DOI: 10.1111/infa.12496

RESEARCH ARTICLE

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES WILEY

Changes in selective attention to articulating mouth across infancy: Sex differences and associations with language outcomes

Itziar Lozano[1,2] | David López Pérez[3] | Zuzanna Laudańska[3] | Anna Malinowska-Korczak[3] | Magdalena Szmytke[4] | Alicja Radkowska[3,4] | Przemysław Tomalski[3]

[1]Department of Cognitive Psychology and Neurocognitive Science, Faculty of Psychology, University of Warsaw, Warsaw, Poland

[2]Universidad Autónoma de Madrid, Faculty of Psychology, Madrid, Spain

[3]Neurocognitive Development Lab, Institute of Psychology, Polish Academy of Sciences, Warsaw, Poland

[4]Neurocognitive Development Lab, Faculty of Psychology, University of Warsaw, Warsaw, Poland

Correspondence
Przemysław Tomalski, Institute of Psychology, Polish Academy of Sciences, Jaracza,1,00-378, Warsaw, Poland.
Email: tomalski@mac.com

Itziar Lozano, Department of Cognitive Psychology and Neurocognitive Science, University of Warsaw, Faculty of Psychology, Stawki, 5/7, 00-183, Warsaw, Poland.
Email: i.lozano-sanchez@uw.edu.pl

Funding information
Polish National Science Centre, Grant/Award Numbers: 2011/03/D/HS6/05655, 2016/23/B/HS6/03860, 2018/30/E/HS6/00214; People Programme Marie Curie Actions of the EU FP7, Grant/Award Number: PCIG10-GA-2011-304255

Abstract
Infants' attention to the mouth is thought to support language acquisition, yet this relation has been scantly tested longitudinally. This study assessed attention to the mouth and the eyes at 5.5 (n = 91; Polish, 49% females) and 11 months, between time-point changes and their associations with language development in infancy (11 months) and toddlerhood (24 months). Sex differences were also explored. Results showed an age-related increase in looking to the mouth, and the magnitude of this change was associated with productive language, but only in toddlerhood. By contrast, looking to the eyes did not change and its duration at 5.5 months correlated with language development at 2 years. Exploratory analyses showed that in females but not males, reduced mouth-looking was related to better language outcomes in toddlerhood. Thus, looking to the mouth in infancy likely plays a long-term role in language acquisition and is potentially modulated by participant sex.

1 | INTRODUCTION

Early language acquisition is a complex learning process involving diverse mechanisms that require a variety of skills (Werker & Hensch, 2015). Especially in infancy, it relies on a wide range of perceptual, attentional, and motor mechanisms that are not specifically tied to language (i.e., not language-specific but domain-general processes) to support their gradual mastering of this higher-order cognitive function (Kuhl, 2004).

A crucial challenge that infants face is learning how to process speech. Speech has been traditionally treated as an auditory signal, but its audiovisual (AV) nature is currently widely acknowledged (Irwin & DiBlasi, 2017). Infants' speech perception is audiovisual from early on (Danielson et al., 2017), as illustrated by the compromised phonological development in blind children (Mills, 1987). Although AV speech processing mainly relies on infants' successful acquisition of perceptual skills (e.g., cross-modal integration or detection of correspondences between auditory and visual phonetic cues; for a review, see Lalonde & Werner, 2021), it relies on other mechanisms as well (e.g., memory, attention, and processing speed; Rose et al., 2009). Given the complexity, dynamics, and unpredictable nature of AV speech (Chandrasekaran et al., 2009), it is crucial for infants to learn how to deal with this input during their daily life communicative experiences with adults.

Attention constitutes a core mechanism for speech processing because it filters infants' access to properties of linguistic input (de Diego-Balaguer et al., 2016). Ultimately, through enhancing or disregarding different characteristics (e.g., prosodic, temporal cues), attention constrains and shapes the trajectory of language development (D'Souza et al., 2020). Selective attention to parts of a display is crucial for the efficient processing of AV speech. Here, it refers to the ability to focus on different features of talking faces at the expense of others (usually, differentially attending to the eyes relative to the mouth; e.g., Birulés et al., 2019; Pons et al., 2015). In previous work, it was generally assumed that selective attention to the mouth involves a relative preference to this part of the face (compared to preference for the eyes). Across the first year of life, infants learn how to modulate their attention to the eyes and the mouth to gain access to socio-communicative (Pons et al., 2019), emotional (Segal & Moulson, 2020), and presumably, also linguistic information. To date, two lines of evidence have addressed the role of selective attention to the mouth in language learning in infancy: cross-sectional studies exploring its changes during the first year of life (Kushnerenko, Tomalski, Ballieux, Ribeiro, et al., 2013; Lewkowicz & Hansen-Tift, 2012; Mercure et al., 2019; Pons et al., 2015; Tomalski et al., 2013; Wilcox et al., 2013) and longitudinal studies testing whether it predicts linguistic outcomes in infancy and toddlerhood (Kushnerenko et al., 2013; Tenenbaum et al., 2015; Tsang et al., 2018; Young et al., 2009).

Cross-sectional studies have revealed three key findings. Firstly, the typical pattern of changes in selective attention across the first year of life follows a trajectory involving more looking at the eyes at 4 months, followed by a shift to the mouth in 8-month-olds, and a balanced attention between the talker's eyes and mouth in 12-month-olds (see Lewkowicz & Hansen-Tift, 2012, though recent studies show that, despite a minor reduction of infants' mouth-looking at 12 months, it is still increasing until around 2 years; e.g., Morin-Lessard et al., 2019). This shift in looking to the mouth at 8 months occurs around the onset of canonical babbling when infants' interest in visual speech increases, suggesting that it is likely facilitating the pairing of phoneme–viseme correspondences. Secondly, early linguistic experience modulates this typical pattern. Infants, who experience greater variability in linguistic input (i.e., bilinguals; Pons et al., 2015), acquire two close languages (i.e., close bilinguals; Birulés et al., 2019), or who are exposed to a reduced AV linguistic experience (e.g., bimodal bilingual infants; Mercure et al., 2019) do not show equal attention to the eyes and the mouth at the end of the first year of life but keep preferentially attending to the mouth. This suggests that the period of time the infants typically need to benefit from visual speech cues can be protracted. Thirdly, the timing of changes in infants' pattern of selective attention to the eyes relative to the mouth (Lewkowicz & Hansen-Tift, 2012) seems to mirror the timing of changes in their perceptual narrowing to AV speech (Danielson et al., 2017; Kushnerenko, Tomalski, Ballieux, Ribeiro, et al., 2013; Pons et al., 2009). As infants become more attuned to their phonological native categories, they tend to preferentially attend to the mouth, until they become proficient enough and show a balanced attention between the talker's eyes and mouth (though see Wilcox et al., 2013). This suggests that infants' attention to the mouth is driven by their level of expertise in a language (the so-called '*language expertise hypothesis*'; e.g., Morin-Lessard et al., 2019). Although indirectly, altogether, these results suggest that selective attention to the mouth is an attentional skill that supports key processes underlying early language development (e.g., perceptual tuning to speech).

Despite this body of research, there is limited direct evidence for the role of selective attention to a speaker's mouth in infants' language acquisition for at least two reasons. Firstly, cross-sectional designs are unable to track within-subject changes in the mouth-relative-to-eyes-looking across infancy. This observation is important because, if selective attention to the mouth constitutes a language learning mechanism constrained to a specific period of life, then it is necessary to identify, within the same infants, what are the crucial periods when its use is relevant for language development. Secondly, none of the existing studies measured directly the relationship between these within-subject attention shifts in infancy and later language outcomes.

Evidence coming from longitudinal studies is more direct—it covers the latter limitation, but it is also considerably scanter. Crucially, no study tested within-subject changes in attention to the mouth. To our knowledge, only five studies have examined to date the predictive role of selective attention to the mouth for language development (Elsabbagh et al., 2014; Kushnerenko, Tomalski, Ballieux, Potton, et al., 2013; Tenenbaum et al., 2015; Tsang et al., 2018; Young et al., 2009), revealing three main findings: Firstly, that the period of 6–12 months of age is crucial for this mechanism, since associations between selective attention to the mouth and linguistic outcomes are only observed within this time window (but not slightly later, such as in 14-month-olds; see Elsabbagh et al., 2014). Tsang et al. (2018) found that a developmental increase in selective attention to the mouth occurs between these two time-points and that mouth-looking within this period is concurrently positively associated with expressive language. Consistently, more fixations to the mouth in a live mother–infant interaction at 6 months also predicted better expressive language outcomes in 24-month-olds and expressive language growth until 24 months (Young et al., 2009). This result has been replicated at older ages: 12-month-olds who looked longer at the speaker's mouth showed higher vocabulary size at 18 and 24 months (Tenenbaum et al., 2015). Altogether, these findings suggest that selective attention to the mouth may not operate as a learning mechanism for language acquisition throughout the entire infancy, but only within a specific period of time (similarly to other mechanisms supporting language acquisition; Werker & Hensch, 2015). Secondly, selective attention to the mouth specifically supports some language outcomes, facilitating expressive (e.g., babbling, phonemic utterances, and word production) but not receptive skills. This specificity has been found even in prelinguistic stages of production. For example, 6-month-olds that look more at the mouth of talking faces show increased vocal imitation (Imafuku et al., 2019), suggesting that relying more on the mouth may facilitate expressive language development through the construction of sensorimotor mappings of speech. Thirdly, increased attention to the mouth during the second half of the first year of life concurrently predicts infants' age-normed expressive skills (Tsang et al., 2018). More specifically, when controlling for age, infants who preferred looking at the speaker's mouth showed higher productive skills compared to those who preferred looking at the eyes—a finding interpreted as an indication that selective attention to the mouth is more strongly associated with infants' development of expressive skills rather than chronological age. Taken together, longitudinal evidence supports the idea that selective attention to the mouth constitutes a mechanism supporting language development.

