



OPEN ACCESS

EDITED BY

Huei-Mei Liu,
National Taiwan Normal University,
Taiwan

REVIEWED BY

Carolyn Palmquist,
Amherst College, United States
Peter Gerhardstein,
Binghamton University, United States
Teresa Mitchell,
Brandeis University, United States

*CORRESPONDENCE

Zsafia Belteki
z.a.belteki@uu.nl

SPECIALTY SECTION

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 18 July 2022

ACCEPTED 03 October 2022

PUBLISHED 28 October 2022

CITATION

Belteki Z, van den Boomen C and
Junge C (2022) Face-to-face contact
during infancy: How the development
of gaze to faces feeds into infants'
vocabulary outcomes.
Front. Psychol. 13:997186.
doi: 10.3389/fpsyg.2022.997186

COPYRIGHT

© 2022 Belteki, van den Boomen and
Junge. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Face-to-face contact during infancy: How the development of gaze to faces feeds into infants' vocabulary outcomes

Zsafia Belteki*, Carlijn van den Boomen and Caroline Junge

Department of Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, Netherlands

Infants acquire their first words through interactions with social partners. In the first year of life, infants receive a high frequency of visual and auditory input from faces, making faces a potential strong social cue in facilitating word-to-world mappings. In this position paper, we review how and when infant gaze to faces is likely to support their subsequent vocabulary outcomes. We assess the relevance of infant gaze to faces selectively, in three domains: infant gaze to different features within a face (that is, eyes and mouth); then to faces (compared to objects); and finally to more socially relevant types of faces. We argue that infant gaze to faces could scaffold vocabulary construction, but its relevance may be impacted by the developmental level of the infant and the type of task with which they are presented. Gaze to faces proves relevant to vocabulary, as gazes to eyes could inform about the communicative nature of the situation or about the labeled object, while gazes to the mouth could improve word processing, all of which are key cues to highlighting word-to-world pairings. We also discover gaps in the literature regarding how infants' gazes to faces (versus objects) or to different types of faces relate to vocabulary outcomes. An important direction for future research will be to fill these gaps to better understand the social factors that influence infant vocabulary outcomes.

KEYWORDS

faces, social development, infancy, gaze, vocabulary

Introduction

Early word learning appears a difficult task, given that an infant's environment is full of objects and sounds (Quine, 1960; Smith and Yu, 2008). Infants acquire words from the people they interact with, who repeatedly expose them to certain word-to-world combinations. The first signs of infants' word comprehension are observable as early as 6 months (Bergelson and Swingley, 2012; Tincoff and Jusczyk, 2012). This precedes the earliest observations of more complex social abilities that are first observed 9 months of age onward, such as joint attention, and which are considered instrumental for word learning (Carpenter et al., 1998; Cleveland et al., 2007). Previous literature has frequently

shown that where infants attend in such social settings can have cascading effects on how their word learning progresses (Moore et al., 1999; Charman et al., 2000; Mundy et al., 2003; Slaughter and McConnell, 2003; Houston-Price et al., 2006). One such powerful cue could be how infants attend to other faces present in the interaction (from now on “gaze to faces”; Hessels, 2019), as infants start attending to faces and to features within faces very early on in life (Grossmann, 2017). The purpose of this position paper is to review studies demonstrating a link between infant gaze to faces and vocabulary outcomes. We will assess the literature on gaze to faces, because it is considered one of the pre-requisites of later developing more complex social abilities [such as following the gaze of a social partner (Reid and Striano, 2005; Gredebäck et al., 2010; Grossmann, 2017)] that are found to correlate with word learning and vocabulary outcomes. Our aims in this position paper are to find new directions for future research to improve our understanding on whether, when, and how gaze to faces feeds into vocabulary acquisition.

While the relevance of being able to attend to social partners has been acknowledged for all facets of early language development, it is likely that gaze to the faces and facial features of social partners helps to bootstrap one aspect of language acquisition in particular: vocabulary acquisition (Tomasello, 1992, 2000; Çetinçelik et al., 2021). Acquiring a vocabulary requires the mapping of auditory spoken words to their matching concepts, both of which may require some level of disambiguation. The words that infants tend to acquire early are concrete objects that are visually present in the interaction, such as “bottle” and “sock” (Kavanaugh and Jirkovsky, 1982; Bergelson and Swingley, 2012; Braginsky et al., 2019). Infants can benefit from attending to the face of their social partner, as it can provide both visual (e.g., their eyes gazing at the named object) and auditory cues (e.g., mouth speaking the word form) to guide word learning. We theorize that infant gaze to faces is important to vocabulary acquisition, not only as it is a precursor to infants’ ability to utilize the cues that faces signal to word learning, but also as it scaffolds the development of those more complex social abilities instrumental in word learning (e.g., gaze following, joint attention, Reid and Striano, 2005). These more complex social abilities have been shown to relate to vocabulary outcomes; however, the relation between the precursor of these complex social abilities and vocabulary outcomes has been contemplated less (Carpenter et al., 1998; Morales et al., 2000; Slaughter and McConnell, 2003). Therefore, we examine the relevance of infant gaze to faces by zooming in on vocabulary as our chosen outcome measure. Note that vocabulary outcomes can refer to both productive and receptive vocabulary. Whenever there are mixed findings in the literature discussed below, we highlight the nature of the vocabulary outcomes (i.e., whether expressive, receptive, or both vocabulary outcomes are affected), as this distinction may help to explain mixed results.

The relevance of gaze to faces to vocabulary is not yet fully understood. One complexity stems from the gaze to faces being

defined in a number of ways. For instance, we can assess it by zooming in on the relevance of specific elements of a face, or we can compare it against other objects present within the same scene, or by contrasting different kinds of faces across different interactions. Moreover, it remains unclear whether this relationship proves stable across development or whether there are any developmental trends. It could be that the relevance of gaze to faces wanes off across development, but holds especially for the youngest age group just starting to build a vocabulary but who cannot yet rely on higher social-cognitive abilities such as joint attention. It is also unclear whether the choice of experimental tasks impacts the relationship. That is, it is possible that how infants utilize and weigh cues in a face may depend on the situation at hand and their developmental stage, making it difficult to understand exactly the relationship between infant gaze to faces and vocabulary outcomes. For instance, in cases where it may be relatively difficult to hear the words correctly, infants may look more at the mouth, whereas in cases where there are multiple novel objects present, infants may benefit more from gazing at the eyes. Hence, the relevance that looking to the mouth will have to infants’ word learning and subsequent vocabulary outcomes will depend on what needs to be disambiguated in the task, and whether the infant is at a developmental level where they can successfully disambiguate this information. Considering this, the developmental stage of the infant and the task they are presented with are important to consider in our review, as they may moderate which aspects of gaze to face are the most related to vocabulary outcomes.

This paper is organized as follows. We will start by presenting an overview of the existing theoretical frameworks pointing to gaze to faces as a possible facilitator to vocabulary outcomes in infancy. Then, we will review relevant literature that assesses how gaze to faces impact vocabulary outcomes, for three specific aspects of gaze to faces: first, to specific elements in the face (the eyes compared to the mouth); then, to faces versus other elements in the environment; finally, to different types of faces. Regarding vocabulary outcomes, we will assess papers look at both productive vocabulary outcomes and receptive vocabulary outcomes. Within each section, We will assess the relevance of this specific type of gaze to faces to vocabulary outcomes and describe how this relation is moderated by developmental trends and experimental tasks. In our discussion, we will evaluate the evidence for and against a relation between gaze to faces and vocabulary outcomes, based on the literature reviewed. We will then conclude by addressing the questions that are left unanswered by the existing literature, followed by recommendations for future research directions.

We therefore aim not only to collate literature showing whether, when, and why infant gaze to other persons’ faces could scaffold infants’ vocabulary outcomes (Locke, 1993) but also to identify gaps in the literature. Our overarching goal will be to complement and bridge existing theoretical frameworks that link social factors to early word learning (Tomasello, 2000;

Werker and Curtin, 2005; Kuhl, 2007; Çetinçelik et al., 2021; Bastianello et al., 2022).

Theoretical frameworks on infancy: Why would gaze to faces facilitate vocabulary?

Research has evidenced the crucial role that the social environment has in language acquisition in the first years of life (for a review, see e.g., Kuhl, 2007). Similarly, theoretical frameworks have attempted to outline the feedforward effects of the social environment on language acquisition (Pinker, 1979; Hollich et al., 2000; Tomasello, 2000; Werker and Curtin, 2005; Chater and Manning, 2006; Kuhl, 2007; Frank et al., 2009a; Gogate and Hollich, 2010; Johnson, 2016). This section aims to outline the theoretical frameworks that link the dynamics of the social environment to vocabulary acquisition to provide the motivation for assessing the feedforward effect that infant gaze to faces may have on their vocabulary outcomes.

