PHONEMIC DISCRIMINATION AND EYE-MOVEMENTS IN INFANTS

SHIR BACH – KAY

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

December 2021

Abstract

The ability to discriminate between different phonemes is a crucial part of language development in the first year of life. While language acquisition is a process that has been studied in both infants and adults in the past, the paradigms that were used to study this sensitive process have a number of shortcomings. To overcome these shortcomings, the present study examined 6-month-old infants' ability to discriminate between two different phonemes by means of an eye-tracking task, the Visual Expectation Cueing Paradigm (VExCP). In this paradigm, one randomly presented phoneme (paired with a central visual stimulus) predicted a visual target on the right side of a monitor screen and the other randomly presented phoneme predicted a visual target on the left side of the screen. If the infants could discriminate between the different phonemes then they would be able to correctly make anticipatory eye movements to the target location at a rate above chance. Results indicated that 6-month-old infants successfully discriminated between the two different phonemes forming an expectation for the phoneme-target location relations, and thereby making correct anticipatory eye-movements to the correct target location at a rate greater than chance. The findings indicate that the VExCP is an appropriate paradigm for the study of phonemic discrimination while overcoming the weaknesses of previously used paradigms.

Table of Contents

List of Figures

Introduction

The attainment of language is a key milestone during early development that enables advances in varied cognitive functions including memory and social interaction, among others. The sequence for language acquisition is similar to many cognitive functions in that there is differential development with the processing of lower-level components preceding and perhaps necessary for, the processing of higher-level cognitive components. For example, developmentally, categorization of objects based on their perceptual properties precedes categorization on the basis of their conceptual framework (Mack & Palmeri, 2011). Additionally, the development of object perception demonstrates this principle as the perception of basic features, such as edges, boundaries, and contour preceding the perception of more complex features, such as texture, colour, and shape (Feldman, 2003; Hauffen et al., 2012). Infants' face perception also shows a similar developmental hierarchy, with perceptual focus on the eyes and hairline developmentally preceding perception of a whole and complete face, and subsequently being able to discriminate faces that differ in spatial distance between the eyes (Tanaka et al., 2014; Bhatt et al., 2005).

Selective attention also shows the same developmental pattern with the functioning of bottom-up attentional mechanisms developmentally preceding the functioning of top-down mechanisms (Adler & Wong-Kee-You, 2015; Comishen, Bialystok & Adler, 2019; Amso & Scerif, 2015; Adler et al., under review). That is, the neural and attentional mechanisms that support bottom-up selection (externally/stimulus driven, from stimuli in the environment, sensory based) are functional in the first few months of life. The mechanisms supporting top-down selection, in contrast, do not seem to begin to develop until after 3 months postnatally and have an exaggerated timeline to reach full maturity, perhaps years (Adler & Wong-Kee-You, 2015; Atkinson, 1984, 2000; Johnson, 1995, 2002; Braddick & Atkinson, 2011). All of these examples illustrate that for many perceptual and attentional capacities, lower systems develop earlier and are necessary for the development of the higher order systems, meaning, earlier developing mechanisms form a foundation on which later developing mechanisms are built. Similarly, for language perception, the development of the capacity to distinguish phonemes occurs prior and supports the development of higher order language development such as the comprehension and production of words (Narayan, 2013).

An initial step in language development in early infancy prior to being able to comprehend and produce language, is the capacity to detect and discriminate the basic units of language, called phonemes (Chládková & Paillereau, 2020). A phoneme can be characterized as the smallest unit of sound in a language that carries meaning, such that changing a phoneme changes the meaning of the word in which it is embedded. As all words are constructed from phonemes (Bialystok, 2001), the ability to discriminate between different phonemes, therefore, is important before infants start to use actual words for communication (in order to differentiate between different but similar sounding words and for associating those different words with objects). For example, the word "cat" has three different phonemes/sounds: /k/ /a/ /t/. Changing the first phoneme from /k/ to /b/ will result in the word "bat" -- /b/ /a/ /t/, which has a different meaning from the word "cat". Thus, as the example illustrates, being able to discriminate between closely related phonemes is necessary for being able to discriminate between words and their meaning.

The capacity to distinguish phonemes is present early in life as infants seem to be born with the ability to distinguish between all phonetic sounds that might occur in any language (Cheour et al., 1998). In a classic study, Eimas et al. (1971), for example, found that infants as young as 1–4-months-old were able to discriminate between the phonemes /b/ and /p/. Even infants with little experience (as young as 1-month-old) were not only responsive to the phonemes they heard but could also make fine discriminations between them. Thus, demonstration of phonemic discrimination at such an early age suggests that this capacity might be innate. By approximately 12 months of age, however, the ability to discriminate any phoneme from any language narrows. Infants then begin to only distinguish sounds/phonemes in languages to which they have had exposure (Larraza, Molnar, & Samuel, 2020). That is, near the end of the first year of life (approximately 10 months of age), infants can still detect phonemic distinction but more narrowly in that they are readily able to discriminate between phonemes but mostly only from the language spoken around them. This narrow distinction is consistent with what is seen with adults; namely, they are able to essentially discriminate only between phonemes that are part of their language (Werker & Tees, 2002). Narrowing is further illustrated by the findings of Kuhl et al., (2006) in which 10- to 12-month-old infants showed a decrease in their ability to distinguish between phonemes that were not part of their native language, compared to the abilities of 6- to 8-month-old infants who were able to discriminate between phonemes that are not part of their native language. Younger infants had an even higher level of correct response percentages in phonemic discrimination compared to older infants on a task with phonemes that were not from their native language. This demonstrates that phonemic discrimination progressively narrow across development, particularly in the first year.

Narrowing of infants' phonemic discrimination over the first year is likely related to general principles of narrowing that govern many perceptual capacities. Initially, after birth, brain pathways exist to process and discriminate any perceptual input, including phonemic, that might be encountered (Chládková & Paillereau, 2020). Through experience with specific perceptual input, these perceptual pathways narrow in the input to which they are sensitive, resulting in the loss of the ability to discriminate, for example, between phonemes that are not part of one's native language and preferentially shaping perceptual discriminatory abilities around that experienced input (Scott & Monesson, 2010). By narrowing perceptual capacities, those abilities and the efficiency of the system (e.g., decrease in reaction time) to the items the individual typically experiences, improve. Concurrently, there is a reduction in the perceptual abilities (e.g., reduction in discriminability capacities) and efficiency to the unused (e.g., the pathways that code perceptual input that is not experienced) pathways (Lewkowicz & Ghazanfar, 2009; Scott, Pascalis, & Nelson, 2007; Scott & Monesson, 2010). This process of perceptual narrowing is most evident during a sensitive period of development (e.g., infancy) when many perceptual systems are first being established and become instantiated through neuroplasticity (Tierney & Nelson, 2009).

Perceptual narrowing is not limited to auditory information (e.g., phoneme discrimination) but seems to be a general developmental process that occurs for a number of sensory registers. Visual information (e.g., face perception), which is important for social behaviour, has been shown to undergo perceptual narrowing as well (Pascalis, De Haan, & Nelson, 2002). For example, 6-month-olds, 9-month-olds, and adults are equally capable of discriminating between two human faces, whereas only 6-month-olds were able to discriminate between two monkey (non-human) faces (Pascalis, De Haan, & Nelson, 2002). The findings by Pascalis et al. (2002) demonstrate that 6-month-olds can discriminate any type of face, whether experienced or not, whereas this capacity subsequently narrows to being able to only discriminate the type of faces with which they have experience.

As already described, infants' language development, similar to the development of face perception and other perceptual capacities, goes through a crucial period where infants' phonemic discriminatory abilities narrow according to their experience (Lewkowicz & Ghazanfar, 2009) – a process that might be a necessary precursor for higher order language skills. In order to more completely understand the course of language development, therefore, it is important to study phonemic discrimination and its narrowing in infants as these developmental processes are occurring. This developmental process takes on particular significance as phonemes are the foundation upon which higher order language components are based (Narayan, 2013; Chládková & Paillereau, 2020). Because the initial stage of language development (the narrowing of phonemic discriminatory abilities) is happening during a very sensitive and rapid time of development, the measures by which data is collected therefore, need to be sensitive and precise. Previous studies that have investigated phonemic discrimination and its development have used a number of different methodologies, paradigms, and measurements (Aslin, 2007). The goal of the present study is to introduce a new paradigm, the Visual Expectation Cueing Paradigm, into the mix that will address some of the shortcomings of previously used paradigms.

One typically used paradigm for assessing phonemic discrimination has been preferential looking. Preferential looking assesses whether, when an infant is simultaneously presented with two different stimuli (e.g., a patterned stimulus with the characteristic of interest paired with a homogenous stimulus), the infant looks at one of them longer. If the infant does preferentially look longer to one of the stimuli that is interpreted as indicating that the infant was able to differentiate between the two stimuli and the critical features (Golinkoff, Song, & Hirsh-Pasek, 2013). As is typical, using the preferential looking paradigm to assess auditory discrimination, no training is required, and head-movement/head-turning or eye gaze are used as the dependent measures. Humphrey and Tees (1980), for example, using preferential-looking, tested 3-, 6-, and 10-month-old infants' ability to discriminate auditory stimuli (slow tempo versus fast-tempo auditory patterns). Infants were presented with two possible visual targets and the auditory stimuli were temporally synchronized to one of the visual targets. Humphrey and Tees found no significant difference in the ability to discriminate between slow/fast-tempo in 3- and 6-month-old infants and only a small effect for the 10-month-old infants. They concluded that this method was not sensitive enough to study the development of sound discrimination, particularly in very young infants. Another study used the preferential-looking paradigm to test phonemic discrimination in 14-month-old infants (Stager & Werker, 1997). In this study, the phonemes /dih/ and /bih/ were associated with two coloured objects. Stager and Werker defined a differentiation between the phonemes based on the looking time at one object compared the other. Analysis revealed no significant different in looking time to the two phoneme-associated coloured objects, indicating that the infants were unable to discriminate between the phonemes. In contrast, when they tested 8-month-old infants the researchers found a significant difference in looking time, indicating that the infants could discriminate between the two different phonemes.

