphasis can be conveyed visually through gestures and facial expressions as well as audibly through tone of voice.

IDS is theorized to bootstrap infant language learning (Vallabha, McClelland, Pons, Werker, & Amano, 2007). Research has shown the prosody of IDS helps to differentiate phoneme and vowel sound discrimination because of the greater emphasis given to certain syllables than in ADS (Adriaans & Swingley, 2012). IDS also helps with word segmentation due to its slower pace and longer pauses, allowing infants to better process distinct words during language learning (Thiessen, Hill, & Saffran, 2005). However, there is evidence that IDS is not significantly acoustically different than ADS, and that the increase in learning seen in other studies may be due to IDS's slower rate of speech instead of its ability to enhance different sounds (McMurray, Kovack-Lesh, Goodwin, & McEchron, 2013). Other researchers have hypothesized it is IDS's greater variability and less acoustically restrictive design that allows more emotional expression in speech, which facilitates infant attention and processing (Trainor, Austin, & Desjardins, 2000). When compared to ADS that is conveying a strong emotion, IDS and ADS do not differ in acoustic features, indicating it may be the affect conveyed in IDS that recruits infant attention better than the more neutral ADS (Trainor, Austin, & Desjardins, 2000).

There is evidence that IDS is a universal, species-typical behavior that can be used by adults to convey social information in the absence of semantic understanding. English-speaking mothers recorded sentences with different "intentions" (i.e. prohibitive, approval, comfort, and attention) in both IDS and ADS, which were then played to non-industrialized, non-literate South American Shuar adults who had no experience with English (Bryant & Barret, 2007). The adults could correctly recognize the intent of speech in both ADS and IDS, but performed significantly better with IDS trials. This has been further tested with tonal (Thai) and non-tonal (Australian English) languages with both adults and infants; a tonal language relies on different tones or pitches in pronunciation to differentiate word meanings. While there are some differences between IDS of tonal and non-tonal languages, such as tonal being slightly more subdued compared to non-tonal, there is evidence for IDS being used universally with young infants from birth to 12 months of age (Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2001).

Recent studies have investigated infants' ability to match audio and visual information in native and non-native language when presented in ADS and IDS. In a study that controlled for temporal synchrony by showing the audio and visual information sequentially, German-learning infants were shown side-by-side silent videos of French and German speakers reciting a script, followed by a blank screen with the audio of one of the languages playing, then additional time viewing the silent talking faces (Kubicek, Gervain, de Boisferon, Pascalis, Lœvenbruck, & Schwarzer, 2014). During trials on which the speakers used ADS, the 12-month-olds did not look significantly longer at either face, regardless if the language was native or non-native. This similarity in looking indicates they were unable to correctly match the audio to the visual stimuli in ADS. When the actors were speaking in IDS, infants looked longer at the German face when hearing German audio, indicating they could correctly match their native language to the native speaker. However, they were unable to match the French speaker when the audio was French IDS, indicating they were unable to match the non-native language to the facial and mouth movements. The researchers hypothesized this may be due to a greater familiarity with native speech enhanced by the more exaggerated auditory and visual components of IDS. As non-native speech is not familiar, the infants had perceptually narrowed to German, even the additional information presented with IDS was not sufficient to facilitate audiovisual intersensory matching.

Similar to Kubicek and colleagues (2014) study, Kubicek, de Boisferon, Dupierrix, Pascalis, Lœvenbruck, Gervain, and Schwarzer (2014) also tested intersensory matching but this time used the additional measure of temporal synchrony and only ADS. Researchers tested 4.5- and 6-month-olds, and found 4.5-month-old infants could match both native (German) and non-native (French) speech when audio and visual information was presented sequentially. However, 6-month-olds only correctly matched native speech, indicating perceptual narrowing for intersensory matching of speech may occur between 4.5 and 6 months of age. In their second experiment, researchers presented the auditory and visual information simultaneously to provide temporal synchrony cues to 6- and 12-month-old infants. In the synchronous condition, 6-month-olds matched both native and non-native speech, but 12-month-olds only matched non-native speech. At first this seems counter-intuitive, but aligns with other eye tracking studies (i.e. Lewkowicz & Hansen-Tift, 2012) on how infants attend to talking faces through development. Around 6 months of age, infants are in the "babbling" phase and attending to the mouth of speakers regardless if the language is native or non-native, and at 12-months-of age infants have significant experience and attend to the eyes of native speakers, but continue to look at the mouths of non-native speakers (Lewkowicz & Hansen-Tift, 2012). Kubicek and colleagues' (2014) results indicate the 12-month-olds were attending to the mouths of non-native speakers and therefore could benefit from the intersensory redundancy provided by the temporal synchrony. However, they were most likely attending to the eyes of the native speakers, meaning they lost out on the additional intersensory information provided by the mouth movements synchronizing with the audio stream.

A recent study investigated the effects of sequential audio and visual information in intersensory matching of native and non-native speech at different age groups. Lewkowicz, Minar, Tift, and Brandon (2015) used adults speaking in IDS in either native or non-native speech and tested 4-, 8- to 10-, and 12- to 14-month-olds. Infants were shown the Spanish and English speaker videos silently, then with the audio track from one of the languages playing. For trials with synchrony, the audio playing matched the monologue being displayed. For trials without synchrony, the audio playing was from a different monologue than the one being displayed. The only group that looked longer at the corresponding face speaking either English or Spanish, indicating they matched the language of the audio to the video of the correct speaker, was the group of 12-14-month-olds. Results indicated that age group could match native audiovisual speech regardless of synchrony, and could match non-native speech with synchronous displays in the second block. This implies 12-14-month-old infants were experienced enough in their native language to not rely on intersensory redundancy to match English trials, and with intersensory redundancy and additional time to process the stimuli the infants could match Spanish trials as well.

The potential interaction of IDS and intersensory redundancy on infant selective attention to audiovisual speech has yet to be investigated. As IDS facilitates audiovisual matching in only some age groups, is dependent on if the language spoken is native or not, and depends on if the audiovisual stimulus is temporally synchronous or not; this study attempts to understand how these components relate in infant development. It is hypothesized that IDS facilitates intersensory perception even after intersensory perceptual narrowing of fluent audiovisual speech has occurred, both in the absence and presence of temporal synchrony. IDS may recruit infants' attention to the mouth of the speaker due the exaggerated mouth movements. When attending to the mouth, infants are able to benefit from intersensory redundancy when processing amodal information (Kubicek, de Boisferon, et al., 2014; Kubicek, Gervain, et al., 2014). The current study investigated whether IDS and intersensory redundancy working in conjunction to direct attention to the mouth aids infants' ability to notice a change in prosody in both native and non-native speech.

Gap in Knowledge

The purpose of this study was to investigate the possibility that intersensory redundancy and prosody modify 12-month-old infants' selective attention to video clips depicting native and non-native speech (Bahrick, et al., 2014). This research contributes to a relatively small body of work examining perceptual narrowing using dynamic multimodal stimuli. Understanding the range of variability that occurs in these tasks in typically developing infants will ultimately aid in the early identification of patterns indicative of atypical development. Another contribution of this study is the use of ecologically valid stimuli to examine intersensory perceptual narrowing in speech perception. Much of past work in the area has utilized static greyscale images of faces or speech paired with basic visual patterns as opposed to talking faces (e.g., Anzures, Wheeler, Quinn, Pascalis, Slater, Heron-Delaney, Tanaka, & Lee, 2012; Vogel, Monesson, & Scott, 2012; Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés 2009). Finally, no study to date has utilized non-native stimuli as increased task difficulty to examine the potential interaction of intersensory redundancy and prosody on infant audiovisual speech processing. Although perceptual narrowing in language occurs, how is this affected by redundancy? If infant-directed speech draws infants' attention to the mouth, and the mouth is the source of intersensory redundancy during audiovisual speech, could monolingual infants capitalize on these phenomena to detect a change in prosody in non-native speech during test trials (de Boisferon, Tift, Minar, & Lewkowicz, 2016; Kubicek, de Boisferon, et al., 2014; Kubicek, Gervain, et al., 2014)? If there is a lack of redundancy, would infants still notice a change in prosody in both native and non-native speech at 12 months of age?

