AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

2023, Vol. 59, No. 8, 1359–1376
https://doi.org/10.1037/dev0001549

Effects of English Versus Spanish Language Exposure on Basic Multisensory Attention Skills Across 3 to 36 Months of Age

Elizabeth V. Edgar[1], James Torrence Todd[2], Bret Eschman[3], Timothy Hayes[2], and Lorraine E. Bahrick[2]
[1] Yale Child Study Center, Yale University School of Medicine, Yale University
[2] Department of Psychology, Florida International University
[3] Department of Psychology, University of Tennessee at Chattanooga

Recent research has demonstrated that individual differences in infant attention to faces and voices of women speaking predict language outcomes in childhood. These findings have been generated using two new audiovisual attention assessments appropriate for infants and young children, the Multisensory Attention Assessment Protocol (MAAP) and the Intersensory Processing Efficiency Protocol (IPEP). The MAAP and IPEP assess three basic attention skills (sustaining attention, shifting/disengaging, intersensory matching), as well as distractibility, deployed in the context of naturalistic audiovisual social (women speaking English) and nonsocial events (objects impacting a surface). Might children with differential exposure to Spanish versus English show different patterns of attention to social events on these protocols as a function of language familiarity? We addressed this question in several ways using children ($n = 81$ dual-language learners; $n = 23$ monolingual-language learners) from South Florida, tested longitudinally across 3–36 months. Surprisingly, results indicated no significant English language advantage on any attention measure for children from monolingual English versus dual English–Spanish language environments. Second, for dual-language learners, exposure to English changed across age, decreasing slightly from 3–12 months and then increasing considerably by 36 months. Furthermore, for dual-language learners, structural equation modeling analyses revealed no English language advantage on the MAAP or IPEP as a function of degree of English language exposure. The few relations found were in the direction of greater performance for children with greater Spanish exposure. Together, findings indicate no English language advantage for basic multisensory attention skills assessed by the MAAP or IPEP between the ages of 3 to 36 months.

Public Significance Statement
We examined the potential impact of home language exposure on basic attention skills when infants exposed to one language (English) or more than one language (both English and Spanish) are presented with faces and voices of women speaking English. We found that basic multisensory attention skills including sustaining attention, shifting and disengaging attention, and intersensory matching of the sights and sounds of women speaking in English are relatively unaffected by the specific language environment (English vs. Spanish) of children across 3–36 months of age.

Keywords: multisensory attention, individual differences, language exposure, English, Spanish

Supplemental materials: https://doi.org/10.1037/dev0001549.supp

Selective attention provides the input for all that we perceive, learn, and remember, and is the foundation for the development of complex skills such as language, cognitive, and social functioning (Bahrick et al., 2020; Bahrick & Lickliter, 2014; Fisher, 2019). In a multisensory environment characterized by constantly changing, overlapping stimulation to all the senses, infants face a significant

This article was published Online First May 18, 2023.
Elizabeth V. Edgar https://orcid.org/0000-0003-4419-1876
We thank the children and families who participated in this research. We also thank all of the research assistants in the Infant Development Lab for their help in transcription. This research was supported by National Institute of Child Health and Human Development grants K02 HD064943 and R01 HD053776 awarded to Bret Eschman.

Elizabeth V. Edgar served as lead for conceptualization, formal analysis, writing–original draft, and writing–review and editing. James Torrence Todd contributed equally to formal analysis and supervision. Bret Eschman served in a supporting role for conceptualization, supervision, and writing–review and editing. Timothy Hayes served in a supporting role for formal analysis, methodology, supervision, and writing–review and editing. Lorraine E. Bahrick served as lead for data curation, investigation, resources, and supervision, contributed equally to conceptualization, and served in a supporting role for formal analysis. James Torrence Todd and Lorraine E. Bahrick contributed equally to methodology and writing–review and editing.
Correspondence concerning this article should be addressed to Elizabeth V. Edgar, Child Study Center, Yale University, 230 South Frontage Road, New Haven, CT 06519, United States, or Lorraine E. Bahrick, Department of Psychology, Florida International University, 11200 SW 8th Street, DM 115, Miami, FL 33199, United States. Email: elizabeth.edgar@yale.edu or bahrick@fiu.edu

challenge in learning to attend to meaningful stimulation from unitary audiovisual events, such as the faces and voices of people speaking, while ignoring competing, irrelevant stimulation, such as a television playing in the background. Selective attention to faces and voices and audiovisual speech is critical for fostering the typical development of basic attention skills and, in turn, the more complex skills such as language that rely on this foundation (e.g., Bahrick, Todd, et al., 2018; Edgar et al., 2022, 2023; Pons et al., 2019; for reviews, see Bahrick et al., 2020; Soto-Faraco et al., 2012).

The development of basic attention skills in the context of audiovisual events (referred to as "multisensory attention skills"; Bahrick, et al., 2020) including efficiently sustaining attention, avoiding distracting stimuli, quickly disengaging from one event and shifting to another, and successfully matching faces and voices of people speaking, requires exposure to a wide array of multisensory stimulation. For example, multisensory stimulation, such as audiovisual speech, provides a rich source of intersensory redundancy (the synchronous co-occurrence of stimulation across two or more senses), properties that are highly salient to young infants, including audiovisual synchrony, common tempo, rhythm, and intensity changes common across faces and voices (for reviews, see Bahrick et al., 2020; Bahrick & Lickliter, 2012). Furthermore, the development of basic attention skills is also scaffolded by face-to-face interaction with caregivers (see Gogate et al., 2001). During face-to-face communication, caregivers typically guide attention by providing intersensory redundancy using temporally coordinated face, voice, and gesture when speaking to infants (e.g., Gogate et al., 2000; Gogate et al., 2015), as well as infant-directed speech (e.g., slow tempo, much repetition, and exaggerated prosody and affect; Fernald, 1989).

The development of basic attention skills (including shifting and sustaining attention, and matching sights and sounds of events) is promoted by early experiences with objects, people, and face-to-face communication. For example, research has found that deeper processing of dynamic faces and voices occurs when infants show greater sustained attention to faces and voices, and when audiovisual speech is synchronous (and provides intersensory redundancy) rather than asynchronous, according to the measure of heart rate and event-related potentials (e.g., Courage et al., 2006; Curtindale et al., 2019; Reynolds et al., 2013; Shaddy & Colombo, 2004). Also, infant attention to audiovisual events develops alongside well-established improvements in attention to silent visual events (Amso & Johnson, 2008; Atkinson et al., 1992; Butcher et al., 2000; Colombo et al., 2001; Richards, 1985). However, the development of basic attention skills, although promoted by language exposure (e.g., language input provided by caregivers), may not depend on exposure to a specific language (e.g., English, Spanish). First, these basic attention skills emerge early in infancy, require no language skills, and are evident for social and nonsocial events alike. Second, parent-infant communication shares many features across a range of language contexts. This includes the use of infant-directed speech and intersensory redundancy evident in synchronized movements of the face, gesture, and voice during speech (see Fernald et al., 1989; Gogate et al., 2015). Intersensory redundancy provided by audiovisual speech is highly salient to young infants. It provides amodal properties such as face–voice synchrony, rhythm, tempo, duration, and intensity changes common across movements of the face and sounds of the voice (Bahrick et al., 2002; Bahrick & Lickliter, 2004; Lewkowicz, 2000, 2003). Furthermore, affect, prosody, gender, and age of speakers are detected by young infants and specified by a range of amodal properties common across language environments (Bahrick et al., 1998, 2019; Bahrick & Lickliter, 2012; Flom & Bahrick, 2007). Thus, at least in early development, it is likely that attention to audiovisual speech is guided by factors that are not language specific, including the detection of redundant amodal properties such as face–voice synchrony, prosody, and affect, common to faces and voices during speech. In sum, although exposure to language is critical for the development of typical attention skills, the relative amount of exposure to a given language (e.g., English vs. Spanish) may have little impact on the development of basic attention skills.

Alternatively, attention skills deployed in the context of audiovisual speech may be influenced by language context. Infant attention to faces and voices may be enhanced by the salience of a familiar language, or in contrast, by the novelty of an unfamiliar language. For example, infants show evidence of differential attention patterns to faces across development as a function of whether the language is familiar versus novel in eye-tracking paradigms. They looked more to the mouth than the eyes later in development when the language was unfamiliar, but showed the opposite pattern (enhanced attention to the eyes than mouth) when the language was familiar (Ayneto & Sebastian-Galles, 2017; Lewkowicz & Hansen-Tift, 2012; Pons et al., 2015). At present, however, it is unknown the extent to which infant attention to audiovisual speech differs as a function of familiarity with the language spoken, particularly in early development, or to what extent attention-language relations change across early childhood. Infants may show (a) no preference or effect of language familiarity, (b) a novelty or familiarity preference at all ages, or (c) a novelty or familiarity preference at some ages (e.g., older ages once language exposure has increased) but not at other ages.

For decades, indices of attention have been used as probes to assess infant and child perception and discrimination of audiovisual events, including differentiating native (familiar) from nonnative (unfamiliar) speech. Visual recovery to a novel event in the habituation paradigm has demonstrated discrimination of native from nonnative speech sounds. For example, young infants show evidence of discriminating between English and Spanish audiovisual speech on the basis of amodal properties such as rhythm and prosody by 5 months of age (Bahrick & Pickens, 1988). Research has also focused on the role of language background (bilingual, monolingual) on higher-order skills such as attentional control and executive functions in older children. For example, some studies have found evidence for a "bilingual advantage" for these skills (Bialystok, 1999; Carlson & Meltzoff, 2008; Morales et al., 2013) and others have not.

Even though basic attention patterns are routinely used as indices of perception and cognition, and as a primary means of assessing discrimination and perception of language as a function of language background, we have little systematic data about the development of multisensory attention skills themselves, how they are affected by language familiarity, or how these change across age. This is, in part, because there have been no fine-grained individual difference tests appropriate for assessing attention to audiovisual events in infants and young children. Given the importance of early attention skills as both an index of perception and as a foundation for later language outcomes, and the growing number of infants who are raised in bilingual language environments, understanding the role of early language exposure in the development of basic attention skills is both timely and important.

The present study explores the role of language familiarity on basic attention skills using two new individual difference measures of basic attention skills developed in our lab, the Multisensory Attention Assessment Protocol (MAAP; Bahrick, Todd, et al., 2018) and the Intersensory Processing Efficiency Protocol (IPEP; Bahrick, Soska, et al., 2018). Both assess attention to audiovisual speech (women telling stories in English) as well as to nonsocial events (objects striking a surface). The MAAP assesses fine-grained individual differences in three foundational "multisensory attention skills" (maintaining attention to audiovisual events, shifting/disengaging attention, and intersensory matching of synchronous sights and sounds). These skills are assessed in both the presence and absence of visual distractors. Distractor looking also provides an overall index of distractibility. The IPEP focuses specifically on intersensory processing by assessing the speed and accuracy of matching sights and sounds from a single event, including the sound-synchronous face and voice of a woman speaking (target event) amidst five sound-asynchronous faces of women speaking (distractor events), and the sound-synchronous sights and sounds of an object impacting a surface amidst five sound-asynchronous objects moving. Infant attention to the face–voice events on both of these protocols is the primary focus of the present paper given that attention to the faces and voices on both protocols has recently been found to predict later language outcomes in children (Edgar et al., 2022, 2003).

Here, we assess the extent to which basic multisensory attention skills assessed by the MAAP and IPEP (e.g., sustained attention in the presence of distractors, shifting/disengaging from distractors to look at faces, or matching faces and voices on the basis of temporal synchrony) are affected by language familiarity at different ages. We addressed these questions by assessing multisensory attention skills in a longitudinal sample of young children raised in South Florida, where English and Spanish are common household languages and children receive varying degrees of exposure to English and Spanish. It is unclear whether children who are exposed to a greater amount of English will have an advantage in attending to the events, given that both the MAAP and IPEP portray women speaking English. Findings would also reveal the extent to which the MAAP and IPEP could be meaningfully used with children from different language backgrounds at different ages.

