



# Inside bilingualism: Language background modulates selective attention to a talker's mouth

Joan Birulés<sup>1</sup> | Laura Bosch<sup>1</sup> | Ricarda Brieke<sup>1</sup> | Ferran Pons<sup>1</sup> | David J. Lewkowicz<sup>2</sup>

<sup>1</sup>Department of Cognition, Development and Educational Psychology, Universitat de Barcelona, Barcelona, Spain

<sup>2</sup>Department of Communication Sciences and Disorders, Northeastern University, Boston, Massachusetts

## Correspondence

Joan Birulés, Department of Cognition, Development and Educational Psychology, Universitat de Barcelona, Barcelona, Spain.  
Email: joanbirules@ub.edu

## Abstract

Previous findings indicate that bilingual Catalan/Spanish-learning infants attend more to the highly salient audiovisual redundancy cues normally available in a talker's mouth than do monolingual infants. Presumably, greater attention to such cues renders the challenge of learning two languages easier. Spanish and Catalan are, however, rhythmically and phonologically close languages. This raises the possibility that bilinguals only rely on redundant audiovisual cues when their languages are close. To test this possibility, we exposed 15-month-old and 4- to 6-year-old close-language bilinguals (Spanish/Catalan) and distant-language bilinguals (Spanish/"other") to videos of a talker uttering Spanish or Catalan (native) and English (non-native) monologues and recorded eye-gaze to the talker's eyes and mouth. At both ages, the close-language bilinguals attended more to the talker's mouth than the distant-language bilinguals. This indicates that language proximity modulates selective attention to a talker's mouth during early childhood and suggests that reliance on the greater salience of audiovisual speech cues depends on the difficulty of the speech-processing task.

## KEYWORDS

audiovisual speech, bilingualism, language development, linguistic distance, multisensory perception, selective attention

## 1 | INTRODUCTION

Whenever infants interact with other people, they not only hear but often also see them speaking. Consequently, infants have access to concurrent auditory and visual speech cues during many social interactions. When such cues are processed together, they provide redundant (i.e., overlapping and equivalent) perceptual cues that enhance the perceptual salience of speech signals. The greater perceptual salience of audiovisual speech cues is illustrated by findings that adults comprehend redundantly specified audiovisual speech better than auditory-only speech when the auditory information is unclear (Chandrasekaran, Trubanova, Stillitano, Caplier, & Ghazanfar, 2009; Risberg & Lubker, 1978; Sumbly & Pollack, 1954; Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998). It is also illustrated by findings that adults deploy their attention more to the source of redundant audiovisual speech, namely the talker's mouth. For instance, studies have

found that adults attend more to a talker's mouth when speech is at a low volume (Lansing & McConkie, 2003), when they have access to facial motion cues that provide linguistic information (Võ, Smith, Mital, & Henderson, 2012; Wass & Smith, 2014), when speech is difficult to understand (Reisberg, McLean, & Goldfield, 1987), and when they have to perform a challenging language detection task (Barenholtz, Mavica, & Lewkowicz, 2016).

If greater attention to redundant audiovisual speech cues facilitates speech processing in adults, is it possible that infants might attend more to redundancy during speech and language acquisition? Indeed, studies have found that multisensory redundancy generally enhances infant perception, learning, and memory (Bahrick & Lickliter, 2012; Bremner, Lewkowicz, & Spence, 2012; Burr & Gori, 2012; Lewkowicz, 1988, 2004, 2010; Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006) and, crucially, that infants detect audiovisual speech redundancy (Soto-Faraco, Calabresi, Navarra,



Werker, & Lewkowicz, 2012). For example, studies have found that infants can perceive the temporal congruence of lip movements and speech sounds (Dodd, 1979; Hillairet de Boisferon, Tift, Minar, & Lewkowicz, 2017; Lewkowicz, 2010; Pons & Lewkowicz, 2014) and that they can perceive the equivalence of auditory and visual speech and non-speech articulations (Kuhl & Meltzoff, 1982; Lewkowicz & Ghazanfar, 2006; Lewkowicz, Leo, & Simion, 2010; Patterson & Werker, 2003; Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009). Finally, evidence also indicates that infants can take advantage of audiovisual speech redundancy in that they, like adults, integrate audible and visible speech (Rosenblum, Schmuckler, & Johnson, 1997) in an automatic manner consistent with the McGurk effect (McGurk & MacDonald, 1976).

If infants can detect the redundancy of audiovisual speech cues then they are likely to attend to these cues in a talker's face. Indeed, evidence indicates that this is the case. In the first study to investigate the development of infant selective attention to talking faces and the effects of early experience on responsiveness, Lewkowicz and Hansen-Tift (2012) presented videos of a talker's face producing fluent-speech English or Spanish monologues to 4-, 6-, 8-, 10-, and 12-month-old monolingual, English-learning infants and English-speaking adults. When the talker spoke in the native language, 4-month-olds attended more to the eyes, 6-month-olds attended equally to the eyes and mouth, 8- and 10-month-olds attended more to the mouth, 12-month-olds attended equally to the eyes and mouth, and adults attended more to the eyes. When the talker spoke in the non-native language, results were identical except that this time 12-month-olds continued to attend more to the talker's mouth.

Lewkowicz and Hansen-Tift (2012) noted that the initial attentional shift from the talker's eyes to the talker's mouth between four and 8 months of age coincides with the emergence of two key milestones: canonical babbling (Oller, 2000) and endogenous attention (Colombo, 2001). Given this, Lewkowicz and Hansen-Tift (2012) concluded that the attentional shift from the eyes to the mouth reflects infants' emerging interest in audiovisual speech and their ability to voluntarily focus their attention on it. This conclusion is consistent with findings that greater attention to the mouth during the first year of life is associated with concurrent expressive language skills (Tsang, Atagi, & Johnson, 2018), greater rates of expressive language growth in the second year (Tenenbaum, Shah, Sobel, Malle, & Morgan, 2013; Tenenbaum et al., 2015; Young, Merin, Rogers, & Ozonoff, 2009), and not with the saliency of low-level visual features (Frank, Vul, & Johnson, 2009).

In addition to noting infants' new interest in audiovisual speech, Lewkowicz and Hansen-Tift (2012) concluded that the emergence of greater attention to a talker's mouth by 8 months of age reflects the greater attentional salience of redundantly specified audiovisual speech and infants' emerging reliance on the greater salience of audiovisual speech—as opposed to auditory speech—during the acquisition of native speech forms. Given this, these researchers also concluded that 12-month-old infants' greater attention to the talker's mouth in response to non-native speech most likely reflected the infants' need to rely more on the maximally salient audiovisual

## RESEARCH HIGHLIGHTS

- Overall, 15-month-old bilinguals attend more to a talker's mouth than eyes, but bilinguals learning rhythmically close languages attend more to the mouth than bilinguals learning distant languages.
- Four to six-year-old bilingual children also attend more to a talker's mouth than eyes but only if they are learning close languages.
- The findings show that bilinguals' reliance on redundant audiovisual speech cues depends on the rhythmical distance of their input languages and, thus, on the difficulty of their speech-processing task.

speech cues to help them disambiguate what, by this age, has become unfamiliar speech due to perceptual narrowing (Maurer & Werker, 2014).