Current Study

The current study used eye-tracking to explore how intersensory redundancy affects infant selective attention and infants' ability to detect a change in prosody in native and non-native speech after perceptual narrowing has begun. Twelve-month-old infants were shown a woman talking in either ADS or IDS, and halfway through each trial the speaker switched to the other prosody. Infants saw both English and Spanish speakers and trials were counter-balanced. Infants saw either all redundant (synchronous) clips or all non-redundant (asynchronous) clips. The distribution of infants' selective attention to the speaker's facial features was analyzed. Infant scanning patterns were also analyzed for evidence of detection of a change in prosody across conditions. Twelve-month-old infants were tested based on the Lewkowicz and Hansen-Tift's (2012) experiment that demonstrated this age was when infants tended to diverge when viewing non-native versus native redundant speech (looking longer at mouths versus looking equally at eyes and mouths). Twelve months is also an age when infants' attention is less focused on processing universal audiovisual information and instead processing unimodal information within bimodal events, making shifts to the mouths of speakers indicative of the increased task difficulty and a reliance on redundancy (Flom & Bahrick, 2007). The major aims of this study were focused on answering the following research questions:

1. Do 12-month-old infants attend to different facial features if there is intersensory redundancy as opposed to when there is a lack of redundancy? Is this affected by how familiar the infant is with the language being spoken?

I predicted monolingual English learning 12-month-old infants would focus equally on the eyes and mouth of the speaker in the temporally synchronous (redundant) English speaking clips (Lewkowicz & Hansen-Tift, 2012). By 12 months of age, English learning infants should be highly familiar with their native speech and not need to focus on the mouth as much for redundant information (Bahrick, Lickliter, Castellanos, & Vaillant-Molina, 2010).

Temporally asynchronous (non-redundant) English video clips should be more difficult for 12-month-olds to process. This may lead infants to rely on intersensory redundancy to process the stimuli, shifting their gaze to the mouth of the English speaker (Bahrick et al., 2010). Non-native Spanish trials should also be more difficult to process for infants due to lack of exposure and perceptual narrowing, which should lead infants to focus on the mouth of the Spanish speaker regardless of redundancy (Lewkowicz & Hansen-Tift, 2012).

2. Can 12-month-old infants notice a change in prosody when the speaker switches from ADS to IDS or vice versa? With the aid of redundancy, can monolingual English learning infants notice a change in prosody in Spanish? Will a lack of redundancy affect their ability to notice this change in either their native or a non-native language?

I predicted infants in viewing the synchronous (redundant) stimuli would demonstrate differential scanning based on prosody changes for both native and non-native speech (Lewkowicz, Minar, Tift, & Brandon, 2015). Infants who viewed the asynchronous (non-redundant) stimuli would only demonstrate differential scanning based on a prosody change in native (English) speech (Kubicek, Gervain, et al., 2014). There would be no significant difference in scanning between the ADS and IDS when viewing asynchronous non-native (Spanish) speech (Kubicek, de Boisferon, et al., 2014).

The intersensory redundancy provided by the synchronous condition should enable infants to detect amodal changes, such as when the speaker changes from one prosody to the other. Since the redundancy is providing two sources of sensory information, audio and visual, infants should be able to differentiate the change, even in non-native speech (Bahrick & Lickliter, 2000). However, the lack of intersensory redundancy in the asynchronous condition will be too difficult for infants to notice the change in non-native speech (Flom & Bahrick, 2007). They will rely on redundancy for the increased difficulty of processing an unfamiliar language ((Bahrick, Krogh-Jespersen, Argumosa, & Lopez, 2014). During English asynchronous trials, their native language will be familiar enough to notice the change in prosody unimodally (Bahrick, Lickliter, & Flom, 2004).

CHAPTER II METHODS

Participants

The final dataset included 16 White infants and 1 Black infant (12 males and 5 females) with a mean age of 365.06 days (*SD* = 6.31, *range* = 357 - 378) recruited from the Child Development Research Group participant database based at the University of Tennessee in Knoxville, Tennessee. Infants were tested within two weeks of their first birthday. The final dataset included infants who were born full-term (no less than 38 weeks gestation) without complications during pregnancy or delivery, had no known visual difficulties or other major health issues, and whose caregivers confirmed had no significant experience with a language other than English. Participants were recruited without regard to race, ethnicity, or gender. Additional infants were tested but excluded due to significant experience with Spanish or a language other than English (*N* = 3), inadequate number of trials (*N* = 7), inability to validate calibration due to technical difficulties (*N* = 3), or fussiness (*N* = 3).

Infants were randomly split into either the synchronous condition or asynchronous condition prior to testing. Within each redundancy condition, counter-balancing was done to control for potential prosody and language order effects. Of the infants in the final dataset, 9 were in the synchronous group and 8 were in the asynchronous group. Of the infants in the synchronous group, 5 saw ADS first and 4 saw IDS first. Of the infants in the asynchronous group, 3 saw ADS first and 5 saw IDS first.

Visual Stimuli

The stimuli were 30 second audiovisual video clips of a woman, either an English speaker or Spanish speaker, talking in either IDS or ADS, followed by an immediate switch to the other prosody for an additional 30 seconds. The stimuli from Lewkowicz and Hansen-Tift's (2012) experiment on native and non-native speech effects on point of gaze in infancy were used. Both the native and the non-native speaker had their hair pulled back from their face and were filmed against a plain black background. Redundancy refers to the temporal synchrony of the video clips of the women speaking. In synchronous conditions, the audio and video were synchronous (thus providing intersensory redundancy) and the video played normally. In asynchronous conditions, the audio was edited to not play in synchrony with the video, although both the audio track and the video onset occurred at the same time. Asynchronous videos were modified using the video editing software iMovie (version 10.1.8. Apple Inc. 2017. Mac OS X 10.12.6). The audio track was divided into two halves, then played in the opposite order. Infants heard the second half of the audio track followed by the first half. This ensured both the video and audio track began and ended at the same time and the speech was normal, but the visual and auditory components were not aligned and were not presented in synchrony. Infants still received the same audio and visual information, but did not have the added benefit of intersensory redundancy in the asynchronous trials.

Apparatus

Infants were seated in their parent's or guardian's lap as a comforting measure throughout testing. Participants were approximately 60 cm away from a 17" Dell UltraSharp 1704FPV color LCD monitor, which displayed the stimuli for testing. There were Dell desktop computer speakers hidden behind black fabric surrounding the monitor to provide sound presented at 55 dB during audiovisual trials without being a visible distraction. Testing took place in a sound-attenuated room that was darkened with light-blocking black curtains. Stimuli were presented through Experiment Builder software (SR Research Ltd., Mississauga, Ontario, Canada) on a Dell Windows PC. A remote eye-tracker (SR Research Ltd., Mississauga, Ontario, Canada, Eye-Link 1000 plus) was placed under the LCD monitor, facing the infant, to record eye scanning patterns with a 500 Hz sampling rate. The calibration of the eye-tracker was performed through a three-point calibration procedure. Animated cartoon characters moving and sounding in concert appeared pseudo-randomly at the top center, bottom-left, and bottom-right of the LCD monitor. Calibration was repeated until the infant attended to all three points, and a validation procedure of the same points confirmed calibration accuracy. A small bullseye pattern sticker was placed on the infant's forehead to aid the eye tracker with pupil localization for calibration and tracking scanning patterns.