One general problem that has been consistently identified with the preferential-looking paradigm is that if the infant does not look longer at one stimulus or the other then it cannot be unambiguously stated whether or not the infant was able to discriminate between the stimuli, as it is possible that the infant can discriminate but just does not prefer one stimulus over the other (Aslin, 2007). Thus, the absence of preference for one stimulus does not necessarily imply that infants are unable to discriminate between the stimuli. Furthermore, as stated by Aslin (2007), preferential-looking paradigms with head-movement as the dependent measure are not sensitive enough to pinpoint the underlining cognitive developmental processes, especially in young infants, because the motor system is not fully developed (e.g., harder to move their heads/takes them longer to move their heads). Consequently, findings based on this measure might be due to infants' underdeveloped motor system, or/and their preference of one visual target stimulus over the other (or no difference in preference between the target stimuli).

Two other related looking paradigms have also been used to assess phonemic discrimination in young infants, habituation and familiarity/novelty-preference. In these paradigms, which do not assume that infants have an inherent preference for one stimulus over the other, infants are trained by repeatedly exposing them to the same stimulus until they habituate to that stimulus by exhibiting a 50% decrease in looking time (habituation paradigm) or become familiarized to it by exhibiting a significant decrease in looking time after a pre-set number of trials (Aslin, 2007). The primary measures of discrimination are whether the participants dishabituate or exhibit a recovery (increase) in looking time when subsequently exposed to a novel stimulus after habituation and whether a greater of looking is allocated to either of simultaneously presented novel and familiar stimuli in familiarity/novelty-preference (Thomas & Gilmore, 2004). Singh, Loh and Xiao (2017), for example, used the habituation paradigm to assess phonemic discrimination in 10- to 11.5-month-old monolingually and bilingually exposed infants. They found that infants in a bilingual environment were able to discriminate between phonemes that are not part of the language spoken to them whereas monolingual infants could not discriminate. In addition, Maye, Werker and Gerken (2002) tested both 6- and 8-month-old infants' discrimination abilities for the phonemes /da/ and /ta/ using a familiarity/novelty-preference paradigm. Both the 6- and 8-month-old infants could discriminate between these phonemes based on significantly longer looking time at the novel target stimulus.

Roder, Bushnell and Sasseville (2000), however, also using familiarity/novelty-preference, showed that 4.5-months-old infants preferred a familiar stimulus (e.g., different faces and different objects) as much as a novel stimulus. The contradictory findings show that familiarity/novelty-preference might not be sensitive enough in detecting whether an infant discriminated between two phonemes or not.

Though these looking paradigms have been instrumental in describing perceptual capacities over the last half century, there are critical problems that should give pause to accepting their theoretical implications (Thomas & Gilmore, 2004). In addition to the issue with preferential-looking mentioned above, familiarization and looking paradigms suffer from a problem in that contrary to the standard assumption that infants will always prefer a novel stimulus (hence, why the paradigm is often simply referred to as the novelty-preference paradigm), infants can also exhibit a preference for the familiar stimulus or exhibit no preference at all (Fiser & Aslin, 2002; Cashon & Cohen, 2000; Schilling, 2000). The determination of the pattern of behaviour exhibited by infants in this paradigm is related to a number of different parameters and factors of the paradigm itself, the stimuli, as well as with the infant (Hunter & Ames, 1988; Order, Bushnell, & Sasseville, 2000). Consequently, determination of infants' discriminatory abilities from this paradigm can be ambiguous, as there are confounding variables that might influence the outcome. Thus, all these paradigms may potentially mischaracterize infants' discriminatory abilities and provide inaccurate assessments of their development.

For speech perception and phonemic discrimination, besides the paradigms outlined above, one of the main paradigms that has been used is the Conditioned Head-Turn (CHT). In this paradigm, infants are trained to associate a speech sound such as a phoneme with a reinforcing target stimulus appearing either to their left or right (Werker et al., 1998). In the training stage, infants will turn their heads either to the left or to the right depending on where the target will appear. Across trials, infants will learn to anticipate the target stimulus prior to its onset upon hearing the associated phoneme. When the phoneme is changed, if infants discriminate the novel phoneme, then they will not anticipate the target stimulus as the novel phonemic sound is not associated with a target location (was not part of the training stage). If they do anticipate when the phoneme is novel that would indicate that they do not discriminate the phonemic change. One study with 6-months-old infants using the CHT paradigm trained infants to produce an anticipatory head turn to the left (to the target – a toy) when they heard a particular phoneme (see more, Kuhl et al., 1992). Similarly, another study used CHT to investigate infants' ability to categorize syllables by their ending sounds by training them to look at a target to the left (turning their head to the left) when they heard words ending with a specific phoneme and turned their head to the other side when they heard a word ending with a different phoneme (see more, Hayes, Slater, & Brown, 2000). Though these studies found evidence of anticipatory head turning and discrimination, the CHT paradigm has potential issues as well.

Similar to looking paradigms, if infants do not anticipate when presented with the novel stimulus, does this indicate that they are able to discriminate? There are a number of factors that might preclude an anticipation, such as less than reliable and efficient motor control, particularly in younger infants (Meng & Martin, 2004; Martin, 2005). An additional factor is that due to initial learning occurring with only a single phoneme-to-target stimulus association, whether they are in fact learning a specific phoneme-target location association is not clear. Instead, the infants might be learning a more generic, nonspecific sound-target location association. If that is the case, then anticipating after a novel phoneme might not reflect a lack of discrimination but the expression of the learning of the nonspecific sound-target association.

For all the paradigms, one of the more problematic issues, which has been pointed out before for other perceptual capacities (e.g., Adler, 2005), is their level of temporal resolution. Most perceptual processes, including phoneme detection and discrimination, occur on the order of hundreds of milliseconds. But the paradigms used to this point measure behaviour on a scale of seconds, an order of magnitude greater than the processing itself. Consequently, these paradigms might not be sensitive to processing that occurs on the hundreds of milliseconds scale and, hence, might not be providing an accurate assessment of that processing. In other words, the longer temporal scale of these paradigms might allow for the influence of cognitive processes beyond perceptual discrimination, including memory for example. Furthermore, because the time scale for measuring behaviour is greater than the processing of interest, the measured behaviour has a greater likelihood being undetected/insignificant as it is happening in millisecond (less than a minute) whereas the measurement is in seconds.

Finally, all these paradigms and their measures have been designed to assess perceptual processing in early infancy and, therefore, are generally not viable for use as the child gets older. That is, comparing the performance of infants in any of these paradigms with the performance of older infants, children, or adults in different paradigms, and thereby extract an accurate developmental timeline would be a dubious undertaking. Though studies using these paradigms have greatly expanded the knowledge base of perceptual capacities and phonemic discrimination in early infancy (Aslin, 2007), because of the myriad of issues that these paradigms have, there is a clear need for a paradigm that can overcome the described problems. To overcome these issues, the present study will assess whether the Visual Expectation Cueing Paradigm (Adler, Comishen, Wong-Kee-You, & Chubb, 2020; Baker, Tse, Gerhardstein, & Adler, 2008;

Comishen & Adler, 2019) has validity in measuring phonemic discrimination via an assessment of infants' eye movements.

The Visual Expectation Cueing Paradigm (VExCP) is a variant of Haith's Visual Expectation Paradigm (VExP) which was designed to investigate infants' future-oriented expectations and the ability to visually anticipate the location of predictable, forthcoming images (Haith, Hazan, & Goodman, 1988). Future-oriented thinking is an important cognitive construct that enables the allocation of resources (e.g., attention) prior to the occurrence of a predicted event and has been theorized to be precursor of cognitive planning (Haith, 1994). In the typical VExP, 3.5- and 5.5-month-old infants' ability to visually anticipate the location of images that appear in a predictable spatial sequence (e.g., left-right-left-right) on a screen relative to a random, unpredictable spatial sequence is assessed. As the infants learn the predictable sequence of the images, they begin to correctly anticipate and look towards the side of the screen before the images appear (see more: Haith et al., 1988). Anticipations were used as the primary measurement for expectations, as expectations are the predictions and anticipatory eye-movements are the observed behaviour which governs by these expectations. In other words, anticipation is a behavioural response that was guided by internal understanding (i.e., expectation) of the predictable sequences of an event.

In contrast to the VExP, the VExCP uses central cues as predictors of the spatial locations of target stimuli. In this paradigm, for infants to successfully anticipate the location of the target stimuli above chance performance they have to be able to discriminate the perceptual information that distinguishes the central cues. The use of auditory cues in the VExCP and multiple target locations provides the framework to present more than one auditory stimulus within the same paradigm and then the capacity to use the timing and direction of eye movements as a discriminatory visual response of those stimulus cues. Consequently, in the VExCP, infants need to be able to discriminate different cues in order to learn the given cue-target location association and correctly anticipate. In order to even further substantiate that visual responses to the targets are due to discrimination of the cues and not learning a spatial rhythm of predictable stimulus presentation, the cues are randomized (Baker, Tse, Gerhardstein, & Adler, 2008). To enable the use of the VExCP with discriminatory auditory cues, each auditory cue instead of visual cue, presented synchronously with a central visual stimulus, predicts either a visual target appearing to the right or the left location.

Using the VExCP paradigm to assess phonemic discrimination in infants overcomes a number of the problems with other (previously discussed) paradigms. Due to the cue randomization and multiple cues requiring a discriminatory response, infants' performance in the VExCP when used with auditory cues (paired with a central visual non-discriminatory stimulus) could not be due to learning a general sound-location association, thereby overcoming a potential shortcoming that might occur with the CHT. Moreover, because eye movements occur in the order of hundreds of milliseconds, the VExCP has greater temporal sensitivity that is potentially consistent with the resolution of the processing mechanisms, overcoming the issue of poor temporal resolution that might hinder previous paradigms. Further, the early maturity of saccadic eye movements (Canfield, Smith, Brezsnyak, & Snow, 1997) allows for a direct comparison with comparable stimuli and measures across development, across all ages from early infancy through adulthood. For example, in Adler and Wong-Kee-You (2015), saccadic eye movements were measured as 3-month-old infants performed two distinct attentional tasks (a spatial cueing task or a visual expectation task). The study showed that 3-month-old infants can perform and complete these stimulus-driven spatial attentional tasks. They found that 3-months-old infants allocate attention and initiate an eye movement towards a stimulus target across all conditions in both tasks. Furthermore, this study also found that caesarean-section delivered infants' stimulus-driven attention was slower compared to vaginally delivered infants. Interestingly, a recent study with adults (Adler et al., under review) who performed the same attentional cueing task as the infants in Adler and Wong-Kee-You (2015), revealed similar results, caesarean-section delivered adults' stimulus-driven attention was slower compared to vaginally delivered adults. Another study by Adler and Gallego (2014) showed similar results between 3-months-old infants and adults. Similarly, infants' and adults eye movement latencies to localize a target were less efficient with increasing set size in feature-absent arrays, compared to feature-present arrays that were more efficient and decreased saccade latencies to the target. These studies demonstrate that assessing eye-movements as the primary measure is a valid and effective measure in assessing cognitive development in infants and in adults, and likely ages in between.