Multisensory Attention Skills: Background

The term "multisensory attention skills," recently introduced by Bahrick and colleagues, describes three basic attention skills (intersensory processing, attention maintenance, and attention shifting/disengaging) that provide essential building blocks for later developmental outcomes (e.g., language; Bahrick et al., 2020; Bahrick, Todd, et al., 2018; Edgar et al., 2022, 2023). In the past, these skills have typically been studied separately in different paradigms, limiting our ability to assess interrelations among them and our understanding of how they lead to important developmental outcomes. Intersensory processing (matching sights and sounds based on audiovisual synchrony) requires attentional control and selectivity in that the perceiver must filter out irrelevant asynchronous stimulation and select the visual source of a sound. Similar to studies with adults (Alsius et al., 2005; Alsius & Soto-Faraco, 2011; for a review, see Soto-Faraco et al., 2019), intersensory processing in infancy is defined as selectively attending to audiovisual events (e.g., audiovisual speech) by detecting temporal synchrony across visual and auditory stimulation while ignoring irrelevant sights and sounds provided by competing events. Attention shifting and/or disengaging involve responding to an audiovisual event by disengaging from the immediate focus of attention at the onset of its sound and/or shifting to attend to the face or object within view. Attention maintenance (i.e., sustained attention) requires focusing on an event while ignoring concurrent distracting events.

Bahrick and colleagues recently developed the first two individual difference protocols for assessing fine-grained differences in these attention skills, opening the door to assessing developmental change and predicting outcomes based on infant attention skills deployed in a multisensory environment. The MAAP (Bahrick, Todd, et al., 2018) and the IPEP (Bahrick, Soska, et al., 2018) both assess attention skills in the context of dynamic, audiovisual social (speech), and nonsocial (object) events, making them relevant to natural learning contexts of infants and children where people and objects move and can be both seen and heard. They have now been successfully used for predicting language outcomes (Bahrick, Todd, et al., 2018; Edgar et al., 2022, 2003), have been adapted for virtual platforms (Eschman et al., 2022), and are becoming increasingly used by researchers in developmental science (e.g., Bruce et al., 2022).

The MAAP assesses individual differences in the three basic multisensory attention skills (intersensory processing, attention maintenance, and shifting/disengaging) in the context of audiovisual social (English audiovisual speech) and nonsocial (objects striking a surface) events. It assesses all three multisensory attention skills during conditions of high and low competing stimulation (presence or absence of a visual distractor event), as well as distractibility as a function of the visual distractor event. The IPEP focuses on just intersensory processing and provides a measure of both speed and accuracy. It requires participants to find a sound-synchronous target event amidst five similar distractor events, and thus provides a more fine-grained measure of intersensory processing. Like the MAAP, the IPEP assesses intersensory processing in the context of social (English audiovisual speech) and nonsocial (objects impacting a surface) events.

Home Language Exposure and Basic Attention Skills

Nationally, 22% of U.S. children (U.S. Census Bureau, 2020) come from homes in which Spanish or another language is spoken part of the time. In predominately Hispanic areas such as Miami, Florida (and the surrounding area), this figure is significantly higher (75%). It is currently not known whether and to what extent the multisensory attention skills measured in the MAAP and IPEP are affected by language environment (degree of English and/or Spanish language exposure). Given that both protocols assess basic attention skills that emerge early in infancy, one might expect little influence of language background on attention to the events, especially for nonsocial events which provide no language content. However, for social events depicting women speaking English, children with greater exposure to English (i.e., more familiar language) may have an advantage (greater attention maintenance, faster shifting/disengaging attention, and better intersensory matching of synchronous faces and voices; reduced distractibility) in attending to social events that are presented in their primary language. Alternatively, infants and children raised in primarily English-speaking homes may have no advantage, or they may have an increasing advantage across age (as they learn more language) in attending to the social events depicting English audiovisual speech relative to infants and children raised in homes where Spanish is spoken most or part of the time. Given the increasing use of the MAAP and IPEP, and the large percentage of children who hear a second language, it is important to address this issue.

Language Exposure and Basic Attention Skills in Infants

A few studies have assessed the extent to which home language exposure affects basic attention skills in infants. Studies assessing attention to audiovisual speech events have shown mixed findings. For example, one study found that home exposure to two languages had no influence on looking behavior (i.e., dwell times, saccades) to audiovisual speech events across 3–15 months of age (Schonberg et al., 2014). In contrast, it has been found that selective attention to particular regions of the face (e.g., greater looking to the mouth over the eyes, or vice versa) differs as a function of whether infants are more or less familiar with the language of the speech events across 4–12 months of age (Ayneto & Sebastian-Galles, 2017; Pons et al., 2015, 2019). Studies assessing attention to audiovisual nonsocial events have also shown mixed findings. Again, home exposure to two languages had no influence on looking behavior (i.e., dwell times, saccades) to object events across 3–15 months of age (Schonberg et al., 2014). Furthermore, studies comparing infants exposed to a monolingual versus a dual-language home environment using the anticipatory looking paradigm have typically found no group differences in attentional control (i.e., anticipatory looking) to nonsocial stimuli (i.e., geometric shapes, stars with smiles) for infants across the first year of life (D'Souza et al., 2020; Kalashnikova et al., 2021; but see Arredondo et al., 2022; Comishen et al., 2019). However, one of these studies also examined basic attention skills in conjunction with anticipatory looking and found that dual-language learners showed faster disengagement and shifting of attention than infants exposed to a monolingual home environment (D'Souza et al., 2020). Thus, research assessing infant attention skills as a function of home language exposure (monolingual vs. dual-language environments) provides differing evidence about whether and at what age basic attention skills are impacted for both social and nonsocial audiovisual events. Despite mixed findings with respect to whether infants show greater visual attention to social and/or nonsocial events in a more or less familiar language, visual attention is often used to, index discrimination of familiar/native versus unfamiliar/nonnative speech sounds across the first year (Bosch & Sebastián-Gallés, 1997; Dehaene-Lambertz & Houston, 1998; Kinzler et al., 2007; Shaw et al., 2015). In the present study, we examine the potential impact of home language exposure on basic multisensory attention skills when infants exposed to one language (English) or more than one language (both English and Spanish) are presented with faces and voices of women speaking English.

Language Exposure and Cognitive Skills in Bilingual Children

Some studies have also found a bilingual advantage in older children (e.g., 4–10 years) who are exposed to two languages, particularly for more complex cognitive skills (e.g., executive functions; working memory). Bilingual preschool and school-aged children have been found to perform better than monolingual children on tasks involving inhibitory control (Bialystok, 1999; Bialystok & Martin, 2004; Carlson & Meltzoff, 2008) and working memory (Morales et al., 2013). In contrast, some studies have found no advantage for bilingual over monolingual school-aged children on a series of tasks assessing executive functions (Dick et al., 2019; Paap & Greenberg, 2013). Thus, research provides conflicting evidence regarding the bilingual advantage for tasks assessing cognitive and executive functions.

However, it is not clear whether in early development, more basic skills such as multisensory attention skills would be affected by language background. Given the mixed findings in the current literature, it is important to determine the extent to which language background biases basic attention skills assessed by these protocols for children who have differential exposure to English.

The Present Study

The goal of the present study is to explore if (and when) the degree of exposure to English versus Spanish language in the home impact tests of multisensory attention skills as assessed by the MAAP and IPEP across 3–36 months of age. First, we assessed group differences in multisensory attention skills assessed by the MAAP and IPEP between monolingual English learners and dual-language learners as a function of the home language environment at each age (3, 6, 12, 18, 24, and 36 months). Second, we focused specifically on dual-language learners and asked whether and how the amount of exposure to English versus Spanish language children receive changes across age between 3 and 36 months. Finally, using the degree of language exposure at each age for dual-language learners, we conducted structural equation modeling (SEM) analyses to ask if there was a relation between language exposure (English vs. Spanish) and multisensory attention skills for the social events (English audiovisual speech) at each age (3, 6, 12, 18, 24, and 36 months). If attention skills assessed by the MAAP and IPEP are affected by language exposure at any age, then we expected greater exposure to English to predict better MAAP and IPEP performance on audiovisual speech events at that age. In contrast, for nonsocial events (analyses presented in the online supplemental materials), we had no such predictions given they presented no language content.

Method

Participants

Children ($N = 104$) participating in an ongoing longitudinal study were assessed at 3, 6, 12, 18, 24, and 36 months. The longitudinal study, entitled "Development of Intermodal Perception of Social and Nonsocial Events," received IRB approval from the Social and Behavioral Review Board of Florida International University (IRB-13-0448-CR06). Participants were recruited from local birth records and enrolled in the study at 3 months of age. All participants were born within 14 days of their due date ($M = -3.28$ days, $SD =$ 6.64), weighed more than 5 pounds at birth, and had APGAR scores of 9 or greater. Demographic information for the sample, reflecting the natural demographics of Miami, FL, can be found in Table 1. Eighty-one children were classified as dual English–Spanish language learners, and 23 were categorized as monolingual English learners.

Table 1
Demographic Information for the Sample (N = 104) of Dual English–Spanish Language Learners (n = 81) and Monolingual English Language Learners (n = 23)

Measure	Dual language		Monolingual	
	n	%	*n*	%
Gender				
Male	42	51.9	11	47.8
Female	39	48.1	12	52.2
Ethnicity				
Hispanic	60	74.1	3	13.0
Non-Hispanic	19	23.5	19	82.6
Did not disclose	2	2.4	1	4.4
Race				
White/European-American	60	74.1	10	43.5
Black/African-American	8	9.9	8	34.8
Asian/Pacific Islander	3	3.7	—	—
More than one race	3	3.7	5	21.7
Other	1	1.2	—	—
Did not disclose	6	7.4	—	—
Maternal education				
High school or equivalent	11	13.5	2	8.7
Some college	15	18.5	1	4.4
Associate's degree	11	13.6	4	17.4
Bachelor's degree	18	22.2	8	34.8
Master's degree or higher	20	24.7	8	34.8
Did not disclose	6	7.4	—	—
Age (*N* = 104)	*M*	*SD*		
3-month visit	3.03	0.18		
6-month visit	5.97	0.20		
12-month visit	12.05	0.25		
18-month visit	18.05	0.42		
24-month visit	24.19	0.37		
36-month visit	36.13	0.64		

The Multisensory Attention Assessment Protocol

Apparatus and Equipment

A 46-in. widescreen monitor (NEC Multisync PV61) was used to present the MAAP. Children were seated approximately 40 in. from the widescreen display. At younger ages (e.g., 3, 6, 12, 18 months), children sat on their caregiver's lap, but at older ages (e.g., 24, 36 months) some sat independently with their caregiver next to them. Caregivers wore black-out glasses, so they were unaware of the side of the screen that depicted the sound-synchronous event. An experimenter was seated behind the child and presented the stimuli to the widescreen monitor from a second computer (Mac Pro Computer with 16 GB of RAM, a 3.33-GHz processor, and a 400-MHz graphics card) using a custom MatLab-based program.

Stimulus Events

The MAAP (Bahrick, Todd, et al., 2018) assesses multisensory attention skills to audiovisual social and nonsocial events via a 3-screen procedure. Social events portray women telling stories using child-directed speech, and nonsocial events portray small wooden objects being dropped into a container in an erratic temporal pattern (see Figure 1). The MAAP has 12 social and 12 nonsocial trials, arranged into four blocks of six trials each (social, nonsocial, social, nonsocial, or vice versa, counterbalanced across participants). The MAAP was designed with separate blocks so that the social and nonsocial trial blocks could be analyzed separately. For the present study, we focused on social trials, and nonsocial trials were used as a comparison (presented in the online supplemental materials). Each trial begins with a 3-s silent visual event of moving geometric shapes presented in the center of the screen (central stimulus), followed by the onset of the two 12-s lateral events. The lateral events (right and left sides of the three-screen display) depict either two social or two nonsocial events (e.g., two different women each telling a different story, or two different object sets being dropped into a container). The sounds and movements of one of the lateral events are synchronous, while the movements of the other lateral event are asynchronous with the soundtrack. For half of the trials the central distractor event (i.e., the morphing geometric shapes) is presented for the duration of the 12-s trial, providing an additional source of competing stimulation (high-competition trials). For the other half of the trials, the central distractor event disappears at the onset of the lateral events (low-competition trials). For an example video, visit: https://nyu.databrary.org/volume/326.