Procedure

Caregivers of the infants were asked what languages were spoken at home and for an estimate of total possible exposure to non-native speech to determine if infants qualified as monolingual English learners. Infants with significant exposure to a language other than English, more than an occasional encounter, were excluded. This included infants who had a bilingual parent or caretaker or whose parents intentionally exposed the infant to a non-native language as part of learning enrichment. The experimenter went over the procedure and obtained informed consent from the parent or guardian. A certificate of participation was given as an incentive for taking part in the study. Caregivers were told they would receive the incentive even if they withdrew from the study before completion. Parents and legal guardians were never separated from their child during participation. The adult was instructed to sit holding the infant in front of the monitor. Once seated, a target sticker was placed on the center of the infant's forehead. The adult was asked to refrain from moving or making noise as the trials ran.

After calibration, an attention getter, e.g. an audiovisual clip of Sesame Street, was played to centrally fixate the participant. A trial consisted of a single speaker, either English or Spanish, speaking for 60 seconds, switching prosodies at 30 seconds. Following the 30 s of speech after the prosody switch, an attention getter was presented in the center of the monitor until the infant was centrally fixated. Then the second trial begin with whichever speaker the infant did not see first, playing for 60 seconds in the same prosody order as the first one. The experiment had 2 blocks comprised of 2 trials each. Both groups of infants viewed an English (native) speaker and a Spanish (non-native) speaker during Block 1, and the same order of language (English first or Spanish first) but a counter-balanced order of prosody for Block 2 (e.g. if infants heard ADS to IDS in Block 1, they heard IDS to ADS in Block 2). Order of language and prosody were counter-balanced across participants. Comparisons of English versus Spanish, pre- and post-change of prosody, and block number were made within subjects. Comparisons of redundancy and prosody order were made between subjects. Preliminary analyses were run to screen for potential effects of language, order, and sex. No significant effects were found, and thus all analyses reported were collapsed across language, order, and sex.

Measurement

To measure selective attention to facial features, both speakers' faces were divided into four areas of interest (AOIs). AOIs were defined as a horizontal rectangle around both eyes, a square around the nose, an oval around the mouth/jaw area, and an oval around the entire face. Proportion of looking was calculated by equating the total duration of fixations on the face to

1.00 and determining what proportion of fixation duration on the face occurred within each AOI (see Figure 1 and Figure 2).

To measure detection of prosody change, the difference between the last four fixations of the habituation prosody (first prosody shown) and the first four fixations of the test prosody (change that occurred after 30 seconds) was measured as an index of habituation/dishabituation. Fixations were calculated three different ways with similar results: as the sum of the last and first four fixations, as the averages of the last and first four fixations, and as the peak fixation of the last and first four fixations.

CHAPTER III RESULTS

Statistical Analysis Approach

Repeated measures ANOVAs were used for factorial analyses, and post-hoc comparisons were run using paired sample t-tests and independent sample t-tests. Effect sizes were reported using η_p^2, and the alpha level was set at .05 for all tests.

Facial Interest Areas

1. Do 12-month-old infants attend to different facial features if there is intersensory redundancy as opposed to when there is a lack of redundancy? Is this affected by how familiar the infant is with the language being spoken?

To analyze what areas of the face infant attention was drawn to depending on condition, a repeated measures ANOVA was conducted with Prosody (2: ADS, IDS) and Language (2: English, Spanish) as within-subjects factors, and Redundancy (2: Synchronous, Asynchronous) as a between-subjects factor for each AOI (3: eyes, nose, mouth).

There was a significant main effect of Prosody on infant selective attention to the eyes, $F(1,15) = 11.00$, $p < .01$, $\eta_p^2 = .42$. Infants looked significantly less at the eyes during infant-directed speech ($M = .15$, $SE = .04$) in comparison to adult-directed speech ($M = .25$, $SE = .05$).

There was a significant interaction of Prosody and Language on infant selective attention to the nose, $F(1,15) = 14.26$, $p < .01$, $\eta_p^2 = .49$. Post-hoc analysis using a paired samples t-test revealed significant difference between how infants looked at the nose during English infant-directed speech ($M = .15$, $SE = .03$) in comparison to English adult-directed speech ($M = .21$, $SE = .03$) ($p < .01$).

There was a significant main effect of Prosody on infant selective attention to the mouth, $F(1,15) = 8.04$, $p = .01$, $\eta_p^2 = .35$. Post-hoc analysis revealed infants looked significantly more at the mouth during infant-directed speech ($M = .50$, $SE = .08$) in comparison to adult-directed speech ($M = .39$, $SE = .08$). For significant results related to AOIs during adult-directed and infant-directed speech collapsed across both languages, see Figure 3.

There was also a significant interaction of Language and Redundancy on infant selective attention to the mouth, $F(1,15)$ 5.98, $p = .03$, $\eta_p^2 = .29$. However, post-hoc analysis using an independent samples t-test were not significant.

During English infant-directed speech, infants looked significantly more at the mouth ($M = .50$, $SE = .07$) than at the nose ($M = .15$, $SE = .03$) ($p = .001$) or eyes ($M = .13$, $SE = .04$) ($p < .01$), but they did not look significantly different at the nose compared to the eyes ($p = .68$). During Spanish infant-directed speech, infants also looked significantly more at the mouth ($M = .49$, $SE = .07$) than the nose ($M = .23$, $SE = .05$) ($p = .03$) or eyes ($M = .18$, $SE = .05$) ($p < .02$), but they did not look significantly different at the nose compared to the eyes ($p = .49$). To see infant-directed speech AOIs separated by language, see Figure 4.

During English trials depicting adult-directed speech, infants did not look at one AOI significantly more than another, i.e. eyes, nose, and mouth were not significantly different in proportion of looking times (eyes to mouth, $p = .52$; eyes to nose, $p = .34$; mouth to nose, $p = .15$). During Spanish trials depicting adult-directed speech, infants looked significantly more at the mouth ($M = .42$, $SE = .08$) than the nose ($M = .15$, $SE = .03$) ($p < .01$), but there was no significant difference in looking to the mouth compared to the eyes ($M = .24$, $SE = .05$) ($p = .15$), or the nose compared to the eyes ($p = .12$). To see adult-directed speech AOIs separated by language, see Figure 5. For visual heat maps of infant looking patterns to the faces by prosodies and language, see Figure 6.

Results show infants paid attention to the mouth of the Spanish speaker regardless of intersensory redundancy or prosody. As predicted, the novelty of the non-native language drew attention to the mouth regardless of other manipulations, which replicates what Lewkowicz and Hansen-Tift (2012) found. For the English trials, infants were drawn to the mouth during IDS and looked more generally over the face during ADS. This also replicates Lewkowicz and Hansen-Tift (2012) and again shows that infant attention is less affected by IR than by language and prosody.

Change of Prosody

2. Can 12-month-old infants notice the change in prosody when the speaker switches from ADS to IDS or vice versa? With the aid of redundancy, can monolingual English learning infants notice the change in Spanish? Will a lack of redundancy affect their ability to notice this change in either their native or a non-native language?