The Anticipatory Eye Movement (AEM) paradigm is another previously used paradigm used to assess auditory categories of 6-month-old infants, which has some similarities to the VExCP. The AEM paradigm was used by McMurray and Aslin (2004) to investigate infants' categorization abilities using anticipations (anticipatory eye movements) as a measurement (See more: McMurray & Aslin, 2004; Baker et al., 2008). The AEM paradigm used stimuli (e.g., motion and occlusion in one display) that can limit the complexity of the stimulus as a whole and, therefore, limit the age of infants who can be tested using this paradigm because the stimulus might not be as challenging to older infants, which is a limitation that the VExCP does not have (Baker et al., 2008). Moreover, in contrast to the VExCP, the AEM paradigm has a training stage that according to McMurray and Aslin (2004) served as a confounding variable influencing the results.

In the present study, the VExCP and eye movements will be used to assess phonemic discrimination. That discrimination of auditory cues is possible within VExCP which has been demonstrated by a recent study that assessed 6-month-old infants' capacity to discriminate randomly tone-scramble cues that differed by whether they contained a major or minor note (Adler et al., 2020). Infants had to discriminate between major versus minor tone-scrambles that cued different visual target locations (right versus left). Results indicated that 6-month-old infants who could discriminate the major and minor tone scrambles, as some adults do, correctly anticipated the cued target location (Adler et al., 2020). The Adler et al. (2020) study demonstrates that auditory stimuli can be used successfully with the VExCP, but whether phonemic stimuli can be used is an open question.

In the present study, the timing and direction of infants' eye movements were used as a measure of (phonemic) discrimination. The VExCP was used, similarly to its use in Adler et al. (2020), and discrimination of two distinct phonemes was required for the infants to learn cue-location associations. In particular, the present study assessed two phonemes as the discriminatory cues, /ma/ and /na/, that are common among many languages. Previous studies have shown that people from different linguistic origins are able to discriminate between the phonemes /ma/ and /na/ compared to phonemes that are unique to a specific language (e.g., phonemes that are present only in English and not in other languages) (Hillenbrand, 1984; Narayan, 2008; Narayan, 2013; Narayan, 2019). Six-month-old infants were consequently tested with the VExCP in which one phoneme cue (e.g., /ma/) predicted a target appearing on one side of the screen and the other phoneme cue (e.g., /na/) predicted a target appearing on the other side of the screen. If infants could discriminate the phonemes, then they would exhibit correct anticipations to the appropriate target location at a rate greater than chance. As the cues are randomly presented, infants will only be able to anticipate correctly if they can discriminate the specific phonemic cues thereby forming an expectation of the phoneme cue-target location associations.

The primary goal of the proposed study was to assess the viability of using the VExCP and the measuring of eye movements to study the development of speech perception and phonemic discrimination. If the VExCP is shown to be valid then this will potentially open up a new, more sensitive paradigm to begin to explore in greater depth the development of speech and language capacities. An exploration that would not be limited to early infancy but could involve direct comparison across ages.

Methods

Participants.

Twenty-three 6-month-old infants were be recruited from multiple mailing lists supplied by a local marketing company (Z-Retail Marketing Company Inc., Toronto, Canada). These lists included the name of the parent, their mailing address, and the expected due date of the infant, all of which remain confidential. An invitation letter to participate was sent to the mailing addresses provided. Parents interested in participating were encouraged to mail back an included prepaid post card, completed an online form, emailed, or were contacted via phone to gain more detailed information about the project and received information regarding eligibility. Upon arrival, parents were informed about the details of the specific study and completed a demographic questionnaire. In addition, informed consent was obtained from the parents of each infant before testing commenced.

The data from 7 participants were excluded due to crying and fussiness (n = 4), inattentiveness to the task and who finished less than 60% of the trials (n = 2), equipment failure (e.g., eye tracker failure) and experimental error (*n* = 1). As a result, 16 infants (10 males, 6 females) ranging in age from 170 to 197 days (M = 184.43, SD = 7.68) were included in the final analysis. The infants' ethnic backgrounds were Caucasian (*n* = 4), Asian (*n* = 3), African (*n* = 2), Native (*n* = 1), Hispanic (*n* = 1), Middle-Eastern (*n* = 1), and Mixed Ethnicity (*n* = 4), and came from families with middle social economic status (SES). Infants came from a monolingual (*n* = 7) background and a bilingual background (*n* = 9). Infants from a monolingual background heard Cantonese as the primary language with no English background (*n* = 2), and the other infants from a monolingual background heard only English (*n* = 5). Infants from a bilingual background heard Portuguese (*n* = 1), French (*n* = 1), Spanish (*n* = 2), Korean (*n* = 2), Ukrainian (*n* = 1), Farsi (*n* = 1), and Mandarin (*n* = 1). In total, 11 infants heard English as the primary language, and 5 infants heard other languages as the primary language. No major birth complications were reported. All infants were born at full-term, in good health, with no apparent visual, auditory, neurological, or other abnormalities as documented by parental report. Infant recruitment and experimental testing protocols followed the guidelines set out and approved by the York University ethics review board.

Stimuli and Apparatus.

The cues and target stimuli were computer-generated images and sounds. A red triangle, approximately 4.5° in diameter, was the computer-generated centre fixation stimulus presented during phonemic cue sounding and used to maintain the starting point for infants' eye movements that is equidistant from both possible target locations. The discriminatory auditory cues were the phonemes /ma/ and /na/. The target stimuli appeared either to the left or right of the centre fixation, with a visual angle of 5.5° degrees, and were random characters from the kids' educational television show "Bubble-Guppies©" (see Figure 1). All stimuli are approximately 4.5° degrees in diameter.

The infants were laid supine in a specialized crib (see Figure 2) and viewed the images on a 19-inch LCD colour monitor (1024 x 768-pixel resolution, an 8-bit/pixel grayscale, and a refresh rate of 75 Hz), mounted 48 centimeters above them. Located on either side of the monitor were 2 speakers that emitted the 2 different phonemes. A 30 x 30 cm infrared-reflecting, visible-transmitting mirror was positioned between the infant and monitor. A remote, pan-tilt infrared eye tracking camera (Model 504, Applied Science Laboratories [www.a-s-l.com], Bedford, MA) was fixed, creating a 45° angle of the mirror to the camera and to the infant (see Figure 2). The eye-tracker recorded infants' eye movements at a temporal resolution of 60 Hz. To minimize outside light entry into the crib, black felt curtains were drawn over and around the crib.

***Figure 1*. Visual stimuli used in the experiment. The image on the top (the red triangle) the centre computer-generated fixation stimulus presented during phonemic cue sounding and used to maintain the starting point for infants' eye movements. The remaining six images were the "Bubble Guppies" heads that served as the target stimuli.**

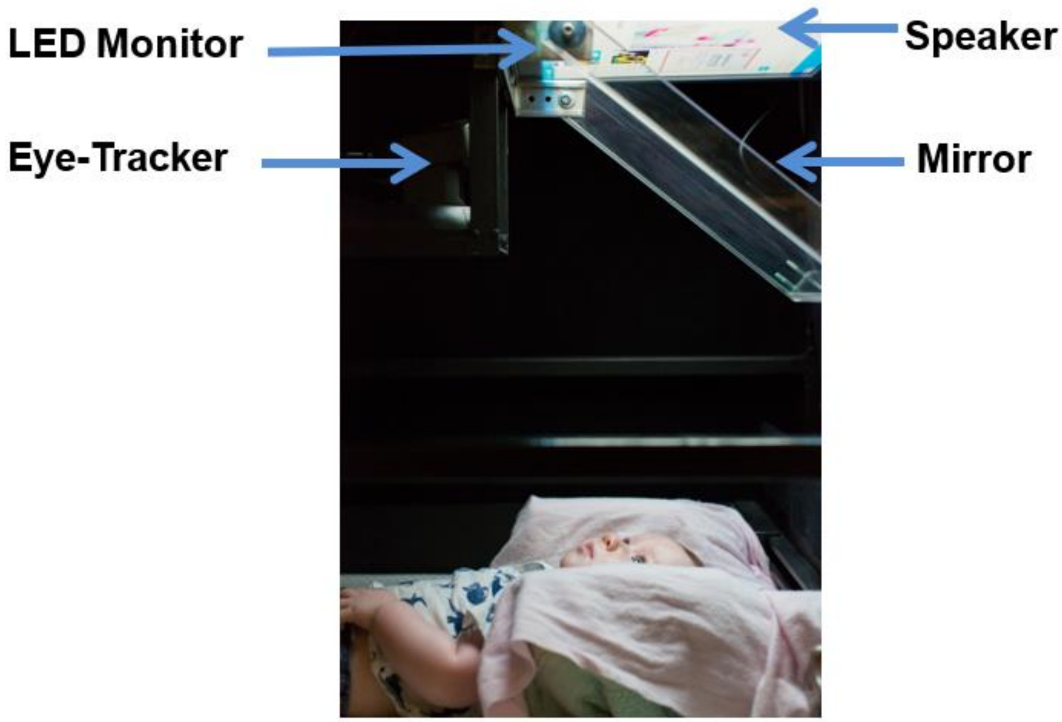

Figure 2. **The specialized crib's setup/testing crib apparatus.**

Bright pupil and corneal reflection, infrared eye tracking were used to capture the infants' eye movements. Infrared light emitted from the diodes on the camera was reflected from the mirror into the infants' eye and then reflected off the infants' retina through the pupil, producing a backlit white pupil. In addition to this, the infrared light produced a point of reflection on the cornea of the infants' eye. Using proprietary software (Applied Sciences Laboratories, Inc. [www.a-s-l.com], Bedford, MA), the eye fixation position was calculated as the relation between the centroid of the backlit pupil and the corneal reflection. The eye tracker was calibrated by having each infant view a continuous loop of varying shapes and colours at three known locations on the screen. All future recorded eye tracker fixation values were filtered through the calibration file to produce measures of eye position as a function of time (i.e., eye movements).