The social environment is composed of many sub-parts of parental warmth and input (Madigan et al., 2019; Anderson et al., 2021). To facilitate word learning, it is beneficial to infants to detect the communicative intent from the partner to the infant, as it highlights potential situations where infants can acquire new words (Floor and Akhtar, 2006). One important source for detecting communicative intent is provided by the faces of social partners. Cues, such as direct gaze (but not averted gaze), can indicate to the infant that their social partner wishes to engage with them (Senju and Csibra, 2008). Indeed, one account on infant language development, the Social-Pragmatic account of language acquisition, holds that infants' ability to recognize communicative intent is crucial to language learning and more specifically their learning of words (Tomasello, 1992, 2000). This theory considers the detection of communicative intent critical, as it provides the foundation of more complex social behaviors (such as joint attention) that are required for language learning. Hence, since the Social-Pragmatic account holds that recognizing communicative intent mediates word learning, it follows that infants' gazes to faces as the more socially informative parts of the environment could be related to their vocabulary outcomes.

Besides detecting the cues that signal communicative intent, infants need to be able to combine and comprehend information that is presented multi-modally in order to map meaning to words. That is, infants need to map the auditory (i.e., words) information to the visual (i.e., concrete objects present in the scene) information. As is highlighted in the Intersensory Redundancy hypothesis, learning is facilitated whenever there is temporal synchrony between two modalities, because the same amodal information is highlighted above other sources of non-synchronous amodal information (Bahrick and Lickliter, 2012). While this account is not specific to word learning, it predicts

that multi-modal information, as is present in dynamic faces, would facilitate word learning by linking information from two modalities through their temporal synchrony. That is, faces provide information from two modalities at the same time: The auditory information that we hear and the visual information from the movements of the eyes and mouth. Therefore, an early face preference is arguably necessary to selectively direct the attention of the infant. This then allows them to gain extensive experience with faces and thereby successfully learn to process multimodal information presented by their environment. For example, viewing a moving mouth and hearing speech involves information from the audio and visual modalities that can be linked based on their temporal synchrony to create a combined auditory and visual signal, i.e., a speaking mouth producing sounds, that benefits word recognition (Hollich et al., 2005). Infants may use this synchronous information not only to improve word processing (the mouth-word relationship) but also to facilitate the word-to-world mappings (using visual cues such as the speaker's gaze to a named object, the eyes-word relationship) (Gogate and Hollich, 2010; Bahrick and Lickliter, 2014). Therefore, how well infants can connect the auditory to visually synchronous information that they receive from faces may relate to their vocabulary outcomes.

Above, we discussed two theories that emphasized that gaze to faces could provide useful cues to guide infants' word learning. The theories posited that gaze to faces could improve awareness of communicative intent, boost auditory word processing, and guide word-object pairings, all of which are instrumental for the learning of words. In both the Social Pragmatic Account and the Intersensory Redundancy hypothesis, it was argued that infants need to be able to flexibly utilize the multiple cues that they receive from social partners to learn words (Tomasello, 2000; Bahrick and Lickliter, 2014). But will there be development when infants start utilizing these cues? There are two theoretical frameworks on infant language learning that emphasize that there is a progression in the kind of cues that infants use to guide their early word learning: the PRIMIR framework (a developmental framework for Processing Rich Information from Multi-dimensional Interactive Representations; Werker and Curtin, 2005) and the ECM-model (Emergentist Coalition Model; Hollich et al., 2000). Both accounts acknowledge that the extent to which infants will be able to utilize cues will depend on the developmental stage of the infant. PRIMIR focuses on explaining development in speech perception and word learning by progression in the developmental level of the child, next to initial biases and language-specific requirements. The ECM-model stresses more the social nature of word learning, as it explains development in word learning through the combined roles of social-pragmatic factors, cognitive constraints, and global attentional mechanisms (Hollich et al., 2000; see also Tomasello, 2000 for the social-pragmatic account). As we are interested in gaze to faces as a social cue likely to facilitate

vocabulary acquisition, we examine this account more closely below.

The Emergentist Coalition model creates an important distinction between cues that are available versus utilized by infants, arguing that younger children rely on only a subset of the cues that older children are able to use. For example, whereas 7–8-month-old infants are shown to have their word recognition disrupted if a familiarized word is produced by a new voice, at 11 months infants no longer show this disruption (Houston and Jusczyk, 2003). Thus, although the social environment makes a number of cues available (such as multiple speakers repeating the same word; McRoberts et al., 2009), infants' utilization of these cues is dependent on their cognitive and social abilities. That is, attending to socially informative information present in faces will relate to word learning if (a) the infant is at a developmental stage where they are able to use the socially informative information for word learning, and (b) the information is relevant to the task at hand, and thus indeed informative. Therefore, whether gaze to faces proves relevant to vocabulary outcomes may depend on experimental parameters, such as ages tested or choice of tasks. For instance, since we know that certain social abilities, such as joint attention and gaze following (Reid and Striano, 2005), become relevant to vocabulary acquisition later in development, we can similarly imagine that the relevance of gaze to face to vocabulary acquisition also changes over development. Across development and across tasks, we may expect different relations between infant gaze to faces and vocabulary outcomes. This is why we next turn to review empirical evidence not only on whether gaze to faces feeds into vocabulary, but also to evaluate how this relation is affected by the developmental stage of the infants and tasks at hand.

How does gaze to faces relate to vocabulary outcomes?

The faces of social partners can provide infants with cues to guide their leaning of words but can be indexed in different ways. We will relate three specific aspects of gaze to faces relevant to vocabulary outcomes: first, gaze to specific facial features – to the eyes versus the mouth: then selective gaze to faces relative to other objects; and finally, selective gaze to more versus less social faces. In each of these three subsections, we will relate gaze to faces to vocabulary outcomes across infants' developmental trajectories and tasks.

How do gaze to the eyes and the mouth relate to vocabulary?

During our literature review, we found multiple studies that related infants' fixations to the eyes and mouth to vocabulary

outcomes (for recent reviews, see Çetinçelik et al., 2021; Bastianello et al., 2022). In what follows next, we first zoom in on whether there is any development across infancy in how infants attend to those facial features. For each facial element, we then consider how this relation to vocabulary is moderated by development and task. Finally, we explain the observed patterns by returning to the theories discussed earlier.

When infants are exposed to faces, they usually fixate first on the eyes of a social partner (Hills et al., 2013). Additionally, they attend to the eyes longer and more frequently than to the mouth (Haith et al., 1977). Infants prefer to look at eyes, even when presented with faces missing various parts (e.g., eyes, mouth, or nose): 2-month-olds fixate equally long to complete faces as to faces with only the eyes are present, and less to faces with only the mouth or the nose present (Maurer, 1985). From 2 to 6 months of age infants fixate more to the eyes than on other features of a social partner including their mouth and body (Jones and Klin, 2013). Some studies indicate that infants maintain the highest proportion of gaze to the eyes (Hunnus and Geuze, 2004; Lewkowicz and Hansen-Tift, 2012; Morin-Lessard et al., 2019). Others find that this higher frequency of fixations to the eyes is only to be present earlier in infancy (Frank et al., 2012). Although some studies indicate infants have a robust preference to attend to eyes, the mixed findings of studies indicate that this does not necessarily hold across development or across tasks.

Although the eyes remain the primary focus of attention, the mouth also increasingly draws attention during the first year of life (Young et al., 2009; Lewkowicz and Hansen-Tift, 2012; Tenenbaum et al., 2013; Elsabbagh et al., 2014). This finding is confirmed by several studies, notably with high participant numbers, and thus with good statistical power. These studies often test infants at multiple time points throughout the first years of life, some looking at as many as four to five separate developmental time points (Lewkowicz and Hansen-Tift, 2012). Infants are observed, in a task where they are gazing at a speaking face, to look longer to the eyes at 4 months, equally to the eyes and mouth at 6 and 8 months, more to the mouth at 10 months and finally more to the eyes at 12 months (Lewkowicz and Hansen-Tift, 2012). Therefore, across different timepoints in the first year of life, infants divide their length of fixations between the eyes and the mouth differentially for the same task.

The developmental stage of the infant also impacts the frequency with which infants attend to different types of eyes and mouths. That is, infants show progression in the types of facial features they prefer to attend to. For example, although both 9- and 10-month-old infants can differentiate between open and closed eyes, it is only from 10 months onward that infants recognize that only open eyes could provide information about where a social partner is looking (Brooks and Meltzoff, 2005). Thus, infant sensitivity to open versus closed eyes is a prerequisite to being able to follow their partner's eye gaze, which is a powerful cue shown to predict subsequent productive

vocabulary growth (e.g., (Brooks and Meltzoff, 2008, 2015). This indicates a developmental progression in the extent to which infants can utilize the cues provided by the eyes to direct their gaze.