Although the previously reviewed longitudinal research moves beyond some of the issues unexplored by cross-sectional studies, several others remain unsolved. Firstly, different methodological approaches (*free-viewing* task, Tsang et al., 2018, still face paradigm, Young et al., 2009) and stimuli (sometimes including both talking faces and objects; Tenenbaum et al., 2015) were used, which may affect infants' spontaneous looking patterns to facial features. In the current study, we used a *free-viewing* paradigm that exclusively measures selective attention to mouth and the eyes of the talking faces (see, e.g., Lewkowicz & Hansen-Tift, 2012). Another important aspect that remains unclear is discovering which factors account for individual differences in infants' developmental shifts in selective attention to the mouth. To our knowledge, only one cross-sectional study has explored changes in this mechanism from an individual differences approach (Morin-Lessard et al., 2019), yet neither bilingualism nor vocabulary size successfully accounted for heterogeneity in looking to the mouth. No evidence for an effect of bilingualism in mouth-looking was observed, and although monolinguals showed a positive association between productive vocabulary and mouth-looking, bilinguals with better comprehension looked less at the mouth, thus showing a different pattern and in the opposite direction, which was interpreted by the authors as inconclusive evidence.

It has been long known that there are sex differences in language development, with girls having a well-documented advantage in several linguistic domains (Marjanovič-Umek & Fekonja-Peklaj, 2017; but see Wallentin, 2009). The exact onset of these differences is unclear. For example, early prelinguistic skills, such as babbling, are influenced by testosterone levels as early as at 5 months, even in the absence of observable sex differences in language development at that point (Quast et al., 2016). Importantly, sex differences do not seem to be constant, but constrained to certain developmental stages (Etchell et al., 2018). Thus, it seems relevant to study if there are sex differences in language precursors and, more specifically, in the underlying language-learning mechanisms. If selective attention to the mouth in AV speech is a candidate mechanism supporting language acquisition, then sex differences may be present prior to sex differences in early linguistic outcomes. A single study has found sex differences in infants' visual attention to the mouth of emotional static faces at 9 months (Kleberg et al., 2019). However, no study has explored to date if sex is a potential factor accounting for individual differences in selective attention to the mouth of AV talking faces.

1.1 | The current study

The present study had three aims. Firstly, to longitudinally examine the pattern of changes in selective attention to the mouth relative to the eyes of a talking face between 5.5 and 11 months. To our knowledge, this is the first longitudinal study investigating changes in this skill at 2 time-points within the first year of life. We selected these time-points to match the periods prior to and subsequent to perceptual narrowing for audiovisual speech (Danielson et al., 2017; Pons et al., 2009). In line with previous studies (e.g., Tsang et al., 2018), we predicted an increase in attention to the mouth between the first and the second half of the first year of life. Secondly, we tested if infants' looking at the mouth relative to the eyes at 5.5 and 11 months is concurrently and prospectively associated with the linguistic outcomes at the end of the infancy period (11 months) and in toddlerhood (24 months). We also examined if the magnitude of change in selective attention to the mouth between those two time-points predicts language outcomes. Our longitudinal design allowed to test whether the role of this mechanism involves a relatively protracted period of infancy or if, instead, it operates at a shorter interval. Importantly, we measured both expressive and receptive language outcomes to test whether there is specificity and continuity across time in the supportive role of selective mouth-looking to expressive skills. In line with previous studies, we hypothesized that there would be a positive relationship between infants' preference for the mouth in AV speech in the second half of the first year of life and their language development (Imafuku & Myowa, 2016; Tenenbaum et al., 2015; Tsang et al., 2018; Young et al., 2009). More specifically, we predicted that the proportion of looking time to the mouth at 11 months (but not at 5.5 months) would be concurrently (11 months) and prospectively (24 months) associated with their performance in expressive and receptive language, meaning that longer looking to the mouth would result in better linguistic outcomes. Similarly, we expected that a greater magnitude of change in infants' preference for the mouth between 5.5 and 11 months would predict better expressive and receptive linguistic outcomes at 11 and 24 months. Thirdly, we aimed to explore if participant sex constitutes a factor that may contribute to individual differences in the development of selective attention to the mouth as a mechanism supporting language acquisition. We also examined potential sex differences in the associations between infants' preference for the mouth and subsequent language outcomes. We consider these analyses exploratory, so we do not anticipate any directional predictions.

2 | METHOD

2.1 | Participants

One hundred and 20 infants were recruited to participate in a longitudinal study on attention and infant–parent interactions at 5.5 months (T1). Out of those, 100 returned for the second visit at 11 months (T2). At T1, 77 participants had useable eye-tracking data for the free-viewing task. An additional 20 came but were excluded due to not contributing with valid looking data (because of either equipment failure, fussiness, crying, excessive movement, or calibration problems). Further 4 infants were excluded for having bilingual exposure at home (more than 20% of exposure to a second language; DeAnda et al., 2016) and additional 19 due to having looking times to the faces below 15% of the total trial length (based on Kleberg et al., 2019). At T2, 75 participants had useable eye-tracking data. An additional 9 did not contribute with valid looking data, further 4 had bilingual exposure, and 12 more had looking times to the speakers' face below 15% of the trial length.

The final sample consisted of 91 infants (45 females; 46 males) that contributed with useable eye-tracking data at either both time-points (n = 61) or only at 5.5 months (n = 16) or 11 months (n = 14). The mean age at 5.5 months was M = 164.7 days, SD = 12.8, range 134–187, while at 11 months it was M = 347.8 days, SD = 9.9, range 330–376. All infants were born at term (>36 weeks gestation age; birthweight M = 3406.5 g, SD = 474.5, range 2380–4600), and did not have any suspected or confirmed major medical conditions, including visual and auditory impairments. Infants' families consisted of mainly middle-class families coming from a city with >1.5 million inhabitants. The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the ethics committee of the University of Warsaw. All families received for their participation, a baby book and a certificate. Data collection was pre-pandemic (2013–2016).

2.2 | Eye-tracking procedure

At both time-points, the infants took part in a bigger project that included other eye-tracking tasks, EEG/ERP tasks, infant–parent interaction, questionnaires, and Mullen Scales of Early Learning

(Mullen, 1995). During the eye-tracking *free-viewing* task, infants sat on their parents' lap or in a high chair approximately 60 cm from the monitor. Eye-tracking data were collected on a Tobii T60 XL eye tracker (Tobii Inc.) with a 24″ monitor (1920 × 1200 pixels), 60 Hz sampling rate, and 0.5° accuracy (values provided by the manufacturer). The testing session began with a five-point infant-friendly calibration routine. The experimental task was presented after infants were successfully calibrated in at least 4 points. The free-viewing dynamic social stimuli were presented in 2 blocks (one video per each block), mixed with other eye-tracking tasks (not reported here) in two pseudo-random orders counterbalanced across participants. The entire eye-tracking procedure lasted <15 min. The stimuli were presented using the Matlab Psychophysics Toolbox (Brainard, 1997) and the eye-tracker was controlled with the Talk2Tobii package (Deligianni et al., 2011). Infants' behavior during the task was recorded using a remote-controlled CCTV camera.

2.3 | Eye-tracking stimuli

The stimuli consisted of two 30-sec-long videos with a female native Polish actress reciting baby rhymes (2 different actresses in total). Speakers were instructed to use auditory infant-directed speech (i.e., higher pitch and prosodic exaggerations) and visual 'gestured-directed speech' (i.e., direct gaze and exaggeration of face and lips movements). Similar AV speech stimuli have been previously used in studies exploring selective attention to talking faces (e.g., Lewkowicz & Hansen-Tift, 2012; Morin-Lessard et al., 2019). Videos were captured at a rate of 25 frames per second.