To analyze whether the infants noticed the change of prosody depending on condition, a 2 x 2 x 2 x 2 x 2 mixed design with Redundancy (synchronous or asynchronous) and Order (ADS first or IDS first) as between-subjects factors and Language (English or Spanish), Habituation (habituation or change trial), and Block (first or second) as within-subjects factors was used. The last four fixations to the face during the habituation trial and the first four fixations to the face during the change trial were compared to determine if infants displayed a recovery of looking to the prosody change (see Figure 1 and Figure 2 for the face AOI). Fixation differences were analyzed separately using the sum of four fixations, the average of four fixations, and the peak fixation of the four with similar statistical significance across analyses. The following results were found using the peak fixation of the final four fixations of the habituation trial and the peak fixation of the first four fixations of the change trial. Fixations were measured in milliseconds.

A significant interaction of Block, Habituation, and Order was found, $F(1,13) = 5.07$, $p = .04$, $\eta_p^2 = .28$. Post-hoc analysis revealed an interaction of Habituation and Order in Block 2, $F(1,13) = 8.31$, $p < .04$, $\eta_p^2 = .49$. Further analysis revealed an effect of Habituation for infants in the Synchronous condition who viewed the English speaker changing from ADS to IDS in Block 2, $F(1,3) = 13.57$, $p < .04$, $\eta_p^2 = .82$. Infants in this category showed a significant difference in time spent looking at the face between viewing the English ADS ($M = 640.50$, $SE = 106.72$) and the change to English IDS ($M = 2644.50$, $SE = 528.90$). For significant results related to infants in the Synchronous redundancy viewing the speakers change prosodies during Block 2, see Figure 7. For infants in the Asynchronous condition, see Figure 8.

These results indicate that infants did not detect a change in prosody in any condition other than those participants in the Synchronous condition viewing English ADS to IDS in the second Block of test trials ($N = 4$). Whether this is due to the majority of infants being unable to notice the change due to age or language experience, or due to infants not being sufficiently habituated is discussed further below.

CHAPTER IV DISCUSSION

The results indicate intersensory redundancy, prior experience with native speech, and presentation order affected infants' detection of prosody changes in their native language. The greater familiarity with English was enhanced by the exaggerated voice and facial expressions provided by IDS, which facilitated infants' ability to notice the change after sufficient experience with the trials, namely by Block 2 of testing. While I predicted infants would be able to discriminate a change of prosody in the synchronous condition, I hypothesized they would be able to do so for both native and non-native speech. Instead, results showed infants were only able to do so in native speech. Additionally, I predicted the infants would be able to discriminate a change in asynchronous native speech, but results do not support this hypothesis. Infants in the asynchronous condition did not notice the prosody change in native or non-native speech.

Infant selective attention to facial features was not affected by redundancy but by prosody. I predicted asynchronous native speech would cause infants to look more at the mouth to process amodal information during a lack of redundancy, but the results did not support this prediction. Infants who viewed native speech only differed in looking patterns based on prosody, not redundancy. However, my prediction that infants would look more at the mouth of the non-native speaker regardless of redundancy was supported by the data.

These selective attention results may be due to the age group tested. The study mentioned earlier by de Boisferon, Tift, Minar, and Lewkowicz (2016) investigated the effects of desynchronization on infant selective attention to native and non-native audiovisual speech in both IDS and ADS. The study indicates the 12-month-olds' visual attention was not disrupted by asynchrony in comparison to synchronous stimuli. Since synchrony did not seem to have an effect on attention, it may be the case that intersensory redundancy does not have an effect at this age as well. Without utilizing the benefits of redundancy, the infants may not have been able to discriminate prosody changes in asynchronous speech.

Given these studies, infant selective attention to features may not be affected by presence or lack of intersensory redundancy at 12 months of age. Infants have sufficient experience with processing audiovisual speech at 12 months and do not rely as heavily on amodal information given by the mouth (Cheour et al., 1998; Mauer & Werker, 2013; Werker & Tees, 1984). The salience of IDS versus the more neutral ADS instead drove the changes in selective attention these data show, regardless of synchrony or language (Fernald & Mazzie, 1991; Graf Estes & Hurley, 2013).

Limitations and Future Research

A few limitations to the current study are worth noting. Sample size could be increased considering the large number of factors analyzed. Additional infants should be recruited for each of the conditions. Infant-directed speech is also highly salient. In comparison to the much more neutral affect characteristic of adult-directed speech, the difference in saliency may have affected infant attention above and beyond the effects of redundancy or language. A study examining differences in affect, such as happiness or anger, conveyed exclusively through either IDS or ADS, could potentially provide a more sensitive test for intersensory facilitation by controlling for the salience of prosody across conditions.

As described earlier, de Boisferon, Tift, Minar, and Lewkowicz (2016) found that 12-month-olds' selective attention to facial features is not as affected by intersensory redundancy as 10-month-olds'. A future direction could use 10-month-old infants in the same experimental design to investigate how synchrony affects attention and ability to discriminate a change of prosody. Given the results of the change of prosody, it seems 12-month-olds required the additional block of exposure before noticing the change. If experimenting with 12-month-olds again or even 10-month-olds, trials should be longer to allow sufficient time to process the stimuli. Perhaps 60 s monologues would provide infants more time to fully process the complex stimuli.

Conclusions

Perceptual narrowing in language development is not cut and dry. While many studies have shown infants have “narrowed down” to their native language by their first year of life (e.g., Cheour et al., 1998; Mauer & Werker, 2013; Werker & Tees, 1984), many other factors influence how and where infants’ attention is directed, affecting what perceptual properties they attend to. Intersensory redundancy, prosody, and prior experience with native language all interact to affect perceptual processing priority. Infants’ ability to discriminate a change of prosody and selective attention to facial features is not dependent on a single factor, such as whether the language is native or not, but rather a combination of many factors influencing one another. Although some infants did pick up on a change in prosody in their native language, infants in other conditions did not show they noticed the change despite the actor still speaking English (i.e. in Block 1 of trials, in the asynchronous condition, and in the IDS to ADS prosody order).

After perceptual narrowing has occurred, monolingual infants do not need to rely on looking at the mouths of speakers to understand audiovisual speech, but this can be moderated by prosody, infant age, intersensory redundancy, and familiarity with a native language. While redundancy was useful to the infants viewing their native language, it did not help infants viewing a non-native language detect a change in prosody. Despite the infants' extensive experience with English, those who viewed videos without intersensory redundancy showed no indication they were able to differentiate the change in prosody. However, infants in both redundancy conditions did look differently at the English speaker's face between IDS and ADS. Infant attention and cognition are complex processes. While some data from this study shows infants paid attention to different facial features depending on prosody, other data clarifies that they did not necessarily discriminate the changes in prosody. Instead, the infants paying more attention to mouths during IDS and the rest of the face during ADS indicates selective attention is affected by prosody, but a higher order processing of the change of prosody may not be happening in most of the conditions. Those viewing synchronous English changing from ADS to IDS in the second block were the only participants in this study to demonstrate differential selective attention based on prosody in addition to significant visual recovery indicating an awareness of the prosody change.

Infants' ability to discriminate a change of prosody relies on familiarity of the language, the age of the infant, the time the infant has to process the stimuli, and whether the stimuli has intersensory redundancy. Future studies should investigate these factors further by controlling for and manipulating different conditions among different age groups. This will help control for differences in infant selective attention versus what could be an indication of a discrimination of prosody. Understanding this network of influences will provide a more comprehensive view of infant development, perception, audiovisual speech processing, and perceptual narrowing of language.

REFERENCES

Adriaans, F., & Swingley, D. (2012). Distributional learning of vowel categories is supported by prosody in infant-directed speech. Proceedings of the Cognitive Science Society, 34, 72-77.