Procedure

The experimenter viewed the child through a hidden front facing camera (SONY FDR-AX33) directly above the widescreen monitor. Hidden with a black curtain behind the widescreen monitor, observers viewed the child through the front facing camera. They coded infant fixations to the left, center, and right sides of the

Figure 1
Static Images of the Dynamic Audiovisual Social and Nonsocial Events from the Multisensory Attention Assessment Protocol (MAAP)

Note. Images of faces published with permission.

screen on a game pad in real time. For additional details regarding the MAAP, see Bahrick, Todd, et al. (2018, pp. 2216–2217).

MAAP Measures

The MAAP provides measures of three multisensory attention skills: intersensory matching, sustained attention and speed of shifting/disengaging, as well as an index of distractibility. Intersensory matching is the proportion of total looking time to the sound-synchronous event (PTLT). PTLT was calculated for each trial by dividing the looking to the audiovisual synchronous event by the total looking time to both lateral events (the synchronous and asynchronous events) and reflects matching based on synchrony detection. PTLTs greater than 0.50 reflect a preference for the sound-synchronous display. Sustained attention is the proportion of available looking time spent fixating the two lateral events (PALT). PALT was calculated for each trial by dividing the total looking time to both lateral events (the synchronous and asynchronous events) by the length of the trial and reflects overall interest in the faces and voices. Speed of shifting/disengaging is the child's reaction time (RT) to look to a lateral event and was calculated for each trial as the latency to shift attention (in seconds) from the central stimulus to either of the two lateral events. It reflects the speed of disengaging from the central distractor event, and on trials where the distractor is turned off as the lateral events begin, it reflects the speed of shifting attention. The three multisensory attention indices provide conceptually distinct indices of attention. Distractibility is the looking time to the central distractor event divided by the total looking time to all three events, the distractor, the synchronous and the asynchronous lateral events (PTLT distractor). This measure reflects the child's overall level of distractibility resulting from the silent central visual event, while viewing the lateral audiovisual events, and can only be calculated on the high-competition trials where the distractor is present. Pearson correlations for the primary and secondary observer were as follows: 0.92 for intersensory matching, 0.94 for sustained attention, 0.91 for RT, and 0.92 for distractibility.

The Intersensory Processing Efficiency Protocol

Apparatus and Equipment

The IPEP (Bahrick, Soska, et al., 2018) apparatus and equipment were identical to that of the MAAP, except a Tobii X120 eye-tracker was used for gaze recording. The Tobii eye-tracker was located directly under the widescreen monitor and tilted upward (20°) toward the child's eyes. An experimenter was seated behind the child and presented the stimuli to the widescreen monitor from a second computer (Mac Pro Computer with 16 GB of RAM, a 3.33-GHz processor, and a 400-MHz graphics card) using Tobii Studio (Version 3).

Stimulus Events

The IPEP assesses intersensory processing skills for audiovisual social and nonsocial events. Social events depict women telling stories in a child-directed manner, and nonsocial events depict small wooden objects or clusters of objects being dropped onto a surface in an erratic temporal pattern (see Figure 2). The IPEP has 24 social and 24 nonsocial trials, arranged into blocks of 12 each (social, nonsocial, social, nonsocial, or vice versa, counterbalanced across participants). Similar to the MAAP, separate blocks of social and nonsocial trials were presented and are designed to be used separately if needed. Each trial begins with a smiley face (one of six, each a different primary color and presented in pseudorandom order across trials) zooming in out for 2 s to attract the child's attention to the center of the screen. The 8-s trials depict six concurrent visual events (arranged in two rows of three; six different women telling a different story or six different objects being dropped onto a surface), accompanied by a single soundtrack. The movements of one event are synchronized with the soundtrack, while the movements of the other five events are asynchronous with the soundtrack. The infant's task on each trial was to visually fixate the sound-synchronous speaker/object (target event) amidst the five asynchronous distractors. For an example video, see https://nyu.databrary.org/volume/336.

Procedure

Unlike the MAAP, the IPEP used eye-tracking to assess infant gaze. The experimenter viewed the child through a front facing camera (SONY FDR-AX33) hidden directly above the widescreen monitor. The experimenter made sure the infant was seated in an optimal position for eye-tracking calibration and viewing the stimuli. Tobii Studio's "Infant" 5-point calibration procedure was used to calibrate the infrared corneal reflection-to-pupil tracking system for each infant. The experimenter calibrated the infant's eye gaze to five

Figure 2

Static Images of the Dynamic Audiovisual Social and Nonsocial Events From the Intersensory Processing Efficiency Protocol (IPEP)

Note. Images of faces published with permission.

points on the widescreen monitor for accurate calculation of infant visual fixations during the procedure. For additional details regarding the IPEP, see Bahrick, Soska, et al. (2018; p. 2230).

Eye-Tracking and Data Processing

The Tobii X120 system was used to sample infant eye gaze at 120 Hz. Trials were considered unusable when infants were inattentive (spent less than 250 ms looking to the screen during a trial) or when data were missing. The number of usable trials ranged from 2 to 48, with an average of 38.66 ($SD = 10.63$). Fixations were derived from the raw gaze data with a Velocity-Threshold Identification (I-VT) filter (for details, see Olsen, 2012). Six areas of interest (AOIs) were created from the 2×3 grid demarcating each of the six concurrent events for each social and nonsocial trial. For additional details regarding eye-tracking and data processing, see Bahrick, Soska, et al. (2018, p. 2231).

IPEP Measures

The IPEP provides three fine-grained measures of intersensory processing skills: accuracy and speed of intersensory matching, and frequency of target selection. Speed of matching is how quickly children locate the sound-synchronous event (RT). RT was calculated as the latency from trial onset to produce a fixation (of at least 50 ms) to the sound-synchronous event. Latency on each trial was then averaged across all trials within each condition (social, nonsocial). Accuracy of intersensory matching is the proportion of total looking time to the sound-synchronous event (PTLT). PTLT was calculated for each trial by dividing the looking time to the sound-synchronous target event AOI by the total looking time to all of the AOIs and has a chance value of 0.167. Frequency of target selection is the proportion of trials on which the sound-synchronous target event was fixated (PTTF). PTTF was calculated by dividing the number of trials on which the sound-synchronous target event was fixated (for at least 50 ms) by the total number of trials.

Degree of English Language Exposure

At each age, parents reported how much English and Spanish were "spoken to your child at home?" on a scale from 0% to 100%. We defined our measure of English language exposure as the mean percentage exposure to English relative to Spanish at each age. Mean exposure to English relative to Spanish ranged from 45% to 60% across age and can be found in Table 2.[1] Parent report estimates of the exposure to one language relative to another have been used in previous research (Hoff et al., 2018; Patterson, 2002; Ribot et al., 2018) and have been found to be reliable and strongly related to measures of the relative amount of language exposure obtained from diaries ($rs = .64$ and .71, respectively, $ps < .001$; Hoff et al., 2012; Lauro et al., 2020) and day-long recordings ($r = .76$, $p < .001$; Orena et al., 2020). Furthermore, parent report estimates of exposure to one language relative to another predict receptive and expressive vocabulary in both monolingual and bilingual samples (DeAnda et al., 2016; Hoff et al., 2012; Lauro et al., 2020; Place & Hoff, 2011).

Table 2

Means and Standard Deviations for Percentage Exposure to English Language (Relative to Spanish Language) at Each Age (3, 6, 12, 18, 24, and 36 Months)

	English language exposure	
Age	*M*	*SD*
3 months	52.23	25.73
6 months	49.64	24.81
12 months	45.47	29.34
18 months	50.17	26.57
24 months	53.87	26.19
36 months	60.21	27.15

Note. At each age, parents reported how much English and Spanish were "spoken to your child at home?" on a scale from 0% to 100%.

Results

Overview

Descriptive statistics for each social MAAP variable (intersensory matching, sustained attention, speed of shifting, distractibility), and social IPEP variable (accuracy and speed of intersensory matching and frequency of target selection) at each age (3, 6, 12, 18, 24, and 36 months) are displayed in Table 3. Given that our research questions involved characterizing relations between home language exposure and multisensory attention skills, our primary analyses focus on multisensory attention skills to social events (i.e., those that provide language). In addition, we briefly summarize overall findings from nonsocial events and provide descriptive statistics and analyses for nonsocial events in the supplement pp. 2–5 and Tables S5–S10 in the online supplemental materials. For high- and low-competition MAAP measures, we collapsed across high- and low-competition conditions, given a lack of significant differences in effects on other variables (see Tables S2–S4 in the online supplemental materials).

To address our main research questions, three types of analyses were conducted for social events to assess effects of language familiarity on multisensory attention skills. First, to assess whether there were group differences in multisensory attention skills between monolingual and dual-language learners, we conducted *t* tests between children exposed to monolingual English environments ($n = 23$) and children exposed to dual English–Spanish environments ($n = 81$) on each MAAP and IPEP measure at each age (3, 6, 12, 18, 24, and 36 months).[2] Second, for dual-language learners, to assess whether the amount of exposure to English versus Spanish was constant across age or changed across 3 and 36 months, we conducted growth curve analyses on the percentage exposure to English (as a continuous variable) versus Spanish. Given that the degree of English language exposure changed across age, we used separate estimates of language exposure at each age in our subsequent analyses of dual-language learners assessing whether language

[1] In analyses where only dual-language learners were used (growth curve and structural equation modeling), we categorized children on the basis of the average exposure to English and Spanish across age.

[2] Omnibus ANOVAs were not possible due to missing data in the sample, and because they were not appropriate for our research question involving the specific ages at which effects of language familiarity are evident.

Table 3

Descriptive Statistics for Performance on Social Trials for Each MAAP Measure (Sustained Attention, Intersensory Matching, Shifting/Disengaging, Distractibility) and Each IPEP Measure (Accuracy and Speed of Intersensory Matching and Frequency of Target Selection) at Each Age (3, 6, 12, 18, 24, and 36 Months)

	3 months		6 months		12 months		18 months		24 months		36 months	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
MAAP												
Sustained attention (PALT)	0.58	0.12	0.63	0.14	0.59	0.14	0.59	0.14	0.64	0.13	0.66	0.12
Intersensory matching (PTLT)	0.50	0.12	0.50	0.08	0.51	0.08	0.51	0.07	0.52	0.05	0.55	0.08
Speed of shifting/disengaging (RT)	2.13	0.85	1.16	0.30	1.10	0.26	1.17	0.38	1.13	0.28	1.21	0.29
Distractibility (PTLT distractor)	0.64	0.18	0.47	0.23	0.46	0.20	0.44	0.21	0.42	0.17	0.36	0.18
IPEP												
Speed of intersensory matching (RT)	2.71	1.06	2.59	0.69	2.57	0.79	2.65	0.95	2.87	0.56	3.01	0.69
Accuracy of intersensory matching (PTLT)	0.18	0.09	0.17	0.04	0.17	0.05	0.17	0.05	0.18	0.05	0.20	0.07
Frequency of target selection (PTTF)	0.41	0.13	0.45	0.11	0.49	0.13	0.52	0.17	0.56	0.14	0.58	0.15

Note. Variables for the MAAP include PALT (proportion of available looking time to lateral events), PTLT (proportion of total looking time to sound-synchronous event), RT (reaction time to shift or disengage to lateral events), and PTLT distractor (proportion of total looking time to central distractor). Variables for the IPEP include RT (reaction time to select the synchronous target), PTLT (proportion of looking time to synchronous target), and PTTF (proportion of times the synchronous target was found).

familiarity had a significant effect on MAAP and IPEP performance at any age. Third, using SEM, we tested path models assessing the extent to which the degree of English language exposure predicted multisensory attention skills tested by the MAAP and IPEP. We approached this question by using a variety of strategies to probe whether the null hypothesis, that the degree of English language exposure has no effect on multisensory attention skills, was a better fit with the data than the alternative hypothesis, that English language exposure does affect performance on tests of multisensory attention skills assessed by the MAAP and IPEP. We describe the logic for these models on pp. 23–26.