There are also noticeable changes in how infants attend to different types of mouths. Infants increasingly also prefer looking at speaking mouths over other types (such as smiling mouths) from 6 to 9 months of age (Tenenbaum et al., 2013, 2015). Another example is that infant gaze to facial elements hinges on the type of speech they hear: a cross-sectional study manipulating whether infants watched speakers' producing a language that was native versus non-native to the infant showed that the looks that infants directed to the eyes and the mouth differed across the native and non-native speech conditions. While 4- to 8-month-olds increasingly devoted more gazes to the speakers' mouths irrespective of the type of speech, 12-month-olds only fixated more to the mouth when they heard a speaker producing a non-native versus native language (Lewkowicz and Hansen-Tift, 2012). This finding is observed cross-linguistically, with infants also undergoing an attentional shift to the mouth when the dominant language in their environment is Japanese (Sekiyama et al., 2021).

Another study manipulated whether or not there was synchrony between audio and visual information: at 10 months, infants' usual pattern of a preference to fixate to a mouth was absent in the desynchronized condition compared to when synchronized audio-visual information was presented (Hillairet de Boisferon et al., 2017). This finding held both when infants were presented with native or non-native speech. Similarly, looking times to the mouth were longer when infants were presented with speaking compared to silent faces – in the latter condition, looking times to the eyes were shown to be significantly longer (Tomalski et al., 2013). All these illustrations thus point to infants becoming increasingly sensitive to those situations that prove maximally informative, and this is mirrored in their differential gazes to specific facial regions.

While above we summarized studies that provide evidence of a development in how infants attend to both the eyes and the mouth, two recent review papers provide ample evidence that gaze to eyes (Çetinçelik et al., 2021) as well as a gaze to the mouth prove relevant to vocabulary outcomes (Bastianello et al., 2022). We complement their reviews by focusing on those studies that assess whether there is the developmental change in the relevance of gaze to eyes versus mouth regions.

The review by Bastianello et al. (2022) confirmed that increased gaze to (speaking) mouths around the first year of life is associated with infants' early expressive language skills across all of the reviewed studies in the paper. Yet, for slightly younger infants (5-month-olds), it is infant gaze to the eyes (over the mouth) that has been shown to predict their receptive vocabulary size at 14 months (Viktorsson et al., 2021). Similarly, in a situation where the face provides mismatched auditory and visual information, it is again longer gaze times to the eyes

and shorter gaze times to the mouth that correlates positively with 6 to 9-month-old infants' later receptive and productive vocabulary outcomes (Kushnerenko et al., 2013). However, in another study presenting infants with mismatched auditory and visual information, infants are found to have increased looking times to the mouth between 6 and 9 months of age (Tomalski et al., 2013). This seemingly contradictory pattern of results can be explained by the observation that while infants increasingly attend to the mouth over the course of development, they may not directly be able to utilize all cues that the speaker's mouth provides (Hollich et al., 2000). Given that integration of multi-sensory information is underdeveloped at birth and develops over the course of the first year (Burnham and Dodd, 2004; Bahrick et al., 2013), it could be that these younger infants could not yet fully utilize the multi-modal information most saliently present in the mouth region (Nardini et al., 2010). In this case, younger infants' capacity to ignore mismatched an unreliable cues and to rely more on the cues of the eyes instead may be predictive of vocabulary outcomes (Kushnerenko et al., 2013). In comparison, older infants may be able to use multimodal (that is, auditory and visual) information more flexibly, including in situations where the modal information is mismatched. Thus, there is development in how informative facial features can be to infants and subsequently in how infants' gazes to the eyes and mouth relate to vocabulary outcomes.

While most research points to a positive link between infants' increased fixation to the mouth or eyes with vocabulary outcomes, it is important to consider that this relation may not only hold for specific ages, but also for specific tasks (Kushnerenko et al., 2013; Altvater-Mackensen and Grossmann, 2015; Ter Schure et al., 2016; Danielson et al., 2017; Bastianello et al., 2022). To illustrate the effect of the task, we compare studies that differed in the complexity of the presented scene, while also linking infant gaze to vocabulary outcomes. In visually complex scenes (live action scenes with many characters performing different activities), 7-month-olds' increased fixations to the mouth is shown to be associated with superior productive vocabulary outcomes at 36 months (Frank et al., 2012). In comparison, in simpler live-action scenes (that contain a single face displaying communicative signals), increased fixations to the mouth in contrast relate to inferior productive vocabulary (Elsabbagh et al., 2014). Why might this be? This could be because the two tasks indicate different abilities when the infants fixate to the mouth. In the simpler scenes, a speaking mouth is the most perceptually salient feature within the scene, whereas in the visually complex scenes, a speaking mouth competes with other perceptually salient elements, but is still the most cue relevant for vocabulary. It is therefore important to consider how the same behavior (e.g., gaze to the mouth) may indicate different abilities depending on the task with which the infant is presented, and which cue (within the task) happens to be the most relevant for word learning. When evaluating how infants' gazes to faces relate to

vocabulary outcomes, it is therefore important to evaluate which cue is likely to be the most informative to the infants' word learning.

Having above reviewed evidence that both gaze to eyes as well as mouth prove relevant to vocabulary outcomes, we now turn to explain why this could be. Infants' fixations to the eyes could be beneficial to their word learning because the eyes of a social partner can signal their communicative intent or provide information about the referent to which a social partner is attending (Tomasello, 1992, 2000). Some studies show that infants first rely more on attentional cues, such as perceptual salience, rather than on gaze-cues from the interlocutor to guide their early word-object mappings (Brooks and Meltzoff, 2005, 2015; Pruden et al., 2006). As eyes are a perceptually salient feature of the face to which infants have a bias to attend to, then eyes may be one of the features that in early infancy draws infants to attend more to faces than to other objects (Di Giorgio et al., 2012). As an example, one study shows that infants who are able to detect when they are being gazed at and who can also subsequently follow the partner's gaze direction to an object may have a considerable advantage in determining the thoughts and intentions of their social partners (Langton et al., 2000). This is advantageous for word learning because understanding the internal state of a social partner increases the probability of an infant correctly discerning which of the many possible referents a social partner is communicating about.

While the eyes provide visual information to whom or about what the speaker is communicating, the mouth provides multimodal (visual and auditory) information about what is being said. In line with the Intersensory Redundancy Hypothesis, the combination of auditory and visual information may benefit word learning in multiple ways (Bahrick and Lickliter, 2012; Lewkowicz and Hansen-Tift, 2012). The development of increased fixations to mouths may benefit word learning because the mouth can provide a combination of auditory and visual cues that aid in the learning of words. This audio-visual information allows listeners to narrow down potential words by segmenting speech streams and locating word boundaries in continuous speech (Hollich et al., 2005; Mitchel and Weiss, 2014). Thus, the synchrony of the visual and auditory modalities (i.e., the movements of the mouth combined with the sounds it produces) makes it easier to narrow down *what* is being said (thereby facilitating infants' receptive vocabulary), *how* it is said (thereby facilitating infants' learning of expressive vocabulary), and *who* said it (highlighting communicative intent) (Bahrick and Lickliter, 2012; Lewkowicz and Hansen-Tift, 2012; Benders, 2013; Altwater-Mackensen and Grossmann, 2015). Infants who are learning words may therefore benefit from fixating to the mouth as a way to reduce several kinds of ambiguities of the visual and/or auditory information they are receiving through the combination of the two modalities.

Besides which facial elements infants selectively attend to, this review further demonstrates that the developmental stage of the infant plays a large role. We synthesized research reporting that progression in how infants attend to eyes versus mouth was relevant to their vocabulary acquisition. According to the Emergentist Coalition model, how infants develop their detection of and subsequently utilize socially more informative cues (such as speaking mouths) is critical to word learning (Hollich et al., 2000). In line with this, infants' development in the utilization of the social cues provided by the eyes is shown to correlate with later receptive and productive vocabulary (Brooks and Meltzoff, 2005, 2008). Additionally, the shift in gaze to specific types of mouths, i.e., speaking ones, is also shown to relate to vocabulary outcomes; for example, 6-month-olds who fixate more to the mother's mouth during live interaction have the superior productive vocabulary at 24 months (Young et al., 2009). This preferential fixation to speaking versus silent mouths could be beneficial to word learning by increasing the likelihood that the infants fixate to a mouth from whom they can learn words. Fixating more frequently to certain eyes and mouths may in turn facilitate word learning by increasing the likelihood of infants' fixating to a partner who is providing more communicative cues (e.g., one with direct gaze and/or a speaking mouth) and thereby increasing the opportunities for word learning.

How does gaze to faces relative to other objects relate to vocabulary?