Anzures, G., Wheeler, A., Quinn, P. C., Pascalis, O., Slater, A. M., Heron-Delaney, M., Tanaka, J. W., & Lee, K. (2012). Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. *Journal of experimental child psychology, 112*(4), 484-495.

Bahrick, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. *Blackwell handbook of infant development, 1,* 120-165.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental psychology*, *36*, 190-201.

Bahrick, L. E., Flom, R., & Lickliter, R. (2002a). Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Developmental psychobiology, 41*(4), 352-363.

Bahrick, L. E., & Lickliter, R. (2002b). Intersensory redundancy guides early perceptual and cognitive development. *Advances in child development and behavior, 30*, 153-189.

Bahrick, L. E., & Lickliter, R. (2004). Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis. *Cognitive, affective, & behavioral neuroscience, 4*(2), 137-147.

Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. *Multisensory development,* 183-206.

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. Current Directions in Psychological Science, 13, 99-102.

Bahrick, L. E., Krogh-Jespersen, S., Argumosa, M. A., & Lopez, H. (2014). Intersensory redundancy hinders face discrimination in preschool children: Evidence for visual facilitation. Developmental psychology, 50(2), 414.

Bahrick, L. E., Lickliter, R., & Castellanos, I. (2013). The development of face perception in infancy: Intersensory interference and unimodal visual facilitation. Developmental Psychology, 49, 1919-1930.

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. Current Directions in Psychological Science, 13, 99-102.

Bahrick, L. E., Lickliter, R., Castellanos, I., & Vaillant-Molina, M. (2010). Increasing task difficulty enhances effects of intersensory redundancy: Testing a new prediction of the intersensory redundancy hypothesis. Developmental Science, 13, 731-737.

Bryant, G. A., & Barrett, H. C. (2007). Recognizing intentions in infant-directed speech: Evidence for universals. Psychological Science, 18(8), 746-751.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature neuroscience*, *1*(5), 351-353.

Colombo, J. (2001). The development of visual attention in infancy. *Annual review of psychology*, *52*(1), 337-367.

de Boisferon, A. H., Tift, A. H., Minar, N. J., & Lewkowicz, D. J. (2016). Selective attention to a talker's mouth in infancy: Role of audiovisual temporal synchrony and linguistic experience. *Developmental science*, *20*(3), 1-12.

Fernald, A. & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant behavior and development*, *10*(3), 279-293.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. Developmental psychology, 27(2), 209.

Flom, R. & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. Developmental Psychology, 43, 238-252.

Graf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. Infancy, 18(5), 797-824.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological science, 18*(12), 1084-1089.

Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2001). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. Infant behavior and development, 24(4), 372-392.

Kubicek, C., de Boisferon, A. H., Dupierrix, E., Pascalis, O., Lœvenbruck, H., Gervain, J., & Schwarzer, G. (2014). Cross-modal matching of audio-visual German and French fluent speech in infancy. PLoS One, 9(2), e89275.

Kubicek, C., Gervain, J., de Boisferon, A. H., Pascalis, O., Lœvenbruck, H., & Schwarzer, G. (2014). The influence of infant-directed speech on 12-month-olds' intersensory perception of fluent speech. Infant Behavior and Development, 37(4), 644-651.

Lewkowicz, D.J. (2010). Infant perception of audio-visual speech synchrony. *Developmental psychology, 46,* 66–77.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. Trends in cognitive sciences, 13(11), 470-478.

Lewkowicz, D. J. & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. Proceeding of the National Academy of Sciences 109.5 (2012): 1431-1436.

Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. Infancy, 15, 46-60.

Lewkowicz, D. J., Minar, N. J., Tift, A. H., & Brandon, M. (2015). Perception of the multisensory coherence of fluent audiovisual speech in infancy: Its emergence and the role of experience. *Journal of experimental child psychology, 130,* 147-162.

Maurer, D. & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. Developmental Psychology, 56, 154-178.

McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence?. Cognition, 129(2), 362-378.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life?. *Science, 296*(5571), 1321-1323.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human perception and performance, 20*(2), 421.

Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *Journal of neuroscience, 26*(18), 4970-4982.

Pons, F., & Lewkowicz, D.J. (2014). Infant perception of audiovisual speech synchrony in familiar and unfamiliar fluent speech. *Acta psychologica*, *149,* 142–147.

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences, 106*(26), 10598-10602.

Reynolds, G.D., Courage, M.L, & Richards, J.E. (2013). The development of attention. In D. Reisberg (Ed.), *Oxford handbook of cognitive psychology,* (pp. 1000 – 1013). Oxford University Press: New York, NY.

Ruff, H. A., & Rothbart, M. K. (2001). *Attention in early development: Themes and variations*. Oxford University Press: New York, NY.

Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. Current Directions in Psychological Science, 16, 197-201.

Slater, A., & Kirby, R. (1998). Innate and learned perceptual abilities in the newborn infant. *Experimental Brain Research*, *123*(1), 90-94.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. Infancy, 7(1), 53-71.

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion?. Psychological science, 11(3), 188-195.

Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the national academy of sciences, 104*(33), 13273-13278.

Vogel, M., Monesson, A., & Scott, L. S. (2012). Building biases in infancy: The influence of race on face and voice emotion matching. *Developmental science, 15*(3), 359-372.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*(1), 49-63.

APPENDIX

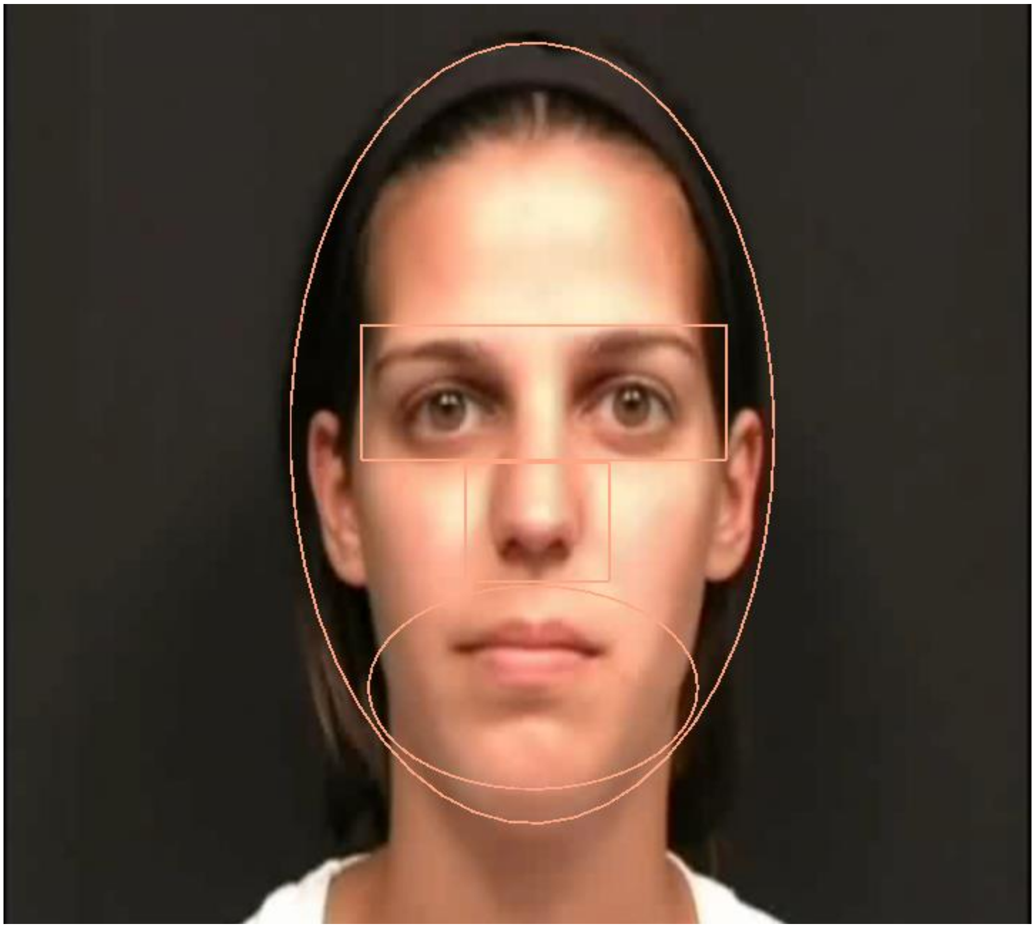

Figure 1. Static image of English speaker AOIs from video clip.