The Lewkowicz and Hansen-Tift (2012) findings were replicated and extended to bilingual infants in a follow-up study by Pons, Bosch, and Lewkowicz (2015). In this study, 4-, 8-, and 12-month-old monolingual and bilingual infants learning either Catalan- or Spanish-learning or both languages were presented with faces that could be seen and heard talking in the infants' dominant language or in English. Results indicated that the monolingual infants exhibited the same developmental pattern of shifting attention as did the monolingual infants in the Lewkowicz and Hansen-Tift (2012) study. The results also showed that the bilingual infants started shifting their attention to a talker's mouth earlier in development than their monolingual counterparts and that they continued to focus on the talker's mouth in response to audiovisual speech in their dominant language at an age (i.e., at 12 months) when their monolingual counterparts no longer do so in response to native speech. Finally, the results showed that bilingual 12-month-old infants attended more to the talker's mouth when she spoke in a non-native language than did their monolingual counterparts, suggesting that bilingual infants rely on redundant audiovisual speech cues to help them with their greater speech-processing needs. The findings from the bilingual infants are consistent with findings showing that 8-month-old Catalan-Spanish bilinguals attend more to the mouth of a person expressing different affective expressions than do monolingual infants (Ayneto & Sebastián-Gallés, 2016) and that 15-month-old Catalan-Spanish bilingual infants do not learn to anticipate a movement that originates in a person's eyes because they attend more to the person's mouth (Fort, Ayneto-Gimeno, Escrichs, & Sebastián-Gallés, 2017).

Together, the findings from the Lewkowicz and Hansen-Tift (2012) and the Pons et al. (2015) studies demonstrate that (a) monolingual infants begin attending selectively to the redundant audiovisual cues inherent in a talker's mouth speaking in a native and non-native language by 8 months of age, (b) that they continue to do so in response to speech in an unfamiliar language at 12 months of age, and (c) that bilingual infants do so even more at this age in

response to their dominant language as well as in response to a non-native language. What is not clear from these findings, however, is whether the rhythmic and phonological distance of the two languages being learned by bilingual infants might differentially affect their deployment of attention to a talker's mouth. Specifically, it is possible that the bilingual Catalan- and Spanish-learning infants in the Pons et al. study deployed more attention to the mouth because Catalan and Spanish are rhythmically and phonologically close languages (Bosch & Sebastián-Gallés, 2001; Ramus, Nespor, & Mehler, 1999).

Evidence indicates that language proximity has an effect on various processes related to the acquisition of auditory-only speech. On the one hand, studies have found that infants acquiring a pair of distantly related languages (English-Spanish) can establish some of their vowel categories earlier than infants acquiring a pair of closely related languages as Catalan and Spanish (Bosch & Sebastián-Gallés, 2003; Sundara & Scutellaro, 2011). Moreover, studies have found that infants acquiring a pair of distantly related languages (i.e., Tagalog and English) can discriminate between them at birth (Byers-Heinlein, Burns, & Werker, 2010) but that infants acquiring rhythmically close languages only discriminate between them at 3–4 months of age (see Bosch & Sebastián-Gallés, 2001; Molnar, Gervain, & Carreiras, 2014; for Catalan-Spanish and Basque-Spanish, respectively). These studies suggest that both similarity in global rhythm and the degree of overlap in phonetic categories reduce the perceptual distance for some language pairs and, as a result, preclude their early differentiation. On the other hand, studies have found that language proximity facilitates vocabulary building and word learning due to the phonological similarity between the words in the two languages being learned (Bosch & Ramon-Casas, 2014; Havy, Bouchon, & Nazzi, 2016). Together, these data suggest that infants learning two close languages might find it more difficult to separate the sounds of each of their two languages. If that is the case, then it is possible that close-language bilingual infants as well as young children take greater advantage of audiovisual redundancy than do their more distant-language bilingual counterparts in situations where they have access to a talking face.

The goal of this study was to test this language-distance hypothesis and, thus, examine the possibility that bilingualism is a heterogeneous phenomenon and that the specific properties of the languages being acquired play an important role in audiovisual speech processing. To do so, we investigated whether the patterns of selective attention found in bilingual infants' and young children's response to a talker's eyes and mouth differ in those learning distant as opposed to close languages. In Experiment 1, we examined selective attention to a talker's eyes and mouth in 15-month-old bilingual infants who were either learning a pair of close or a pair of distant languages while in Experiment 2 we did the same except in 4- to 6-year-old children. Experiment 1 allowed us to test the language-distance hypothesis in infancy to further expand on the already discussed findings while Experiment 2 allowed us to extend this hypothesis to a later point in development where language acquisition is far more advanced.

## 2 | EXPERIMENT 1

### 2.1 | Method

#### 2.1.1 | Participants

We recruited 47 infants from a maternity hospital for this experiment. All were healthy, full-term infants with no history of hearing problems according to parental report. The Language Exposure Assessment Tool (DeAnda, Bosch, Poulin-Dubois, Zesiger, & Frienda, 2016) was used to establish an estimate of daily exposure to the language(s) being learned by the infants. To adequately represent the infants' bilingual environments, parents were instructed to indicate word productions in any of the infants' two languages.

To establish the relative proximity of the two languages being acquired by our bilingual participants, we considered factors such as rhythmic typology, phonetic-phonological category attributes, phonotactic structures, morphological complexity, lexical stress patterns, and the cognate status of words (Bosch, n.d.). Based on these criteria, we defined close bilinguals as those individuals who were acquiring languages that are close in terms of several of these factors and distant bilinguals as those acquiring languages that are further apart in terms of these factors. In this regard, it should be noted that robust metrics to unequivocally categorize languages along the proximity-distance axis are not available. Some variables, such as word-form similarity and phonological overlap have proven to be adequate to establish proximity at the lexical level (see Floccia et al., 2018). Here, however, we have adopted a more general approach that takes into account language parallelism in several dimensions, from global rhythm, phonology, and morphosyntactic structure to lexical stress patterns and word-form similarity. Participants in the close bilingual group were acquiring Catalan and Spanish, both of which are syllable-timed languages. Participants in the distant bilingual group were acquiring either Catalan or Spanish and either Russian, German, Swedish or Arabic (Grabe, 1994; Ramus et al., 1999). The latter languages belong to the stress-timed typology and, thus, differ from Spanish and Catalan not only in global rhythm attributes but also in terms of their segmental, morphological and lexical stress attributes.