2.4 | Eye-tracking data analysis

For each trial, the areas of interest (AOI) were extracted by using Tobii Studio. AOI dimensions slightly varied across the trial to adjust to natural speakers' movements, which is an acknowledged challenge when using dynamic social stimuli (Hessels et al., 2018). To overcome this issue: (1) AOIs were as big as possible to cover the speakers' facial features despite changes in zoom, (2) eyes and mouth regions had equivalent sizes throughout the entire trial to keep regions as constant as possible, and (3) the shape of AOI that better fits with the shape of the moving facial features was chosen (rectangle for the eyes and oval for the mouth, to cover it despite talkers' variations in the size of openings when vocalizing). Three AOIs were drawn for each speaker (see Figure 1): eyes region (rectangular shape of maximum dimension 260 × 100 pixels in speaker 1 and 482 × 200 pixels in speaker 2), mouth region (oval shape of maximum dimension 260 × 100 pixels in speaker 1 and 482 × 200 pixels in speaker 2), and face region (oval shape of maximum dimension 305 × 390 pixels in speaker 1 and 610 × 810 pixels in speaker 2). Actresses' hands (squared shape of maximum dimension 415 × 380 pixels in speaker 1 and 490 × 280 in speaker 2) showed up during some episodes of the clip. Episodes of the trial when hands occluded the face of the actresses were excluded from the analyses (one fragment of 11 s in actress 1 and two fragments of 1.5 and 3.4 s in actress 2).

After extracting AOIs, the total fixation duration was computed for each of them using an in-house MATLAB script (Supplementary Methods). The eye-tracking measure was the proportion of total looking time (henceforth, PTLTs) for the eyes (PTLT Eyes) and mouth (PTLT Mouth). It was calculated for each trial (actress 1 and actress 2), each participant, and at each time-point (5.5 and 11 months) by dividing the total looking time to the eyes and mouth, respectively, by the total looking time to the face (as in prior studies exploring selective attention to facial features in audiovisual speech; Berdasco-Muñoz et al., 2019; Imafuku et al., 2019). Our rationale for this was that, when

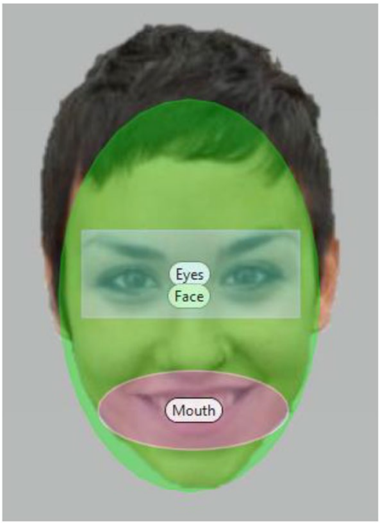

FIGURE 1 Example of stimuli and areas of interest used. This face is an example of how the original stimulus looked like

TABLE 1 Model structure comparisons and fit indices of the main analysis

	Predictor(s)	AIC	−2LL	Df[a]		Δχ2	Δdf	*p*-value
Model 0	Only intercept	14.685	8.685	3				
Model 1	M0 + Time-point	12.654	4.654	4	0 versus 1	4031	1	<0.001
Model 2	M1 + AOI	6.115	−3.885	5	1 versus 2	8539	1	<0.001
Model 3	M2 + Sex	3.856	−8.144	6	2 versus 3	4259	1	<0.001
Model 4	M3 + Time-point × AOI	−14.656	−28.656	7	3 versus 4	20512	1	<0.001
Model 5	M4 + Time-point × Sex	−13.196	−29.196	8	4 versus 5	540	1	<0.001
Model 6	M5 + AOI × Sex	−15.752	−33.752	9	5 versus 6	4556	1	<0.001
Model 7	**M6 + Time-point × AOI × Sex**	**−15.763**	−35.763	10	6 versus 7	2011	1	**< 0.001**

Note: Bold numbers highlight the model with the best fit.

Abbreviation: AIC, Akaike's Information Criterion; −2LL, −2 log likelihood; Df, degrees of freedom; Δχ2, Change in the -2LL between two models; Δdf, Change in the dfs between two models.

[a]Df is equivalent to the number of parameters of each estimated model.

encountering AV speech, infants' attention mainly competes between these two AOIs relative to the entire face, instead of between themselves. For each time-point and AOI, there was consistency of PTLTs between actresses (Supplementary Analysis 1).

All statistical analyses were performed using IBM SPSS Statistics 27 (IBM Inc., 2020). We used linear mixed-effect models to analyze our whole sample despite having multiple missing data points. Fitting of the model was done in three steps. First, we compared the structure of several models until we found the one with the best fit. To make them comparable, we kept all predictors fixed and used the maximum likelihood (ML) estimation (as recommended by Field, 2018) and Compound Symmetry (CS) covariance matrix structure. Using a forward variable selection, we estimated eight models (Table 1) by adding Time-point (5.5 and 11 months), AOI (eyes vs. mouth), and Sex (female vs. male) and all possible two (Time-point × AOI, Time-point × Sex, AOI × Sex) and three-way (Time-point × AOI × Sex) interactions one at a time and then assessed the change in likelihood-ratio test ($\Delta\chi^2$) and AIC. Including each of these predictors significantly improved the model fit (all *p*s. < 0.01). Thus, we selected the full model that kept all predictors (Model 7), which also had the lowest Akaike Information Criteria (AIC = −15.76).

Second, based on the selected model, we estimated three models to find the type of effects with the best fit, starting from an all-fixed effect to a maximal random effects structure (Table 2). We selected the model that (1) had the lowest AIC (−15.76) and (2) converged (Model 7.1), which was the same one selected above, with all predictors, intercept, and slope remaining as fixed. Yet maximizing random effects is usually recommended in mixed models (Barr et al., 2013), it has been also nuanced that this is only advisable as long as the model converges (Bates et al., 2015; Matuschek et al., 2017).

Third, we replicated the selected model, but using Restricted Maximum likelihood (REML) as a model estimator, since not all measures met the assumptions of normality (significant Shapiro–Wilk's test and inspection of Q–Q plots). The final coded model included the proportion of total looking time as the outcome variable, with Time-point (5.5 months and 11 months), AOI (eyes vs. mouth), and Sex (female vs. male) and all possible two-way and three-way interactions as fixed effect factors (see full syntax and data in https://osf.io/t93x8/). Time-point and AOI were entered as within-subjects factors and Sex as a between-subjects factor. Intercept and slopes for subjects were also specified as fixed and REML was used as the model estimator. The same model specifications were used in both main and Supplementary Analyses 1, 2, and 3.

2.5 | Language outcome measures

We used several instruments to measure expressive and receptive language development, including both parent report and direct assessment (Table 3). At 11 months, we assessed infants' vocabulary comprehension and production with a Polish adaptation of the *MacArthur-Bates Communicative*

TABLE 2 Model comparisons to find the type of effects with the best fit. Random intercepts and/or slopes were set for subjects

Predictors	Model 7.1	Model 7.2*	Model 7.3*
Intercept	Fixed	Random	Random
Slope	Fixed	Fixed	Random
Time-point	Fixed	Fixed	Fixed
AOI	Fixed	Fixed	Fixed
Sex	Fixed	Fixed	Fixed
Time-point × AOI	Fixed	Fixed	Fixed
Time-point × Sex	Fixed	Fixed	Fixed
AOI × Sex	Fixed	Fixed	Fixed
Time-point × AOI × Sex	Fixed	Fixed	Fixed
Evaluation			
AIC	**−15.76**	−13.76	−13.76
-2LL	−35.76	−35.76	−35.76
dfs	10	11	11
Convergence	Yes	No	No

Note: Bold numbers highlight the model with the best fit.

*ID covariance structure was used for random intercepts and slopes.

TABLE 3 Language Outcomes at 11 (infancy) and 24 months (toddlerhood) for the entire Group and for each Sex*

		All infants		Females		Males			
	N	*M* (*SD*)	Range	*M* (*SD*)	Range	*M* (*SD*)	Range	*t*-test	Sig.
Infancy									
MSEL receptive language	71	13.00 (1.18)	10–16	13.25 (1.19)	11–16	12.79 (1.15)	11–16	−1.63	0.10
MSEL expressive language	72	12.06 (1.56)	8–16	12.30 (1.44)	9–15	11.87 (1.65)	9–15	−1.16	0.24
CDI-WG- words said	64	5.43 (6.89)	0–29	6.48 (8.04)	1–26	4.67 (5.91)	1–26	−1.03	0.30
CDI-WG- words understood	64	70.17 (58.38)	4–274	76.33 (71.25)	5–274	65.67 (47.42)	5–274	−0.71	0.47
Toddlerhood									
MSEL receptive language	74	27.05 (3.34)	19–39	27.55 (3.49)	19–39	26.47 (3.11)	20–33	−1.39	0.16
MSEL expressive language	71	22.25 (3.91)	12–30	23.65 (3.63)	14–30	20.63 (3.62)	12–27	−3.49	0.001**
CDI-WS	70	59.14 (33.09)	3–100	71.16 (31.00)	5–100	45.66 (30.42)	3–97	−3.46	0.001**
OTSR words understood	75	9.97 (6.09)	0–27	10.70 (6.10)	0–27	9.14 (6.06)	0–26	−1.10	0.27

Note: OTSR (Picture vocabulary test). KIRMIK: Krótki Inwentarz Rozwoju Mowy i Komunikacji (*Short inventory of speech and Communication Development*).