In the examination of the literature that preceded the writing of this review, we could not find empirical evidence that assessed the impact that infants' preferential gaze to faces relative to objects had on their vocabulary outcomes. However, we theorize that infants' gazes to faces relative to other objects is a pre-requisite to more complex social abilities (such as gaze following), which have been shown to be linked to vocabulary outcomes (Carpenter et al., 1998; Morales et al., 2000; Slaughter and McConnell, 2003). We theorize this because it appears logical that before infants can follow adults' gaze correctly to labeled objects, infants first require ample experience with faces. This experience may be facilitated by the presence of a preference for faces over other objects. In this section we therefore review only the evidence of whether there are developmental changes in infant gaze to faces, and whether there are task-related changes. We then continue to speculate how this gaze to faces relative to other objects may relate to infants' vocabulary outcomes and how this relation may change across development and tasks.

Throughout infancy, infants are shown to have a preferential bias to attend to faces compared to other objects, looking longer to and orienting more frequently to face-like stimuli compared to non-face-like stimuli (Johnson et al., 1991;

Valenza et al., 1996). This is observed across a number of visual attention tasks: some studies reporting a face-preference involve infants' free viewing of clips that contain faces (e.g., video clips from the TV show Sesame Street) (Frank et al., 2009a, 2014; Franchak et al., 2016). Other studies look at preferential biases for faces by presenting images of faces together with other static objects (e.g., birds and cars) (Gliga et al., 2009; Elsabbagh et al., 2013). In the case of both paradigms, studies calculate the percentage of trials where the first fixation is directed to a certain category of object and/or calculate the average number of fixations to an area of interest. These similarities in how studies define infants' gaze to faces make it easier to cross-compare findings across studies (Hessels, 2019). In both paradigms, studies find that infants tend to direct their first gazes to faces as well as fixate more frequently to faces, compared to other stimuli on the screen (Gliga et al., 2009; Elsabbagh et al., 2013). Based on the qualities (e.g., sufficient participant numbers) and reoccurring findings of the studies, the finding that infants prefer to attend to faces (compared to other objects) appears robust.

Although all tasks show that infants primarily attend to faces, studies differ in the proportions of face preference, possibly based on differences in study parameters, such as children's age and children's opportunities to explore. To illustrate the mixed findings for the effect of age, there is a set of studies that recorded fixations to all available visual stimuli in a natural interaction for infants between 1 and 24 months of age. These studies report that across age, there is a decline in the frequency of gaze to faces, coupled with an increasing frequency of gaze to hands, meaning that age has an effect on the frequency of gazes to faces versus hands (Jayaraman et al., 2015; Fausey et al., 2016). In this paradigm, infants freely move about in their home environment while their gaze patterns are measured with head-mounted eye-trackers. Contrastingly, other studies (using different tasks) suggest that older infants fixate to faces more than younger infants do when watching video clips (6–24-month-olds; Franchak et al., 2016; 3–9-month-olds; Frank et al., 2014). The mixed findings across different tasks make it important to consider which aspects of the task or situation led to differences in infants' observed frequency of fixations to faces. Factors may include aspects of the methodology such as contrasting stimuli presented or whether the infants' movements during eye-tracking are restricted (as in Frank et al., 2014; Franchak et al., 2016) or not (as in Jayaraman et al., 2015; Fausey et al., 2016). Overall, it appears that the frequency of fixations that infants direct to faces hinges on their capacity for movement. Infants' capacity for movement depends both on their developmental stage (i.e., whether they can walk or sit up) and the task at hand (i.e., whether the procedure restricts their movements or not). Regarding developmental stage, older infants (who are more mobile) are shown to receive a more mixed visual input on faces and hands, whereas younger infants receive more input from faces. Regarding tasks, infants'

preference for faces may be stronger, when their movement and visual input is restricted. When drawing conclusions on the development and ubiquity of face preferences, it is therefore important to consider how the methodological choices may impact the behaviors that infants display during the procedure.

In the preceding paragraphs, we have considered how the developmental stage and task at hand influence how infants fixate to faces. Will this type of gaze to faces (face preference) also impact their vocabulary? As we have seen in the theoretical accounts, fixating to faces can be a facilitator of word learning because the cues provided by the face allow infants discern what their partner intends to communicate about and to whom their communication is directed to, e.g., to the infant, thereby guiding the infants' learning of words. Indeed, the bias to fixate to faces (over other objects) is arguably an important prerequisite to infants' vocabularies because it directs infant gaze to the cues provided by the face (Tomasello, 1992, 2000). Yet it is likely that there are developmental patterns as we have seen that face preference changes with development. A higher proportion of fixations to faces compared to other objects could be particularly important in younger infants, who have less developed social and cognitive abilities than older infants (Hollich et al., 2000), which in turn may compromise their ability to direct and maintain their gaze to objects in their environment, and as a result make them more dependent on a social partner to guide their learning (Colombo and Cheatham, 2006; Reynolds et al., 2013). Future research could investigate whether the relevance of preferential fixation to faces to vocabulary outcomes is more substantial early in vocabulary development, but declines with age. Furthermore, future research should consider how the choice of paradigm affects the relation between preferential fixation to faces and vocabulary outcomes.

How does gaze to more social versus less social faces relate to vocabulary?

When we are considering the types of faces that might prove informative, we note that current studies have not correlated infants developing a preference for more social faces to their vocabulary outcomes. Just as in the preceding subsection, we therefore first evaluate whether there are potential meaningful changes in this type of gaze to faces before, we speculate whether this could impact infant vocabulary outcomes.

Infants gradually begin to preferentially attend to certain types of faces over others, that is, faces that contain potentially more social cues (Smith and Gasser, 2005; Frank et al., 2009b, 2014; Slater et al., 2010). From 3 months onward, infants begin to prefer natural face images to unnatural ones (Turati et al., 2005). Around this age, they are also shown to preferentially fixate to upright over inverted faces (Chien, 2011; Elsabbagh et al., 2013). By 6 months, infants direct their first gazes to upright and not to inverted faces (Gliga et al., 2009).

Additionally, infants are shown to develop a preference to fixate to speaking faces above silent dynamic faces from 2 to 8 months, such that older infants increase their looks to speaking faces and decrease their looking away rates (Bahrick et al., 2016). Thus, the development of preference for specific faces appears to be robust across tasks; infants' increased sensitivity to certain faces moves toward faces to which infants are frequently exposed in their daily environment. Infants' gazes also increase toward more communicative faces over the course of development, while their gazes to less communicative faces remain similar (Bahrick et al., 2016). This preference to specific types of faces could be an index of perceptual learning, which refers to an increased sensitivity to specific faces frequently present in the environment and decreased sensitivity to others (Maurer and Werker, 2014).

When we are considering the relevance of attending to faces that are more social than other faces, we can only speculate that preferential gaze to more social faces could relate to vocabulary outcomes. Social faces are those types of faces that provide more explicit cues indicating communicative intent directed at the child (e.g., *via* direct gaze), and which are present often in the child's environment. As is theorized in the Social Pragmatic account of word learning, the understanding of communicative intent is key to the learning of words (Tomasello, 1992, 2000). Whether infants gaze to more social faces (that contain more explicit cues) arguably increases their opportunities for word learning and subsequently relates to their vocabulary outcomes. The relation between gaze to social faces and vocabulary outcomes may hold in particular for younger infants. Whereas older infants are shown to be able to learn words by overhearing conversations, it is less clear whether younger infants have similar capacities (Akhtar et al., 2001). Younger infants may rely more on communicative intent recognized by more visual engagement with the speaker (Tomasello, 2000). Therefore, infants' capacities to attend to more socially relevant faces could relate more to vocabulary outcomes when they are younger.

There are other task-related factors which arguably influence how socially relevant certain types of faces are, and which further could impact vocabulary outcomes. For example, whether or not a person's face is physically present or whether this person is reciprocating the infants' behaviors could ultimately influence vocabulary outcomes. Whereas some studies find that infants can learn from digitally presented faces (i.e., faces that are not physically present) that direct their gaze at a target object (Houston-Price et al., 2006), others find that there is a "video-deficit," and that infants' learning is hindered when the tutor appears on video instead of being physically present (Anderson and Pempek, 2005). Findings are mixed regarding the extent to which infants can use the cues provided by physically versus digitally presented faces to guide their word learning (O'Doherty et al., 2011; Roseberry et al., 2014; Troseth et al., 2018; Tsuji et al., 2020). In part, these mixed findings may also depend on the difficulty of the word-learning task with

which the infant is presented (Gogate and Madhvilatha, 2017). More difficult word-learning tasks may benefit from the faces of social partners being physically present and reciprocating. Further research is needed to clarify in which tasks infants can use the face to guide their learning of words. Additionally, further research is required as to how the difficulty of the word-learning task impacts infants' abilities to learn from faces that are not physically present and/or not reciprocating.

To summarize, based on the existing literature, it remains unclear how infants' fixation toward more socially relevant faces impacts their vocabulary outcomes. Further research is needed to elucidate how flexibly infants can learn word-world pairings from the faces of social partners and the extent to which this flexibility depends on the developmental stage of the infant, and the task at hand.