Participants were divided into two groups based on their linguistic environment: Close bilinguals (Catalan/Spanish) and distant bilinguals (Catalan or Spanish and a rhythmically and/or phonetically distant language, described below). As in previous bilingual studies (Bosch & Sebastián-Gallés, 2001; see also Byers-Heinlein, 2015), we required that an infant's daily exposure to each of the input languages range between 50%–50% and 25%–75% of exposure time to each input language. Nine additional infants were tested but not included in the final data analyses due to: crying (4), failure to complete the calibration phase of the procedure (1), or failure to obtain a minimum of 9 s of data during a 45 s test trial (this equals to 20% of the test trial, after Frank, Vul, & Saxe, 2012) (4).



The final sample was composed of 38 infants (Mean age = 15 months, 5 days, Range = 14 months, 8 days to 15 months, 13 days). This included 20 close bilingual infants (Spanish-Catalan bilinguals: 11 infants were Catalan dominant and nine were Spanish dominant; Mean age = 14 months, 29 days, Range = 14 months, 20 days to 15 months, 7 days, 11 boys) and 18 distant bilingual infants (Spanish or Catalan-Other bilinguals: nine Catalan and nine Spanish, and Other refers to one Swedish, six German, four Russian, three Arabic, three French and one Rumanian infant; Mean age = 15 months, 12 days; Range = 14 months, 26 days to 15 months, 27 days; five boys). Please note that despite the fact that French and Rumanian are also Romance languages, each of these languages is nevertheless substantially different from Spanish and Catalan at the phonological, morphological, and lexical stress levels.

### 2.1.2 | Stimuli

The stimuli were identical to those used by Pons et al. (2015) and consisted of 45 s audiovisual video clips in which one of two female actors recited a prepared monologue. One of the actors (a highly proficient Catalan-Spanish bilingual) recited a Spanish or a Catalan version of the monologue, whereas the other actor (a native speaker of American English) recited an English version of the monologue. To elicit maximal attention, the actors recited the monologues in an infant-directed manner.

### 2.1.3 | Apparatus and procedure

Infants were seated on an infant seat while the parents sat behind them. Testing took place in a dimly lit and sound attenuated room and the stimuli were presented on a 17-inch computer monitor using Tobii Studio software (Tobii Technology AB). Eye gaze was recorded with a Tobii X120 stand-alone eye tracker at a sampling rate of 60 Hz. We used the Tobii eye tracker's five-point calibration routine to calibrate each participant's gaze. The experiment started with the calibration routine. Once calibration was successfully completed, we presented two videos, one in the infants' dominant native language (Catalan or Spanish) and one in a non-native language (English). The order of the videos was counterbalanced across infants. While the infants watched the videos, the eye-tracker monitored their gaze at two areas of interest (AOI), the eyes and the mouth. The AOIs used here were identical to those used by Pons et al. (2015).

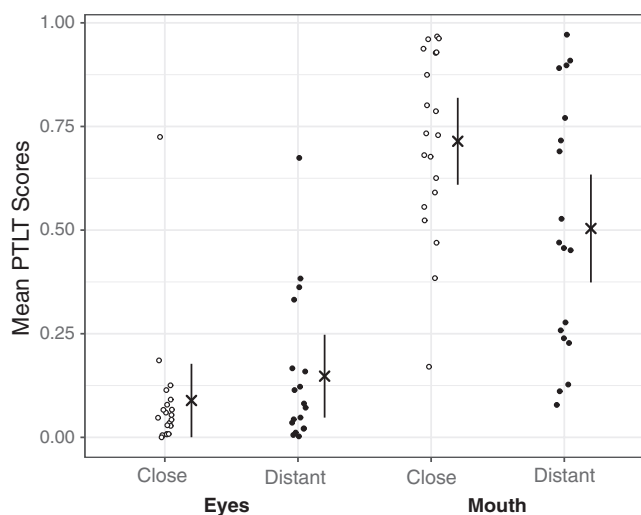
## 2.2 | Results and discussion

To ensure that overall attention levels did not vary across groups, we first analyzed the total amount of looking at the face with a one-way ANOVA with Linguistic Distance (close, distant) as a between-subjects factor. This analysis revealed that the two groups did not differ [close bilinguals  $M = 26.65$  s,  $SD = 8.95$ , distant bilinguals  $M = 29.62$ ,  $SD = 7.38$ ,  $F(1,36) = 1.23$ ,  $p = 0.27$ ]. Therefore, the pattern of shifting attention could not be attributable to differences in overall attention.

Next, we conducted analyses to determine whether attention to the eyes and mouth differed. This entailed comparing the proportion of total looking time deployed to the eyes and mouth AOIs out of the total amount of time infants looked at the face. These proportion-of-total-looking-time (PTLT) scores were computed for each participant by dividing the amount of time s/he looked at each AOI, respectively, by the total amount of time s/he looked at the face (please note that 78% of looking time to the face consisted of looking at the eye and mouth AOIs). We then compared the PTLT scores by way of a mixed, repeated-measures ANOVA, with AOI (eyes, mouth) and Test Language (native, non-native) as within-subjects factors and Linguistic Distance (close, distant) and Dominant Language (Catalan, Spanish) as between-subjects factors. We performed this analysis on untransformed scores as well as on arcsin transformed scores. Both analyses yielded the same results. Given that previous studies that have investigated infant selective attention to talking faces have relied on untransformed data, we only present the results from the analysis of the untransformed PTLT scores.

Results revealed a main effect of Test Language [ $F(1,34) = 4.97$ ,  $p = 0.032$ ,  $\eta^2 = 0.13$ ], a main effect of AOI [ $F(1,34) = 63.18$ ,  $p < 0.01$ ,  $\eta^2 = 0.65$ ], a significant AOI  $\times$  Test Language interaction [ $F(1,34) = 6.93$ ,  $p = 0.01$ ,  $\eta^2 = 0.17$ ], and a significant AOI  $\times$  Linguistic Distance interaction [ $F(1,34) = 4.97$ ,  $p = 0.037$ ,  $\eta^2 = 0.13$ ]. Even though none of the other interactions reached statistical significance, they are presented here to provide a complete picture of our results. These include the Test Language  $\times$  Linguistic Distance [ $F(1,34) = 0.03$ ,  $p = 0.87$ ] interaction, the Test Language  $\times$  Dominant Language [ $F(1,34) = 1.05$ ,  $p = 0.31$ ] interaction, the AOI  $\times$  Dominant Language [ $F(1,34) = 1.85$ ,  $p = 0.18$ ] interaction, the Test Language  $\times$  AOI  $\times$  Linguistic Distance [ $F(1,34) = 0.23$ ,  $p = 0.63$ ] interaction, the Test Language  $\times$  Linguistic Distance  $\times$  Dominant Language [ $F(1,34) = 0.11$ ,  $p = 0.74$ ] interaction, the Dominant Language  $\times$  AOI  $\times$  Linguistic Distance [ $F(1,34) = 0.001$ ,  $p = 0.98$ ] interaction, the Test Language  $\times$  AOI  $\times$  Dominant Language [ $F(1,34) = 0.00$ ,  $p = 0.98$ ] interaction, and the Test Language  $\times$  AOI  $\times$  Dominant Language  $\times$  Linguistic Distance [ $F(1,34) = 0.00$ ,  $p = 0.95$ ] interaction.