*These statistics correspond to infants of our sample with valid eye-tracking data at 11 months and complete language outcomes at this same age.

Development Inventory—Words and Gestures (CDI-WG, Fenson et al., 2007) (IRMIK, Smoczyńska et al., 2015) and receptive and expressive language with a Polish translation of the corresponding MSEL subscales (Mullen, 1995).

At 24 months, we measured receptive language performance with a Polish version of the MSEL Receptive Language Scores (Mullen, 1995) and the Polish picture vocabulary test of comprehension based on choosing one colorful picture depicting the named word out of other competing distractor pictures (OTSR; Haman et al., 2012). To assess expressive language, we used two instruments: a short Polish version (KIRMIK; Krajewski & Smoczyńska, 2015) of the CDI Words and Sentences and a Polish translation of the MSEL Expressive Language subscale (Mullen, 1995). For all language outcomes, we used raw scores.

Since one of our predictions focused on exploring sex differences in selective attention to the mouth, we also investigated sex differences in language outcomes in these two periods. Female and male infants did not significantly differ at 11 months (all *p*s < 0.05; see Table 3); however, at 24 months females showed higher expressive language skills than males in both MSEL Expressive Language, $t(69) = -3.49$, $p = 0.001$, and CDI-WS, $t(68) = -3.46$, $p = 0.001$.

3 | RESULTS

3.1 | Developmental changes and sex differences in selective attention to the articulating mouth

We ran a 2 × 2 × 2 linear mixed-effects model on the proportion of total looking time, with Time-point (5.5 vs. 11 months) and AOI (eyes vs. mouth) as within-subject factors and participant Sex (female vs. male) as a between-subjects factor (see full descriptive statistics in Table S1). There was a significant Time-point × AOI interaction, $F(1, 234.55) = 22.17$, $p < 0.001$, $r = 0.27$, indicating that changes between time-points were different for the two AOIs. The interaction between AOI and Sex was also significant, $F(1, 234.55) = 4.60$, $p = 0.03$, $r = 0.16$, indicating that females and males differed in the looking time to each area of interest. Finally, there were no significant interactions between Time-point and Sex, $F(1, 255.07) = 0.54$, $p = 0.46$ or a three-way interaction $F(1, 234.55) = 1.96$, $p = 0.16$.

To determine the source of the significant Time-point × AOI interaction (see Figure 2; and Figure S1 for individual trajectories), we ran two post-hoc comparisons (Bonferroni-corrected). First, we assessed differences in the proportion of total looking time to the eyes and the mouth within each timepoint. At 5.5 months, infants looked more to the eyes than the mouth, $F(1, 234.55) = 29.7$, $p < 0.001$. However, at 11 months, they looked equally long at either area, $F(1, 234.55) = 1.52$, $p = 0.21$. Secondly, we assessed differences between time-points in the proportion of looking time for each AOI. There was a significant increase in looking time to the mouth between 5.5 months ($M = 0.19$, $SD = 0.20$) and 11 months ($M = 0.41$, $SD = 0.28$), $F(1, 244.35) = 24.33$, $p < 0.001$, whereas the proportion of total looking time to the eyes between these time-points ($M = 0.44$, $SD = 0.23$ at 5.5 months; $M = 0.35$, $SD = 0.28$ at 11 months) remained constant, $F(1, 244.35) = 3.74$, $p = 0.05$.

To determine the source of the significant AOI × Sex interaction (see Figure 3), we ran two post-hoc comparisons (Bonferroni-corrected). First, we assessed differences between females and males in the proportion of looking time for each AOI. Females ($M = 0.38$, $SD = 0.27$) did not differ from

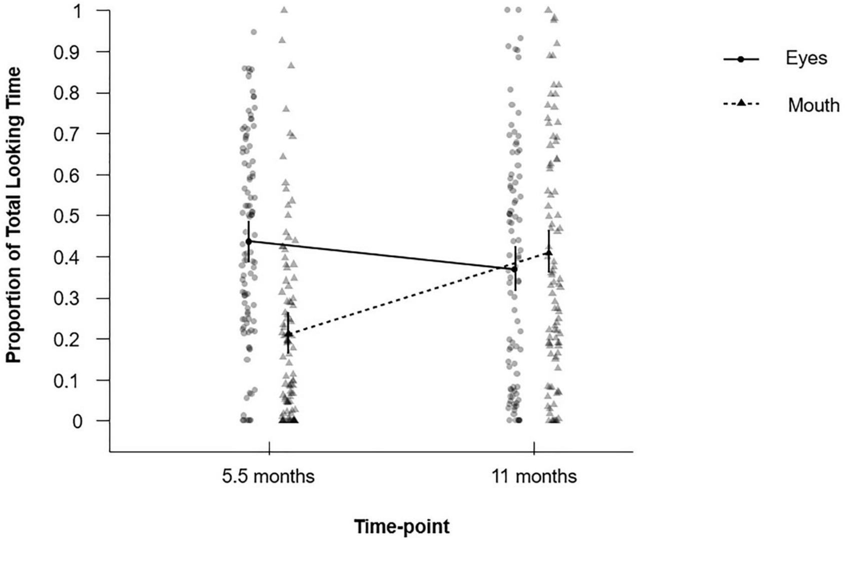

FIGURE 2 Mean proportion of total looking time by areas of interest (AOI) (eyes and mouth) as a function of time-point (5.5 and 11 months of age). Error bars represent one standard error of the mean. Black dots represent means, while gray triangles and dots represent individual data points

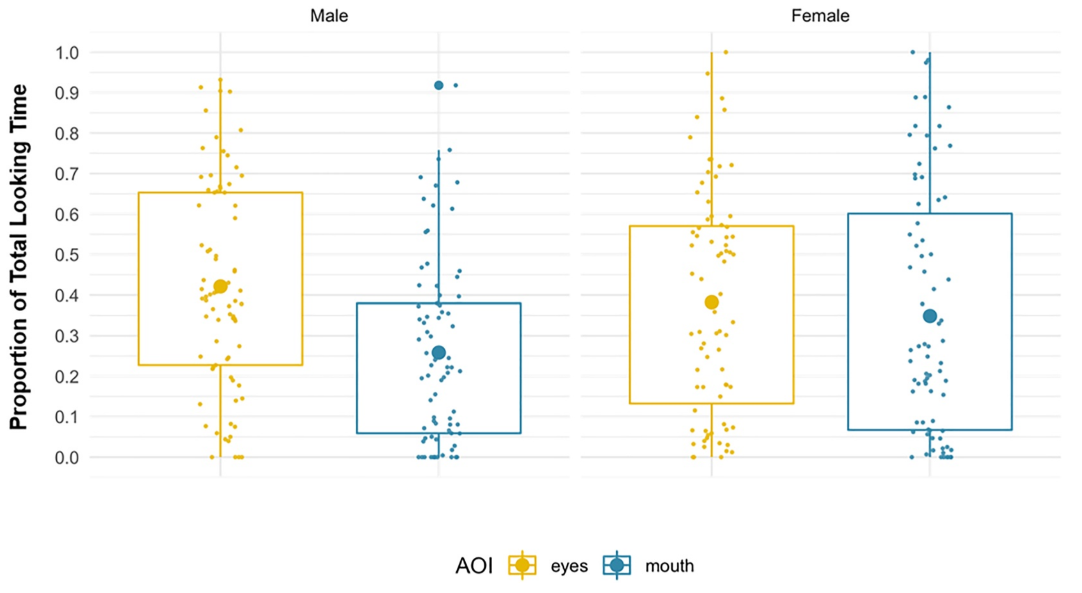

FIGURE 3 Boxplots showing the proportion of total looking time by areas of interest (AOI) (eyes and mouth) as a function of sex (female and male). Big dots represent means, while small ones represent individual data points

males ($M = 0.42$, $SD = 0.25$) in the proportion of total looking time to the eyes, $F(1, 292.36) = 1.13$, $p = 0.28$, but females looked more ($M = 0.35$, $SD = 0.30$) than males ($M = 0.26$, $SD = 0.22$) to the mouth, $F(1, 292.36) = 7.96$, $p = 0.005$. Secondly, we assessed differences in the proportion of total looking time to the eyes and the mouth within each Sex. Whereas in females, looking time to the eyes and to the mouth did not significantly differ from each other, $F(1, 234.55) = 0.32$, $p = 0.57$, males looked more to the eyes than to the mouth, $F(1, 234.55) = 13.18$, $p < 0.001$. See full descriptive statistics in Tables S2 and S3.