Multisensory Attention Skills for Social Events in Children From Monolingual English Versus Dual English–Spanish Language Environments: Group-Level Analyses

We first asked whether there were group differences in multisensory attention skills as a function of language background (monolingual English vs. dual English–Spanish language environments) at each age (3, 6, 12, 18, 24, and 36 months). To answer this question, we conducted independent sample *t* tests between children exposed to monolingual English environments ($n = 23$) and children exposed to dual English–Spanish language environments ($n = 81$) for each MAAP variable (accuracy of intersensory matching, duration of sustained attention, speed of shifting/disengaging, distractibility) and each IPEP variable (accuracy and speed of intersensory matching and frequency of target selection) at each age. Given that multiple *t* tests were conducted (one for each of seven variables at six different ages), we used a familywise significance value of $p < .008$ (.05 divided by 6; two-tailed) to evaluate results (see Table 4).

Results indicated no significant group differences in performance between children exposed to monolingual English environments and children exposed to dual English–Spanish language environments on the three IPEP variables (accuracy and speed of intersensory matching and frequency of target selection) at any of the six ages (after correcting for familywise error rate; $ps > .008$). Furthermore, there were no significant group differences in performance on three of the MAAP variables (duration of sustained attention, speed of shifting/disengaging, and distractibility) and for the fourth variable (intersensory matching of faces and voices) there was only one significant difference at one of the six ages (18 months). Children exposed to monolingual English language environments showed significantly *lower* accuracy than children exposed to dual English–Spanish language environments, $t(61) = -2.89$, $p = .005$ (a finding in the opposite direction predicted if language familiarity facilitated performance). No significant group differences were evident at any of the other five ages (3, 6, 12, 24, and 36 months; $ps > .008$) for this variable.[3] Strikingly, out of 42 comparisons (seven variables at six ages each), there was no evidence for an advantage for monolingual English language learners on the MAAP or IPEP for the social events depicting women speaking English at any age tested.

Change in English Exposure Across 3–36 Months for Dual-Language Learners: Growth Curve Analyses

Prior to assessing relations between the degree of English language exposure and multisensory attention skills, we asked whether the amount of exposure to English relative to Spanish is constant across age or if it changes across age for dual-language learners in South Florida. To address this question, we conducted a series of growth curve analyses. Some children did not participate in all visits and thus, missing data ranged from 6.2% (English language exposure at 3 months and sustained attention at 3 months) to 49.4% (speed of shifting/disengaging at 24 months; see Table S1 in the online supplemental materials). To assess whether data were missing in a systematic way, we tested for mechanisms of missingness.

[3] Without correcting for familywise error, there were four cases in which there was a significant group difference between children exposed to monolingual English environments and children exposed to dual English–Spanish language environments: intersensory matching of faces and voices on the MAAP at 6 and 18 months, speed of shifting/disengaging on the MAAP at 12 months, and frequency of target selection on the IPEP at 12 months. All significant differences were in the opposite direction predicted if language familiarity facilitated performance (see Table 4).

Table 4

t-Tests Between Monolingual English and Dual English–Spanish Language Learners for Performance on Social Trials for Each MAAP Measure (Sustained Attention, Intersensory Matching, Shifting/Disengaging, Distractibility) and Each IPEP Measure (Accuracy and Speed of Intersensory Matching and Frequency of Target Selection) at Each Age (3, 6, 12, 18, 24 and 36 Months)

	Monolingual			Dual language			
Measure	*M*	*SD*	*n*	*M*	*SD*	*n*	*t* test
MAAP							
PALT							
3 months	0.61	0.10	31	0.58	0.13	53	0.93
6 months	0.62	0.12	27	0.62	0.15	51	0.29
12 months	0.56	0.17	24	0.59	0.14	39	−0.81
18 months	0.60	0.13	23	0.60	0.12	40	0.13
24 months	0.62	0.17	17	0.63	0.12	34	−0.11
36 months	0.65	0.13	17	0.67	0.13	43	−0.57
PTLT							
3 months	0.48	0.13	26	0.50	0.13	43	−0.72
6 months	0.53	0.11	27	0.49	0.07	51	2.0*[a]
12 months	0.51	0.08	23	0.51	0.08	39	−0.31
18 months	0.47	0.09	23	0.52	0.05	40	−2.89**
24 months	0.53	0.06	17	0.52	0.05	34	0.84
36 months	0.55	0.07	17	0.56	0.08	43	−0.40
RT							
3 months	1.96	0.65	31	2.07	0.81	50	−0.64
6 months	1.28	0.32	27	1.16	0.35	51	1.38
12 months	1.27	0.45	24	1.07	0.21	39	2.47*[a]
18 months	1.17	0.39	23	1.18	0.41	40	−0.12
24 months	1.22	0.51	16	1.16	0.31	34	0.56
36 months	1.10	0.14	16	1.24	0.30	43	−1.74
PTLT distractor							
3 months	0.61	0.19	31	0.66	0.17	51	−1.21
6 months	0.49	0.25	27	0.48	0.23	51	0.19
12 months	0.48	0.20	24	0.44	0.18	39	0.90
18 months	0.49	0.20	22	0.44	0.20	40	0.89
24 months	0.40	0.21	17	0.43	0.15	33	−0.51
36 months	0.40	0.20	19	0.36	0.17	45	0.76
IPEP							
RT							
3 months	2.62	1.07	27	2.84	1.02	53	−0.89
6 months	2.62	0.98	23	2.57	0.67	50	0.25
12 months	2.68	0.88	25	2.47	0.72	37	1.001
18 months	2.46	1.03	18	2.68	0.92	41	−0.83
24 months	2.92	0.58	22	2.94	0.10	33	−0.09
36 months	2.83	0.61	20	3.02	0.71	43	−1.02
PTLT							
3 months	0.17	0.06	27	0.17	0.07	53	0.23
6 months	0.17	0.04	23	0.16	0.05	50	0.43
12 months	0.16	0.05	25	0.17	0.04	37	−0.24
18 months	0.18	0.05	18	0.18	0.05	41	0.06
24 months	0.20	0.06	22	0.18	0.05	33	1.05
36 months	0.17	0.05	20	0.20	0.07	43	−1.41
PTTF							
3 months	0.40	0.11	27	0.41	0.13	53	−0.22
6 months	0.43	0.13	23	0.45	0.11	50	−0.73
12 months	0.42	0.11	25	0.49	0.13	37	−2.26*[a]
18 months	0.50	0.18	19	0.53	0.15	41	−0.54
24 months	0.50	0.17	23	0.56	0.15	33	−1.43
36 months	0.51	0.15	20	0.57	0.15	44	−1.50

Note. Variables for the MAAP include PALT (proportion of available looking time to lateral events), PTLT (proportion of total looking time to sound-synchronous event), RT (reaction time to shift or disengage to lateral events), and PTLT distractor (proportion of total looking time to central distractor). Variables for the IPEP include RT (reaction time to select the synchronous target), PTLT (proportion of looking time to synchronous target), and PTTF (proportion of times the synchronous target was found).
[a] Did not meet significance cutoff when controlling for familywise error (six families for six ages: $p = .05/6 = .008$).
* $p < .05$. ** $p < .01$.

Table 5
Model Fit Indices for No-Growth, Linear Growth, and Quadratic Growth Curve Models Assessing Change in the Degree of English Language Exposure Across Age (3, 6, 12, 18, 24, and 36 Months)

Fit index	No-growth	Linear growth	Quadratic growth
Chi-square	47.35	39.75	16.62
df	19	16	13
p	.0003	.0008	.87
CFI	—	0.16	0.87
TLI	—	0.67	0.81
RMSEA [90% CI]	0.14 [0.09–0.18]	0.14 [0.08–0.19]	0.06 [0–0.13]
SRMR	0.10	0.06	0.04
AIC	3,372.93	3,371.33	3,354.21
BIC	3,392.18	3,397.81	3,387.90

Note. Fit indices include the comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error approximation (RMSEA), standardized root mean squared residual (SRMR), Akaike information criterion (AIC), and Bayesian information criterion (BIC).

Various techniques were used (e.g., correlations, logistic regressions) for analyses, which supported the conclusion that data were missing at random (MAR; Rubin, 1976). Therefore, we used full information maximum likelihood (FIML) estimation for all growth curve analyses (this section) and SEM analyses (subsequent section).[4]

Growth curve analyses were conducted on the proportion of English language exposure (relative to Spanish language exposure) from 3 to 36 months and assessed whether a no-growth model or a, linear, or quadratic growth model was a better fit with the data. Model fit indices appear in Table 5, model comparisons are in Table 6, and model parameters are in Table 7.[5] Likelihood ratio tests indicated that the quadratic growth model fit the data significantly better than both the linear growth model, $\chi^2(3) = 23.13$, $p < .001$, and the no-growth model, $\chi^2(6) = 30.73$, $p < .001$. Model fit indices for the quadratic model indicate fair-to-good fit for the model. The chi-square value was small and nonsignificant, indicating low levels of misfit, $\chi^2(13) = 16.62$, $p = .87$. A variety of other fit indices also indicated fair (comparative fit index [CFI]) or good (root mean square error approximation [RMSEA], standardized root mean squared residual [SRMR]) fit.[6] At 3 months (i.e., set as the intercept), the mean percentage of English language exposure relative to Spanish was 52.50, $SE = 2.82$. For each month from 3 to 12 months (i.e., the portion depicting the linear slope), English language exposure decreased by 1.74%, $SE = 0.64$, $p = .007$. Each month between 18 and 36 months (i.e., the portion depicting the increasing slope), English language exposure increased by 0.23%, $SE = 0.06$, $p < .001$ (see Figure 3).

Table 6
Model Comparisons (Assessed by the Likelihood Ratio Test) Among the No-Growth, Linear Growth, and Quadratic Growth Curve Models Assessing Change in the Degree of English Language Exposure Across Age (3, 6, 12, 18, 24, and 36 Months)

Likelihood ratio test	χ^2 difference	*df*	*p*
Linear growth and no-growth	7.60	3	.06
Quadratic growth and linear growth	23.13	3	<.001
Quadratic growth and no-growth	30.73	6	<.001

For children from bilingual families in South Florida, it appears that English language exposure changes across 3–36 months of age. Exposure to English language decreases (very slightly) across the first year, and then increases across the second and third years. Given that English language exposure is not constant across the ages employed in our main analyses, we use English language exposure at each age (rather than a single estimate across age) in subsequent analyses.