The significant Test Language main effect reflects infants' greater total PTLT (i.e., PTLT to the eyes AOI + PTLT to the mouth AOI) when exposed to the face talking in the non-native language than to the face talking in the native language. The significant AOI main effect was due to the fact that, overall, infants looked more at the mouth than the eyes. The significant Test Language  $\times$  AOI interaction was explored by way of separate repeated measures ANOVAs; one compared PTLT to the eyes in the native and the non-native language condition while the other compared PTLT to the mouth in the two test language conditions. The results from the analysis of PTLT<sub>eyes</sub> scores revealed that infants looked equally at the eyes in both the native and non-native language conditions [ $F(1,37) = 3.43$ ,  $p = 0.07$ ,  $\eta^2 = 0.08$ ]. In contrast, and as expected, the results from the analysis of the PTLT<sub>mouth</sub> scores revealed that infants looked longer at the mouth when they were exposed to the non-native than native language [ $F(1,37) = 7.42$ ,  $p = 0.01$ ,  $\eta^2 = 0.17$ ]. The fact that the effects



**FIGURE 1** Distribution of mean proportion-of-total-looking-time (PTLT) scores to the eyes and mouth as a function of Linguistic Distance (Close and Distant Bilinguals), collapsed across languages. Dots represent each infant's Mean PTLT score; Crosses with error bars represent Mean PTLT score and standard error of the mean for each group

of the Test Language were not affected by Linguistic Distance indicates that the distance between the non-native language (English) with the languages being acquired by the infants did not affect their attention.

Finally, the most interesting result from the standpoint of our language distance hypothesis was the  $\text{AOI} \times \text{Linguistic Distance}$  interaction. Figure 1 displays the mean PTLT scores for each AOI, collapsed across the two languages, as a function of linguistic distance. As can be seen, even though both the close and the distant bilingual groups looked more at the mouth than the eyes [ $t(19) = 7.99$ ,  $p < 0.01$ ,  $d = 3.27$ ;  $t(17) = 3.65$ ,  $p < 0.01$ ,  $d = 1.42$ , respectively], the close bilingual group looked more to the mouth than the eyes than did the distant bilingual group (difference score: PTLT to the eyes minus the PTLT to the mouth [ $t(34) = 2.17$ ,  $p = 0.037$ ,  $d = 0.70$ ]).

In sum, the current experiment yielded three principal findings. First, we found that, overall, 15-month-old bilingual infants attended more to the mouth than the eyes when they were exposed to a talking face. Second, we found that the infants attended more to the mouth than the eyes when the talker spoke in a non-native than native language and that they did so regardless of the linguistic distance between English and the bilinguals' second language. Finally, as predicted, we found evidence consistent with our language distance hypothesis. Close bilingual infants attended longer to a talker's mouth than did distant bilingual infants.

### 3 | EXPERIMENT 2

The purpose of this experiment was to investigate whether the greater amount of selective attention deployed to a talker's mouth by close as opposed to distant bilingual infants extends into early

childhood. Studies of phonological development indicate that the basic phonemes of the system are already established and are produced contrastively by around 5 years of age (Bosch Galceran, 2004) and that speech is mostly error-free by 6 years of age (Dodd, Holm, Hua, & Crosbie, 2003). If phonological expertise alone mediates selective attention to a talker's mouth then children might exhibit a reduction in selective attention to the mouth. If, however, the distance between the two languages being learned also contributes to the mouth bias in school-age children, then they may also focus more on a talker's mouth presumably because this helps them comprehend and/or disambiguate the two languages. Thus, the specific question that we addressed here was whether selective attention to a talker's face is modulated by the distance of the languages being learned during the preschool period.

We expected that young bilingual children learning two close languages might take greater advantage of redundant audiovisual cues than bilingual children learning two distant languages. This prediction would not only be consistent with our infant findings but also with findings showing that adults also take advantage of the redundancy of audiovisual speech when their speech-processing task is challenging. For example, Barenholtz et al. (2016) presented two videos of a talker uttering two different audiovisually-specified sentences and tracked their eyes gaze. In an immediately following phase of the experiment, participants were asked to determine which of the two previous sentences corresponded to an auditory-only sentence that was extracted from one of the previous videos. Results indicated that monolingual speakers attended more to the mouth of the talker uttering non-native language sentences than native language sentences, whereas bilingual speakers who were familiar with both languages exhibited equivalent looking to the eyes and mouth. This shows that adults devote more attention to a talker's mouth when a speech-encoding task requires them to process speech in an unfamiliar language. In some ways, bilingual children who are not yet fully competent in the two languages they are learning are confronted with a similar task as were the adults in the Barenholtz et al. (2016) study. Therefore, it is reasonable to expect that bilingual children also may take advantage of the greater salience of the redundant audiovisual cues available in a talker's mouth. In addition, it is reasonable to expect that those children learning close languages may continue the strategy of taking greater advantage of redundancy given that they have a more difficult task (i.e., language disambiguation) than do children learning two distant languages.

To our knowledge, only two studies to date have investigated selective attention to talking faces in children. In one of these studies, Pons et al. (2018) employed the same free viewing method used by Lewkowicz and Hansen-Tift (2012) to examine selective attention to a talker's eyes and mouth in 7-year-old children with Specific Language Impairment and in their typically developing (TD) peers. Findings indicated that, similar to infants, close language bilingual TD children deployed more attention to the talker's mouth. In a second study, Byers-Heinlein, Morin-Lessard, Poulin-Dubois, and Segalowitz (2014) also employed a free-viewing method with





distant bilingual infants and children ranging in age from 5 months to 6 years of age. This study also revealed greater attention to the mouth throughout this developmental period, with attention to the mouth being most accentuated at 20 months of age and then slowly transitioning to equal attention to the eyes and mouth by 5 years of age. No studies to date have investigated whether the linguistic proximity of young bilingual children's two languages modulates their attentional deployment to a talker's eyes and mouth. Therefore, we tested this possibility in 4- to 6-year-old bilingual children learning pairs of either close or distant languages. We hypothesized that, like in infants, language distance would affect children's relative deployment of selective attention to a talker's eyes and mouth.

## 3.1 | Method

### 3.1.1 | Participants

We recruited 39 children from a private school. None of the children had a history of hearing problems according to parental report. Parents completed an online language questionnaire to establish the language background of the participants. All participants came from Spanish, Catalan or Russian homes and they had been exposed to their second language early in life (i.e., before entering school). As a result, the participants were early sequential bilinguals, and could be divided into two groups based on the proximity of their two languages; 17 close bilinguals (Catalan and Spanish) and 15 distant bilinguals (Russian and Spanish). After entering school at 3 years of age, participants also had exposure to English as a third language. Considering the linguistic background of these children, and to ensure that they had comparable competence in Spanish (L1/L2) and English (L3), participants' competence in these two languages was obtained from the school (teachers' formal assessment) and compared with a Mann-Whitney U test. The test showed that the children's competence level in the two languages used in the experiment was equivalent across the two groups (for Spanish  $U = 76.5$ , n.s.; for English  $U = 161$ , n.s.).