Specificity of sex differences. To rule out that the sex differences in PTLTs to the mouth found were explained by low-level visual scanning of AOIs, we conducted the former analysis on the number of fixations (see Supplementary Analysis 2). Results were consistent with those for PTLTs, except that males showed higher number of fixations to the eyes than the mouth, while females showed equal to both. Moreover, to explore if sex differences were not specific to mouth-looking but related to the overall attention to the face, we examined sex differences in the total absolute looking time to the face (Supplementary Analysis 3, Tables S10–S11). Our results suggest that sex differences were specific to mouth-looking. Finally, to eliminate sex differences in visual scanning as a possible confound, we confirmed that males and females did not differ in the total number of fixations to the face (Supplementary Analysis 4, Tables S8–S9).

3.2 | Associations between selective attention to the mouth and language outcomes

The relationship between selective attention to the mouth and language development was tested using correlational analyses (Table S12; Supplementary Analyses 5). Note that the sample size varied for each analysis depending on the validity of eye-tracking data and availability of language outcomes.

Additional correlation analyses between eyes-looking and language outcomes were run to test the specificity of these associations to the mouth AOI.

Mouth-looking and language outcomes in infancy. First, we ran correlations between infants' mouth-looking at each time-point (5.5 and 11 months) and receptive and expressive language scores at 11 months (MSEL and CDI-WG). No significant associations were found between looking at the mouth at 5.5 months and language outcomes at 11 months (all *ps* > 0.35). Likewise, there were no significant associations between mouth-looking at 11 months and any concurrent language outcomes (all *ps* > 0.19). Finally, the magnitude of the shift to the mouth from 5.5 to 11 months (defined as PTLT mouth at 11 months minus PTLT mouth at 5.5 months) was not related significantly to language scores at this age (all *p*s > 0.39).

Mouth-looking and language outcomes in toddlerhood. Next, we explored associations between the proportion of looking time at the mouth at 5.5 and 11 months and linguistic outcomes at 24 months. Neither mouth-looking at 5.5 months nor at 11 months predicted any productive linguistic skills (CDI-WS and MSEL Expressive Language scale; all *p*s > 0.10). No significant associations were found between mouth-looking at either time-point and receptive language skills at 24 months (OTSR and MSEL Receptive Language; all *ps* > 0.28). Unexpectedly, infants' preference for looking to the eyes at 5.5 months was associated with higher receptive, $r(47) = 0.37$, $p = 0.01$, and expressive language skills in toddlerhood, $r(42) = 0.35$, $p = 0.02$; Figure S3. Finally, the magnitude of change in mouth-looking from 5.5 to 11 months predicted language outcomes at 24 months. Infants with greater increase in looking time to the mouth had better expressive language skills in MSEL, $r(38) = 0.33$, $p = 0.04$. The magnitude of change was not associated with any other language measure (all *p*s > 0.60).

3.3 | Sex differences in the relation between selective attention to the mouth and language development

Prior analyses revealed sex differences in infants' looking to the mouth, along with sex differences in expressive language at 2 years (Table 3). To explore the longitudinal associations between sex differences in infants' selective attention to the mouth and its later effects on their language outcomes, we split our data set by sex and conducted the same correlation analyses, but separately for females and males (Table S12; Supplementary Analyses 6).

First, we explored associations between infants' attention to the mouth and language outcomes at 11 months. In line with previous analysis for the entire group, mouth-looking at 5.5 months was not associated with any linguistic outcome at 11 months for females or males. Likewise, there were no sex-specific correlations between mouth-looking at 11 months and language outcomes at this age.

Our second analysis revealed sex-specific correlations between looking at the mouth at 5.5 and linguistic outcomes at 24 months. Female infants' mouth-looking at 5.5 months was negatively associated with expressive, $r(22) = -0.53$ $p = 0.01$, and receptive language skills in MSEL, $r(24) = -0.43$, $p = 0.03$, but not for other receptive measures (CDI-WS and OTSR; all *ps* > 0.54). This indicates that those females who looked less at the mouth at 5.5 months showed better language comprehension and productive outcomes in toddlerhood (Figures S4 and S5; see also Supplementary Analysis 7 for control on the specificity of this result to the mouth AOI instead of the entire face). By contrast, for male infants we did not find any associations between mouth-looking at 5.5 months and outcomes at 24 months.

Third, attention to the mouth at 11 months did not predict any linguistic outcome at 24 months, for females or males. To clarify the sex specificity of these associations, we ran the same correlation analyses using a measure of the PTLT Difference Score (defined as PTLT to the eyes minus PTLT to the mouth). The results replicated those shown above (Supplementary Analyses 8), further supporting female-specific association of lower mouth-looking at 5.5 months with better language comprehension and production at 2 years.

Finally, we further explored whether there were sex differences in how the magnitude of change in mouth preference relates to language outcomes (Supplementary Analyses 9 and Figure S6). Female infants with lower gain in looking time to the mouth between 5.5 and 11 months had higher expressive language skills at 11 months, $r(27) = -0.43$, $p = 0.02$, while this relationship did not hold for male infants ($p = 0.40$).

4 | DISCUSSION

4.1 | Developmental increases of selective attention to the mouth during infancy

Our first goal was to longitudinally examine changes in selective attention to the mouth relative to the eyes of a talking face between 5.5 and 11 months of age. As expected, infants showed an increase in looking time to the mouth between these two timepoints, while, by contrast, looking time to the eyes did not change and remained relatively high. A general increase in looking time to the entire face between these two time-points was also observed. Furthermore, while 5.5-month-olds showed preferential looking to the eyes rather than to the mouth, 11-month-olds showed no preference for either AOI. Our results add several important findings to the literature on the development of selective attention to AV speech.

First, they help to better understand discrepant results from prior cross-sectional studies. While the increase in infants' preference for the speaker's mouth between the first and second half of the first year replicates results obtained by Mercure et al. (2019) and Tomalski et al. (2013) with syllables, the lack of increase in eyes-looking between 5.5 and 11 months is novel. Interestingly, our analysis provides a new comparison to the one usually run in prior cross-sectional studies (i.e., comparing the change across time-points of attention to the eyes and to the mouth separately, instead of only within-age preferences to each AOI; e.g., Lewkowicz & Hansen-Tift, 2012; Pons et al., 2015). It shows an increase in looking that seems mouth-specific, since there were no between time-point changes in preference to the eyes, instead, the eyes-looking remained high throughout this period.

The pattern we observed indicates that during the second half of the first year, infants increasingly attend to the mouth of a talking face but without reducing their looking to the eyes. This suggests that developmental patterns of selective attention to the eyes and mouth change independently. Together with our finding of increased attention to the face between 5.5 and 11 months, it is likely that infants gradually learn to distribute their attention within the speaker's face by first (at 5.5 months) attending to the eyes to seek socio-communicative cues (i.e., eye gaze) and then (at 11 months), by relying on mouth movements to maximize the extraction of linguistic articulatory information (Lewkowicz & Hansen-Tift, 2012). However, 11-month-olds also maintained interest in the eyes, which may indicate that they provide information crucial for both social interactions (e.g., gaze following or joint attention; Schietecatte et al., 2012) and language acquisition. Such a possibility is supported by the fact that increased preference for the mouth relative to the eyes at 6–12 months is associated with linguistic outcomes at this age (Tsang et al., 2018) and preference for the eyes at 12 months correlates with concurrent higher socio-communicative abilities (Pons et al., 2019). Altogether, this suggests that changes in selective attention to the eyes and mouth across the second half of the first year of life may occur independently but their functional role in development is closely interrelated, both supporting the development of social, communicative, and linguistic skills.

Second, our results on preference within each time-point replicate prior cross-sectional findings. We found a preference for the eyes relative to the mouth at 5.5 months (Lewkowicz & Hansen-Tift, 2012; Pons et al., 2015) and a balanced attention between the speaker's eyes and mouth at 11 months, replicating prior findings (Lewkowicz & Hansen-Tift, 2012; Pons et al., 2019; Sekiyama et al., 2021; but see Wilcox et al., 2013).

Although it has been argued that the developmental pattern of selective attention to the mouth might be language-specific (Berdasco-Muñoz, et al., 2019), our findings with Polish 11-month-olds are consistent with prior studies conducted with English (e.g., Lewkowicz & Hansen-Tift, 2012), Japanese (Sekiyama et al., 2021), French (Morin-Lessard et al., 2019), Spanish, and Catalan infants (e.g., Pons et al., 2015). Surprisingly, this points to how generalizable is the timing of changes in the pattern of facial scanning across rhythmically and phonologically different languages. Given the complexity of their mother tongue, one may speculate that Polish infants need to rely on the speaker's mouth for a protracted period of time relative to infants with other language backgrounds, thus either slowing the timing of onset of balanced attention between the talker's eyes and mouth or, alternatively, starting the shift to the mouth earlier. Yet, this is not consistent with the lack of clear preference for any of these AOIs we observed at 11 months. Future studies could further explore cross-linguistic differences in selective attention to the mouth throughout the first year.