Multisensory Attention Skills for Social Events and Language Exposure in Dual-Language Learners: Individual Difference Analyses

We asked if there was a relationship between the degree of English language exposure and multisensory attention skills assessed by the MAAP and IPEP for social events in the children exposed to dual English–Spanish language environments. Would children with greater exposure to English demonstrate a language familiarity effect by showing enhanced attention skills to the women speaking English in the MAAP and IPEP? We approached this question using SEM analyses of two types: (a) analyses assessing the fit of null models, indicating that there is no effect of language familiarity on multisensory attention skills, and (b) analyses assessing the fit of freely estimated models, indicating there is a significant effect of language familiarity on multisensory attention skills. To evaluate fit, we focused on the global fit statistics of the two models. Global fit statistics indicated good fit for all null and freely estimated models (*ps* > .05). We used a chi-square difference test to determine whether the null or freely estimated model was a better fit to the data. We reasoned that if the null model fit and the freely estimated model had equivalent fit (i.e., no significant difference) then we could adopt the more parsimonious model as our final model (West et al., 2012).[7] Adopting this model indicates a failure to reject the null hypothesis, supporting the conclusion that there is no effect of English language familiarity on performance of attention indices assessed by the MAAP and IPEP. However, if the freely estimated model fit the

[4] FIML maximizes statistical power by borrowing information from the observed, available data (Enders, 2010). It has been shown to be appropriate for missing data rates around 50% (see Enders, 2010; Graham & Schafer, 1999) and to produce unbiased parameter estimates for data that are MAR.

[5] The no-growth model was used to calculate the comparative fit index (CFI) and the Tucker–Lewis index (TLI) for both the linear and quadratic models.

[6] The CFI was 0.87, indicating that our hypothesized model reduces 87% of the approximation error of the baseline model. The RMSEA was 0.06, 90% CI [0.00–0.13], indicating a 0.06 increase in standardized covariance residual per degree of freedom due to approximation error (The lower value of the confidence interval was ideally at the value of 0, and the interval contained the value of 0.05, indicating good fit. Finally, the SRMR demonstrated that the average residual correlation was 0.04, indicating good fit.

[7] When two nested models exhibit equivalent fit, it is standard practice to prefer the more parsimonious nested model that estimates fewer parameters (West et al., 2012; in this case, the null model where pathways are constrained to zero). Moreover, when using SEM to compare nested models, a less parsimonious model that estimates more parameters will always fit the data at least slightly better than a nested model that estimates fewer parameters. Thus, when the chi-square difference between the fit of our two models is non-significant, this indicates that freely estimating the pathways failed to significantly improve model fit (i.e., equivalent model fit), and the null model can be adopted.

Table 7

Model Parameters for the No-Growth, Linear Growth, and Quadratic Growth Curve Models Assessing Change in the Degree of English Language Exposure Across Age (3, 6, 12, 18, 24, and 36 Months)

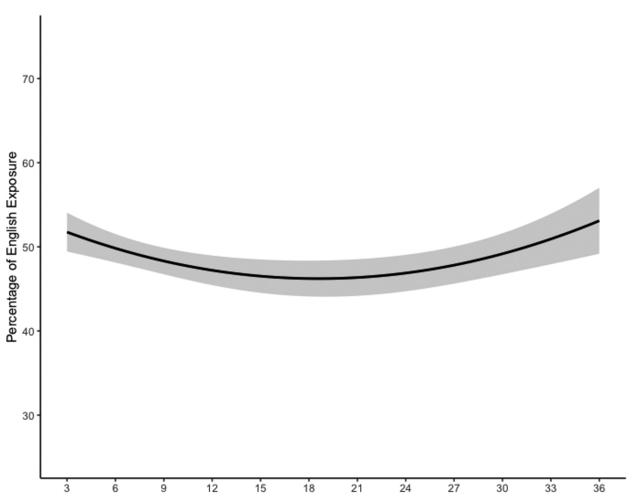
Parameter	No-growth			Linear growth			Quadratic growth		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Estimates									
Intercept	51.94	2.65	<.001	50.49	2.78	<.001	52.50	2.82	<.001
Linear slope	—	—	—	0.55	0.23	.02	−1.74	.64	.007
Quadratic slope	—	—	—	—	—	—	0.23	.06	<.001
Variances									
Intercept	529.17	89.44	<.001	553.33	97.67	<.001	600.29	104.27	<.001
Linear slope	—	—	—	0.60	0.59	.31	9.21	3.27	.005
Quadratic slope	—	—	—	—	—	—	0	0	0
3-month English exposure	100.29	26.01	<.001	99.12	28.96	.001	50.63	29.54	.09
6-month English exposure	144.87	32.39	<.001	130.75	30.75	<.001	138.47	30.82	<.001
12-month English exposure	328.90	65.98	<.001	335.52	67.24	<.001	317.19	64.10	<.001
18-month English exposure	204.44	46.74	<.001	207.58	47.19	<.001	165.92	43.15	<.001
24-month English exposure	207.54	48.24	<.001	205.01	49.55	<.001	160.90	45.96	<.001
36-month English exposure	328.60	66.70	<.001	233.66	64.43	<.001	211.86	105.75	.05
Covariances									
Intercept and linear slope	—	—	—	−4.66	5.63	.41	−26.65	16.63	.11
Intercept and quadratic slope	—	—	—	—	—	—	1.92	1.43	.18
Linear slope and quadratic slope	—	—	—	—	—	—	−0.66	0.17	<.001

data significantly better than the null model, we accepted the freely estimated model. We concluded that there is an effect of English language familiarity on that specific MAAP and/or IPEP variable at one (or more) of the six ages. This approach was used for each of the seven MAAP and IPEP measures at each of the six ages.

For both the null models and freely estimated models, we tested seven SEM models (one for each of the seven variables assessed by the MAAP and IPEP) characterizing paths between English language exposure and multisensory attention skills for social events at each age (a total of 42 contemporaneous pathways for each model type). All models included (a) degree of English language exposure and (b) a single multisensory attention skill measure (e.g., duration of sustained attention) at each of the six ages (3, 6, 12, 18, 24, and 36 months). Our focus was on contemporaneous pathways from English language exposure at a given age to attention skills at that same age (see Figure 4). For example, for sustained attention, we estimated pathways from English language exposure to sustained attention at each age (e.g., 3-month English language exposure predicts 3-month sustained attention, 6-month English language exposure predicts 6-month sustained attention, etc.; for an example, see Figure 4). However, in each model there were 51 pathways tested in total (six contemporaneous, 10 autoregressive, 20 prospective, and 15 cross-lagged pathways). In the freely estimated model, all 51 pathways were freely estimated. In the null models, the six contemporaneous pathways were fixed to 0 and the remaining 45 were freely estimated.

Figure 3

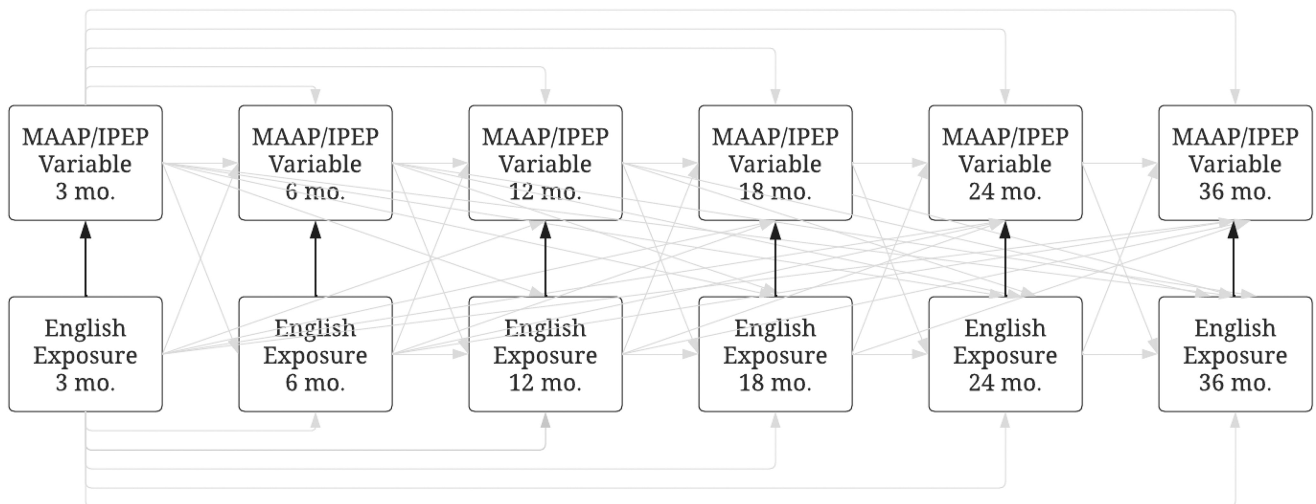
Quadratic Growth in the Degree of English Language Exposure in the Home Across Age (3, 6, 12, 18, 24, and 36 Months) With 95% Confidence Intervals



For null models, we conducted an SEM panel model in which pathways from degree of English language exposure to one of the multisensory attention skills (e.g., sustained attention) at each of the six ages were fixed so that the parameter estimate (and variance) was zero (however, we let all other pathways be freely estimated). If the null model and the freely estimated model for sustained attention had equivalent model fit (i.e., a nonsignificant chi-square difference test), we assumed the pathways from English language exposure to sustained attention at each age were not significantly greater than chance (i.e., not significantly greater than zero). For freely estimated models, the SEM panel models were identical to the null models, with the exception that there were no constraints on the pathways from English exposure to sustained attention at each of the six ages. If the freely estimated model fit the data significantly better than the null model (i.e., a significant chi-square difference test), we took two additional steps to reveal the ages and directionality of these effects. We examined the path coefficients from English exposure to sustained attention to assess their significance (i.e.,

Figure 4

Contemporaneous Paths of Interest for Structural Equation Models Depicting English Language Exposure Predicting Multisensory Attention Skills as Assessed by the Multisensory Attention Assessment Protocol (MAAP) and the Intersensory Processing Efficiency Protocol (IPEP)



Note. In order to estimate the unique relations between degree of English language exposure and multisensory attention skills, all autoregressive (e.g., 3-month English language exposure to 6-month English language exposure; $n = 10$ pathways), prospective (e.g., 3-month English language exposure to 12-month English language exposure; $n = 20$ pathways), and cross-lagged pathways (e.g., 3-month English language exposure to 6-month sustained attention; $n = 15$ pathways), were also included and freely estimated when testing each null and each freely estimated model. Also note that the contemporaneous pathway should be interpreted holding constant effects at prior ages. For example, the contemporaneous pathway between 12-month sustained attention and 12-month English language exposure 6-month sustained attention and English exposure, as well as 3-month sustained attention and English exposure.

did English exposure predict sustained attention) at each age, and evaluated the direction (positive, negative) of the significant path coefficient (e.g., did English exposure predict greater or lower sustained attention).

Global fit statistics appear in Table 8. Model comparisons (chi-square difference tests) between the null and freely estimated models for each variable from the MAAP and IPEP are depicted in Table 9. Model parameters for the paths of interest (contemporaneous paths: English language exposure predicting multisensory attention skills) in the freely estimated models are depicted in Table 10. (Also see Tables S11 and S12 in the online supplemental materials for all model parameters.) We next present results for MAAP and IPEP models in turn.[8]

MAAP: Null Versus Freely Estimated Models

The null models (i.e., that there is no relation between degree of English language exposure and multisensory attention) showed excellent fit for all four MAAP variables (accuracy of intersensory matching, duration of sustained attention, speed of shifting/disengaging, distractibility). For each of these variables, chi-square difference tests indicated that the null models exhibited fit equivalent with that of the freely estimated models χ^2s(6) < 8.94, *ps* > .18 (see Table 9). Therefore, we concluded the pathway between degree of English language exposure and each MAAP and IPEP variable at each age were not significantly greater than 0. Thus, English language exposure does not appear to affect indices of multisensory attentions skills to faces and voices on the MAAP at any age between 3 and 36 months.