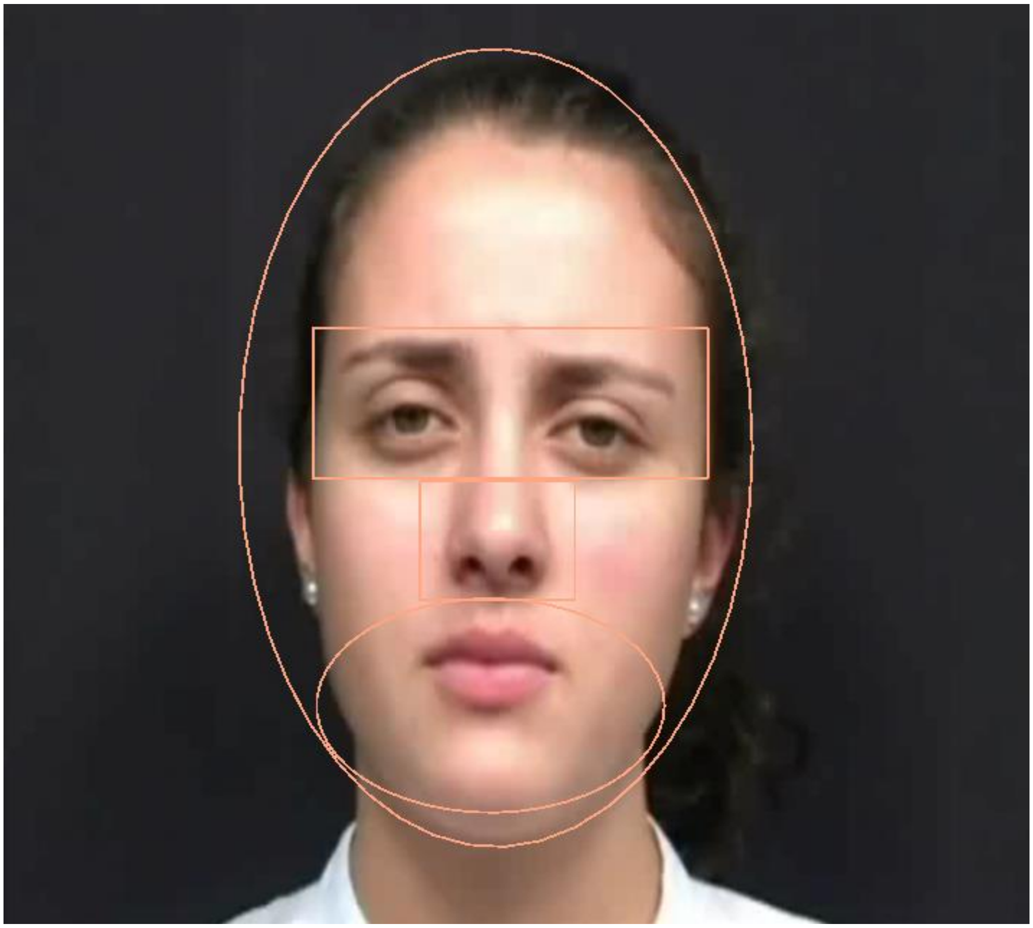

Figure 2. Static image of Spanish speaker AOIs from video clip.

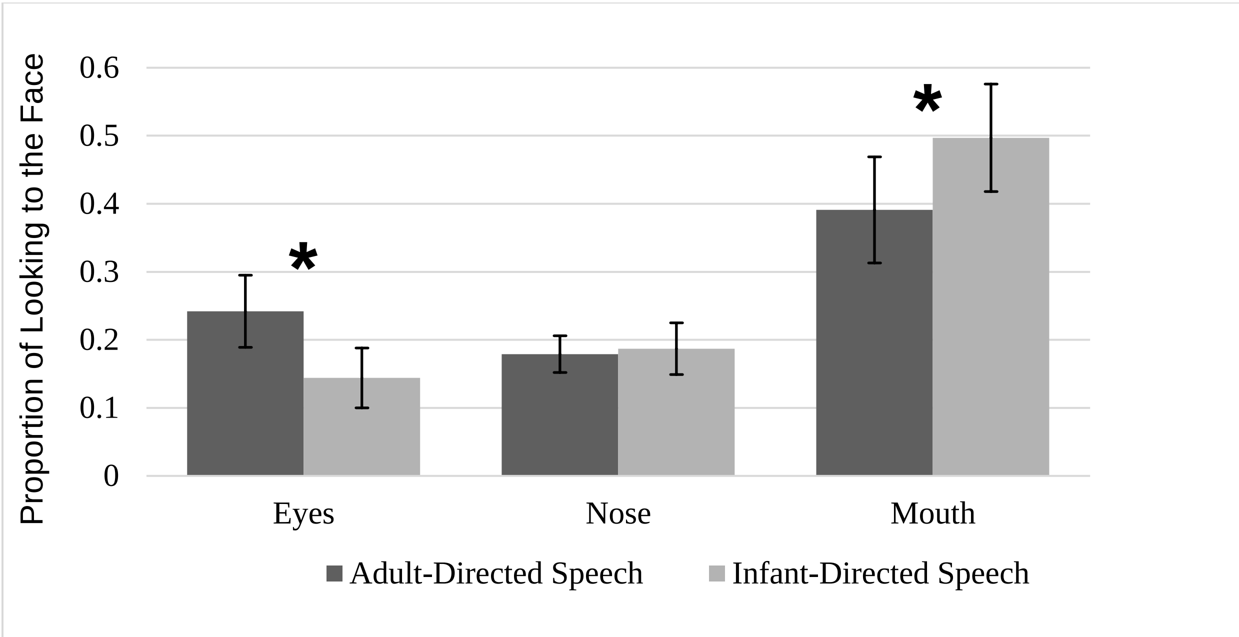

Figure 3. AOIs during adult-directed and infant-directed speech, collapsed across both languages.