Throughout the experimental session, two Dell computers were used. One computer generated and presented the stimuli using the program Direct RT (Empirisoft Inc., New York; www.empirisoft.com/DirectRT.aspx). The stimuli that were generated and displayed from this computer were relayed to the LCD monitor that is above the crib. The same computer was connected to another LCD screen, which allowed the experimenter to simultaneously view what the infant is seeing. The second computer was used to control the eye-tracker and collect the eye-tracker data. The stimulus-generating computer sent a unique, time-stamped numerical code through a parallel port to the data-collecting computer indicating the onset of a trial and the type and stimulus parameters of the trial. Synchronization of the unique code with the eye movement data in the data file allowed coordination of the eye movement sequences to specific stimuli and their onsets.

Procedure.

Following a successful calibration, each infant was exposed to 60 experimental trials. Each experimental trial began with the cue being displayed at the centre of a computer screen with a greyscale background for a duration of 750 milliseconds (msec). Simultaneously with the centre cue (i.e., the fixation), a /ma/ or /na/ cue sound was played for 500 msec (i.e., the centre cue was on 250 msec before the cue sound), followed by an ISI (inter stimulus interval) of 2500 msec during which the screen was blank, after which a target was presented for 1500 msec (see Figure 3). The target stimuli were randomly selected from six pictures of characters from the kids' educational television show "Bubble-Guppies©". Target selection was randomized and independent of the phoneme, meaning, the same character could appear after either /ma/ or /na/, though the side of its appearance was dependent on the phoneme. This is to prevent infants from learning a generic association between a specific picture to a specific phoneme, and to ensure that the infants are discriminating between the two phonemes and as a result learning a phoneme target location association and not merely preferring one picture over the other. Thus, the target was counterbalanced to appear equally either to the right or to the left depending on the phoneme cue type and experimental conditions. Every infant was exposed to 30 trials of each phoneme (30 with /ma/ and 30 with /na/) regardless of the infant's experimental condition. The total time of each trial was 4750 msec and the total time of each experimental session during which eye-movement data was collected was 4.75 minutes.

Infants were randomly assigned to one of two conditions. In the Predictable condition (the experimental condition), the location of the target was predicted by the phonemic cue (either /ma/ or /na/) such that one cue predicted a target appearing to the left and the other cue predicted a target appearing to the right. Which phonemic cue was presented on any given trial was randomized with just the criterion that each phonemic cue was presented on an equal number of trials. Additionally, the phoneme cue-location relation was counterbalanced across participants. Infants in the Predictable condition were randomly assigned to a Predictable A group or to a Predictable B group. In the Predictable A group, the /ma/ phoneme cue predicted a picture appearing on the left, whereas the /na/ phoneme cue predicted a picture appearing to the right. In the Predictable B group, the opposite relation held, and the /ma/ phoneme cue predicted a picture appearing to the right, whereas the /na/ phoneme cue predicted a picture appearing to the left. In the Unpredictable control condition, no relation existed between the phonemes and the location of the target picture. Consequently, target pictures appeared randomly to the left or to the right independent from which phoneme cue was presented. The Unpredictable (control) condition served as a baseline measure for the percentage of correct anticipations infants made by chance – when there were no predictable relations upon which to form an expectation. In total, 16 infants were assigned to each condition (6[1] in the control/unpredictable condition, 5 in predictable A and 5 in predictable B of the Predictable condition).

[1] Please note that due to the covid-19 pandemic and the sudden stoppage of in-person studies, only 6 participants were assigned to the Unpredictable (control) group. The goal is to have 10 participants in the Unpredictable group as well. The extra 4 participants will be included after in-person studies resume at York University.

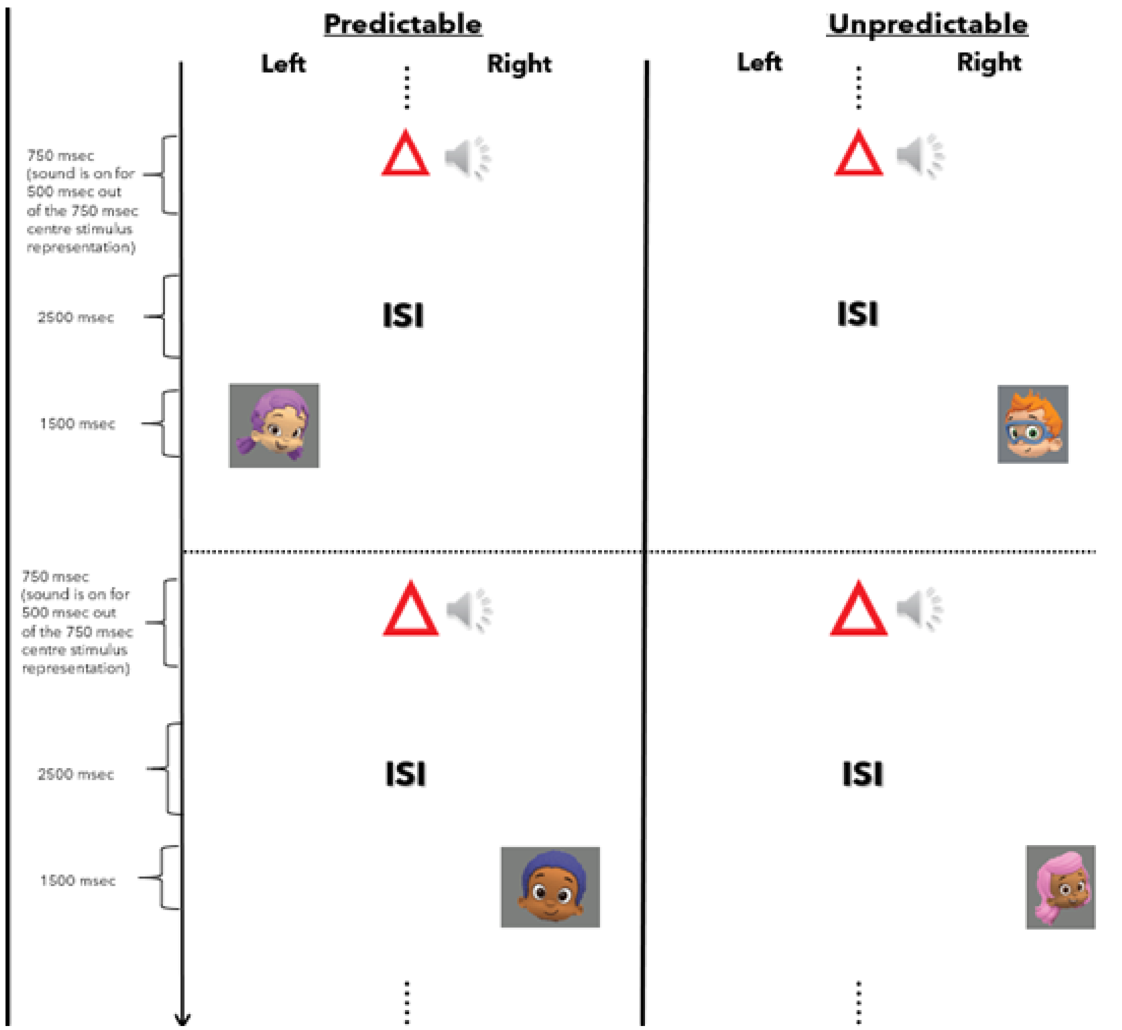

***Figure 3*. Schematic of trial sequence for the two conditions (Predictable and Unpredictable). Two trials from the Predictable and two trials from the Unpredictable conditions are shown. The timing of each step (fixation, phoneme, ISI and the picture at the target location) are also indicated. In the Predictable condition the phoneme predicted the subsequent target location, whereas in the Unpredictable condition there was no relation between the phoneme and the subsequent target location. The centre visual fixation stimulus was on for duration of 750 msec including the phoneme sound that had a duration of 500 msec, followed by ISI with a blank screen for 2500 msec, after which a visual stimulus at the target location that was presented for a duration of 1500 msec.**

Data Reduction and Analysis.

The raw digital data recorded by the eye tracker were imported into a MATLAB toolbox called ILAB for analysis (Gitelman, 2002). ILAB separates individual eye movements into their horizontal and vertical components while displaying the components on a trial-by trial basis. ILAB also displays the scan path of the eye, which allowed eye movements to be analyzed based on their timing, direction, and distance relative to the stimuli shown on screen.

For an eye movement to be included in the final data sample, it had to meet several criteria. First, the infants had to fixate on the centre stimulus prior to initiating an eye-movement to the target location for a trial to be considered valid. Second, for an eye movement to be counted as anticipatory it needed to be initiated between 133 msec after offset of the cue and prior to 133 msec after target onset. This latency value was chosen as the anticipation cut-off because it has been previously determined that 6-month-olds cannot initiate eye movements in reaction to the onset (or offset) of a stimulus faster than 133 msec (Canfield et al., 1997). If the eye movement was initiated 133 msec after target onset until 133 msec after target offset, it was considered reactive in nature. Third, for an infant's data to be included in the final sample, they must have looked at the target stimulus location on a minimum of 60% of the trials or 36 of the 60 trials. This ensured that the infants paid adequate attention to the task (e.g., Adler & Haith, 2003; Adler & Orprecio, 2006). Finally, the eye movement to the target needed to trace a path that was more than 50% of the distance between the centre fixation and the target location. The 50% criterion has been used in previous studies using infants' eye movements (e.g., Adler & Orprecio, 2006: Adler et al., 2020) and is taken as an indication that the eye movement was intentional and not random.