Seven additional children were tested but were not included in the final data analysis because another language was spoken at home that was not Catalan, Spanish or Russian (4) or because we failed to obtain a minimum of 2 s of looking per each 10 s trial (3) (Frank et al., 2012). Thus, the final sample consisted of 32 children (Mean age = 5 years, 8 months, Range = 4 years, 2 months to 6 years, 9 months) of whom 17 were close bilinguals (Spanish-Catalan, Mean age = 5 years, 8 months, Range = 4 years, 2 months to 6 years, 8 months, 11 boys) and 15 were distant bilinguals (Spanish-Russian, Mean age = 5 years, 9 months, Range = 4 years, 4 months to 6 years, 9 months, nine boys).

### 3.1.2 | Stimuli

We first conducted a pilot test with the stimuli presented in Experiment 1 (two videos in infant-directed speech, 45 s-long each) and found that the children did not sustain sufficient attention for

videos this long. Therefore, we made similar stimuli for this experiment but ones that were more appropriate for 4- to 6-year-old children. The new stimuli consisted of shorter video clips (10 s) and were uttered in child-directed speech. In each video, one of two female actors uttered a short monologue (part of "the Snowman" story by Briggs, 1978) in their native language (one in Spanish and the other in American English) in a child-directed manner. Note that in this experiment, English was not an unfamiliar language for the participants: it was non-native but familiar (L3).

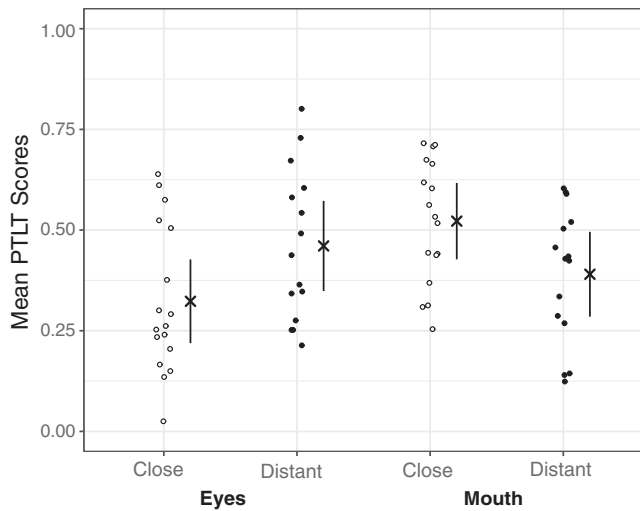
### 3.1.3 | Procedure

Children were seated on an adjustable chair, 60 cm in front of a Tobii 17-inch computer monitor, in a small and dimly lit room of their school. Like in Experiment 1, the stimuli were presented with Tobii Studio software (Tobii Technology AB) and eye gaze was recorded using a Tobii X120 stand-alone eye tracker at a sampling rate of 60 Hz. We used the Tobii eye tracker's nine-point calibration routine to calibrate each participant's gaze. After the calibration was completed, each participant watched the two video clips, one in Spanish and the other in English. The order of the videos was counterbalanced across children. While the children watched the videos, the eye-tracker monitored their gaze to the same two AOIs, namely the eyes and the mouth as in Experiment 1.

## 3.2 | Results and discussion

We used a mixed, repeated-measures ANOVA, with AOI (eyes, mouth) and Test Language (Spanish, English) as within-subjects factors and Linguistic Distance (close, distant) as the between-subjects factor to analyze the PTLT scores. Results yielded a main effect of Test Language [ $F(1,30) = 7.87$ ,  $p < 0.01$ ,  $\eta^2 = 0.21$ ] and an AOI  $\times$  Linguistic Distance interaction [ $F(1,30) = 5.13$ ,  $p = 0.03$ ,  $\eta^2 = 0.15$ ]. The remaining effects did not reach significance. These include the AOI main effect [ $F(1,30) = 1.17$ ,  $p = 0.29$ ,  $\eta^2 = 0.04$ ], the Test Language  $\times$  Linguistic Distance interaction [ $F(1,30) = 0.06$ ,  $p = 0.81$ ,  $\eta^2 = 0.00$ ], the Test Language  $\times$  AOI [ $F(1,30) = 0.09$ ,  $p = 0.77$ ,  $\eta^2 = 0.00$ ], and the Test Language  $\times$  AOI  $\times$  Linguistic Distance interaction [ $F(1,30) = 0.88$ ,  $p = 0.35$ ,  $\eta^2 = 0.03$ ].

The Test Language main effect reflects children's greater total PTLT (i.e., PTLT to the eyes AOI + PTLT to the mouth AOI) when exposed to the face talking in the non-native language than to the face talking in the native language. The AOI  $\times$  Linguistic Distance interaction indicates that the differential amount of attention deployed to the eyes and mouth depended on children's linguistic background. Figure 2 displays the mean PTLT scores for each AOI, collapsed across the two languages, as a function of linguistic distance. Follow-up paired t-tests showed that the close language group looked more at the mouth than the eyes [ $t(16) = 2.51$ ,  $p = 0.023$ ,  $d = 1.18$ ], whereas the distant language bilingual group looked equally at the eyes and mouth [ $t(14) = 0.78$ ,  $p = 0.45$ ,  $d = 0.92$ ]. The absence of a Test Language  $\times$  AOI interaction could be due to the fact that the non-native language was familiar to the participants in



**FIGURE 2** Distribution of mean proportion-of-total-looking-time (PTLT) scores to the eyes and mouth as a function of Linguistic Distance (Close and Distant Bilinguals), collapsed across languages. Dots represent each child's Mean PTLT score; Crosses with error bars represent Mean PTLT score and standard error of the mean for each group

Experiment 2 whereas the non-native language was unfamiliar to the infants in Experiment 1.

Overall, the results from this experiment show that, as a group (i.e., regardless of the distance of their two languages), 4- to 6-year-old bilingual children looked equally at a talker's eyes and mouth. Nonetheless, when language distance was taken into account, the findings showed that the distant-language children attended equally to the eyes and mouth whereas the close-language children attended longer to the mouth than the eyes. The different patterns of attention found in the two groups of children indicate that the distance of the languages being learned continues to affect the deployment of attention to the eyes and mouth of talking faces into early childhood.