IPEP: Null Versus Freely Estimated Models

The null model showed excellent fit for one IPEP variable (speed of intersensory matching). For this variable, chi-square difference tests indicated that the null model exhibited equivalent fit with that of the freely estimated model, $\chi^2(6) < 5.46$, $p = .49$ (see Table 9), allowing us to assume that the pathways from English exposure to speed of intersensory matching for faces and voices were not significantly greater than 0 at any age (3, 6, 12, 18, 24, 36 months). In contrast, the freely estimated models fit the data significantly better than the null models for two of the IPEP variables: accuracy of intersensory matching of faces and voices, $\chi^2(6) = 13.65$, $p = .03$, and frequency of fixating the speaking face, $\chi^2(6) = 12.87$, $p = .05$ (frequency of target selection). Therefore, we assessed the freely estimated path coefficients from English exposure to accuracy of intersensory matching and frequency of target selection (see Table 10). For both of these variables, English exposure was only a significant predictor at one of the six ages tested (3 months). English language exposure at 3 months of age predicted accuracy of intersensory matching of faces and voices at 3 months, $b = -0.12$, $SE = 0\ .04$, $p = .003$, as well as frequency of fixating the speaking face at 3 months, $b = -0.21$, $SE = 0.06$, $p = .001$. A 10% increase in English language exposure is

[8] All SEM models were also conducted controlling for maternal education level. All findings were similar in direction and magnitude to those from the main SEM models. Furthermore, there were no significant differences in maternal education between monolingual and dual-language learning participants, $t(95) = -0.33$, $p = .75$.

Table 8

Global Fit Statistics for the Freely Estimated and Null (Fixed) Models for Social Events for Each MAAP Measure (Sustained Attention, Intersensory Matching, Shifting/Disengaging, Distractibility) and Each IPEP Measure (Accuracy and Speed of Intersensory Matching and Frequency of Target Selection)

	Sustained attention		Intersensory matching		Speed of shifting/ disengaging		Distractibility	
MAAP fit indices	Free	Fixed	Free	Fixed	Free	Fixed	Free	Fixed
Chi-square	19.28	22.84	17.36	26.57	14.88	23.46	23.24	28.97
df	15	21	15	21	15	21	15	21
p	.2	.35	.30	.19	.46	.32	.08	.11
CFI	0.99	1	0.99	0.98	1	0.99	0.99	0.99
TLI	0.94	0.98	0.97	0.94	1	0.98	0.95	0.96
RMSEA [90% CI]	0.06 [0–0.13]	0.03 [0–0.10]	0.04 [0–0.12]	0.06 [0–0.12]	0.0 [0–0.10]	0.04 [0–0.11]	0.07 [0.00–0.13]	0.06 [0.00–0.11]
SRMR	0.04	0.05	0.05	0.06	0.04	0.06	0.03	0.03
AIC	6,239.15	6,230.71	5,834.12	5,831.33	3,712.37	3,708.96	−4,136.77	−4,139.64
BIC	6,418.21	6,395.92	6,013.71	5,996.55	3,891.95	3,874.18	−4,125.15	−4,125.15
	Speed of matching		Accuracy of matching		Frequency of selection			
IPEP fit indices	Free	Fixed	Free	Fixed	Free	Fixed		
Chi-square	28.78	34.24	14.48	28.13	16.94	29.81		
df	15	21	14	21	15	21		
p	.02	.03	.49	.14	.32	.10		
CFI	0.96	0.96	1	0.98	0.99	0.97		
TLI	0.82	0.87	1.01	0.93	0.97	0.92		
RMSEA [90% CI]	0.11 [0.04–0.17]	0.09 [0.02–0.14]	0 [0–0.10]	0.07 [0–0.12]	0.04 [0–0.12]	0.07 [0–0.13]		
SRMR	0.05	0.06	0.04	0.11	0.04	0.11		
AIC	4,225.18	4,218.64	5,582.45	5,584.10	6,208.83	6,209.70		
BIC	4,404.76	4,383.85	5,762.04	5,749.32	6,388.41	6,374.91		

Note. Fit indices include the comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error approximation (RMSEA), standardized root mean squared residual (SRMR), Akaike information criterion (AIC), and Bayesian information criterion (BIC).

associated with a 1.20-unit *decrease* in accuracy of matching faces and voices, and a 0.21-unit *decrease* in the frequency of fixating the speaking face. These results are in the opposite direction expected if English language exposure provided an advantage in intersensory processing of English audiovisual speech. Thus, greater English exposure at 3 months of age (but not at the other five ages) impaired performance on two of the IPEP variables at 3 months of age: accuracy of intersensory matching and frequency of target selection.

Table 9

Model Comparisons Between Freely Estimated Models and Null Models (Fixing Paths From English Language Exposure to Multisensory Attention Skills to 0) for Each Nonsocial Measure From the MAAP (Sustained Attention, Intersensory Matching, Shifting/Disengaging, Distractibility) and IPEP (Accuracy and Speed of Intersensory Matching and Frequency of Target Selection)

Measure	χ^2 difference	*df*	*p*
MAAP			
Sustained attention	3.56	6	.74
Intersensory matching	8.94	6	.18
Speed of shifting/disengaging	8.58	6	.20
Distractibility	5.74	6	.45
IPEP			
Speed of intersensory matching	5.46	6	.49
Accuracy of intersensory matching	13.65	6	.03
Frequency of target selection	12.87	6	.05

Multisensory Attention Skills for Nonsocial Events

Given that nonsocial events on the MAAP and IPEP present no language content, there was no reason to expect to find a relation between English language exposure and multisensory attention skills for object events. Performance under this condition should reflect general attention patterns rather than a language familiarity or novelty effect. Thus, although attention patterns in the nonsocial condition do not address the research questions of the present study, we include analyses of data from this condition in the supplement for purposes of comparison with social events (see supplement pp. 2–5, Tables S5–S10 in the online supplemental materials). In general, results for nonsocial events paralleled those of social events on the MAAP and IPEP. Group-level analyses (*t* tests between monolingual English and dual English–Spanish learners) indicated no significant group differences for the three IPEP variables (accuracy, speed, and frequency of intersensory matching) and for three of the four MAAP variables (accuracy of intersensory matching, speed of shifting/disengaging, and distractibility; see Table S6 in the online supplemental materials). For the fourth MAAP variable (sustained attention to object movements and sounds), there was only one significant difference at one of the six ages (3 months). Children exposed to monolingual English language environments showed greater duration of attention to object movements and sounds than dual English–Spanish language learners, $t(84) = 2.79$, $p = .007$. No significant group differences were evident at any of the other five ages (6, 12, 18, 24, and 36 months; $ps > .008$) for this variable. Thus, there were no significant group differences for 41 of 42 comparisons (seven variables at six ages each).

Table 10

Model Parameters for the Contemporaneous Pathways in the Freely Estimated Models: English Language Exposure Predicting Multisensory Attention Skills for Each Measure From the MAAP (Sustained Attention, Intersensory Matching, Shifting/Disengaging, Distractibility) and IPEP (Accuracy and Speed of Intersensory Matching and Frequency of Target Selection)

	Intersensory matching			Sustained attention			Speed of shifting/ disengaging			Distractibility		
Pathway	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
3 months MASk on 3 months English	−0.05	0.06	.43	−0.02	0.05	.74	−0.01	0.004	.04	0.002	0.06	.97
6 months MASk on 6 months English	0.11	0.08	.16	0.13	0.13	.33	0.001	0.003	.65	0.11	0.19	.57
12 months MASk on 12 months English	0.04	0.06	.55	0.02	0.11	.88	0.001	0.002	.66	−0.17	0.15	.24
18 months MASk on 18 months English	0.03	0.07	.66	0.02	0.13	.85	−0.01	0.003	.13	−0.24	0.18	.18
24 months MASk on 24 months English	−0.10	0.07	.13	0.16	0.17	.35	0.002	0.003	.51	−0.17	0.19	.36
36 months MASk on 36 months English	−0.14	0.07	.04	−0.11	0.09	.23	0.002	0.002	.30	0.18	0.15	.26
	Speed of matching			Accuracy of matching			Frequency of selection					
Pathway	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*			
3 months IP on 3 months English	−0.06	0.13	.62	−0.12	0.04	.003	−0.21	0.06	.001			
6 months IP on 6 months English	0.26	0.22	.24	0.04	0.04	.38	0.06	0.11	.61			
12 months IP on 12 months English	−0.20	0.19	.29	−0.02	0.03	.62	−0.11	0.09	.22			
18 months IP on 18 months English	−0.003	0.26	.99	−0.05	0.04	.20	−0.02	0.15	.87			
24 months IP on 24 months English	0.43	0.27	.11	0.09	0.05	.10	−0.11	0.16	.50			
36 months IP on 36 months English	0.17	0.23	.47	−0.04	0.06	.51	0.07	0.12	.58			

Note. MASk = multisensory attention skills; IP = intersensory processing.

Results of SEM analyses were also similar to those of social events. English language exposure was not related to any of the four MAAP variables (accuracy of intersensory matching, duration of sustained attention, and speed of shifting/disengaging to object events, or distractibility) or to speed of intersensory matching for objects and sounds on the IPEP (Tables S8 and S9 in the online supplemental materials). Furthermore, similar to results of the social models, the freely estimated models fit the data significantly better than the null models for the same two IPEP variables (accuracy of intersensory matching and frequency of target selection for objects and sounds; *p*s = .02) at just one of the six ages (12 months) and predicted a decrease in performance (*p*s = .001; see Tables S8 and S10 in the online supplemental materials). Overall, similar to findings of the social events, basic attention skills assessed by the MAAP and IPEP show no evidence of an advantage for either dual-language learners or infants and young children with a greater degree of English language exposure between 3 and 36 months of age.

Discussion

Recent research has shown that individual differences in what we call "multisensory attention skills" assessed in infancy using the MAAP and IPEP predict language outcomes in early childhood (Bahrick, Todd, et al., 2018; Edgar et al., 2022, 2003). Although the basic attention skills assessed by the MAAP (accuracy of intersensory matching, duration of sustained attention, speed of shifting/disengaging, distractibility) and IPEP (accuracy and speed of intersensory matching and frequency of target selection) require no language and can be assessed in infants as young as 3 months of age, it is not known if language familiarity impacts basic attention skills (i.e., if basic attention skills are enhanced in the context of speech events depicting a more familiar language), and in turn influences performance on these protocols. Given that a large number of children in the United States (22%) and an even greater number in South Florida (75%) come from homes in which another language (e.g., Spanish) is spoken, it is important to determine if familiarity with English impacts performance on these tests of basic attention skills. Might children with a high degree of exposure to English show heightened performance on the MAAP and IPEP (e.g., longer maintenance of attention, faster shifting/disengaging attention, better intersensory matching, less distractibility) compared to children with a higher degree of exposure to another language (e.g., Spanish)?

The present study examined whether individual differences in the degree of exposure to English versus Spanish in the home would influence children's multisensory attention skills. We addressed this question using a longitudinal sample of 104 children, 23 of whom were from households in which only English was spoken and 81 were from households in which English and Spanish were spoken (dual-language learners). Several novel findings emerged. We found: (a) no language familiarity effect—virtually no difference in multisensory attention skills to audiovisual speech events between children exposed to monolingual English versus dual English–Spanish language environments at any age, (b) similar patterns of attention as a function of English language exposure for both social (speech) and nonsocial (object) events, (c) that for dual-language learners, the degree of English versus Spanish language exposure in the home changes as a function of age, and (d) that greater English exposure provides no advantage on measures of basic multisensory attention skills for the audiovisual speech events in the MAAP or IPEP. Findings are each elaborated below.

No Evidence of an Advantage in Basic Multisensory Attention Skills for English Audiovisual Speech Events in Children With Greater English Language Exposure

Our main research question focused on characterizing the nature of the relation between the degree of English language exposure and multisensory attention skills to audiovisual social events (which portray women telling stories in English) assessed by the MAAP and IPEP at 3, 6, 12, 18, 24, and 36 months of age. We explored this question two ways: (a) assessing group differences between children from monolingual English backgrounds versus children from dual-language (English and Spanish) backgrounds, and (b) assessing individual differences in dual-language learners using SEM analyses.