In the above three sections of this paper, we discussed how three aspects of infant gaze to faces relate to vocabulary outcomes: gaze to different elements within faces; gaze to faces relative to other objects; and gaze to different types of faces. We will now discuss our findings in the literature and make suggestions for future research.

Discussion

This position paper first aimed to assess how infant gaze to faces may feed into their vocabulary outcomes. We reviewed the literature on three aspects of infant gaze to faces: gaze to the eyes compared to the mouth; gaze to faces compared to objects; and gaze to more socially relevant faces. Several studies were found that related infant gaze to facial elements and vocabulary outcomes. Here, we observed that the relationship between infants' gaze to eyes and the mouth with their vocabulary outcomes was impacted by the developmental stage of the infant, and the task at hand. However, to the best of our knowledge, no studies explored how gaze to faces (compared to other objects) or gaze to more social (compared to less social) faces relate to vocabulary outcomes. We will now discuss how gaze to faces could influence vocabulary outcomes, pointing out how different strands of future research can tackle the further assessment of these cascading effects. We will then discuss some of the limitations of the review and point to possible future directions.

A number of studies related gaze to the eyes versus mouth to vocabulary outcomes. Studies examined both how infant gaze to the eyes (Carpenter et al., 1998; Mundy and Gomes, 1998; Morales et al., 2000; Brooks and Meltzoff, 2005, 2008) and to the mouth relates to vocabulary outcomes (Young et al., 2009; Elsabbagh et al., 2014; Tenenbaum et al., 2015), and how these relations change over development. That is, infants increasingly attend to the mouth over the course of the first year of life, and this developmental shift longitudinally predicts infants' vocabulary outcomes later in development

(Elsabbagh et al., 2014; Tenenbaum et al., 2015). Developing a capacity to switch between fixating to the eyes or mouth of a social partner is shown to have feedforward effects on vocabulary outcomes because the eyes and mouth provide meaningful yet different cues to word-object pairings in their environment. Arguably, infants who can selectively attend to the eyes and the mouth at points in time when such a facial feature is maximally socially informative will receive more cues to guide their word learning. Subsequently, these infants will have larger vocabulary outcomes compared to other infants. Yet, although there is substantial research linking gaze to the eyes or to mouth with word learning, additional clarification is needed on when and in which situations it is that infants develop an appropriate ability to socially encode information from the more relevant facial features, such that their gazes predicts and facilitates their word learning.

Whether there is a relationship between infants' preference to gaze to faces (over other objects) and their vocabulary outcome remains theoretical, with no empirical evidence looking into whether face-preference has a feedforward effect on vocabulary outcomes. Additionally, no studies have attempted to correlate how infants' preferential fixation to certain types of social faces may relate to their later vocabulary outcomes. Infant preference for faces (compared to objects) and their preference for more compared to less social faces may facilitate word learning by increasing the probability of infants attending to the relevant social cues that the face provides to guide their word learning (Tomasello, 1992, 2000).

Gaze to faces has frequently been theorized and empirically shown to have feedforward effects on vocabulary outcomes (Çetinçelik et al., 2021). Feedforward effects (also defined in the literature as cascading effects) are the cumulative consequences of the many interactions and transactions occurring in developing systems (Masten and Cicchetti, 2010; Sameroff, 2010; D'Souza and Karmiloff-Smith, 2011; Junge et al., 2020). Preferential gaze to faces is an early instance of selective attention, that may direct infants' gazes to faces and provide infants with extensive experience of face stimuli. This extensive experience with faces could be a precursor to more complex social abilities (e.g., gaze following) that have frequently been shown to relate to vocabulary outcomes (Çetinçelik et al., 2021). Subsequently, infants' early face preference could feed into word learning by feeding into infants' complex social abilities that directly relate to their word learning. Understanding these feedforward effects of preferential gaze to faces on vocabulary outcomes may guide us toward the mechanisms and constraints leading to the acquisition of words and the subsequent vocabulary outcomes (D'Souza and Filippi, 2017; Kidd and Donnelly, 2020). In light of this, studies correlating infants' early preferential gaze to faces with their vocabulary outcomes will give us insight into whether and why infants can learn their first words from the information that faces provide.

As yet, confirmative research is required to substantiate our hypotheses and further explain the nature of this feedforward effect. Regarding this nature, it is possible that gaze to faces has some direct feedforward effects on vocabulary outcomes. Alternatively, it is possible that face-preference serves as a mediator to vocabulary outcomes, as it scaffolds more sophisticated social abilities, such as joint attention (Junge et al., 2020). The strength of the relationship between infant gaze to faces and vocabulary outcomes could change as a function of the developmental stage, with face-preference facilitating word learning earlier in development more than later in development. The relation between gazes to faces and vocabulary outcomes is therefore likely to be stronger when assessing younger infants, because older infants have access to a larger range of social mechanisms, e.g., joint attention, that are made up of multiple smaller social domains. Therefore, we recommend that future research explores the relationship between the three domains that are addressed in this review (gazes to eyes and mouth; gazes to faces versus objects; and gazes to specific face types) and vocabulary outcomes in younger infant groups, who have fewer social capabilities at their disposal for word learning.

Additionally, the task given to the infant is also likely to affect the feedforward effect. When examining how infants attend to faces it is therefore important to consider both how the developmental stage and task given to the infant influence their processing of a face that they could attend to and how this processing may subsequently impact infants' vocabulary outcomes.

Limitations

This review has some limitations. First, as outlined above, the links made between some aspects of infant gaze to faces and vocabulary outcomes in this review remain theoretical and require more empirical evidence. Although we did not systematically review experimental findings, it appears there is insufficient research to draw clearer conclusions about how specific aspects of infant gaze to faces feed into vocabulary outcomes. Additional research is needed to confirm these links, including large-scale longitudinal studies, experimental studies with different paradigms, and intervention studies to illuminate whether, how, and when infant gaze to (aspects of) faces impacts vocabulary outcomes (Masten and Cicchetti, 2010).

Second, it is also important to take into account that the research from which we draw our theoretical links involve primarily (if not completely) samples of infants and parents from societies that are Western, educated, industrialized, rich, and democratic (that is, WEIRD societies: Henrich et al., 2010). The beliefs, traditions, and day-to-day lives of individuals from non-WEIRD societies may differ from WEIRD societies to the extent whereby the links observed in one society may not be comparable to the other. For example, whereas in

Boston, MA, America (a WEIRD society) mothers are shown to frequently visually engage with their infants, in Gusii, Kenya (a non-WEIRD society) mothers engage more frequently through holding and touching instead of direct gazing at their infant (Akhtar and Gernsbacher, 2007). Our hypothesis that gaze to faces proves relevant in a number of ways should therefore not be interpreted as the most significant or only successful facilitators of word learning.

Our aim was to disentangle how an early developing social cue (infants' attending to faces) related to vocabulary outcomes. Of course, this does not mean that gaze to faces is the only potential cue that relates to vocabulary outcomes. The literature documents a myriad of factors that impacts early word learning, ranging from the infant level to familial risks, to the environment (Kidd and Donnelly, 2020). For instance, at the infant level there exist many possible predictors: infants' non-verbal cognitive skills (Colombo et al., 2008; Rose et al., 2009), general auditory abilities (Benasich and Tallal, 2002), and speech perception abilities (Fernald and Marchman, 2012; Cristia et al., 2014; Ference and Curtin, 2015; Wang et al., 2021). Moreover, we note that there are several other perceptual processes developing during infancy that contribute to more complex forms of social processing, such as joint attention (Lewkowicz and Ghazanfar, 2006; Scherf and Scott, 2012; Hadley et al., 2014; Happé and Frith, 2014; Pascalis et al., 2014). In this review, we have chosen to restrict the scope to fixation to faces as a proxy of early social behavior observable in infants from birth onward. It remains to be seen how all these potential factors hold together when explaining individual variations in early vocabulary.

Future research directions

Based on the research findings and hypotheses compiled in this review, there are a number of gaps in the literature and subsequent directions that future research can take to further elucidate the role of infants attending to faces in being relevant to vocabulary outcomes. One line of studies could use repeated multiple measurements to investigate how and when across development gaze to faces (over objects) or gaze to more (compared to less) socially informative types of faces are indeed related to vocabulary outcomes. This could shed light on the developmental processes that gaze to faces comprises. For example, how infants attend to faces (compared to objects) may be more predictive of vocabulary outcomes earlier in development and become less predictive when infants start utilizing additional cues (such as direct and averted gaze) to guide their learning. Similarly, gaze to social faces may (or may not) become less predictive of vocabulary outcomes as infants start to direct more of their gaze to the surrounding environment or to other social stimuli, e.g., the hands (Smith and Yu, 2013; Deák et al., 2014; Jayaraman et al., 2015; Fausey et al., 2016).