## 4 | GENERAL DISCUSSION

Previous studies have demonstrated a modulatory effect of learning two close languages on selective attention to a talking face. These studies have found that bilingual infants learning Catalan and Spanish deploy more attention to a talker's mouth than do their monolingual counterparts (Ayneto & Sebastián-Gallés, 2016; Fort et al., 2017; Pons et al., 2015). Here, we investigated whether the greater attention previously shown in close-language bilingual infants would extend to bilingual infants and bilingual children learning other language pairs. To test this possibility, we measured eye gaze in 15-month-old bilingual infants and 4- to 6-year-old bilingual children while they heard and saw a talker speaking in their native language as well as in a non-native language. In each age group, some of the participants were learning close-language pairs (i.e., Catalan and Spanish) whereas other participants were learning distant-language

pairs (e.g., Spanish and Russian or Spanish and German). We predicted that greater attention to the talker's mouth may not only be driven by bilingualism per se but also by the distance of the two languages being acquired. This prediction was supported by our findings in that, both, bilingual infants learning close-language pairs and bilingual children learning close-language pairs attended more to a talker's mouth whereas bilingual infants learning distant-language pairs and bilingual children learning distant-language pairs did not.

Language distance is usually defined in terms of rhythmic, phonetic, phonotactic, morphological, and lexical attributes. Presumably, the closer that two specific languages are in terms of these attributes, the more challenging it is to distinguish between them. For example, in the auditory domain, studies have found that newborn infants can detect the rhythmical attributes of different languages (Ramus, Hauser, Miller, Morris, & Mehler, 2000) and that they can use this information to discriminate between languages that are rhythmically distant (Byers-Heinlein et al., 2010). In contrast, studies have found that it is not until between 3.5 and 4 months of age that infants can discriminate rhythmically close languages (Bosch & Sebastián-Gallés, 2001; Molnar et al., 2014). These findings suggest that the challenge for infants learning two languages, especially when they are close, is to disambiguate between the sounds of each language so that they can build distinct phonological representations for each language. One way in which close-language bilinguals can accomplish this is by deploying more attention to a talker's mouth. By doing so, they gain access to highly salient audiovisual speech cues which are known to facilitate speech processing in adults (Chandrasekaran et al., 2009; Risberg & Lubker, 1978; Sumby & Pollack, 1954; Vatikiotis-Bateson et al., 1998).

Our interpretation of the bilingual infants' and children's behavior in the current study is based on initial findings from studies of infant selective attention to talking faces demonstrating that once infants become interested in speech, both monolingual and bilingual infants begin to deploy more attention to a talker's mouth. In the first of these studies, Lewkowicz and Hansen-Tift (2012) found that monolingual infants begin deploying their attention to a talker's mouth by 8 months of age regardless of language, continue to do so at 10 months of age, and that they deploy more attention to a talker's mouth in response to non-native speech but not in response to native speech at 12 months of age. Lewkowicz and Hansen-Tift (2012) interpreted the attentional shift to the mouth as reflecting infants' attempt to gain access to the highly salient audiovisual speech cues located there and posited that this is likely to facilitate their ability to process and produce speech. This interpretation received additional support from a study by Pons et al. (2015) who found that bilingual infants deploy more attention to a talker's mouth than do monolingual infants. Pons et al. (2015) suggested that the reason that bilinguals attend more to the mouth is because they rely on the highly salient audiovisual speech cues to help them keep their two languages apart. The findings from the current study extend the previous two sets of findings by demonstrating that greater attention to a talker's mouth is not a characteristic of bilingualism per se. Rather, when the results from the previous two studies and the current one



are considered together, the conclusion that emerges is that greater attention to a talker's mouth in bilingual infants is due, in part, to the distance of the two languages being acquired. When the languages are close, bilingual infants and children attend more to a talker's mouth than when the languages are more distant.

Our finding that, overall, the infants attended to the talker's mouth more than the eyes and that, in addition, they attended more to the mouth when the language uttered by the talker was non-native is interesting in light of the earlier findings cited above. Specifically, Lewkowicz and Hansen-Tift (2012) found that 8- and 10-month-old infants attended more to the mouth than eyes both when a talker spoke in a native and a non-native language but that 12-month-old infants no longer did so when the talker spoke in their native language. The authors attributed equivalent attention to the mouth in the case of native speech to the fact that by 12 months of age infants have essentially acquired their initial native-speech phonological expertise. As a result, by this age, infants no longer have to rely on the greater perceptual salience of redundant audiovisual speech to disambiguate what has now become familiar speech. Although this interpretation might, at first blush, seem at odds with our finding that 15-month-old infants attended more to a talker's mouth, Pons et al. (2015) found that bilingual 12-month-old infants attended more to a talker's mouth than did monolingual infants and that they exhibited greater attention to a talker's mouth regardless of language. Other studies of bilingual infants also have found that they attend more to a talker's mouth than eyes than do monolingual infants (Ayneto & Sebastián-Gallés, 2016; Fort et al., 2017). Finally, a study of 14- and 18-month-old monolingual infants' selective attention to talking faces (Hillairet de Boisferon, Tift, Minar, & Lewkowicz, 2018) has found that, at these ages, monolingual infants attend more to a talker's mouth in response to native and non-native speech. Considered together, these findings suggest that infants continue to rely on the redundantly specified audiovisual speech cues when they are cognitively challenged by having to acquire two languages as well as when they face the challenging task of having to acquire a brand-new lexicon during the latter part of the second year of life (Hillairet de Boisferon et al., 2018).

Our results from the bilingual children provide new insights into the role of audiovisual redundancy in speech processing during development. Whereas our bilingual infants attended more to a talker's mouth regardless of language, and regardless whether they were learning close or distant languages, our bilingual children only did so when they were learning close language pairs. One reason why our bilingual children did not exhibit a preference for the talker's mouth may be due to the fact that phonological development is largely finished by 6 years of age (Bosch Galceran, 2004) and that children may also be interested in overall facial expressions rather than just the social information located in a talker's eyes. In addition, it should be noted that infants, but not adults, deploy greater attention to a talker's mouth spontaneously (i.e., in the absence of any specific task) but that adults only do so when their task requires them to explicitly process linguistic input (Barenholtz et al., 2016). This shows that young children's attentional strategy is more like that of adults

in that they only deploy greater attention to a talker's mouth when language processing becomes challenging (i.e., when they are learning close-language pairs). This interpretation is consistent with other evidence indicating that multisensory processing and integration are very much task-dependent processes (Murray, Lewkowicz, Amedi, & Wallace, 2016).