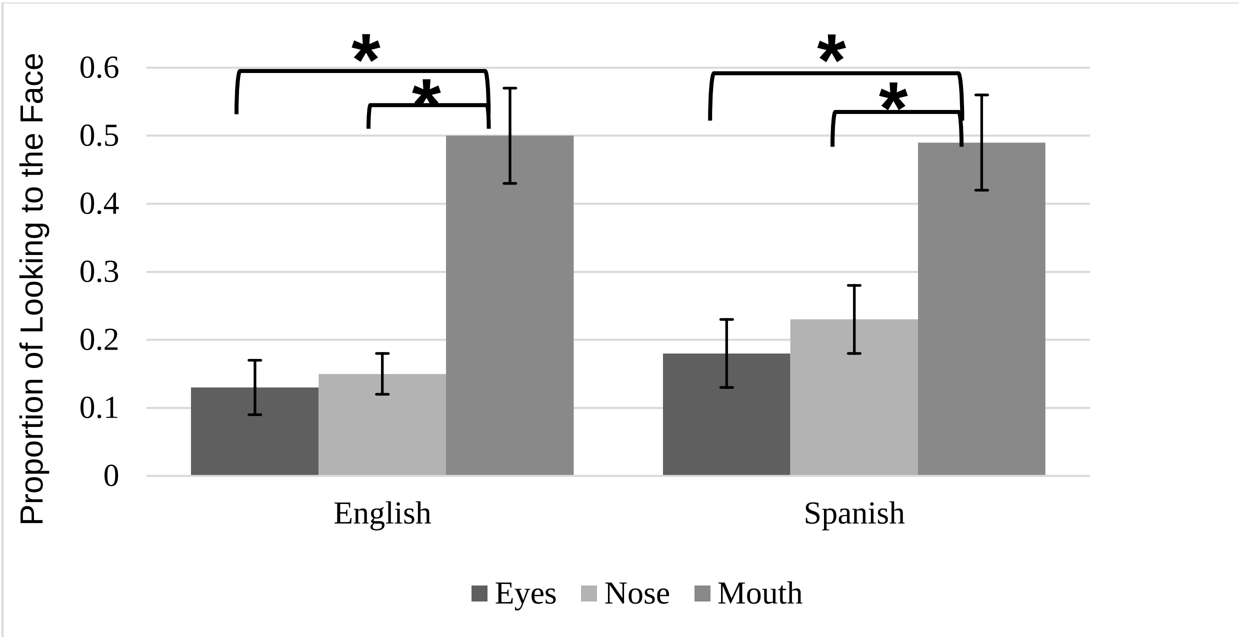

Figure 4. AOIs during infant-directed speech by language.

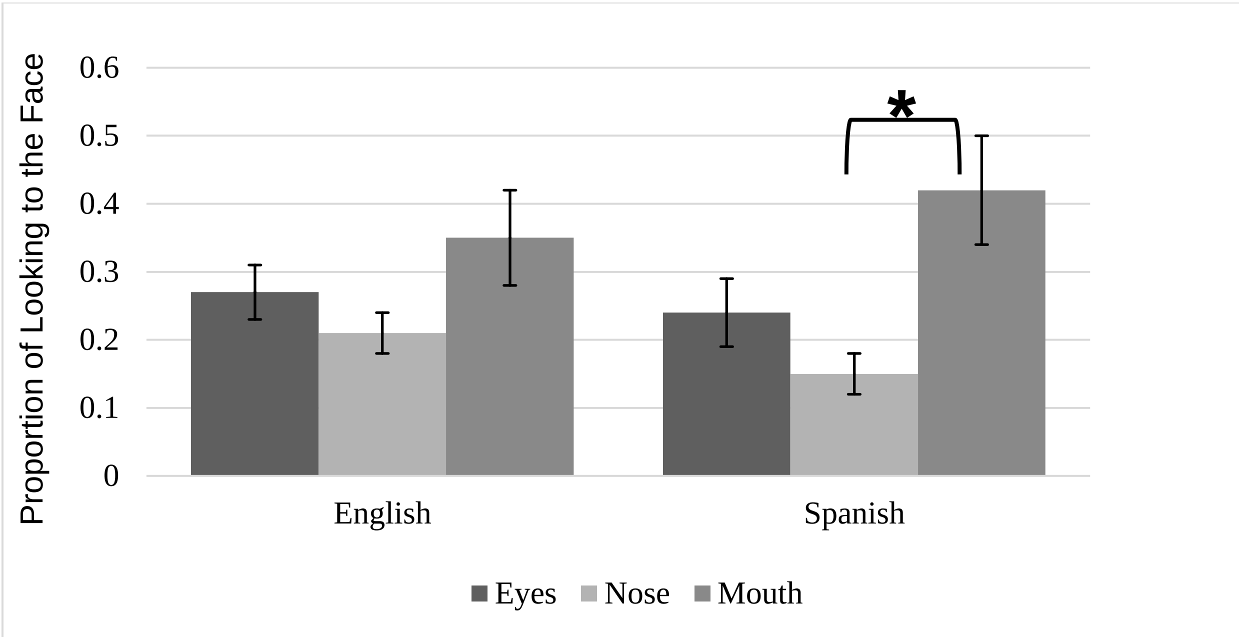

Figure 5. AOIs during adult-directed speech by language.

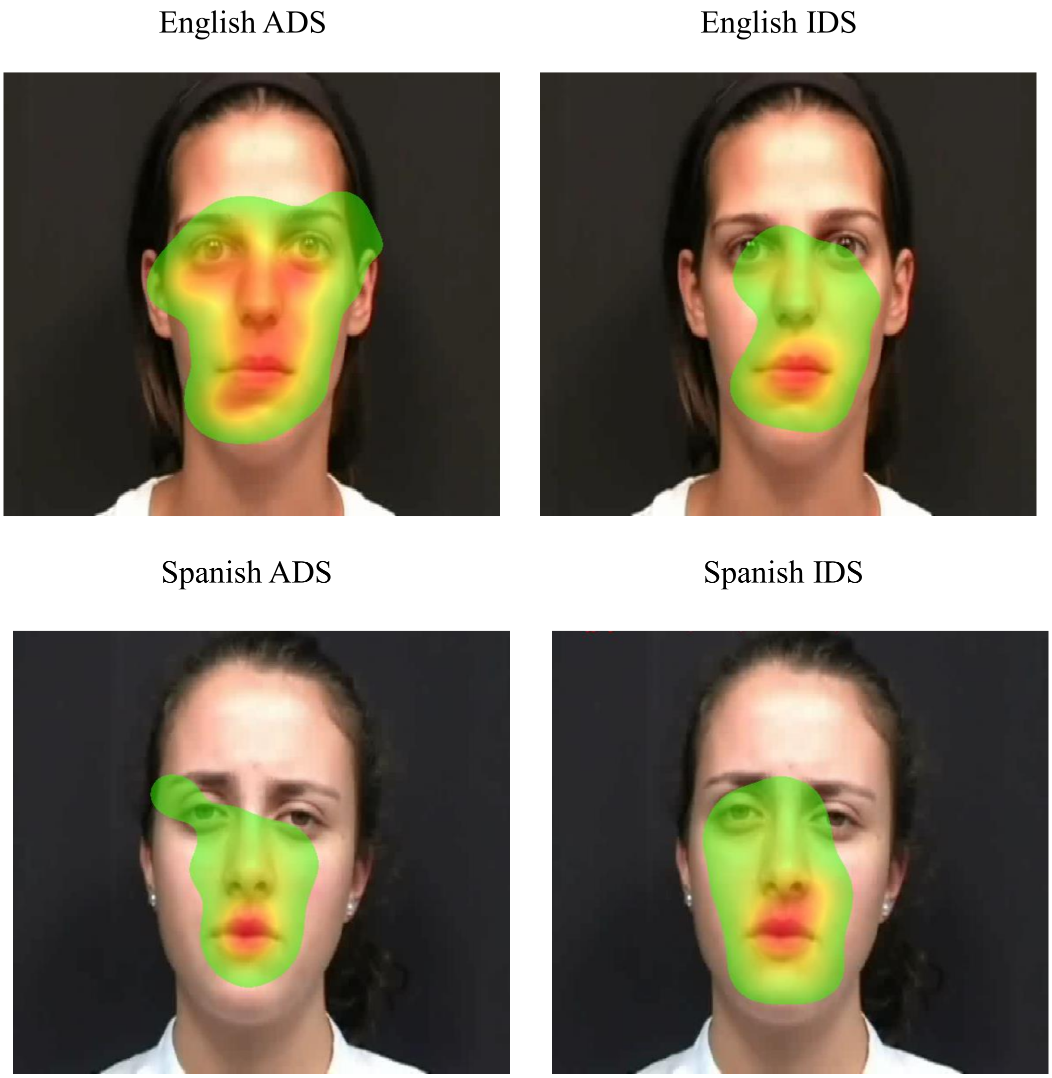

Figure 6. Visual heat maps of infant looking patterns to the faces by prosodies.