Infants' eye movement data was analyzed in terms of three dependent measures. First, a total anticipation measure was calculated by taking all valid eye movements that were made to the targets and then taking the percentage of the eye-movements that were anticipations (correct and incorrect, out of the total valid eye-movements). Second, a correct anticipation measure was calculated as the percent of all anticipations that correctly localized target locations. Finally, the mean latencies of all valid reactive eye movements towards the target after its onset that were not anticipatory were calculated. Reactive latency has been included as a dependent measure because although past studies have shown a dissociation between anticipatory and reactive eye movements (e.g., Adler & Haith, 2003), the dissociation is not entirely exclusive (Haith et al., 1988; Haith & McCarty, 1990; Haith, Wentworth, & Canfield, 1993). Therefore, though correct anticipations are the primary measure of phonemic discrimination and learning the phoneme-target location relation, reactive eye movements might serve as a secondary confirmatory measure. That is, the initiation/facilitation of reactive eye-movements towards the target stimuli might also be an index of an underlying expectation. Moreover, the majority of valid eye-movements exhibited by the infants were reactive eye-movements, thus, it is important to not discard them but include them in the analysis.

Results

Before delving into the primary measures of interest, it was necessary to ensure that there was no possible difference between the two predictable groups (Predictable A and Predictable B) that were counterbalance across participants. A two-sample t-test was performed to compare the percent of total anticipation between Predictable A and Predictable B. There was no significant difference in percent of total anticipation between Predictable A (M = 36.50%, SD = 10.13) and Predictable B (M = 30.90%, SD = 6.28); t(1.05), p = 0.32. Another two-sample t-test was performed to compare the percent of correct anticipation between Predictable A and Predictable B. Again, there was no significant difference in percent of correct anticipation between Predictable A (M = 78.91%, SD = 4.55) and Predictable B (M = 79.40%, SD = 6.78); t(-0.13), p = 0.89. A final two-sample t-test was performed to compare the mean reactive eye-movement latencies between Predictable A and Predictable B. There was no significant difference in mean reactive latencies between Predictable A (M = 385.36, SD = 89.84) and Predictable B (M = 342.66, SD = 88.14); t(0.75), p = 0.47. Because there was no significant difference between Predictable A and Predictable B in the 3 dependent variables, Predictable A and B were pooled together for future analyses into one Predictable group (with total 10 participants; 5 from Predictable A and 5 from Predictable B).

Total Anticipations.

Even though the percent of correct anticipations is the primary measure, there is a need to ensure that any possible difference was not due to the total number of anticipations that were made by the infants (both correct and incorrect). The percent of total anticipations is shown in Table 1. A 2-way unbalanced analysis of variance (ANOVA) was performed to analyze the effect of predictability (Predictable, Unpredictable) as a between-participant factor and Phonemes (/ma/, /na/) as a within-participant factor on the percent of total anticipations. The analysis showed that there was a significant main effect of Predictability on the percent of total anticipations, $F(1,28) = 23.23$, $p < 0.001$, indicating that 6-month-old infants made more anticipations (correct and incorrect) in the Predictable condition (M = 33.70%, SD = 8.48) than the infants in the Unpredictable condition (M = 18.35%, SD = 5.56). In contrast, the main effect of Phoneme on the percent of total anticipations was not significant, $F(1,28) = 1.93$, $p > 0.05$, indicating that infants made similar total anticipations for both phonemes. In addition, the interaction between Predictability and Phonemes was also not significant, $F(1,28) = 0.09$, $p > 0.05$. The results from this analysis indicate that the number of anticipations were influenced by predictability as the infants could learn the phoneme – target location relations when those relations were predictable, form expectations for those relations, and then use expectations to make anticipations to predictable targets. The results also show, however, that the number of anticipations was not affected by the type of phoneme nor by the interaction between phoneme and predictability, indicating that infants formed equivalent expectations for the two distinct phoneme – target location relations, thereby exhibiting comparable levels of anticipations with the two phonemes.

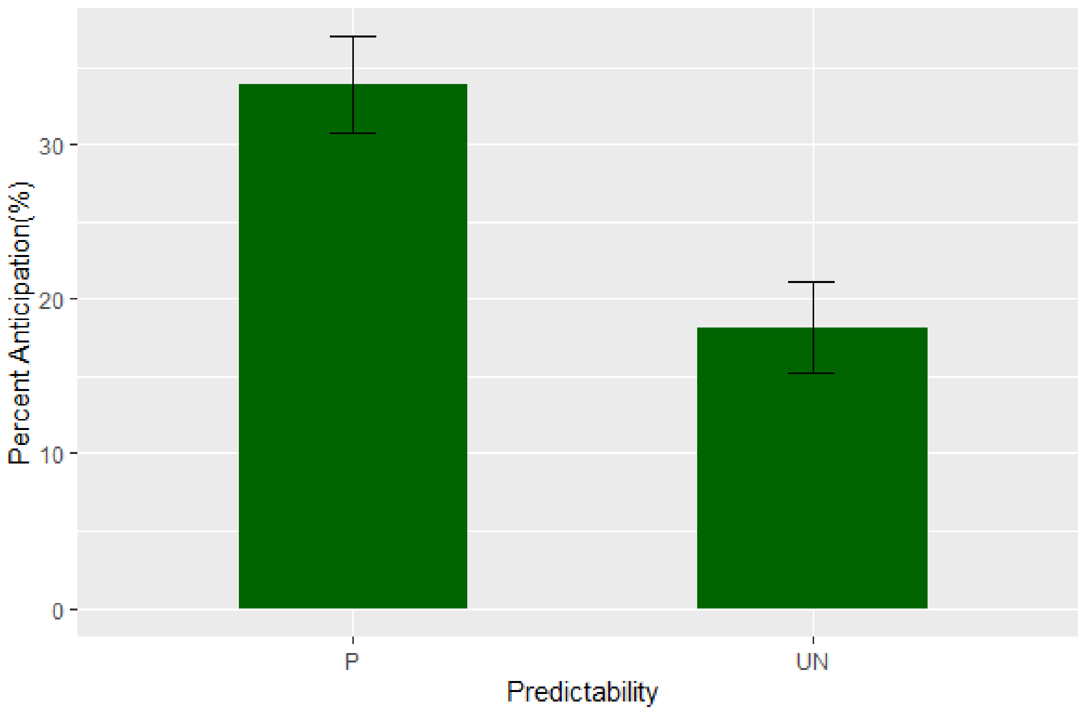

Figure 4. **Mean percent of total anticipations produced by the 6-month-old infants to the target as a function of phoneme and predictability. There is a significant main effect of predictability on the total anticipations, as there are more total anticipations in the Predictable group. No significant main effect of phoneme, nor for interaction between predictability and phoneme. Error bars represent +/- 1 standard error of the mean.**

Correct Anticipations.

To determine whether infants could discriminate between the different phonemes, and as a result successfully form expectations, the percent of correct anticipations that predicted the target's location was analyzed. If infants can discriminate between the two phonemes then they would form expectations, thereby, exhibiting more correct anticipations when the phoneme – target location relations were predictable. When the phoneme – target location was unpredictable, in contrast, infants had no predictability as phonemes were independent from the target's location upon which to form expectations and, therefore, would exhibit fewer correct anticipations.

A 2-way unbalanced ANOVA was performed with the effect of Predictability (Predictable, Unpredictable) as a between-participant factor and Phonemes (/ma/, /na/) as a within-participant factor on the percent of correct anticipations. The analysis revealed that the main effect of Predictability on the percent of correct anticipation was significant, $F(1,28) = 6.06$, $p < 0.02$, indicating that 6-month-old infants made more correct anticipations in the Predictable condition ($M = 79.16\%$, $SD = 5.45$) than the infants in the Unpredictable condition ($M = 55.83\%$, $SD = 6.97$). This indicates that 6-month-old infants were able to discriminate between the phonemes, form expectations, and use the phonemic cue to make correct anticipations to the target's location (see Figure 5). In contrast, the main effect of Phoneme on the percent of correct anticipations was not significant, $F(1,28) = 2.99$, $p > 0.05$, indicating that infants made a similar percentage of correct anticipations for both phonemes. In addition, the interaction of Predictability with Phonemes was not significant, $F(1,28) = 2.91$, $p > 0.05$.

Table 1. *Mean responses made by the 6-month-old infants (standard error)*

Predictability	**Total Anticipations (%)**	**Correct Anticipations (%)**	**Reactive Latencies (msec)**
Predictable	33.70 (2.68)	79.16 (1.72)	364.01 (27.47)
Unpredictable	18.35 (2.27)	55.83 (2.85)	488.58 (32.64)

While the above analysis revealed a significant difference in percent of correct anticipations due to predictability type (Predictable versus Unpredictable conditions), it does not establish whether infants made correct anticipations at a rate that was significantly greater than chance (i.e., 50%). If the infants could discriminate between the two phonemes thereby learning the association between the specific phoneme and its target location when it was predictable and use this information to form expectations, then the correct anticipations they initiate should occur at a rate greater than chance (i.e., greater than 50%). In contrast, if they could not discriminate the phonemes and target location, then they would not form any expectations and anticipatory responding would be random and not different from chance. Due to the main effects of phoneme (both in total anticipation and in correct anticipations) not being significant, data from the two phoneme cues was collapsed leaving predictability as the only variable for the following analysis. To determine, therefore, whether infants in either Predictability type exhibited correct anticipations at a rate greater than chance, t-tests comparing correct performance to 50% for each Predictability type were performed. A one-sample t-test revealed that infants in the Predictable condition made correct anticipations at a rate significantly greater than chance, $t(9) = 16.92$, $p < 0.01$, Cohen's $d = 5.35$, indicating that infants in this condition discriminated the two phonemes which enabled the infants to form expectations for each phoneme – target location relation (see

Figure 5). In contrast, a one-sample t-test indicated that infants in the Unpredictable condition made correct anticipations at a rate not significantly different than chance, $t(5) = 2.05$, $p > 0.05$, indicating that though they could discriminate between the phonemes there were no predictable phoneme – target location relations upon which to form expectations and guide their anticipations.