Group Differences: Monolingual English Versus Dual-Language Learners

Overall, there was no evidence of an English language advantage for monolingual English language learners over dual English–Spanish language learners on any measure of the MAAP or IPEP at any of the ages assessed. Using a group-difference approach, we compared performance of monolingual English versus dual English–Spanish learners at each age (3, 6, 12, 18, 24, and 36 months) for each measure of the MAAP (accuracy of intersensory matching, duration of sustained attention, speed of shifting/disengaging, degree of distractibility) and IPEP (accuracy and speed of intersensory matching and frequency of target selection), a total of 42 comparisons. Overall, across the 42 comparisons, results revealed no evidence of an advantage in attending to English audiovisual speech for children exposed to monolingual English language environments over dual-language environments at any age, for any measure. To the contrary, only one group-difference was evident (intersensory matching at 18 months) and it was in the opposite direction expected if there were a language familiarity advantage. Thus, surprisingly, English language learners showed no advantage over dual-language learners in basic multisensory attention skills assessed by the MAAP and IPEP. These findings are also consistent with previous research demonstrating a lack of group differences in attentional control between infants exposed to a monolingual versus a dual-language home environment using the anticipatory looking paradigm with nonsocial events (D'Souza et al., 2020; Kalashnikova et al., 2021).

Individual Differences: Dual-Language Learners

Overall, there was no evidence that dual English–Spanish language learners showed an English language advantage for any of the multisensory attention skills at any age. Using an individual difference approach, we focused specifically on dual-language learners to assess the nature of the relation between the degree of English language exposure and performance on each of the multisensory attention measures of the MAAP and IPEP at each age. Given that the MAAP and IPEP assess basic attention skills that do not depend on exposure to a specific language, we were unsure if we would find a language familiarity effect—better performance for children with greater English exposure. We explored the relation between degree of English exposure in dual-language learners and multisensory attention skills by assessing the fit and then comparing fit indices of two models, a null model (that there is no relationship) and a freely estimated model (that there is a relationship). Models were tested for each of the measures assessed by the MAAP (attention maintenance, shifting/disengaging, intersensory matching, distractibility) and the IPEP (accuracy and speed of intersensory matching and frequency of target selection) at each of six ages (3, 6. 12, 18, 24, 36 months). Consistent with results of our group-level analyses, individual difference analyses revealed that the null models and the freely estimated models had equivalent fit for all of the MAAP measures and one of the IPEP measures (speed of intersensory matching). Although, for two of the IPEP measures (accuracy of intersensory matching and frequency of target selection) the freely estimated models showed significantly better fit than the null models, results were in the opposite direction expected if there were a language familiarity effect. *Lower* levels of English exposure (greater Spanish exposure) were associated with greater intersensory processing (accuracy of intersensory matching, frequency of target selection) of faces and voices at just one age (3 months). These findings are consistent with previous research demonstrating that home exposure to a second language did not influence looking behavior (i.e., dwell times, saccades) to audiovisual speech across the first year and a half (Schonberg et al., 2014).

Thus, surprisingly, across two analytic approaches, we found no evidence of an English language advantage for any of the attention skills assessed for audiovisual speech events by the MAAP or IPEP at any age. Furthermore, the finding that the null models and the freely estimated models had equivalent fit indicates support for the null hypothesis, that there is no relationship between degree of English language exposure and attention to English speech events between the ages of 3 and 36 months.

No Evidence of an Advantage in Basic Multisensory Attention Skills for Object Events for Dual English–Spanish Language Learners or for Children With Greater English Language Exposure

We did not expect to find a relation between English language exposure and multisensory attention for nonsocial events on the MAAP and IPEP given that they depict object movements and impact sounds and no language content. Consistent with this expectation, our supplemental analyses revealed virtually no difference in multisensory attention skills as a function of language background, paralleling results for social events. Group-level analyses found no difference in multisensory attention skills between monolingual English and dual English–Spanish language learners in the vast majority of comparisons across seven measures at six ages each (i.e., in 41 of 42 comparisons). Furthermore, for dual-language learners, SEM analyses revealed no relation between degree of English language exposure and performance on the MAAP and IPEP for the majority of measures, paralleling results with social events. The null models and the freely estimated models had equivalent fit for all of the MAAP and IPEP measures except two. In the two exceptions (similar to analyses for social events), greater Spanish and lower English exposure was associated with better performance. These findings are consistent with recent research demonstrating that infants exposed to bilingual language environments showed faster latency to shift attention to a target event than infants exposed to monolingual-language environment when a visual cue was presented on the opposite side of the screen from the target (Arredondo et al., 2022). Thus, as expected, there was no evidence of an advantage in multisensory attention skills for nonsocial events for children with greater English language exposure on the MAAP and IPEP. Furthermore, multisensory attention skills assessed by the MAAP and IPEP to social (English audiovisual speech) and nonsocial events (objects impacting a surface) show remarkably similar pattens as a function of language background.

The Nature of Change in Exposure to English Versus Spanish Across 3–36 Months of Age

We also assessed whether degree of exposure to English versus Spanish remains constant across age or whether it shows evidence of developmental change. Findings demonstrated significant evidence of U-shaped change across 3–36 months of age. English language exposure decreased slightly across the first year, and then increased across the second and third years. To the best of our knowledge, although prior studies have reported the degree of English versus Spanish language exposure in children in dual-language environments (e.g., Hoff et al., 2018), no studies have previously assessed the nature of this change across age. Hoff et al. (2018) modeled growth in both English and Spanish child expressive vocabulary size across 30–60 months and found that the percentage of English language exposure was related to English (but not Spanish) expressive vocabulary size. Similar to the present study, Hoff et al. (2018) used a separate measure of the percentage of English language exposure at each age as a predictor of growth in child expressive vocabulary size. These findings, together with findings from the present study, provide evidence that English language exposure in dual English–Spanish language learners is dynamic across age. Thus, as in the present study, when possible, a separate estimate of English/Spanish language exposure should be obtained for participants of each age in research assessing this variable.

Little Evidence for a Bilingual Advantage for Multisensory Attention Skills

In some studies, prior research has indicated a bilingual advantage for attentional control and executive function skills (Comishen et al., 2019) whereas other studies have indicated no advantage (D'Souza et al., 2020; Kalashnikova et al., 2021; Schonberg et al., 2014). However, across two different types of analytic approaches (group differences between monolingual English language learners and dual English–Spanish learners and analyses of degree of English language exposure for dual-language learners), we found very little evidence of a bilingual advantage for multisensory attention skills to audiovisual speech events and object events. Group-difference analyses showed no advantage for English language learners over dual-language learners in basic multisensory attention skills assessed by the MAAP and IPEP. Furthermore, we found minimal evidence of a Spanish language advantage in multisensory attention skills for dual-language learners when using individual difference (SEM) analyses. There was virtually no relationship between the degree of English versus Spanish language exposure and attention to English language speech events in the MAAP and IPEP between the ages of 3 and 36 months. Together, these findings are most consistent with prior studies indicating no bilingual attention advantage for infants, indicating no differences in dwell time or saccades (Schonberg et al., 2014) or anticipatory looking (D'Souza et al., 2020; Kalashnikova et al., 2021). They are also consistent with previous research examining older children, indicating no bilingual advantage in a series of tasks assessing executive functions in school-aged children (Dick et al., 2019; Paap & Greenberg, 2013). Thus, across both group-level and individual difference analyses, the present study indicates virtually no evidence of a bilingual advantage in basic attention skills assessed by the MAAP and IPEP between 3 and 36 months of age.

Conclusions and Broader Implications

In sum, our findings revealed no evidence of an English language advantage on tests of multisensory attention skills as assessed by the MAAP and IPEP across 3–36 months of age in monolingual English or dual English–Spanish language learners. Our data fit best with statistical models supporting a null hypothesis, that there is no relation between the degree of exposure to English and performance on each multisensory attention skill at each age. These findings indicate that basic multisensory attention skills, including sustaining attention, shifting and disengaging attention, and intersensory matching of the sights and sounds of speech or object events, as well as distractibility, are relatively unaffected by the specific language environment (English vs. Spanish) of children across 3–36 months of age.

Our findings also have a number of theoretical implications. First, they are consistent with the proposal that attention to audiovisual speech in early development is guided by detection of amodal properties (e.g., synchrony, rhythm, tempo, prosody, etc.) at the expense of modality-specific or language-specific information, (e.g., familiar words, syntax, etc.). This is also consistent with the body of research supporting the Intersensory Redundancy Hypothesis (Bahrick & Lickliter, 2000, 2012) as well as research demonstrating that perceptual development proceeds in order of increasing specificity, from detection of global information to increasingly more specific information (Bahrick, 1987, 2001; Gibson, 1969). Finally, these findings indicate that multisensory attention skills are domain general skills (similar patterns of findings for social and nonsocial events) and their development is not reliant on a specific language environment. Our findings also have practical implications. They support the conclusion that the MAAP and IPEP can be used effectively with children from various English and Spanish language backgrounds across the first 3 years of life.

Limitations and Future Directions

The are several limitations to the present study. First, our study focused on infants and young children learning English or both English and Spanish in South Florida. Although our findings indicate that the basic attention skills assessed by the MAAP and IPEP are relatively unaffected by the degree of English language exposure, it is not known how these attention skills would be affected by language learning environments other than English or English–Spanish. Therefore, it is critical for future research to test the generalizability of these findings to infants learning languages other than Spanish. Second, the social events of the MAAP and IPEP are currently only available depicting women speaking English. However, in the absence of a Spanish version of the MAAP or IPEP, we are unable to test language familiarity effects in both directions. If attention skills are unaffected by language familiarity, we would also expect no differences between language groups (monolingual English, monolingual Spanish) on either the English or Spanish versions of the MAAP and IPEP. Finally, our study focused on infants and young children ages 3 through 36 months. It is not known if older children, as they become more skilled in language, show differences in attention to the speech events depicted by the MAAP and IPEP as a function of language familiarity.

References

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*(9), 839–843. https://doi.org/10.1016/j.cub.2005.03.046

Alsius, A., & Soto-Faraco, S. (2011). Searching for audiovisual correspondence in multiple speaker scenarios. *Experimental Brain Research*, *213*(2–3), 175–183. https://doi.org/10.1007/s00221-011-2624-0

Amso, D., & Johnson, S. P. (2008). Development of visual selection in 3- to 9-month-olds: Evidence from saccades to previously ignored locations. *Infancy*, *13*(6), 675–686. https://doi.org/10.1080/15250000802459060

Arredondo, M. M., Aslin, R. N., Zhang, M., & Werker, J. F. (2022). Attentional orienting abilities in bilinguals: Evidence from a large infant sample. *Infant Behavior and Development*, *66*, Article 101683. https://doi.org/10.1016/j.infbeh.2021.101683

Atkinson, J., Hood, B., Wattam-Bell, J., & Braddick, O. (1992). Changes in infants' ability to switch visual attention in the first 3 months of life. *Perception*, *21*(5), 643–653. https://doi.org/10.1068/p210643

Ayneto, A., & Sebastian-Galles, N. (2017). The influence of bilingualism on the preference for the mouth region of dynamic faces. *Developmental Science*, *20*(1), 1–11. https://doi.org/10.1111/desc.12446

Bahrick, L. E. (1987). Infants' intermodal perception of two levels of temporal structure in natural events. *Infant Behavior and Development*, *10*(4), 387–416. https://doi.org/10.1016/0163-6383(87)90039-7

Bahrick, L. E. (2001). Increasing specificity in perceptual development: Infants' detection of nested levels of multimodal stimulation. *Journal of Experimental Child Psychology*, *79*(3), 253–270. https://doi.org/10.1006/jecp.2000.2588

Bahrick, L. E., Flom, R., & Lickliter, R. (2002). Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Developmental Psychobiology*, *41*(4), 352–363. https://doi.org/10.1002/dev.10049

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, *36*(2), 190–201. https://doi.org/10.1037/0012-1649.36.2.190

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, and Behavioral Neuroscience*, *4*(2), 137–147. https://doi.org/10.3758/CABN.4.2.137

Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 183–206). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199586059.003.0008

Bahrick, L. E., & Lickliter, R. (2014). Learning to attend selectively: The dual role of intersensory redundancy. *Current Directions in Psychological Science*, *23*(6), 414–420. https://doi.org/10.1177/0963721414549187

Bahrick, L. E., Lickliter, R., & Todd, J. T. (2020). The development of multisensory attention skills: Individual differences, developmental outcomes, and applications. In J. J. Lockman & C. S. Tamis-LeMonda (Eds.), *The Cambridge handbook of infant development* (pp. 303–338). Cambridge University Press.