From a broader perspective, our findings not only provide crucial insights into the role of audiovisual redundancy in infant speech processing per se but also into the development of speech, language, and social responsiveness. With specific regard to the latter point, our findings suggest that once infants and young children gain direct access to what are, in effect, more intelligible speech signals, they gain access to more meaningful communicative information. This can have obvious cascading effects on other domains of cognitive development. For example, those infants who can quickly and efficiently deploy their attention to redundant audiovisual speech cues and, in the process, detect and learn native speech forms from their social partners, are infants who can profit maximally from their social partners' attempts to teach them canonical speech forms (Goldstein & Schwade, 2008; Kuhl & Rivera-Gaxiola, 2008). This, in turn, means that infants with this capability are more likely to establish reliable patterns of social reciprocity with their interlocutors than those infants who do not possess this capability. In addition, those infants who rapidly detect the redundant audiovisual speech cues in a talker's mouth are in a better position to extract and learn specific words from their interlocutors when they reach their second year of life. This enables them to better learn associations between newly acquired words and the objects being referenced by their social partners during joint-attention episodes. Indeed, this general theoretical scenario is supported by various empirical findings. For example, 12-month-old infants who attend more to a talker's mouth exhibit higher productive vocabulary scores at 18 and 24 months of age (Tenenbaum et al., 2015). In addition, relative to TD children, children with autism spectrum disorder exhibit less attention to faces. This is especially the case when they must exercise endogenous attentional control and when their social partners attempt to engage them by speaking to them (Chawarska, MacAri, & Shic, 2012). Most importantly from the current standpoint, children with autism spectrum disorder attend less to a talker's mouth than do TD children (Chawarska et al., 2012; Nakano et al., 2010) suggesting that measures of selective attention to a talker's mouth hold great promise for the detection and diagnosis of developmental disabilities.

In conclusion, the results from the present study show that linguistic distance plays an important role in mediating selective attention to talking faces throughout early childhood. The present results, together with previous ones (Lewkowicz & Hansen-Tift, 2012; Pons et al., 2015), also suggest that close-language bilinguals develop an attentional strategy that takes greater advantage of the greater perceptual salience of redundant audiovisual speech by 12 months of age, a point in development when bilinguals begin attending more to audiovisual speech cues in a talker's mouth than monolinguals. By adopting this strategy, close-language bilinguals can begin relying on the greater salience of redundant audiovisual cues to more easily



disambiguate and separate the languages they are learning. Given the present findings and interpretations, future studies exploring cognitive and linguistic abilities in bilingual populations should go beyond the comparison of monolinguals and bilinguals and embrace systematic comparisons of different groups of bilingual language learners. Doing so will help build a more accurate picture of the language learning process.

## ACKNOWLEDGEMENTS

This work was supported by Spanish Ministerio de Ciencia e Innovación Grant PSI2014-55105-P. Special thanks to all the infants and children who participated and to St. Peter's School (Barcelona). We thank Jéssica Sanchez-Galán for her assistance.

## CONFLICT OF INTEREST STATEMENT

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

## ORCID

David J. Lewkowicz  <http://orcid.org/0000-0001-6685-3178>

Joan Birulés  <http://orcid.org/0000-0001-9708-4922>

Ferran Pons  <http://orcid.org/0000-0001-5919-8590>

Laura Bosch  <http://orcid.org/0000-0002-6536-4855>

## REFERENCES

- Ayneto, A., & Sebastián-Gallés, N. (2016). The influence of bilingualism on the preference for the mouth region of dynamic faces. *Developmental Science*, 20(1), e12446. <https://doi.org/10.1111/desc.12446>
- Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. J. Bremner, D. J. Lewkowicz & C. Spence (Eds.), *Multisensory development* (pp. 183–206). New York, NY: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199586059.003.0008> <https://doi.org/10.1093/acprof:oso/9780199586059.001.0001>
- Barenholtz, E., Mavica, L., & Lewkowicz, D. J. (2016). Language familiarity modulates relative attention to the eyes and mouth of a talker. *Cognition*, 147, 100–105. <https://doi.org/10.1016/j.cognition.2015.11.013>
- Bosch, L. (n.d.). Language proximity and speech perception in young bilinguals: Revisiting the trajectory of infants from Spanish–Catalan contexts. In M. Gibson & J. Gil (Eds.), *Romance phonetics and phonology*. Oxford, UK: Oxford University Press.
- Bosch Galceran, L. (2004). *Evaluación fonológica del habla infantil*. Barcelona, Spain: Masson.
- Bosch, L., & Ramon-Casas, M. (2014). First translation equivalents in bilingual toddlers' expressive vocabulary: Does form similarity matter? *International Journal of Behavioral Development*, 38(4), 317–322. <https://doi.org/10.1177/0165025414532559>
- Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2(1), 29–49. [https://doi.org/10.1207/S15327078IN0201\\_3](https://doi.org/10.1207/S15327078IN0201_3)
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life\*. *Language and Speech*, 46(2–3), 217–243. <https://doi.org/10.1177/00238309030460020801>
- Bremner, A. J., Lewkowicz, D. J., & Spence, C. (2012). *Multisensory development*. Oxford, UK: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199586059.001.0001> <https://doi.org/10.1093/acprof:oso/9780199586059.001.0001>
- Briggs, R. (1978). *The snowman*. London, UK: Hamish Hamilton.
- Burr, D., & Gori, M. (2012). Multisensory integration develops late in humans. In M. Murray & M. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 1–18). Boca Raton, FL: CRC Press/Taylor & Francis. <https://doi.org/nbk92864>
- Byers-Heinlein, K. (2015). Methods for studying infant bilingualism. In J. W. Schwieter (Ed.), *The Cambridge Handbook of Bilingual Processing* (pp. 133–154). Cambridge, UK: Cambridge University Press.
- Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21(3), 343–348. <https://doi.org/10.1177/0956797609360758>
- Byers-Heinlein, K., Morin-Lessard, E., Poulin-Dubois, D., & Segalowitz, N. (2014). Monolingual and bilingual infants' attention to talking faces from 5–57 months. In Poster session presented at: Boston University Conference on Child Language Development, Boston, MA.
- Chandrasekaran, C., Trubanova, A., Stillitano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, 5(7), e1000436. <https://doi.org/10.1371/journal.pcbi.1000436>
- Chawarska, K., MacAri, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 53(8), 903–913. <https://doi.org/10.1111/j.1469-7610.2012.02538.x>
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, 52(1), 337–367. <https://doi.org/10.1146/annurev.psych.52.1.337>
- DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Frienda, M. (2016). A tutorial on expository discourse: Structure, development, and disorders in children and adolescents. *Journal of Speech, Language, and Hearing Research*, 59(December), 1–15. <https://doi.org/10.1044/2016>
- Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology*, 11(4), 478–484. [https://doi.org/10.1016/0010-0285\(79\)90021-5](https://doi.org/10.1016/0010-0285(79)90021-5)
- Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics and Phonetics*, 17(8), 617–643. <https://doi.org/10.1080/0269920031000111348>
- Floccia, C., Sambrook, T. D., Delle Luche, C., Kwok, R., Goslin, J., White, L., ... Plunkett, K. (2018). Analyses and results for study 1: Estimating the effect of linguistic distance on vocabulary development. *Monographs of the Society for Research in Child Development*, 83(1), 43–60. <https://doi.org/10.1111/mono.12350>
- Fort, M., Ayneto-Gimeno, A., Escrichs, A., & Sebastián-Gallés, N. (2017). Impact of bilingualism on infants' ability to learn from talking and non-talking faces. *Language Learning*, 68, 31–57. <https://doi.org/10.1111/lang.12273>
- Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, 110(2), 160–170. <https://doi.org/10.1016/j.cognition.2008.11.010>
- Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy*, 17(4), 355–375. <https://doi.org/10.1111/j.1532-7078.2011.00086.x>
- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to babbling facilitates vocal learning. *Psychological Science*, 19(5), 515–523. <https://doi.org/10.1111/j.1467-9280.2008.02117.x>
- Grabe, E. (1994). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology*, 7, 515–546. <https://doi.org/10.1515/9783110197105.2.515>