Alternatively, the pattern of equal looking to the eyes and the mouth at 11 months could be explained by individual differences. The observed increased within-group variability in infants' looking preferences suggests that while some infants may be shifting to looking back to the eyes, others may still preferably rely on the mouth. Ultimately, this would explain an average lack of preference to the eyes or the mouth at a group level. This mixed pattern may reflect individual differences in the timing of AV perceptual narrowing in the first year of life (Pons et al., 2009; Ter Schure et al., 2016). While infants who are more advanced in phonological attunement to their native language would rely less on the mouth at 11 months, those who are still learning to map viseme–phoneme native categories would keep looking at the mouth until later on, in line with the so-called *'language expertise hypothesis'* (e.g., Lewkowicz & Hansen-Tift, 2012; Morin-Lessard et al., 2019).

4.2 | Eyes-looking and gains in mouth-looking in infancy are associated with language outcomes in toddlerhood

Our second goal was to investigate whether selective attention to the mouth in infancy is related to language outcomes. We predicted that greater preference for the mouth relative to the eyes at 11 months (but not at 5.5 months) would be positively be associated with language skills at both 11 and 24 months. Contrary to this prediction, when analyzing the entire sample, looking at the mouth at 5.5 and 11 months was not associated with any language outcomes. This is inconsistent with prior longitudinal studies, which showed positive associations of mouth-looking at 6 (Young et al., 2009) and 12 months (Tenenbaum et al., 2015) with expressive language at 2 years. One possible explanation for nonreplicating these results is the differences between the experimental tasks. Alternatively, slight differences in the ages examined may have led to not tracking the predictive role of mouth-looking precisely when this mechanism operates. However, we consider this latter possibility less plausible because, although mouth-looking at a certain age was not associated with language outcomes, infants with greater increase in attention to the mouth between 5.5 and 11 months had better expressive linguistic skills at 2 years. In other words, the gain in mouth-looking was associated with better productive language. We consider this a crucial result that holds several implications.

Firstly, it points to the second half of the first year of life as a crucial time window, in which gaining visual access to the mouth facilitates language acquisition, particularly word production. Accordingly, increased mouth-looking early into toddlerhood seems to be a marker of the onset of first-words production (Habayeb et al., 2020). Our results could reflect the onset of this increasing mouth-looking trajectory that continues at 14 and 18 months (Hillairet de Boisferon et al., 2018) until two years. What is insufficiently clear is how the increased mouth-looking facilitates expressive language acquisition. The trajectory of gains in mouth-looking may enhance the underlying processes (e.g., perceptual narrowing) and mechanisms (e.g., statistical learning). For example, visual information about speech articulation enhances phoneme discrimination at 6 months (Teinonen et al., 2008), which predicts later word-learning (Tsao et al., 2004). Secondly, this result points to a specificity in the association between mouth-looking and expressive language development only (Tenenbaum et al., 2015; Young et al., 2009; but see Imafuku & Myowa, 2016). Given the close and bidirectional relationship between infants' speech perception and production (DePaolis et al., 2011), this specificity might seem surprising. One potential explanation is that mouth-looking may also indirectly support receptive skills, but at a finer-grained level than the one we measured. Thirdly, the substantial time gap we observed between the gain in mouth-looking and its functional role for expressive language suggests the possibility of deferred effects. Infants who showed greater increase in mouth-looking during the second half of the first year ended up benefiting from it in toddlerhood but not earlier. Thus, the benefits could be long-term and constrained to a specific period in toddlerhood, as infants may need enough experience with increased looking at the mouth of a talking face before profiting from it. In support of this, looking at the mouth at 18 months did not concurrently predict expressive skills (Hillairet de Boisferon et al., 2018). Alternatively, the tools we used may not be sensitive enough to measure productive skills. Accordingly, one study measuring vocalizations during observation of AV speech found associations between 6-months-olds' looking at the speaker's mouth and vocal imitation of vowels (Imafuku et al., 2019). Future longitudinal studies should explore if gains in mouth-looking during the second half of the first year are associated with production of vocalizations.

Unexpectedly, we found that increased preference for the eyes (instead of the mouth) at 5.5 months positively correlated with receptive and expressive language in toddlerhood. To our knowledge, this is the first study finding such a relationship in infancy (but see Sekiyama et al., 2021 and Viktorsson et al., 2021 for partial replication in early toddlerhood). The role of eyes-looking in the context of AV speech processing has been usually underestimated, and, when explored, it has been linked to other socio-communicative skills rather than language acquisition (Pons et al., 2019). Our results suggest that attending to the eyes does have a functional role in infants' language acquisition from a very early age. Eye-gaze may act as an attention-grabbing cue that draws attention to social partners and their speech signals (Çetinçelik et al., 2021). Since the impact of increased preference for the eyes in language development was not observable in our data until after a protracted period of time, we propose that its role in engaging attention to faces as communicative targets would be key not only at 5.5 months but also later, to establish the basis for infants' long-term language learning. This fits with our result that eyes-looking remained high in the second half of the first year. It seems puzzling, though, that we did not find associations between preference for the eyes at 11 months and language outcomes in toddlerhood (neither concurrently). High variability in preference for the eyes at 11 months may have obscured such relations. Alternatively, the contribution of selective attention to the eyes to language learning may vary throughout development. While early on (5.5. months) it could facilitate engagement with the speaker, later on it may reflect a possibility of freedom from attending to the mouth due to sufficient language expertise (see Sekiyama et al., 2021). Future studies should explore if increased eyes-looking during the second half of infancy indirectly supports language development by freeing attention to follow other relevant social cues (e.g., gaze following).

A final observation on our results taken together is that whereas gains in mouth-looking from 5.5 to 11 months were associated with expressive skills only in toddlerhood (as in Tsang et al., 2018), the increased preference for the eyes at 5.5 months was associated with both receptive and expressive language (which, to our knowledge, has not been previously reported). This points to the need for separating the functional role of selective attention to the eyes and mouth for these two subcomponents of language development.

4.3 | A unique mouth-looking trajectory in infancy is associated with language outcomes only in female toddlers

Our third aim was to explore sex differences in the development of selective attention to the mouth as well as in its supportive role for language acquisition. Our results suggest a developmental trajectory of mouth-looking that is likely exclusive to female infants. Sex differences were found in the proportion of total looking time to the mouth, with females showing higher selective attention to the mouth than males at both 5.5 and 11 months. Crucially, sex differences were exclusive to selective attention to the mouth, as supported by follow-up analyses showing that females did not differ from males in the proportion of total looking time to the eyes, but looked more than males to the mouth. Notably, our control analyses indicated that these results cannot be explained by sex differences in face-looking or in low-level measures of scanning (number of fixations).

Consistently, we also found sex-specific longitudinal associations between selective attention to the mouth in infancy and language outcomes in toddlerhood that were statistically significant only for female infants. Females with lower selective attention to the mouth at 5.5 months had better expressive and receptive linguistic skills in toddlerhood. In contrast, no such negative associations were found for males. In light of our prior result for the entire sample showing that gains in mouth preference between 5.5 and 11 months were associated with better language outcomes in toddlerhood, the directionality of these associations is surprising. These findings suggest that at 5.5 months, females show a trade-off between attention to the eyes and the mouth (i.e., increased preference to the eyes at the cost of reduced preference to the mouth) that benefits future language acquisition. In contrast, males would only benefit from an early increased selective attention to the eyes, as indicated by prior analyses with participant sex collapsed.

The early trade-off in females may reflect simultaneous scanning of both the eyes and mouth, since learning language requires engaging with the speaker's faces (mainly attending to their eyes), attending to their speech (by prioritizing mouth-looking), and following external events they refer to (usually cued by the eyes). By contrast, the lack of such a pattern in males may indicate either a later onset of this same trajectory that we failed to identify at a later time-point or an altogether different trajectory (more ‘eyes-looking-based’; as shown in Figure S2). Longitudinal studies with more dense follow-ups and larger samples are needed to explore these possibilities. However, the idea of a male-differential trajectory is further supported by findings on the number of fixations (Supplementary Analysis 2). Although both sexes increased their number of fixations on the mouth (but not the eyes) across infancy, males showed higher number of fixations on the eyes than the mouth, while females showed equal number of fixations on both. This reinforces the view that females might be more skilled from early on in simultaneously monitoring linguistic cues from both the eyes and mouth, which, in the long-term, benefits their scanning efficiency for AV speech (see Ross et al., 2015).

Another puzzling result was a lack of association between preference for the eyes at 5.5 months and/or gains in mouth-looking and language outcomes in toddlerhood when splitting our sample by sex. These two contrastive findings may indicate that female and male infants could share the functional role of selective attention to the eyes in language development, but not the supportive role that selective attention to the mouth provides. Rather, at least the initial role of mouth-looking in infancy for language development would be female-specific, while gains in mouth-looking during the second half of infancy would be shared by both sexes. In support, we found that females, but not males, with a lower magnitude of change in mouth-looking had better expressive skills at 11 months (Supplementary Analysis 9). This suggests that females with an early advantage in mouth-looking (and whose magnitude of increase is smaller) are ahead in this trajectory, thus, showing better expressive skills toward the end of infancy.