Another line of research should focus on explaining the individual variation in early vocabulary (Bates et al., 1994; Frank et al., 2017). Understanding the sources of individual variation ultimately informs both theory-forming as well as practitioners aiming to maximize children's word-learning potential (Kidd and Donnelly, 2020). To assess whether variation in early gaze to faces predicts word learning, as we hypothesize here, we need more empirical evidence testifying that there is such a link, and why such a link would exist. For example, it could well be that infants who look more at their partners may in turn receive more social responses from their partners that prolong the length of the interaction. Lengthening the interaction may increase the time window in which word learning can occur and improve the quality of the interaction may increase the probability of correct word-object pairings being made. Short-term, this could lead to more word learning opportunities when the infant is engaged in communication with a partner and long-term it could lead to observable differences in the vocabulary size and content of the infant. Subsequent studies could then use intervention-designs focusing on fostering infants' fixation to faces to promote early vocabulary.

Finally, research could focus on well-controlled laboratory studies to carefully examine how the task within which the infant is engaged could impact the extent to which gaze to faces is predictive of vocabulary outcomes. For example, it has been shown that in some experimental settings, 1-year-olds hardly look at the faces of their social partner, but instead coordinate joint attention between themselves and a social partner by attending to objects held by themselves or their partner (Smith and Yu, 2013). In situations where the infant and the partner are handling the objects instead of looking to a faraway/out-of-reach object, there may be a lower ambiguity of the word-referent pairing, and thus less reliance on facial cues, than when an object being referred to is not in direct reach of the infant or social partner (Deák et al., 2014). Taking into account the interaction contexts (or tasks) in which infants learn words will expand our understanding of how and why visual social cues, such as those present in the faces of their social partners, affect word learning.

Future research could thus take into account how developmental, task-related and individual differences in infants attending to faces have feedforward effects on vocabulary outcomes. There are ample opportunities and directions for future research.

Conclusion

Overall, infant gaze to faces could have an important effect on early word learning through the constrictions that facial cues provide on the natural variability of environments. A face-preference appears to be an initial bias that aids infants' gazes to social stimuli early in development, when they have less

attentional control. It may also feed into later developing, more complex social abilities, such as gaze following, that have been found to relate to word learning and vocabulary outcomes. Gazes to specific features of the face, on the other hand, develop over time and may constrict information relating to words, referents, and word-object pairings. Infants' gazes to the eyes may aid in their discerning of communicative intent and in determining where in the environment the gaze of a social partner lies. Gaze to mouth movements provide multi-modal information to aid the processing of speech and learning of words, as well as reinforcing child-directed speech. Combined, these processes provide numerous cues to facilitate the creation of word-object pairings. There are a number of studies that have shown how infants gazes to the eyes and mouth relate to vocabulary outcomes. However, whether infant gaze to faces (compared to objects) as well as to more socially relevant types of faces relate to vocabulary outcomes remains speculative and could depend on an infant's developmental level, which affects their ability to correctly discern and use such cues to guide their learning. Developmental level, as well as the task (i.e., situational factors), is therefore important to consider when evaluating correlations between infant gaze to faces and their vocabulary outcomes. Although this review hypothesizes that infant gaze to faces relates to their vocabulary outcomes, and finds some evidence in favor of our hypothesis, future empirical studies could examine the feedforward effects on vocabulary outcomes more directly.

Author contributions

ZB, CB, and CJ: conceptualization, investigation, writing—original draft, and writing—review and editing. CB and CJ:

funding acquisition and supervision. ZB: project administration. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Horizon 2020 Marie Skłodowska-Curie Action (MSCA) Innovative Training Network (ITN) – European Training Network (ETN), grant number: 814302 – SAPIENS) and the Consortium on Individual Development (CID). CID is funded through the Gravitation Program of the Dutch Ministry of Education, Culture, and Science and the NWO (grant number: 024.001.003).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Akhtar, N., and Gernsbacher, M. A. (2007). Joint attention and vocabulary development: A critical look. *Lang. Linguist. Compass* 1, 195–207. doi: 10.1111/j.1749-818X.2007.00014.x
- Akhtar, N., Jipson, J., and Callanan, M. A. (2001). Learning words through overhearing. *Child Dev.* 72, 416–430. doi: 10.1111/1467-8624.00287
- Altwater-Mackensen, N., and Grossmann, T. (2015). Learning to match auditory and visual speech cues: Social influences on acquisition of phonological categories. *Child Dev.* 86, 362–378. doi: 10.1111/cdev.12320
- Anderson, D. R., and Pempek, T. A. (2005). Television and very young children. *Am. Behav. Sci.* 48, 505–522. doi: 10.1177/0002764204271506
- Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., and Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Dev.* 92, 484–501. doi: 10.1111/cdev.13508
- Bahrack, L. E., and Lickliter, R. (2012). "The role of intersensory redundancy in early perceptual, cognitive, and social development," in *Multisensory development*, eds A. J. Bremner, D. J. Lewkowicz, and C. Spence (Oxford: OUP Oxford), 183–206.
- Bahrack, L. E., and Lickliter, R. (2014). Learning to attend selectively: The dual role of intersensory redundancy. *Curr. Direct. Psychol. Sci.* 23, 414–420. doi: 10.1177/0963721414549187
- Bahrack, L. E., Lickliter, R., and Castellanos, I. (2013). The development of face perception in infancy: Intersensory interference and unimodal visual facilitation. *Dev. Psychol.* 49, 1919–1930. doi: 10.1037/a0031238
- Bahrack, L. E., Todd, J. T., Castellanos, I., and Sorondo, B. M. (2016). Enhanced attention to speaking faces versus other event types emerges gradually across infancy. *Dev. Psychol.* 52, 1705–1720. doi: 10.1037/dev0000157
- Bastianello, T., Keren-Portnoy, T., Majorano, M., and Vihman, M. (2022). Infant looking preferences towards dynamic faces: A systematic review. *Infant Behav. Dev.* 67:101709. doi: 10.1016/j.infbeh.2022.101709
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., et al. (1994). Developmental and stylistic variation in the composition of early vocabulary. *J. Child Lang.* 21, 85–123. doi: 10.1017/S0305000900008680
- Benasich, A. A., and Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behav. Brain Res.* 136, 31–49. doi: 10.1016/S0166-4328(02)00098-0

- Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behav. Dev.* 36, 847–862. doi: 10.1016/j.infbeh.2013.09.001
- Bergelson, E., and Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci.* 109, 3253–3258. doi: 10.1073/pnas.1113380109
- Braginsky, M., Yurovsky, D., Marchman, V. A., and Frank, M. C. (2019). Consistency and variability in children's word learning across languages. *Open Mind* 3, 52–67. doi: 10.1162/opmi_a_00026
- Brooks, R., and Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Dev. Sci.* 8, 535–543. doi: 10.1111/j.1467-7687.2005.00445.x
- Brooks, R., and Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *J. Child Lang.* 35, 207–220. doi: 10.1017/S030500090700829X
- Brooks, R., and Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *J. Exp. Child Psychol.* 130, 67–78. doi: 10.1016/j.jecp.2014.09.010
- Burnham, D., and Dodd, B. (2004). Auditory–visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Dev. Psychobiol.* 45, 204–220. doi: 10.1002/dev.20032
- Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., and Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, 1–143. doi: 10.2307/1166214
- Çetincelik, M., Rowland, C. F., and Snijders, T. M. (2021). Do the eyes have it? A systematic review on the role of eye gaze in infant language development. *Front. Psychol.* 11:3627. doi: 10.3389/fpsyg.2020.589096
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., and Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cogn. Dev.* 15, 481–498. doi: 10.1016/S0885-2014(01)00037-5
- Chater, N., and Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends Cogn. Sci.* 10, 335–344. doi: 10.1016/j.tics.2006.05.006
- Chien, S. H.-L. (2011). No more top-heavy bias: Infants and adults prefer upright faces but not top-heavy geometric or face-like patterns. *J. Vision* 11, 13–13. doi: 10.1167/11.6.13
- Cleveland, A., Schug, M., and Striano, T. (2007). Joint attention and object learning in 5- and 7-month-old infants. *Infant Child Dev.* 16, 295–306. doi: 10.1002/icd.508
- Colombo, J., and Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Adv. Child Dev. Behav.* 34, 283–322. doi: 10.1016/s0065-2407(06)80010-8
- Colombo, J., Shaddy, J. D., Blaga, O. M., Anderson, C. J., Kannass, K. N., and Richman, W. A. (2008). “Early attentional predictors of vocabulary in childhood,” in *Infant pathways to language*, eds J. Colombo, P. McCordle, and L. Freund (Hove: Psychology Press), 143–168.
- Cristia, A., Seidl, A., Junge, C., Soderstrom, M., and Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. *Child Dev.* 85, 1330–1345. doi: 10.1111/cdev.12193
- D'Souza, D., and Filippi, R. (2017). Progressive modularization: Reframing our understanding of typical and atypical language development. *First Lang.* 37, 518–529. doi: 10.1177/0142723717720038
- D'Souza, D., and Karmiloff-Smith, A. (2011). When modularization fails to occur: A developmental perspective. *Cogn. Neuropsychol.* 28, 276–287. doi: 10.1080/02643294.2011.614939
- Danielson, D. K., Bruderer, A. G., Kandhadai, P., Vatikiotis-Bateson, E., and Werker, J. F. (2017). The organization and reorganization of audiovisual speech perception in the first year of life. *Cogn. Dev.* 42, 37–48. doi: 10.1016/j.cogdev.2017.02.004
- Deák, G. O., Krasno, A. M., Triesch, J., Lewis, J., and Sepeta, L. (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Dev. Sci.* 17, 270–281. doi: 10.1111/desc.12122
- Di Giorgio, E., Turati, C., Altoè, G., and Simion, F. (2012). Face detection in complex visual displays: An eye-tracking study with 3- and 6-month-old infants and adults. *J. Exp. Child Psychol.* 113, 66–77. doi: 10.1016/j.jecp.2012.04.012
- Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., and Johnson, M. H. (2014). What you see is what you get: Contextual modulation of face scanning in typical and atypical development. *Soc. Cogn. Affect. Neurosci.* 9, 538–543. doi: 10.1093/scan/nst012
- Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., and Johnson, M. H. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behav. Brain Res.* 251, 147–154. doi: 10.1016/j.bbr.2012.07.030
- Fausey, C. M., Jayaraman, S., and Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition* 152, 101–107. doi: 10.1016/j.cognition.2016.03.005
- Ference, J., and Curtin, S. (2015). The ability to map differentially stressed labels to objects predicts language development at 24 months in 12-month-olds at high risk for Autism. *Infancy* 20, 242–262. doi: 10.1111/infia.12074
- Fernald, A., and Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Dev.* 83, 203–222. doi: 10.1111/j.1467-8624.2011.01692.x
- Floor, P., and Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy* 9, 327–339. doi: 10.1207/s15327078inf0903_4
- Franchak, J. M., Heeger, D. J., Hasson, U., and Adolph, K. E. (2016). Free viewing gaze behavior in infants and adults. *Infancy* 21, 262–287. doi: 10.1111/infia.12119
- Frank, M. C., Amso, D., and Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *J. Exp. Child Psychol.* 118, 13–26. doi: 10.1016/j.jecp.2013.08.012
- Frank, M. C., Braginsky, M., Yurovsky, D., and Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *J. Child Lang.* 44, 677–694. doi: 10.1017/S0305000916000209
- Frank, M. C., Goodman, N. D., and Tenenbaum, J. B. (2009a). Using speakers' referential intentions to model early cross-situational word learning. *Psychol. Sci.* 20, 578–585. doi: 10.1111/j.1467-9280.2009.02335.x
- Frank, M. C., Vul, E., and Johnson, S. P. (2009b). Development of infants' attention to faces during the first year. *Cognition* 110, 160–170. doi: 10.1016/j.cognition.2008.11.010
- Frank, M. C., Vul, E., and Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy* 17, 355–375. doi: 10.1111/j.1532-7078.2011.00086.x
- Gliga, T., Elsabbagh, M., Andravizou, A., and Johnson, M. (2009). Faces attract infants' attention in complex displays. *Infancy* 14, 550–562. doi: 10.1080/15250000903144199
- Gogate, L., and Hollich, G. (2010). Invariance detection within an interactive system: A perceptual gateway to language development. *Psychol. Rev.* 117, 496–516. doi: 10.1037/a0019049
- Gogate, L., and Madhavilatha, M. (2017). The origins of verb learning: Preverbal and postverbal infants' learning of word–action relations. *J. Speech Lang. Hear. Res.* 60, 3538–3550. doi: 10.1044/2017_JSLHR-L-17-0085
- Gredebäck, G., Fikke, L., and Melinder, A. (2010). The development of joint visual attention: A longitudinal study of gaze following during interactions with mothers and strangers. *Dev. Sci.* 13, 839–848. doi: 10.1111/j.1467-7687.2009.00945.x
- Grossmann, T. (2017). The eyes as windows into other minds. *Perspect. Psychol. Sci.* 12, 107–121.
- Hadley, H., Rost, G. C., Fava, E., and Scott, L. S. (2014). A mechanistic approach to cross-domain perceptual narrowing in the first year of life. *Brain Sci.* 4, 613–634. doi: 10.3390/brainsci4040613
- Haith, M. M., Bergman, T., and Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science* 198, 853–855. doi: 10.1126/science.918670
- Happé, F., and Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *J. Child Psychol. Psychiatry* 55, 553–577. doi: 10.1111/jcpp.12162
- Henrich, J., Heine, S. J., and Norenzayan, A. (2010). Most people are not WEIRD. *Nature* 466, 29–29. doi: 10.1038/466029a
- Hessels, R. S. (2019). How does gaze to faces support face-to-face interaction? A review and perspective. *Psychon. Bull. Rev.* 27, 856–881. doi: 10.31219/osf.io/8zta5
- Hillairet de Boisferon, A., Tift, A. H., Minar, N. J., and Lewkowicz, D. J. (2017). Selective attention to a talker's mouth in infancy: Role of audiovisual temporal synchrony and linguistic experience. *Dev. Sci.* 20:10.1111/desc.12381. doi: 10.1111/desc.12381
- Hills, P. J., Cooper, R. E., and Pake, J. M. (2013). First fixations in face processing: The more diagnostic they are the smaller the face-inversion effect. *Acta Psychol.* 142, 211–219. doi: 10.1016/j.actpsy.2012.11.013
- Hollich, G., Hirsh-Pasek, K., and Golinkoff, R. M. (2000). *Breaking the language barrier: An emergentist coalition model for the origins of word learning*. Chicago: University of Chicago Press.