- Havy, M., Bouchon, C., & Nazzi, T. (2016). Phonetic processing when learning words: The case of bilingual infants. *International Journal of Behavioral Development*, 40(1), 41–52. <https://doi.org/10.1177/0165025415570646>
- Hillairt de Boisferon, A., Tift, A. H., Minar, N. J., & Lewkowicz, D. J. (2017). Selective attention to a talker's mouth in infancy: Role of audiovisual temporal synchrony and linguistic experience. *Developmental Science*, 20(3), 1–12. <https://doi.org/10.1111/desc.12381>
- Hillairt de Boisferon, A., Tift, A. H., Minar, N. J., & Lewkowicz, D. J. (2018). The redeployment of attention to the mouth of a talking face during the second year of life. *Journal of Experimental Child Psychology*, 172(April), 189–200. <https://doi.org/10.1016/j.jecp.2018.03.009>
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218, 1138–1141. <https://doi.org/10.1126/science.7146899>
- Kuhl, P. K., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual Review of Neuroscience*, 31(1), 511–534. <https://doi.org/10.1146/annurev.neuro.30.051606.094321>
- Lansing, C. R., & McConkie, G. W. (2003). Word identification and eye fixation locations in visual and visual-plus-auditory presentations of spoken sentences. *Perception & Psychophysics*, 65(4), 536–552. <https://doi.org/10.3758/BF03194581>
- Lewkowicz, D. J. (1988). Sensory dominance in infants: I. Six-month-old infants' response to auditory-visual compounds. *Developmental Psychology*, 24(2), 155–171. <https://doi.org/10.1037/0012-1649.24.2.155>
- Lewkowicz, D. J. (2004). Perception of serial order in infants. *Developmental Science*, 7(2), 175–184. <https://doi.org/10.1111/j.1467-7687.2004.00336.x>
- Lewkowicz, D. J. (2010). Infant perception of audio-visual speech synchrony. *Developmental Psychology*, 46(1), 66–77. <https://doi.org/10.1016/j.actpsy.2013.12.013>
- Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences of the United States of America*, 103(17), 6771–6774. <https://doi.org/10.1073/pnas.0602027103>
- 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>
- Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy*, 15(1), 46–60. <https://doi.org/10.1111/j.1532-7078.2009.00005.x>
- Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, 56(2), 154–178. <https://doi.org/10.1002/dev.21177>
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 691–811. <https://doi.org/10.1038/264746a0>
- Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, 19(3), 326–337. <https://doi.org/10.1111/inf.12041>
- Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory processes: A balancing act across the lifespan. *Trends in Neurosciences*, 39(8), 567–579. <https://doi.org/10.1016/j.tins.2016.05.003>
- Nakano, T., Tanaka, K., Endo, Y., Yamane, Y., Yamamoto, T., Nakano, Y., ... Kitazawa, S. (2010). Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 277(1696), 2935–2943. <https://doi.org/10.1098/rspb.2010.0587>
- Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science*, 9(5), 454–464. <https://doi.org/10.1111/j.1467-7687.2006.00512.x>
- Oller, D. K. (2000). The emergence of the speech capacity. *Journal of Child Language*, 30(3), 731–734. <https://doi.org/10.1121/1.1388001>
- Patterson, M. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science*, 6(2), 191–196. <https://doi.org/10.1111/1467-7687.00271>
- 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., & Lewkowicz, D. J. (2014). Infant perception of audio-visual speech synchrony in familiar and unfamiliar fluent speech. *Acta Psychologica*, 149, 142–147. <https://doi.org/10.1016/j.actpsy.2013.12.013>
- Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26), 10598–10602. <https://doi.org/10.1073/pnas.0904134106>
- Pons, F., Sanz-Torrent, M., Ferinu, L., Birulés, J., & Andreu, L. (2018). Children with SLI can exhibit reduced attention to a talker's mouth. *Language Learning*, 68, 180–192. <https://doi.org/10.1111/lang.12276>
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349–351. <https://doi.org/10.1126/science.288.5464.349>
- Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265–292. [https://doi.org/10.1016/S0010-0277\(99\)00058-X](https://doi.org/10.1016/S0010-0277(99)00058-X)
- Reisberg, D., McLean, J., & Goldfield, A. (1987). Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 97–113). Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Risberg, A., & Lubker, J. (1978). Prosody and speech-reading. *Quarterly Progress and Status Report Prosody and Speechreading*, 4, 1–16. Retrieved from [http://www.speech.kth.se/prod/publications/files/qpsr/1978/1978\\_19\\_4\\_001-016.pdf](http://www.speech.kth.se/prod/publications/files/qpsr/1978/1978_19_4_001-016.pdf)
- Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophysics*, 59(3), 347–357. <https://doi.org/10.3758/BF03211902>
- 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, UK: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199586059.003.0009>
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26(2), 212–215. <https://doi.org/10.1121/1.1907309>
- Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, 39(4), 505–513. <https://doi.org/10.1016/j.wocn.2010.08.006>
- Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., & Morgan, J. L. (2013). Increased focus on the mouth among infants in the first year of life: A longitudinal eye-tracking study. *Infancy*, 18(4), 534–553. <https://doi.org/10.1111/j.1532-7078.2012.00135.x>
- Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Shah, R. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language*, 42(6), 1173–1190. <https://doi.org/10.1017/S0305000914000725>



- Tsang, T., Atagi, N., & Johnson, S. P. (2018). Selective attention to the mouth is associated with expressive language skills in monolingual and bilingual infants. *Journal of Experimental Child Psychology*, 169, 93–109. <https://doi.org/10.1016/j.jecp.2018.01.002>
- Vatikiotis-Bateson, E., Eigsti, I. M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics*, 60(6), 926–940. <https://doi.org/10.3758/BF03211929>
- Võ, M. L.-H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision*, 12(13), 3. <https://doi.org/10.1167/12.13.3>
- Wass, S. V., & Smith, T. J. (2014). Individual differences in infant oculomotor behavior during the viewing of complex naturalistic scenes. *Infancy*, 19(4), 352–384. <https://doi.org/10.1111/infa.12049>
- Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science*, 12(5), 798–814. <https://doi.org/10.1111/j.1467-7687.2009.00833.x>

**How to cite this article:** Birulés J, Bosch L, Brieke R, Pons F, Lewkowicz DJ. Inside bilingualism: Language background modulates selective attention to a talker's mouth. *Dev Sci*. 2018;e12755. <https://doi.org/10.1111/desc.12755>