Ultimately, the unique developmental trajectory of female infants appeared to be more efficient or beneficial for language development in the long term but not at the end of the first year. We found sex differences in productive language at 2 years of age (but not earlier), with females outperforming males. This is consistent with robust evidence of a female advantage in language learning across development, especially in the expressive domain: from the early sex-hormone mediated effects on babbling observed in infancy (Quast et al., 2016) to a higher vocabulary size in toddlerhood (Lutchmaya et al., 2001). The lack of sex differences in receptive language at the same age may indicate that the mechanism of selective attention to the eyes shared by both females and males at 5.5 months gives a long-term advantage to both sexes in this subcomponent of language.

Taken together, our results point to the trajectory of selective attention to the mouth and eyes (but especially the mouth) being a very early mechanism of language acquisition that mediates sex differences and, more specifically, the female advantage in language development (Adani & Cepanec, 2019). We are aware that these results were motivated by exploratory analyses and, thus, need further replication to be considered firm. Future preregistered research should continue exploring these associations under more specific a priori directional hypotheses. The potential implications of our results extend to atypical development, since in several neurodevelopmental disorders, the risk of presenting language atypicalities is higher in males than females (e.g., ASD, Howe et al., 2015; Specific Language Impairment, Whitehouse, 2010). We propose that the trajectory of female infants we observed may act as a protective factor from early on. Conversely, deviating from this trajectory may act as a risk-marker in infants for the development of language difficulties at a later stage. Thus, we consider it of particular interest to explore potential sex differences in this trajectory in developmental disorders.

4.4 | Limitations

We note some limitations of our study. First, in some fragments of the clips presented, actresses showed their hands. Although we excluded them from the analyses, we acknowledge this issue as a limitation. Second, the age range of our participants was relatively high at all time-points assessed, which may have increased between-infants' variability in the measures of preferences. Third, sample sizes in correlations split by sex were small. Thus, the result that females who showed reduced mouth-looking in infancy had better language outcomes in toddlerhood might be considered tentative. We suggest that this effect needs further exploration with better-powered samples that test its generalizability. However, we observed it only in females despite the sample size being modest in both sexes, and crucially, it did not seem to be driven by a few outliers (Figure S5), so we consider it unlikely that it was due to under-powered analyses. Finally, the lack of consistency between results in the MSEL expressive subscale and CDI inventory in toddlerhood is inconsistent with prior findings from English-speaking samples (e.g., Tenenbaum et al., 2015). One possibility is that, due to data loss in some CDI-WS questionnaires, the sample size for this instrument was significantly lower than for MSEL, thus reducing available statistical power.

5 | CONCLUSIONS

Our findings suggest that infants follow a trajectory of gradual increase of selective attention to the mouth while not at the expense of exploring the eyes, which may indicate that changes in selective attention to the eyes and mouth occur independently in development and support distinct subcomponents of language development. Whereas gains in mouth-looking would support expressive language, increased eyes-looking would boost both receptive and expressive skills. The effect of both mechanisms on language development may be constrained to toddlerhood. This finding may help to target vulnerabilities in mechanisms underlying language development before atypicalities in language outcomes become observable. A crucial contribution of our study is that we identified sex differences in attention to the mouth, which likely drives the early female advantage in expressive language. Only the longitudinal effects of selective attention to the mouth on language development in toddlerhood were modulated by sex, while those of early increased attention to the eyes were not sex-specific. Therefore, we may have targeted an early female protective factor in language acquisition.

ACKNOWLEDGMENTS

We would like to thank all infants and families that participated in this study for their generous contribution. We thank Mateusz Płatos and Maciej Rudziński for assistance with stimulus preparation and scripting and Magdalena Ratkowska, Sonia Ramotowska, Katarzyna Malec, Agnieszka Wyrzykowska, Marta Młodożeniec, Alicja Niedźwiecka and Karolina Marczuk for assistance with testing. The data collection was funded by a Polish National Science Center grant (2011/03/D/HS6/05655, PT) with additional support by the People Program (Marie Curie Actions) of the EU FP7 (grant no. PCIG10-GA-2011-304255, PT). The data analysis was further supported by the Polish National Science Center (2016/23/B/HS6/03860; 2018/30/E/HS6/00214, PT). IL conducted main statistical analyses, contributed to interpreting the results, wrote the first draft and edited subsequent versions of the manuscript. DLP contributed to the pre-processing of eye-tracking data and commented on earlier versions of the manuscript. ZL contributed to the interpretation of the results and editing the manuscript. MS conducted statistical analyses and contributed to the coordination of the study. AR conducted statistical analyses and contributed to the coordination of the study. AMK managed and conducted the data collection. PT conceived and designed the study, oversaw its execution, secured the funding, conducted main statistical analyses, contributed to the coordination of the study, the interpretation of the results and editing subsequent versions of the manuscript.The authors declare no conflicts of interest with regard to the funding source for this study.

ORCID

Itziar Lozano https://orcid.org/0000-0003-4280-4747
David López Pérez https://orcid.org/0000-0002-1235-6376
Zuzanna Laudańska https://orcid.org/0000-0001-6790-9559
Anna Malinowska-Korczak https://orcid.org/0000-0003-2788-4865
Magdalena Szmytke https://orcid.org/0000-0003-4714-0601
Alicja Radkowska https://orcid.org/0000-0002-0909-3368
Przemysław Tomalski https://orcid.org/0000-0002-0390-5759

REFERENCES

Adani, S., & Cepanec, M. (2019). Sex differences in early communication development: Behavioral and neurobiological indicators of more vulnerable communication system development in boys. *Croatian Medical Journal*, *60*(2), 141–149. https://doi.org/10.3325/cmj.2019.60.141

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv. http://arxiv.org/abs/1506.04967

Berdasco-Muñoz, E., Nazzi, T., & Yeung, H. H. (2019). Visual scanning of a talking face in preterm and full-term infants. *Developmental Psychology*, *55*(7), 1353–1361. https://doi.org/10.1037/dev0000737

Birulés, J., Bosch, L., Brieke, R., Pons, F., & Lewkowicz, D. J. (2019). Inside bilingualism: Language background modulates selective attention to a talker's mouth. *Developmental Science*, *22*(3), e12755. https://doi.org/10.1111/desc.12755

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Çetinçelik, M., Rowland, C. F., & Snijders, T. M. (2021). Do the eyes have it? A systematic review on the role of eye gaze in infant language development. *Frontiers in Psychology*, *11*, 3627. https://doi.org/10.3389/fpsyg.2020.589096

Chandrasekaran, C., Trubanova, A., Stillittano, 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

Danielson, D. K., Bruderer, A. G., Kandhadai, P., Vatikiotis-Bateson, E., & Werker, J. F. (2017). The organization and reorganization of audiovisual speech perception in the first year of life. *Cognitive Development*, *42*, 37–48. https://doi.org/10.1016/j.cogdev.2017.02.004

de Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal attention as a scaffold for language development. *Frontiers in Psychology*, *7*, 44. https://doi.org/10.3389/fpsyg.2016.00044

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

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology*, *47*(6), 1499–1503. https://doi.org/10.1037/a0025659

DePaolis, R. A., Vihman, M. M., & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior and Development*, *34*(4), 590–601. https://doi.org/10.1016/j.infbeh.2011.06.005

D'Souza, D., D'Souza, H., Jones, E. J. H., & Karmiloff-Smith, A. (2020). Attentional abilities constrain language development: A cross-syndrome infant/toddler study. *Developmental Science*, *23*(6), e12961. https://doi.org/10.1111/desc.12961

Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., & Johnson, M. H., & BASIS team. (2014). What you see is what you get: Contextual modulation of face scanning in typical and atypical development. *Social Cognitive and Affective Neuroscience*, *9*(4), 538–543. https://doi.org/10.1093/scan/nst012

Etchell, A., Adhikari, A., Weinberg, L. S., Choo, A. L., Garnett, E. O., Chow, H. M., & Chang, S. E. (2018). A systematic literature review of sex differences in childhood language and brain development. *Neuropsychologia*, *114*, 19–31. https://doi.org/10.1016/j.neuropsychologia.2018.04.011

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *MacArthur–Bates communicative development inventories*. Brookes.

Field, A. (2018). Multilevel linear models. In A. Field (Ed.), *Discovering statistics using IBM SPSS statistics* (pp. 936–990). SAGE.

Habayeb, S., Tsang, T., Saulnier, C., Klaiman, C., Jones, W., Klin, A., & Edwards, L. A. (2020). Visual traces of language acquisition in toddlers with autism spectrum disorder during the second year of life. *Journal of Autism and Developmental Disorders*, *51*(7), 2519–2530. https://doi.org/10.1007/s10803-020-04730-x

Haman, E., Fronczyk, K., & Łuniewska, M. (2012). Obrazkowy test słownikowy – Rozumienie (OTSR). Pracownia Testów Psychologicznych i Pedagogicznych.