- Hollich, G., Newman, R. S., and Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Dev.* 76, 598–613. doi: 10.1111/j.1467-8624.2005.00866.x
- Houston, D. M., and Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *J. Exp. Psychol.* 29, 1143–1154. doi: 10.1037/0096-1523.29.6.1143
- Houston-Price, C., Plunkett, K., and Duffy, H. (2006). The use of social and salience cues in early word learning. *J. Exp. Child Psychol.* 95, 27–55. doi: 10.1016/j.jecp.2006.03.006
- Hunnus, S., and Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy* 6, 231–255. doi: 10.1207/s15327078in0602_5
- Jayaraman, S., Fausey, C. M., and Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLoS One* 10:e0123780. doi: 10.1371/journal.pone.0123780
- Johnson, E. K. (2016). Constructing a proto-lexicon: An integrative view of infant language development. *Annu. Rev. Linguist.* 2, 391–412. doi: 10.1146/annurev-linguistics-011415-040616
- Johnson, M. H., Dziurawiec, S., Ellis, H., and Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40, 1–19. doi: 10.1016/0010-0277(91)90045-6
- Jones, W., and Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature* 504, 427–431. doi: 10.1038/nature12715
- Junge, C., Valkenburg, P. M., Deković, M., and Branje, S. (2020). The building blocks of social competence: Contributions of the consortium of individual development. *Dev. Cogn. Neurosci.* 45:100861. doi: 10.1016/j.dcn.2020.100861
- Kavanaugh, R. D., and Jirkovsky, A. M. (1982). Parental speech to young children: A longitudinal analysis. *Merrill Palmer Q.* 28, 297–311.
- Kidd, E., and Donnelly, S. (2020). Individual differences in first language acquisition. *Annu. Rev. Linguist.* 6, 319–340. doi: 10.1146/annurev-linguistics-011619-030326
- Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Dev. Sci.* 10, 110–120. doi: 10.1111/j.1467-7687.2007.00572.x
- Kushnerenko, E. V., Tomalski, P., Ballieux, H., Potton, A., Birtles, D., Frostick, C., et al. (2013). Brain responses and looking behavior during audiovisual speech integration in infants predict auditory speech comprehension in the second year of life. *Front. Psychol.* 4:432. doi: 10.3389/fpsyg.2013.00432
- Langton, S. R., Watt, R. J., and Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends Cogn. Sci.* 4, 50–59. doi: 10.1016/s1364-6613(99)01436-9
- Lewkowicz, D. J., and Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proc. Natl. Acad. Sci.* 103, 6771–6774. doi: 10.1073/pnas.0602027103
- Lewkowicz, D. J., and Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proc. Natl. Acad. Sci.* 109, 1431–1436. doi: 10.1073/pnas.1114783109
- Locke, J. L. (1993). "The role of the face in vocal learning and the development of spoken language," in *Developmental neurocognition: Speech and face processing in the first year of life*, eds B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, and J. Morton (Berlin: Springer), 317–328. doi: 10.1007/978-94-015-8234-6_26
- Madigan, S., Prime, H., Graham, S. A., Rodrigues, M., Anderson, N., Khoury, J., et al. (2019). Parenting behavior and child language: A meta-analysis. *Pediatrics* 144:e20183556.
- Masten, A. S., and Cicchetti, D. (2010). Developmental cascades. *Dev. Psychopathol.* 22, 491–495. doi: 10.1017/S0954579410000222
- Maurer, D. (1985). "Infants' perception of facedness," in *Social perception in infants*, eds T. N. Field and N. Fox (New York, NY: Ablex).
- Maurer, D., and Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Dev. Psychobiol.* 56, 154–178. doi: 10.1002/dev.21177
- McRoberts, G. W., McDonough, C., and Lakusta, L. (2009). The role of verbal repetition in the development of infant speech preferences from 4 to 14 months of age. *Infancy* 14, 162–194. doi: 10.1080/15250000802707062
- Mitchel, A. D., and Weiss, D. J. (2014). Visual speech segmentation: Using facial cues to locate word boundaries in continuous speech. *Lang. Cogn. Neurosci.* 29, 771–780. doi: 10.1080/01690965.2013.791703
- Moore, C., Angelopoulos, M., and Bennett, P. (1999). Word learning in the context of referential and salience cues. *Dev. Psychol.* 35, 60–68. doi: 10.1037/0012-1649.35.1.60
- Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., et al. (2000). Responding to joint attention across the 6- through 24-month age period and early language acquisition. *J. Appl. Dev. Psychol.* 21, 283–298. doi: 10.1016/S0193-3973(99)00040-4
- Morin-Lessard, E., Poulin-Dubois, D., Segalowitz, N., and 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, 1640–1655. doi: 10.1037/dev0000750
- Mundy, P., and Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behav. Dev.* 21, 469–482. doi: 10.1016/S0163-6383(98)90020-0
- Mundy, P., Fox, N., and Card, J. (2003). EEG coherence, joint attention and language development in the second year. *Dev. Sci.* 6, 48–54. doi: 10.1111/1467-7687.00253
- Nardini, M., Bedford, R., and Mareschal, D. (2010). Fusion of visual cues is not mandatory in children. *Proc. Natl. Acad. Sci.* 107, 17041–17046. doi: 10.1073/pnas.1001699107
- O'Doherty, K., Troseth, G. L., Shimpi, P. M., Goldenberg, E., Akhtar, N., and Saylor, M. M. (2011). Third-party social interaction and word learning from video. *Child Dev.* 82, 902–915. doi: 10.1111/j.1467-8624.2011.01579.x
- Pascalis, O., Loevenbruck, H., Quinn, P. C., Kandel, S., Tanaka, J. W., and Lee, K. (2014). On the links among face processing, language processing, and narrowing during development. *Child Dev. Perspect.* 8, 65–70. doi: 10.1111/cdep.12064
- Pinker, S. (1979). Formal models of language learning. *Cognition* 7, 217–283. doi: 10.1016/0010-0277(79)90001-5
- Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., and Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Dev.* 77, 266–280. doi: 10.1111/j.1467-8624.2006.00869.x
- Quine, W. V. O. (1960). *Word and object: An inquiry into the linguistic mechanisms of objective reference*. Hoboken, NJ: John Wiley.
- Reid, V. M., and Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *Eur. J. Neurosci.* 21, 1763–1766. doi: 10.1111/j.1460-9568.2005.03986.x
- Reynolds, G. D., Courage, M. L., and Richards, J. E. (2013). *The development of attention*. In *The Oxford handbook of cognitive psychology*. Oxford: Oxford University Press, 1000–1013. doi: 10.1093/oxfordhdb/9780195376746.013.0063
- Rose, S. A., Feldman, J. F., and Jankowski, J. J. (2009). A cognitive approach to the development of early language. *Child Dev.* 80, 134–150. doi: 10.1111/j.1467-8624.2008.01250.x
- Roseberry, S., Hirsh-Pasek, K., and Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Dev.* 85, 956–970. doi: 10.1111/cdev.12166
- Sameroff, R. (2010). A unified theory of development: A dialectic integration of nature and nurture. *Child Dev.* 81, 6–22. doi: 10.1111/j.1467-8624.2009.01378.x
- Scherf, K. S., and Scott, L. S. (2012). Connecting developmental trajectories: Biases in face processing from infancy to adulthood. *Dev. Psychobiol.* 54, 643–663. doi: 10.1002/dev.21013
- Sekiya, K., Hisanaga, S., and 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. doi: 10.1016/j.cortex.2021.03.023
- Senju, A., and Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Curr. Biol.* 18, 668–671. doi: 10.1016/j.cub.2008.03.059
- Slater, A., Quinn, P. C., Kelly, D. J., Lee, K., Longmore, C. A., McDonald, P. R., et al. (2010). The shaping of the face space in early infancy: Becoming a native face processor. *Child Dev. Perspect.* 4, 205–211. doi: 10.1111/j.1750-8606.2010.00147.x
- Slaughter, V., and McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *J. Genet. Psychol.* 164, 54–71. doi: 10.1080/00221320309597503
- Smith, L. B., and Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition* 106, 1558–1568. doi: 10.1016/j.cognition.2007.06.010
- Smith, L. B., and Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Lang. Learn. Dev.* 9, 25–49. doi: 10.1080/15475441.2012.707104

- Smith, L., and Gasser, M. (2005). The development of embodied cognition: Six lessons from babies. *Artif. Life* 11, 13–29. doi: 10.1162/1064546053278973
- Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., and Morgan, J. L. (2013). Increased focus on the mouth among infants in the first year of life: A longitudinal eye-tracking study. *Infancy* 18, 534–553. doi: 10.1111/j.1532-7078.2012.00135.x
- Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., and Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development*. *J. Child Lang.* 42, 1173–1190. doi: 10.1017/S0305000914000725
- Ter Schure, S. M. M., Junge, C. M. M., and Boersma, P. P. G. (2016). Semantics guide infants' vowel learning: Computational and experimental evidence. *Infant Behav. Dev.* 43, 44–57. doi: 10.1016/j.infbeh.2016.01.002
- Tincoff, R., and Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy* 17, 432–444. doi: 10.1111/j.1532-7078.2011.00084.x
- Tomalski, P., Ribeiro, H., Ballieux, H., Axelsson, E. L., Murphy, E., Moore, D. G., et al. (2013). Exploring early developmental changes in face scanning patterns during the perception of audiovisual mismatch of speech cues. *Eur. J. Dev. Psychol.* 10, 611–624. doi: 10.1080/17405629.2012.728076
- Tomasello, M. (1992). The social bases of language acquisition. *Soc. Dev.* 1, 67–87. doi: 10.1111/j.1467-9507.1992.tb00135.x
- Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics* 10, 401–413. doi: 10.1075/prag.10.4.01tom
- Troseth, G. L., Strouse, G. A., Verdine, B. N., and Saylor, M. M. (2018). Let's chat: On-screen social responsiveness is not sufficient to support toddlers' word learning from video. *Front. Psychol.* 9:2195. doi: 10.3389/fpsyg.2018.02195
- Tsuji, S., Jincho, N., Mazuka, R., and Cristia, A. (2020). Communicative cues in the absence of a human interaction partner enhance 12-month-old infants' word learning. *J. Exp. Child Psychol.* 191:104740. doi: 10.1016/j.jecp.2019.104740
- Turati, C., Valenza, E., Leo, I., and Simion, F. (2005). Three-month-olds' visual preference for faces and its underlying visual processing mechanisms. *J. Exp. Child Psychol.* 90, 255–273. doi: 10.1016/j.jecp.2004.11.001
- Valenza, E., Simion, F., Cassia, V. M., and Umiltà, C. (1996). Face preference at birth. *J. Exp. Psychol.* 22, 892–903. doi: 10.1037/0096-1523.22.4.892
- Viktorsson, C., Portugal, A. M., Li, D., Rudling, M., Sanchez, M. S., Tammimies, K., et al. (2021). Infants' attention to eyes is an independent, heritable trait predicting later verbal competence. *bioRxiv* [preprint]. doi: 10.1101/2021.10.23.465424
- Wang, Y., Seidl, A., and Cristia, A. (2021). Infant speech perception and cognitive skills as predictors of later vocabulary. *Infant Behav. Dev.* 62:101524. doi: 10.1016/j.infbeh.2020.101524
- Werker, J. F., and Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Lang. Learn. Dev.* 1, 197–234. doi: 10.1080/15475441.2005.9684216
- Young, G. S., Merin, N., Rogers, S. J., and 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. *Dev. Sci.* 12, 798–814. doi: 10.1111/j.1467-7687.2009.00833.x