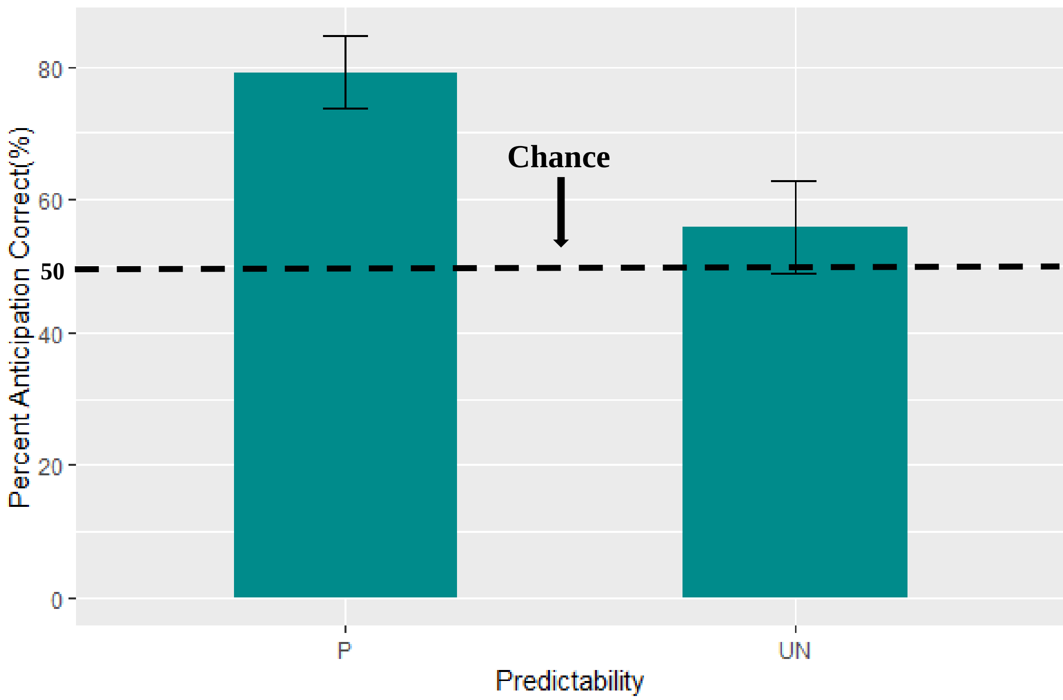

Figure 5. **Mean percent of correct anticipations made by the 6-month-old infants to the targets as a function of predictability. A significant main effect of predictability on the correct anticipations was exhibited, as there were more correct anticipations in the Predictable condition. The Predictable condition made percent of correct anticipation at a rate significantly greater than chance, whereas the percent of correct anticipations in the Unpredictable condition was not significantly different than chance. Error bars represent +/- 1 standard error of the mean.**

Reactive Latencies.

Infants do not exhibit anticipations on every trial, as shown by the total anticipation's values. Instead, they exhibit eye movements to the target's location in reaction to the onset of the target stimulus. How long after the target onset infants reactively initiate an eye movement, that was not an anticipatory eye movement, to the target's location is measured as reactive latencies. To assess the effects of predictability and phonemes on reactive latencies, a 2-way unbalanced ANOVA was performed on mean reactive latencies with Predictability (Predictable, Unpredictable) as a between-participant factor and Phonemes (/ma/, /na/) as a within-participant factor. The analysis revealed that the main effect of Predictability on the mean reactive latencies was significant, $F(1,28) = 9.77$, $p < 0.01$, indicating that 6-month-old infants had faster reactive latencies in the Predictable condition ($M = 364.01$, $SD = 86.87$) compared to the infants in the Unpredictable condition ($M = 488.58$, $SD = 79.96$). The main effect of Phoneme on mean reactive latencies, in contrast, was not significant, $F(1,28) = 2.11$, $p > 0.05$, indicating that infants had similar mean reactive latencies for both phonemes regardless of predictability (see Figure 7). Moreover, there was no significant interaction of Predictability with Phonemes, $F(1,28) = 2.42$, $p > 0.05$. These findings further support the anticipation results indicating that infants can discriminate between the phonemes, which enabled the infants to learn and form expectations when the phoneme-target location relations were predictable. As a consequence, the formed expectations facilitated reactive eye movements to the target after its onset as cued by each discriminated phoneme.

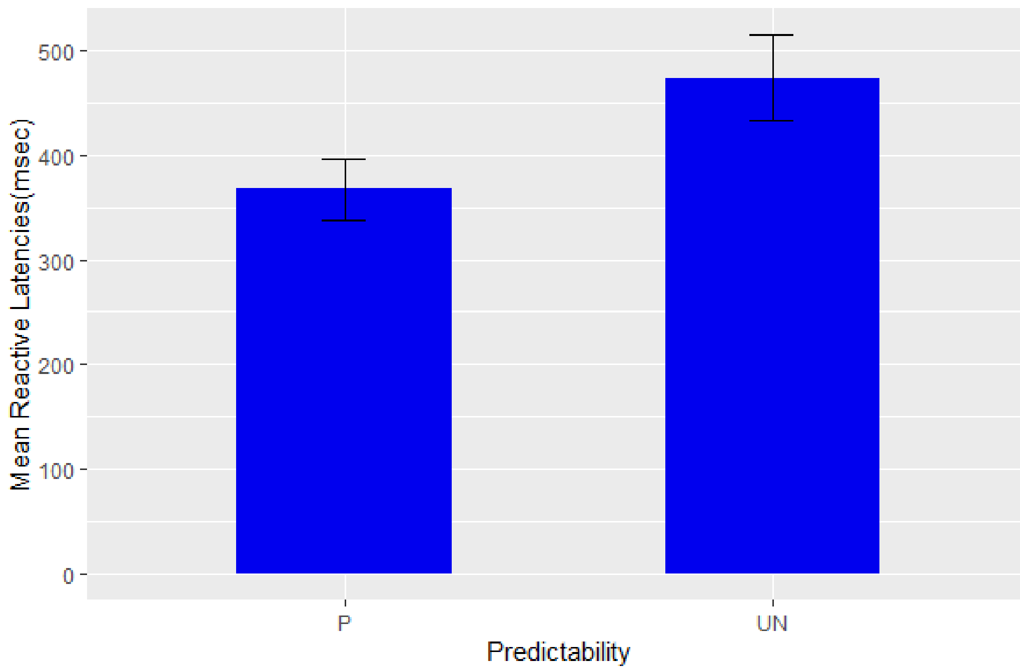

Figure 6. **Mean reactive latencies exhibited by the 6-month-old infants to the targets as a function of predictability. There is a significant main effect of predictability on the mean reactive latencies, as the infants in the Predictable group had faster reactive latencies. Error bars represent +/- 1 standard error of the mean.**

Discussion

Phonemic discrimination is considered to be a lower-level component (compared to, comprehension of words, for example) and an initial step in language development (Chládková & Paillereau, 2020). That is, the capacity of phonemic discrimination and its course of development provide the foundation upon which higher level components of language such as comprehension and production are built (Narayan, 2013; Chládková & Paillereau, 2020). A number of different methodological paradigms, preferential-looking, habituation and familiarity/novelty-preference, and CHT, have been used to document what is known up to now about phonemic discrimination and its development (Aslin, 2007). Though these paradigms have been successful, each suffers from a number of weaknesses. In preferential-looking, for example, one problem is that if the infant does not look longer at one stimulus over the other one cannot conclude that a discrimination occurred. In habituation and familiarity/novelty-preference, for example, one problem is that infants do not necessarily look longer toward the novel stimulus and there are a number of methodological, stimulus, and participants parameters that mediate whether they do so. In CHT (conditional head-turning), one problem is that infants might be learning a generic, nonspecific sound-target location association and their anticipations do not necessarily require discrimination between the phoneme cues. Furthermore, temporal resolution is a common problem present in all these paradigms, as they might not be sensitive to processing that occurs on the hundreds of milliseconds scale because they assess performance on the scale of seconds and thus might not be providing an accurate assessment of that processing. Finally, these previous paradigms likely do not provide the means for comparable assessment of processing and behaviour across development as the paradigms were designed to assess perceptual processes in early infancy and not developmental trajectories beyond that age range.

Due to the potential weaknesses of the paradigms that have been used previously, the purpose of the current study was to investigate if a new eye movement methodology, the Visual Expectation Cueing Paradigm (VExCP), could be used to study phonemic discrimination. The goal of the present study, therefore, was to introduce a new paradigm into the mix that would solve some of the shortcomings of the previously used paradigms. The VExCP has a better temporal resolution than previously used paradigms as it measures eye movements that are happening on a milliseconds scale. In addition, the VExCP uses saccadic eye movements which will potentially allow a direct comparison with comparable stimuli and measures across development, from early infancy through adulthood. Furthermore, the VExCP has the framework to present more than one stimulus, to connect between visual and auditory stimuli, and to substantiate that anticipations occurred due to the discrimination between the phonemic cues and the learning of cue-target location relation and not due to learning a nonspecific sound-target location association. To that end, the current study showed that infants were able to discriminate between the phonemes, consequently forming expectations when the phoneme cue – target location relations were predictable and, on that basis, initiate anticipatory eye movements to the correct target location. In addition, the study also revealed that there were no differences in performance to the two different phonemes across any of the measures (total anticipations, correct anticipation, and reactive latencies), indicating that infants processed each phoneme equivalently. Therefore, by having a distinct directional response for two different phonemes, the VExCP also has the capacity to evaluate not only whether infants can discriminate different phonemes but also the relative processing of two phonemes at the same time by measuring differential responding to the two phonemes.

VExCP Phonemic Discrimination and Correct Anticipations.

Findings from this experiment suggest that the VExCP is a valid paradigm that can be used in assessing phonemic discriminations while overcoming weaknesses in previously used paradigms. That the VExCP is sensitive to and can assess phonemic discrimination was evident from the results which showed that 6-month-old infants made correct anticipatory eye movements to a target after hearing a phoneme cue at a level (nearly 80%) significantly greater than chance when phoneme-target location relations were predictable. In contrast, when the phoneme-target location relations were not predictable, infants correctly anticipated the target only at a chance level (approximately 56%). For infants to perform at these levels in each condition required them to be able to discriminate the phoneme cues. When the phoneme-target relations were predictable, only by discriminating the two phoneme cues could they learn which phoneme predicted where the associated target would appear, form an expectation for that relation, and then correctly anticipate its appearance. When the phoneme-target relations were unpredictable, discriminating the phoneme cues highlighted that there was no associated target location after hearing each phoneme providing no basis to form an expectation for where the target would appear, resulting in random anticipatory eye movements that were correct only by chance. The VExCP, therefore, was sensitive to 6-month-old infants' capacity to discriminate phonemes and demonstrates that the paradigm is potentially a valid approach to investigating the development of phonemic discrimination that lacks the weaknesses of other paradigms.