Hessels, R. S., Benjamins, J. S., Cornelissen, T. H. W., & Hooge, I. T. C. (2018). A validation of automatically-generated Areas-of-Interest in videos of a face for eye-tracking research. *Frontiers in Psychology*, *9*, 1367. https://doi.org/10.3389/fpsyg.2018.01367

Hillairet 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*, 189–200. https://doi.org/10.1016/j.jecp.2018.03.009

Howe, Y. J., O'Rourke, J. A., Yatchmink, Y., Viscidi, E. W., Jones, R. N., & Morrow, E. M. (2015). Female autism phenotypes investigated at different levels of language and developmental abilities. *Journal of Autism and Developmental Disorders*, *45*(11), 3537–3549. https://doi.org/10.1007/s10803-015-2501-y

Imafuku, M., Kanakogi, Y., Butler, D., & Myowa, M. (2019). Demystifying infant vocal imitation: The roles of mouth looking and speaker's gaze. *Developmental Science*, *22*(6), e12825. https://doi.org/10.1111/desc.12825

Imafuku, M., & Myowa, M. (2016). Developmental change in sensitivity to audiovisual speech congruency and its relation to language in infants. *Psychologia*, *59*(4), 163–172. https://doi.org/10.2117/psysoc.2016.163

Irwin, J., & DiBlasi, L. (2017). Audiovisual speech perception: A new approach and implications for clinical populations. *Language and Linguistics Compass*, *11*(3), 77–91. https://doi.org/10.1111/lnc3.12237

Kleberg, J. L., Nyström, P., Bölte, S., & Falck-Ytter, T. (2019). Sex differences in social attention in infants at risk for autism. *Journal of Autism and Developmental Disorders*, *49*(4), 1342–1351. https://doi.org/10.1007/s10803-018-3799-z

Krajewski, G., & Smoczyńska, M. (2015). *Krótki Inwentarz Rozwoju Mowy i Komunikacji (KIRMIK) (dla dzieci w wieku 24–36 miesięcy)*. Instytut Badań Edukacyjnych.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843. https://doi.org/10.1038/nrn1533

Kushnerenko, E., Tomalski, P., Ballieux, H., Ribeiro, H., Potton, A., Axelsson, E. L., Murphy, E., & Moore, D. G. (2013). Brain responses to audiovisual speech mismatch in infants are associated with individual differences in looking behaviour. *European Journal of Neuroscience*, *38*(9), 3363–3369. https://doi.org/10.1111/ejn.12317

Kushnerenko, E. V., Tomalski, P., Ballieux, H., Potton, A., Birtles, D., Frostick, C., & Moore, D. G. (2013). Brain responses and looking behavior during audiovisual speech integration in infants predict auditory speech comprehension in the second year of life. *Frontiers in Psychology*, *4*, 432. https://doi.org/10.3389/fpsyg.2013.00432

Lalonde, K., & Werner, L. A. (2021). Development of the mechanisms underlying audiovisual speech perception benefit. *Brain Sciences*, *11*(1), 49. https://doi.org/10.3390/brainsci11010049

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*, *109*(5), 1431–1436. https://doi.org/10.1073/pnas.1114783109

Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2001). Foetal testosterone and vocabulary size in 18-and 24-month-old infants. *Infant Behavior and Development*, *24*(4), 418–424. https://doi.org/10.1016/S0163-6383(02)00087-5

Marjanovič-Umek, L., & Fekonja-Peklaj, U. (2017). Gender differences in children's language: A meta-analysis of slovenian studies. *Center for Educational Policy Studies Journal*, *7*(2), 97–111. https://doi.org/10.25656/01:14598

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mercure, E., Kushnerenko, E., Goldberg, L., Bowden-Howl, H., Coulson, K., Johnson, M. H., & MacSweeney, M. (2019). Language experience influences audiovisual speech integration in unimodal and bimodal bilingual infants. *Developmental Science*, *22*(1), e12701. https://doi.org/10.1111/desc.12701

Mills, A. E. (1987). The development of phonology in the blind child. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 145–162). Erlbaum.

Morin-Lessard, E., Poulin-Dubois, D., Segalowitz, N., & Byers-Heinlein, K. (2019). Selective attention to the mouth of talking faces in monolinguals and bilinguals aged 5 months to 5 years. *Developmental Psychology*, *55*(8), 1640–1655. https://doi.org/10.1037/dev0000750

Mullen, E. M. (1995). *Mullen scales of early learning*. American Guidance Service.

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*, 80–84. https://doi.org/10.1016/j.infbeh.2018.12.003

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences*, *106*(26), 10598–10602. https://doi.org/10.1073/pnas.0904134106

Quast, A., Hesse, V., Hain, J., Wermke, P., & Wermke, K. (2016). Baby babbling at five months linked to sex hormone levels in early infancy. *Infant Behavior and Development*, *44*, 1–10. https://doi.org/10.1016/j.yhbeh.2018.03.008

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2009). A cognitive approach to the development of early language. *Child Development*, *80*(1), 134–150. https://doi.org/10.1111/j.1467-8624.2008.01250.x

Ross, L. A., Del Bene, V. A., Molholm, S., Frey, H. P., & Foxe, J. J. (2015). Sex differences in multisensory speech processing in both typically developing children and those on the autism spectrum. *Frontiers in Neuroscience*, *9*, 185. https://doi.org/10.3389/fnins.2015.00185

Schietecatte, I., Roeyers, H., & Warreyn, P. (2012). Can infants' orientation to social stimuli predict later joint attention skills? *British Journal of Developmental Psychology*, *30*(2), 267–282. https://doi.org/10.1111/j.2044-835X.2011.02039.x

Segal, S. C., & Moulson, M. C. (2020). Dynamic advances in emotion processing: Differential attention towards the critical features of dynamic emotional expressions in 7-month-old infants. *Brain Sciences*, *10*(9), 585. https://doi.org/10.3390/brainsci10090585

Sekiyama, K., Hisanaga, S., & Mugitani, R. (2021). Selective attention to the mouth of a talker in Japanese-learning infants and toddlers: Its relationship with vocabulary and compensation for noise. *Cortex*, *140*, 145–156. https://doi.org/10.1016/j.cortex.2021.03.023

Smoczyńska, M., Krajewski, G., Łuniewska, M., Haman, E., Bulkowski, K., & Kochańska, M. (2015). *Inwentarze rozwoju mowy i komunikacji (IRMIK) słowa i gesty, słowa i zdania: Podręcznik*. Instytut Badań Edukacyjnych.

Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*, *108*(3), 850–855. https://doi.org/10.1016/j.cognition.2008.05.009

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. 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

Ter Schure, S., Junge, C., & Boersma, P. (2016). Discriminating non-native vowels on the basis of multimodal, auditory or visual information: Effects on infants' looking patterns and discrimination. *Frontiers in Psychology*, *7*, 525. https://doi.org/10.3389/fpsyg.2016.00525

Tomalski, P., Ribeiro, H., Ballieux, H., Axelsson, E. L., Murphy, E., Moore, D. G., & Kushnerenko, E. (2013). Exploring early developmental changes in face scanning patterns during the perception of audiovisual mismatch of speech cues. *European Journal of Developmental Psychology*, *10*(5), 611–624. https://doi.org/10.1080/17405629.2012.728076

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

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*(4), 1067–1084. https://doi.org/10.1111/j.1467-8624.2004.00726.x

Viktorsson, C., Portugal, A. M., Li, D., Rudling, M., Sanchez, M. S., Tammimies, K., Taylor, M. J., Ronald, A., & Falck-Ytter, T. (2021). Infants' attention to eyes is an independent, heritable trait predicting later verbal competence. *bioRxiv*. https://doi.org/10.1101/2021.10.23.465424

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language*, *108*(3), 175–183. https://doi.org/10.1016/j.bandl.2008.07.001

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, *66*(1), 173–196. https://doi.org/10.1146/annurev-psych-010814-015104

Whitehouse, A. J. O., Mattes, E., Maybery, M. T., Sawyer, M. G., Jacoby, P., Keelan, J. A., & Hickey, M. (2010). Is there a sex ratio difference in the familial aggregation of specific language impairment? A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *53*(7), 1015–1025. https://doi.org/10.1111/j.1469-7610.2011.02523.x

Wilcox, T., Stubbs, J. A., Wheeler, L., & Alexander, G. M. (2013). Infants' scanning of dynamic faces during the first year. *Infant Behavior and Development*, *36*(4), 513–516. https://doi.org/10.1016/j.infbeh.2013.05.001

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

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lozano, I., López Pérez, D., Laudańska, Z., Malinowska-Korczak, A., Szmytke, M., Radkowska, A., & Tomalski, P. (2022). Changes in selective attention to articulating mouth across infancy: Sex differences and associations with language outcomes. *Infancy*, *27*(6), 1132–1153. https://doi.org/10.1111/infa.12496