Bahrick, L. E., McNew, M. E., Pruden, S. M., & Castellanos, I. (2019). Intersensory redundancy promotes infant detection of prosody in infant-directed speech. *Journal of Experimental Child Psychology*, *183*, 295–309. https://doi.org/10.1016/j.jecp.2019.02.008

Bahrick, L. E., Netto, D., & Hernandez-Keif, M. (1998). Intermodal perception of adult and child faces and voices by infants. *Child Development*, *69*(5), 1263–1275. https://doi.org/10.1111/j.1467-8624.1998.tb06210.x

Bahrick, L. E., & Pickens, J. N. (1988). Classification of bimodal English and Spanish language passages by infants. *Infant Behavior and Development*, *11*(3), 277–296. https://doi.org/10.1016/0163-6383(88)90014-8

Bahrick, L. E., Soska, K. C., & Todd, J. T. (2018). Assessing individual differences in the speed and accuracy of intersensory processing in young children: The Intersensory Processing Efficiency Protocol. *Developmental Psychology*, *54*(12), 2226–2239. https://doi.org/10.1037/dev0000575

Bahrick, L. E., Todd, J. T., & Soska, K. C. (2018). The Multisensory Attention Assessment Protocol (MAAP): Characterizing individual differences in multisensory attention skills in infants and children and relations with language and cognition. *Developmental Psychology*, *54*(12), 2207–2225. https://doi.org/10.1037/dev0000594

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, *70*(3), 636–644. https://doi.org/10.1111/1467-8624.00046

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, *7*(3), 325–339. https://doi.org/10.1111/j.1467-7687.2004.00351.x

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33–69. https://doi.org/10.1016/S0010-0277(97)00040-1

Bruce, M., Panneton, R., & Taylor, C. (2022). Multisensory integration and maternal sensitivity are related to each other and predictive of expressive vocabulary in 24-month-olds. *Journal of Experimental Child Psychology*, *214*, Article 105304. https://doi.org/10.1016/j.jecp.2021.105304

Butcher, P. R., Kalverboer, A. F., & Geuze, R. H. (2000). Infants' shifts of gaze from a central to a peripheral stimulus: A longitudinal study of development between 6 and 26 weeks. *Infant Behavior and Development*, *23*(1), 3–21. https://doi.org/10.1016/S0163-6383(00)00031-X

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*(2), 282–298. https://doi.org/10.1111/j.1467-7687.2008.00675.x

Colombo, J., Richman, W. A., Shaddy, D. J., Greenhoot, A. F., Colombo, J., Richman, W. A., Shaddy, D. J., Greenhoot, A. F., & Maikranz, J. M. (2001). Heart rate-defined phases of attention, look duration, and infant performance in the paired-comparison paradigm. *Child Development*, *72*(6), 1605–1616. https://doi.org/10.1111/1467-8624.00368

Comishen, K. J., Bialystok, E., & Adler, S. A. (2019). The impact of bilingual environments on selective attention in infancy. *Developmental Science*, *22*(4), Article e12797. https://doi.org/10.1111/desc.12797

Courage, M., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development*, *77*(3), 680–695. https://doi.org/10.1111/j.1467-8624.2006.00897.x

Curtindale, L. M., Bahrick, L. E., Lickliter, R., & Colombo, J. (2019). Effects of multimodal synchrony on infant attention and heart rate during events with social and nonsocial stimuli. *Journal of Experimental Child Psychology*, *178*, 283–294. https://doi.org/10.1016/j.jecp.2018.10.006

DeAnda, S., Arias-Trejo, N., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). Minimal second language exposure, SES, and early word comprehension: New evidence from a direct assessment. *Bilingualism: Language and Cognition*, *19*(1), 162–180. https://doi.org/10.1017/S1366728914000820

Dehaene-Lambertz, G., & Houston, D. (1998). Faster orientation latencies toward native language in 2-month-old infants. *Language and Speech*, *41*(1), 21–43. https://doi.org/10.1177/002383099804100102

Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., Riedel, M. C., Laird, A. R., & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the nationally representative ABCD study. *Nature Human Behaviour*, *3*(7), 692–701. https://doi.org/10.1038/s41562-019-0609-3

D'Souza, D., Brady, D., Haensel, J. X., & D'Souza, H. (2020). Is mere exposure enough? The effects of bilingual environments on infant cognitive development. *Royal Society Open Science*, *7*(2), Article 180191. https://doi.org/10.1098/rsos.180191

Edgar, E. V., Todd, J. T., & Bahrick, L. E. (2022). Intersensory matching of faces and voices in infancy predicts language outcomes in young children. *Developmental Psychology*, *58*(8), 1413–1428. https://doi.org/10.1037/dev0001375

Edgar, E. V., Todd, J. T., & Bahrick, L. E. (2023). Intersensory processing of social events at 6 months predicts language outcomes at 18, 24, and 36 months of age. *Infancy*. https://doi.org/10.1111/infa.12533.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Eschman, B., Todd, J. T., Sarafraz, A., Edgar, E. V., Petrulla, V., McNew, M., Gomez, W., & Bahrick, L. E. (2022). Remote data collection during a pandemic: A new approach for assessing and coding multisensory attention skills in infants and young children. *Frontiers in Psychology*, *12*, Article 731618. https://doi.org/10.3389/fpsyg.2021.731618

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, *60*(6), 1497–1510. https://doi.org/10.2307/1130938

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477–501. https://doi.org/10.1017/S0305000900010679

Fisher, A. V. (2019). Selective sustained attention: A developmental foundation for cognition. *Current Opinion in Psychology*, *29*, 248–253. https://doi.org/10.1016/j.copsyc.2019.06.002

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*(1), 238–252. https://doi.org/10.1037/0012-1649.43.1.238

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts.

Gogate, L., Maganti, M., & Bahrick, L. E. (2015). Cross-cultural evidence for multimodal motherese: Asian Indian mothers' adaptive use of synchronous words and gestures. *Journal of Experimental Child Psychology*, *129*, 110–126. https://doi.org/10.1016/j.jecp.2014.09.002

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 878–894. https://doi.org/10.1111/1467-8624.00197

Gogate, L. J., Walker-Andrews, A. S., & Bahrick, L. E. (2001). The intersensory origins of word-comprehension: An ecological–dynamic systems view. *Developmental Science*, *4*(1), 1–18. https://doi.org/10.1111/1467-7687.00143

Graham, J. W., & Schafer, J. L. (1999). On the performance of multiple imputation for multivariate data with small sample size. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 1–29). SAGE Publications.

Hoff, E., Burridge, A., Ribot, K. M., & Giguere, D. (2018). Language specificity in the relation of maternal education to bilingual children's vocabulary growth. *Developmental Psychology*, *54*(6), 1011–1019. https://doi.org/10.1037/dev0000492

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27. https://doi.org/10.1017/S0305000910000759

Kalashnikova, M., Pejovic, J., & Carreiras, M. (2021). The effects of bilingualism on attentional processes in the first year of life. *Developmental Science*, *24*(2), Article e13011. https://doi.org/10.1111/desc.13011

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12577–12580. https://doi.org/10.1073/pnas.0705345104

Lauro, J., Core, C., & Hoff, E. (2020). Explaining individual differences in trajectories of simultaneous bilingual development: Contributions of child and environmental factors. *Child Development*, *91*(6), 2063–2082. https://doi.org/10.1111/cdev.13409

Lewkowicz, D. J. (2000). Infants' perception of the audible, visible, and bimodal attributes of multimodal syllables. *Child Development*, *71*(5), 1241–1257. https://doi.org/10.1111/1467-8624.00226

Lewkowicz, D. J. (2003). Learning and discrimination of audiovisual events in human infants: The hierarchical relation between intersensory temporal synchrony and rhythmic pattern cues. *Developmental Psychology*, *39*(5), 795–804. https://doi.org/10.1037/0012-1649.39.5.795

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

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187–202. https://doi.org/10.1016/j.jecp.2012.09.002

Olsen, A. (2012). *The Tobii I-VT Fixation Filter: Algorithm description*. Tobii Technology.

Orena, A. J., Byers-Heinlein, K., & Polka, L. (2020). What do bilingual infants actually hear? Evaluating measures of language input to bilingual-learning 10-month-olds. *Developmental Science*, *23*(2), Article e12901. https://doi.org/10.1111/desc.12901

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Patterson, J. L. (2002). Relationships of expressive vocabulary to frequency of reading and television experience among bilingual toddlers. *Applied Psycholinguistics*, *23*(4), 493–508. https://doi.org/10.1017/S0142716402004010

Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development*, *82*(6), 1834–1849. https://doi.org/10.1111/j.1467-8624.2011.01660.x

Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychological Science*, *26*(4), 490–498. https://doi.org/10.1177/0956797614568320

Pons, F., Bosch, L., & Lewkowicz, D. J. (2019). Twelve-month-old infants' attention to the eyes of a talking face is associated with communication and social skills. *Infant Behavior and Development*, *54*(December), 80–84. https://doi.org/10.1016/j.infbeh.2018.12.003

Reynolds, G. D., Zhang, D., & Guy, M. W. (2013). Infant attention to dynamic audiovisual stimuli: Look duration from 3 to 9 months of age. *Infancy*, *18*(4), 554–577. https://doi.org/10.1111/j.1532-7078.2012.00134.x

Ribot, K. M., Hoff, E., & Burridge, A. (2018). Language use contributes to expressive language growth: Evidence from bilingual children. *Child Development*, *89*(3), 929–940. https://doi.org/10.1111/cdev.12770

Richards, J. E. (1985). The development of sustained visual attention in infants from 14 to 26 weeks of age. *Psychophysiology*, *22*(4), 409–416. https://doi.org/10.1111/j.1469-8986.1985.tb01625.x

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592. https://doi.org/10.1093/biomet/63.3.581

Schonberg, C., Sandhofer, C. M., Tsang, T., & Johnson, S. P. (2014). Does bilingual experience affect early visual perceptual development? *Frontiers in Psychology*, *5*, Article 1429. https://doi.org/10.3389/fpsyg.2014.01429

Shaddy, D. J., & Colombo, J. (2004). Developmental changes in infant attention to dynamic and static stimuli. *Infancy*, *5*(3), 355–365. https://doi.org/10.1207/s15327078in0503_6

Shaw, K., Baart, M., Depowski, N., & Bortfeld, H. (2015). Infants' preference for native audiovisual speech dissociated from congruency preference. *PLoS ONE*, *10*(4), Article e0126059. https://doi.org/10.1371/journal.pone.0126059

Soto-Faraco, S., Calabresi, M., Navarra, J., Werker, J. F., & Lewkowicz, D. J. (2012). The development of audiovisual speech perception. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 207–228). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199586059.003.0009

Soto-Faraco, S., Kvasova, D., Biau, E., Ikumi, N., Ruzzoli, M., Morís-Fernández, L., & Torralba, M. (2019). *Multisensory interactions in the real world* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781108578738

U.S. Census Bureau. (2020). *2020 ACS 1-year experimental data tables*. https://www.census.gov/programs-surveys/acs/data/experimental-data/1-year.html

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). Guilford Press.

Received July 18, 2022
Revision received January 18, 2023
Accepted February 22, 2023 ■