Issue of Total Anticipations?

Though findings from correct anticipations support the conclusion that infants discriminated the two phoneme cues and that the VExCP is a valid method for investigating the development of phonemic discrimination, the finding of a difference between the predictability conditions for total anticipations possibly calls that into question. For total anticipations, unexpectedly, infants in the Predictable condition made significantly more anticipations than the infants in the Unpredictable condition which is inconsistent with previous studies using the same paradigm, as those studies did not find significant effect of condition on total anticipations (Comishen & Adler, 2019; Comishen, Bialystok, & Adler, 2019). Nevertheless, the total anticipations exhibited by the infants are consistent with the range of anticipations exhibited in other previous studies that have used the VExCP (e.g., Comishen and Adler, 2019; Adler, Comishen, Wong-Kee-You, & Chubb, 2020).

One might argue that by exhibiting more total anticipations, infants in the Predictable condition had more opportunities to learn the predictable pattern and make more correct anticipations. As a consequence, that infants in the Predictable condition exhibited more correct anticipations than infants in the Unpredictable condition may not be due to discriminating the phoneme cues and forming an expectation when the phoneme-target location relations were predictable than when it was not. Instead, the greater number of correct anticipations in infants who viewed predictable phoneme-target location relations may be due to having more total anticipations, which increases the likelihood of a fraction of them being correct. Moreover, there was no significant main effect or interaction of phonemes on total anticipations, therefore, there is no evidence that anticipations were more likely to one phoneme than the other and that difference contributed in any way to more correct anticipation when phonemes predicted target location. Thus, the possibility exists that evidence of more correct anticipations when the phoneme-target location relations were predictable might be due to greater opportunity to make correct anticipations. As a result, a conclusion that the paradigm was detecting infants discriminating the phonemes might be premature.

Correct anticipations results, however, are based on a percentage of the total anticipations. As a consequence, the score is a relative value that factors in the total anticipation level. If infants who viewed predictable phoneme-target locations relation were not discriminating the phoneme cues and forming expectations for the relations, then their anticipatory responding would be entirely random. Consequently, even if they made more total anticipations and as a consequence more correct anticipations than infants who viewed unpredictable phoneme-target location relations, their correct anticipatory responding would still be approximately 50%, not different from chance. So consider, for example, an infant who viewed predictable phoneme-target location relations, made 15 total anticipations, 8 of which were correct, and an infant who viewed unpredictable phoneme-target relations, made 7 total anticipation, 4 of which were correct. Though the infant who viewed predictable relations made more total and correct anticipations than the infant who viewed unpredictable relations, both will have made the same relative number of correct anticipations, 53% versus 57% respectively. Clearly, this was not the case as infants who viewed predictable relations made a significantly greater percentage of correct anticipations than infants who viewed unpredictable relations and at rate significantly greater than chance (50%).

Therefore, although the total anticipations between the two experimental conditions in the current study was not consistent with previous studies using the VExCP, infants made a greater percentage of correct anticipations at a rate greater than chance when viewing predictable phoneme-target location relations was not due to exhibiting a greater number of total anticipations. Infants' performance, instead, was due to them discriminating between the different phonemes, enabling them to learn the relation between each distinct cue and associated-target location, upon which the initiation of their correct anticipations was based.

Reactive Latencies and Expectations.

In the original framework for understanding visual expectations and the VExCP, both anticipatory and reactive eye movements were understood to be behavioural manifestations of the same cognitive concept of expectations (Haith, 1994; Haith et al., 1988). That is, when infants viewed a predictable alternating (e.g., left-right-left-right) sequence of pictures, they exhibited both a greater percentage of anticipations and faster reactive eye movements relative to infants who viewed a completely random, unpredictable sequence of pictures. Later expectation studies (Adler & Haith, 2003: Adler, Haith, Arehart, & Lanthier, 2008), however, began to question the unification of both behavioural measures as manifestations of the same expectation mechanism. Adler and Haith (2003), for example, demonstrated that when the expectation was defined by the predictability of event content, a dichotomy was exhibited in which infants' anticipations were greater for content predictable events than for content unpredictable events. Reactive latencies, however, did not differ as a function of content predictability. A similar finding was observed when temporal information was the distinguishing predictability factor (Adler et al., 2008).

Adler and Haith (2003) hypothesized that the dichotomy in the exhibition of an effect by the two measures, anticipations and reactive latencies, reflects differential sensitivity to the level of processing of the parameter that is the focus of the expectation. In early studies of visual expectations, the parameter that was manipulated was the spatial predictability of events, which is information processed at a lower level, such as the superior colliculus which codes a spatial map (e.g., the "where" information) and is involved in the generation of reactive eye movements (Krauzlis, Basso, & Wurtz, 2000; Krauzlis & Dill, 2002). In later studies, the parameters manipulated, such as content (e.g., the "what" information), are likely processed at a higher level through the ventral stream pathway (Farah, 2000) which has projections to the frontal eye fields that have been implicated in the generation of anticipatory eye movements (Hanes, Patterson, & Schall, 1998). When the predictability of higher-order event content (or temporal) information but not spatial information is manipulated, effects are seen for anticipatory but not reactive eye movements. When the predictability of lower-order spatial information is manipulated, effects are seen for both reactive and anticipatory eye movements (see Adler & Haith, 2003, for a more detailed discussion).

The current study found effects for both reactive and anticipatory eye movements, which suggests that there is a low-level component to the information infants are processing. Consistent with the hypothesis of Adler and Haith (2003), spatial predictability of the visual information (i.e., predictability of target location) in relation to distinct phonemic information was manipulated in the current study and that likely supported the findings of an effect on reactive latencies. Interestingly, however, previous studies that have used the VExCP and have manipulated content (Baker et al., 2008) or temporal (Comishen & Adler, 2019) information in conjunction with spatial predictability, did not find an effect on reactive latencies. Why an effect on reactive latencies was found, therefore, is not entirely clear. Perhaps the difference is that the previous VExCP studies manipulated only visual information which has direct connections to the eye movement system, so eye movement generation would be controlled by the higher order visual information. In the current study, auditory information, which is not directly connected to the eye movement system, in addition to low level visual information was manipulated, so eye movement generation was controlled by the lower order visual information. Yet, despite the discrepancy relative to previous studies[2], the current study through both correct anticipations and reactive latencies measures demonstrates that the infants discriminated the phoneme cues and that the VExCP paradigm is sensitive to their discrimination.

Implications and future directions.

As previously mentioned, common paradigms used in the study of phonemic discrimination have some shortcomings that the VExCP was able to overcome, as illustrated by the results of the present study. When studying infants, the task and measures used are typically different from those used with older children and adults, therefore, comparing behaviour and functioning across the lifespan can be difficult at the very least and often impossible. As eye-movements mature in early infancy (Canfield et al., 1997), they provide the potential to use them in studies that use the same task and behavioural measures across development (Adler & Gallego, 2014). The VExCP, having been shown to be sensitive to phonemic discrimination in the present study, will potentially therefore, enable the study of different aspects of language acquisition and its developmental trajectory throughout the lifespan.

The VExCP might also help to solve some inconsistencies that were found in studies using the more typical paradigms. For example, Maye et al (2002) and Rorder et al (2000), both used the familiarity/novelty-preference paradigm and found different results which led to different conclusions. The former found that 6-8-months old infants could discriminate between different phonemes whereas the latter found that 4.5-months old infants preferred familiar stimuli as much as novel ones, indicating a lack of discrimination (see more: Maye et al., 2002;

[2] In the near future, when restrictions for in-person studies will be lifted, the addition of 4 more participants might change the results in the reactive latencies and might match what previous studies have found.

Roder et al., 2000). Using the VExCP in studies like those might help in disambiguating the nature of phonemic discrimination in early development.

As the visual and auditory systems start to develop in early infancy (Cheour et al., 1998; Tierney & Nelson, 2009), which result in similar developmental timelines for language and visual perception. The combination of visual and auditory information that constitutes the protocol of the VExCP used in the current study might also have implications for the study of other aspects of language development. Early language processing and acquisition is usually related to visual capacities of object perception and identification (Landau, 2008; Landau, 2017). For example, discrimination between phonemes facilitates the infant's encoding of a language-object relation by distinguishing between different words that represent the distinction between objects (e.g., object association might facilitate language development as first words tend to be objects (Smith, Jones & Landau, 1996). By linking phonemes with visual information and eye movements, the current VExCP study illustrates the possibility that the paradigm can pair and connect visual and auditory stimuli together in the study of the connection between the auditory and visual systems as they develop throughout the first year of life, and potentially further children and adults. This paradigm, for example, might help to begin examining the relation between phonemic discrimination and object perception (Dessalegn & Landau, 2008; Landau, 2017).

The VExCP might also be helpful in assessing perceptual narrowing. Illustrated by the results, this paradigm was successful in assessing phonemic discrimination in infants. As phonemic discrimination is an ability that is sensitive to perceptual narrowing (Werker & Tees, 2002; Kuhl et al., 2006), this paradigm might help in assessing perceptual narrowing in language development throughout the first years of life. The VExCP has greater temporal resolution compared to previously used paradigms, as it uses eye movements as a measurement, which can be used comparatively across ages and, thus, might better enable the timing within development at which infants, for example, lose the ability to differentiate between all phonemes. This will deepen the understanding of language development and language acquisition in general, which will contribute to the study neurodevelopment more broadly. Hence, the assessment and investigation of phonemic discrimination with the VExCP might be suitable for investigating the nature of the contribution of components of the information-processing stream (e.g., visual attention) to phonemic discrimination, the phoneme (word) – object relation, and perceptual narrowing.

Conclusion.