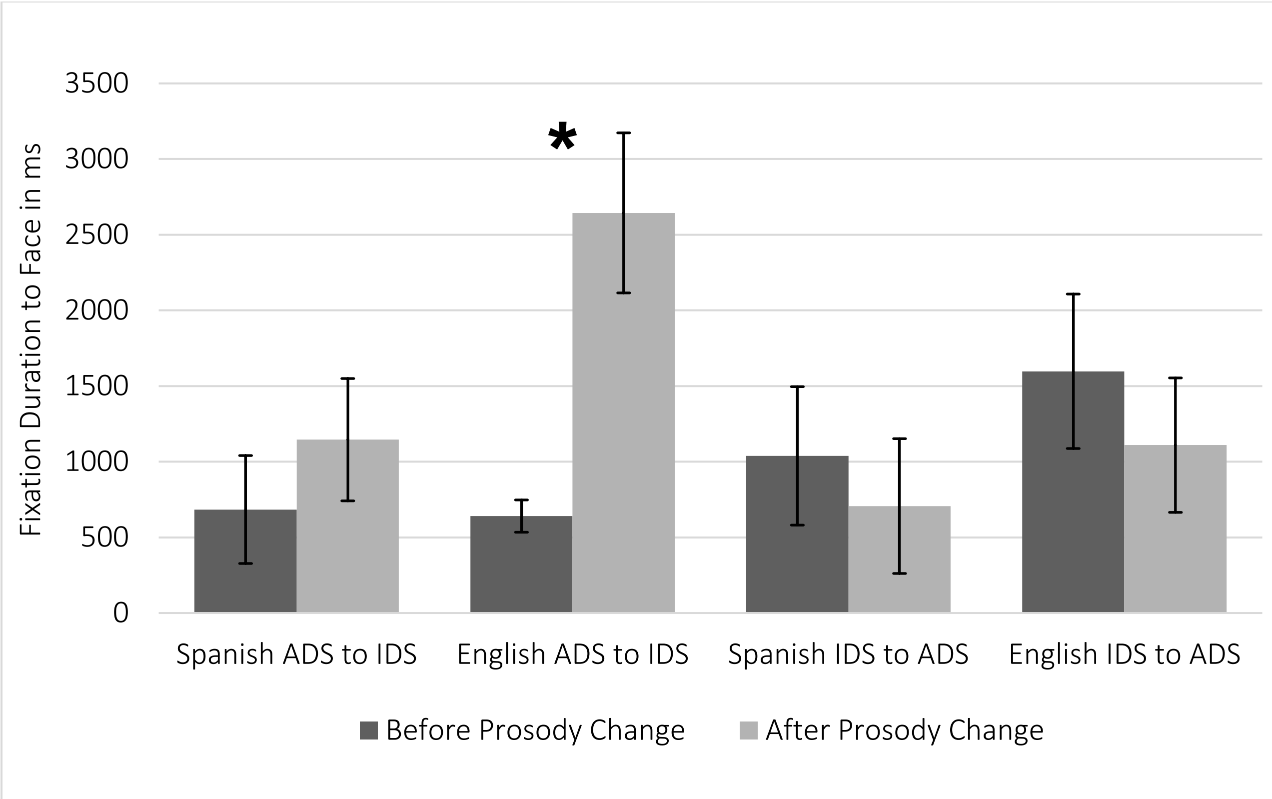

Figure 7. Infants in the Synchronous condition viewing the speakers change prosodies speech during Block 2.

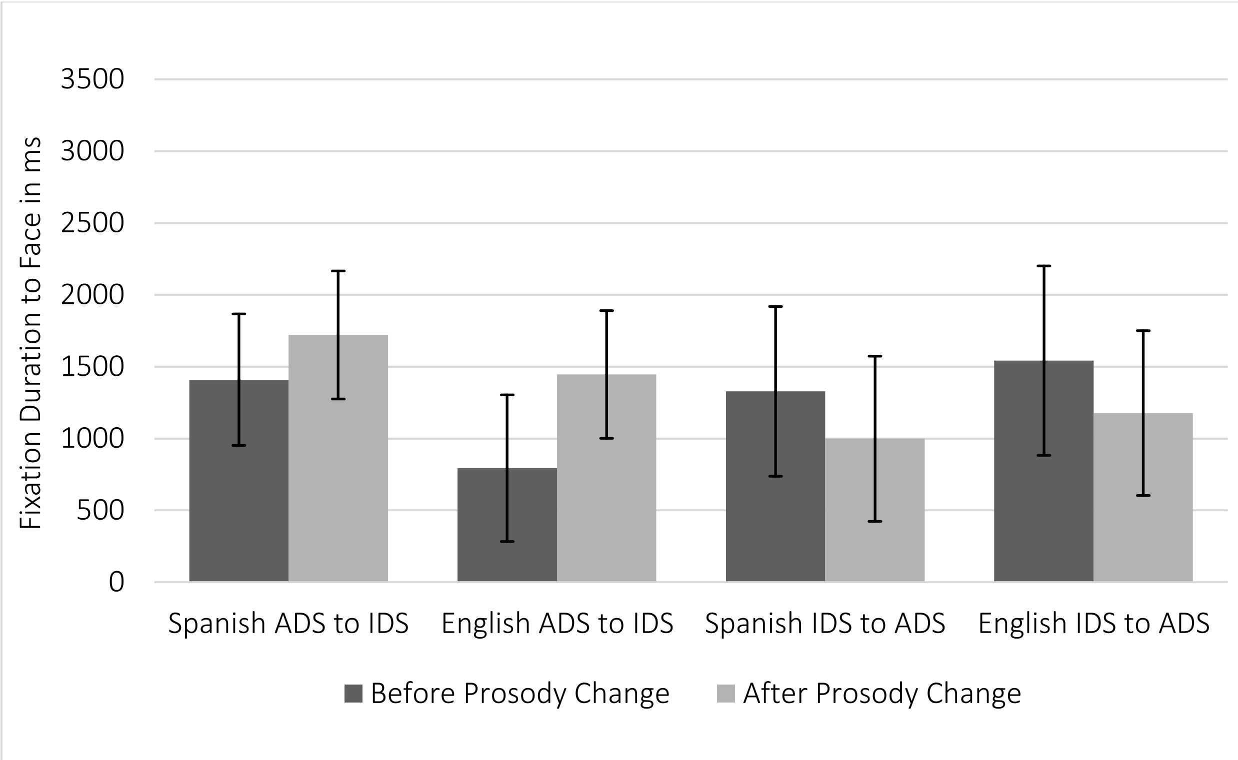

Figure 8. Infants in the Asynchronous condition viewing the speakers change prosodies during Block 2.

VITA

Kelly Roth attended the University of Tennessee in Knoxville, Tennessee, where she received her Bachelor of Arts degree in Psychology Honors. She graduated cum laude through the Chancellor's Honors program in three years. She continued her education at UT by pursuing a master's degree in Experimental Psychology with a minor in Statistics. She has since applied and been accepted to continue her studies as a doctoral student in the Experimental Psychology program at UT.