In conclusion, the Visual Expectation Cueing Paradigm has been shown to be sensitive to phonemic discrimination and potentially could be used to study developmental trends in phonemic discrimination and language acquisition. What makes the VExCP more sensitive is that, compared to previously used paradigms that have used head-turning and looking time, the paradigm uses eye-movements that are measured on a milliseconds scale instead of seconds and as a result has a greater temporal resolution. Furthermore, because eye movements are an early maturing system that is relatively stable across development, the VExCP and eye movement measures have the potential to shed light on various aspects of language development across ages with comparable stimuli and measures. Future studies using this paradigm will be able to potentially investigate the development of perceptual narrowing, phoneme and word to object associations, and language acquisition.

References

Adler, S.A., Comishen, K.J., Wong-Kee-You, A.M.B., & Chubb, C. (2020). Sensitivity to major versus minor musical modes is bimodally distributed in young infants. *The Journal of the Acoustical Society of America, 147*(6), 3758-3764. doi: 10.1121/10.0001349.

Adler, S. A., & Gallego, P. (2014). Search asymmetry and eye movements in infants and adults. Attention, *Perception, & Psychophysics, 76*(6), 1590-1608.

Adler, S. A., & Haith, M. M. (2003). The nature of infants' visual expectations for event content. *Infancy*, *4*(3), 389-421.

Adler, S. A., Haith, M. M., Arehart, D. M., & Lanthier, E. C. (2008). Infants' visual expectations and the processing of time. *Journal of Cognition and Development*, *9*(1), 1-25.

Adler, S. A., & Orprecio, J. (2006). The eyes have it: visual pop-out in infants and adults. *Developmental Science, 9*(2), 189-206.

Adler, S.A., Solomon-Harris, L., Comishen, K., Wong-Kee-You, A.M.B., Turner, G., Spreng, N., & Stevens, W.D. (under review). Caesarean-section birth is related to atypical visual attention and brain Networks in adulthood. *JAMA.*

Adler, S. A., & Wong-Kee-You, A. M. (2015). Differential attentional responding in caesarean versus vaginally delivered infants. *Attention, Perception, & Psychophysics, 77*(8), 2529-2539.

Amso, D., & Scerif, G. (2015). The attentive brain: insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience, 16*(10), 606-619.

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48-53.

Aslin, R. N. (2012). Infant eyes: A window on cognitive development. *Infancy, 17*(1), 126-140.

Aslin, R. N., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *Trends in cognitive sciences*, *9*(3), 92-98.

Atkinson, J. (1984). Human visual development over the first 6 months of life. A review and a hypothesis. *Human Neurobiology, 3,* 61–74.

Atkinson, J. (2000). *The developing visual brain*. Oxford: Oxford University Press.

Baker, T. J., Tse, J., Gerhardstein, P., & Adler, S. A. (2008). Contour integration by 6-month-old infants: discrimination of distinct contour shapes. *Vision research*, *48*(1), 136-148.

Bhatt, R. S., Bertin, E., Hayden, A., & Reed, A. (2005). Face processing in infancy: Developmental changes in the use of different kinds of relational information. *Child development*, *76*(1), 169-181.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition.* Cambridge University Press.

Braddick, O., & Atkinson, J. (2011). Development of human visual function. *Vision Research, 51,* 1588–1609.

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development, 62*, 1-145.

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., Snow, K. L., Aslin, R. N., Haith, M. M., ... & Adler, S. A. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the society for research in child development, i*-160.

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.

Chládková, K., & Paillereau, N. (2020). The what and when of universal perception: A review of early speech sound acquisition. *Language Learning, 70*(4), 1136-1182.

Comishen, K. J., Bialystok, E., & Adler, S. A. (2019a). The impact of bilingual environments on selective attention in infancy. *Developmental Science*, *22*(4), e12797.

Comishen, K. J., & Adler, S. A. (2019b). The development of infants' expectations for event timing. *Timing & Time Perception*, *7*(3), 219-242.

Dessalegn, B., & Landau, B. (2008). More than meets the eye: The role of language in binding and maintaining feature conjunctions. *Psychological science*, *19*(2), 189-195.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303-306.

Farah, M. J. (2000). *The cognitive neuroscience of vision*. Blackwell Publishing.

Feldman, J. (2003). What is a visual object?. *Trends in Cognitive Sciences, 7*(6), 252-256.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, *99*(24), 15822-15826.

Gitelman, D. R. (2002). ILAB: a program for postexperimental eye movement analysis. *Behavior Research Methods, Instruments, & Computers, 34*(4), 605-612.

Golinkoff, R. M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned?. *Perspectives on Psychological Science, 8*(3), 316-339.

Haith, M. M. (Ed.). (1994). *The development of future-oriented processes*. University of Chicago Press.

Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child development*, 467-479.

Haith, M. M., & McCarty, M. E. (1990). Stability of visual expectations at 3.0 months of age. *Developmental Psychology, 26*(1), 68-74.

Haith, M. M., Wass, T. S., & Adler, S. A. (1997). Infant visual expectations: advances and issues. *Monographs of the Society for Research in Child Development*, *62*(2), 150-160.

Haith, M. M., Wentworth, N., & Canfield, R. L. (1993). The formation of expectations in early infancy. *Advances in infancy research, 8,* 251-298.

Hanes, D. P., Patterson, W. F., & Schall, J. D. (1998). Role of frontal eye fields in countermanding saccades: visual, movement, and fixation activity. *Journal of neurophysiology, 79*(2), 817-834.

Hauffen, K., Bart, E., Brady, M., Kersten, D., & Hegdé, J. (2012). Creating objects and object categories for studying perception and perceptual learning. *Journal of visualized experiments: JoVE,* (69).

Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language, and Hearing Research*.

Hayes, R. A., Slater, A., & Brown, E. (2000). Infants' ability to categorise on the basis of rhyme. *Cognitive Development, 15*(4), 405-419.

Hillenbrand, J. (1984). Speech perception by infants: Categorization based on nasal consonant place of articulation. *The Journal of the Acoustical Society of America, 75*(5), 1613-1622.

Horowitz, F. D. (1975). Visual attention, auditory stimulation, and language discrimination in young infants. *Monographs of the Society for Research in Child Development.*

Humphrey, K., & Tees, R. C. (1980). Auditory-visual coordination in infancy: Some limitations of the preference methodology. *Bulletin of the Psychonomic Society, 16*(3), 213-216.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*.

Johnson, M. H. (1995). The development of visual attention: A cognitive neuroscience perspective. In M. S. Gazzinga (Ed.), *The cognitive neurosciences* (pp. 735–747). Cambridge: MIT Press.

Johnson, M. H. (2002). The development of visual attention: A cognitive neuroscience perspective. In M. H. Johnson, Y. Munakata, & R. Gilmore (Eds.), *Brain development and cognition:* A reader (pp. 134–150). Oxford: Blackwell Press.

Krauzlis, R. J., Basso, M. A., & Wurtz, R. H. (2000). Discharge properties of neurons in the rostral superior colliculus of the monkey during smooth-pursuit eye movements. *Journal of Neurophysiology*, *84*(2), 876-891.

Krauzlis, R. J., & Dill, N. (2002). Neural correlates of target choice for pursuit and saccades in the primate superior colliculus. *Neuron*, *35*(2), 355-363.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental science, 9*(2), F13-F21.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606-608.

Landau, B. (2017). Update on “what” and “where” in spatial language: A new division of labor for spatial terms. *Cognitive science*, *41*, 321-350.

Larraza, S., Molnar, M., & Samuel, A. G. (2020). Phonemic contrasts under construction? Evidence from Basque. *Infancy, 25*(3), 304-318.

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

Mack, M. L., & Palmeri, T. J. (2011). The timing of visual object categorization. *Frontiers in Psychology*, 2, 165.

Martin, J. H. (2005). The corticospinal system: from development to motor control. *The Neuroscientist, 11*(2), 161-173.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101-B111.

McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, *6*(2), 203-229.

Meng, Z., Li, Q., & Martin, J. H. (2004). The transition from development to motor control function in the corticospinal system. *Journal of Neuroscience*, *24*(3), 605-614.

Murphy, W. D., Shea, S. L., & Aslin, R. N. (1989). Identification of vowels in “vowelless” syllables by 3-year-olds. *Perception & Psychophysics, 46*(4), 375-383.

Narayan, C. R. (2008). The acoustic–perceptual salience of nasal place contrasts. Journal of *Phonetics, 36*(1), 191-217.

Narayan, C. (2013). Developmental perspectives on phonological typology and sound change. *Origins of sound change: Approaches to phonologization*, 128-146.

Narayan, C. R. (2019). An acoustic perspective on 45 years of infant speech perception, Part 1: Consonants. *Language and Linguistics Compass*, *13*(10), e12352.

Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, *13*(3), 407-420.

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.

Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy, 1*(4), 491-507.

Schöner, G., & Thelen, E. (2006). Using dynamic field theory to rethink infant habituation. *Psychological Review*, *113*(2), 273.

Scott, L. S., & Monesson, A. (2010). Experience-dependent neural specialization during infancy. *Neuropsychologia, 48*(6), 1857-1861.

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*(4), 197-201.

Singh, L., Loh, D., & Xiao, N. G. (2017). Bilingual infants demonstrate perceptual flexibility in phoneme discrimination but perceptual constraint in face discrimination. *Frontiers in psychology, 8*, 1563.

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism?. *Cognition, 60*(2), 143-171.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*(6640), 381-382.

Sundara, M., Ngon, C., Skoruppa, K., Feldman, N. H., Onario, G. M., Morgan, J. L., & Peperkamp, S. (2018). Young infants’ discrimination of subtle phonetic contrasts. *Cognition*, *178*, 57-66.

Tanaka, J. W., Quinn, P. C., Xu, B., Maynard, K., Huxtable, N., Lee, K., & Pascalis, O. (2014). The effects of information type (features vs. configuration) and location (eyes vs. mouth) on the development of face perception. *Journal of experimental child psychology, 124*, 36-49.

Thomas, H., & Gilmore, R. O. (2004). Habituation assessment in infancy. *Psychological Methods*, *9*(1), 70.

Tierney, A. L., & Nelson III, C. A. (2009). Brain development and the role of experience in the early years. *Zero to three, 30*(2), 9.

Werker, J. F., Shi, R., Desjardins, R., Pegg, J. E., Polka, L., & Patterson, M. (1998). Three methods for testing infant speech perception. *Perceptual Development: Visual, Auditory, and Speech Perception in Infancy*, 389-420.

Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *25*(1), 121-133.