



HAL
open science

On the Links Among Face Processing, Language Processing, and Narrowing during Development

Olivier Pascalis, H el ene Loevenbruck, Paul C. Quinn, Sonia Kandel, James W. Tanaka, Kang Lee

► **To cite this version:**

Olivier Pascalis, H el ene Loevenbruck, Paul C. Quinn, Sonia Kandel, James W. Tanaka, et al.. On the Links Among Face Processing, Language Processing, and Narrowing during Development. *Child Development Perspectives*, 2014, 8 (2), pp.65-70. 10.1111/cdep.12064 . hal-00964721

HAL Id: hal-00964721

<https://hal.science/hal-00964721>

Submitted on 21 Aug 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

On the Links Among Face Processing, Language Processing, and Narrowing During Development

Olivier Pascalis,¹ H el ene Loevenbruck,^{1,2} Paul C. Quinn,³ Sonia Kandel,¹
James W. Tanaka,⁴ and Kang Lee⁵

¹Universit e Grenoble Alpes, CNRS,

²Grenoble Images Parole Signal Automatique, CNRS,

³University of Delaware,

⁴University of Victoria,

⁵University of Toronto

In CHILD DEVELOPMENT PERSPECTIVES, Volume 8, Issue 2, June 2014, Pages:
65–70



On the Linkage between Face Processing, Language Processing, and Narrowing during Development

Journal:	<i>Child Development Perspectives</i>
Manuscript ID:	2013-032.R1
Wiley - Manuscript type:	Review (invited)
Keywords:	narrowing, face, speech
Abstract:	<p>From the beginning of life, face and language processing are crucial for establishing social communication. Studies on the development of face and language processing systems have yielded interesting similarities such as the observation of perceptual narrowing occurring across both domains. This article reviews several functions of human communication, and then describes how the tools used to accomplish those functions are modified by perceptual narrowing, concluding that narrowing is a characteristic common to all forms of social communication. We argue that during evolution, social communication has engaged different perceptual and cognitive systems--face, facial expression, gesture, vocalization, sound, and oral language--which have emerged at different times. These systems are interactive and linked to some extent. Narrowing can in this framework be viewed as a mechanism for infants to adapt to their native social group.</p>

On the Linkage between Face Processing, Language Processing,
and Narrowing during Development

For Review Only

ABSTRACT

From the beginning of life, face and language processing are crucial for establishing social communication. Studies on the development of face and language processing systems have yielded interesting similarities such as the observation of perceptual narrowing occurring across both domains. This article reviews several functions of human communication, and then describes how the tools used to accomplish those functions are modified by perceptual narrowing, concluding that narrowing is a characteristic common to all forms of social communication. We argue that during evolution, social communication has engaged different perceptual and cognitive systems--face, facial expression, gesture, vocalization, sound, and oral language--which have emerged at different times. These systems are interactive and linked to some extent. Narrowing can in this framework be viewed as a mechanism for infants to adapt to their native social group.

On the Linkage between Face Processing, Language Processing,
and Narrowing during Development

Social life requires relationships with other group members, acknowledgment of their status, and communication between individuals. Depending on the species studied, communication can be achieved through vocalization, language, faces and their expressions, or some combination.

Similarities observed across species may provide insights into the relationship between different social communication tools and networks. These observations lead to the hypothesis that the emergence of communicative tools happened during evolutionary time and our present systems reflect some aspects of this evolution.

In humans, faces and language are essential for communication, but they have been traditionally studied as separate areas with scarce interaction between the two domains even when their links are acknowledged. They have in some frameworks even been conceived as independent cognitive modules. If faces provide an early channel of communication prior to the comprehension of gestural or oral language, in newborns, postnatal exposure to the mother's voice-face combination is required for recognition of the mother's face (Sai, 2005). Coulon, Guellai, and Streri (2011) found that recognition of dynamic faces is observed only if sound was present. Face processing thus seems to be facilitated by voice processing, even at an early age.

Subsequently, beginning in early childhood, most conversations take place in a face-to-face context. Although auditory information alone is sufficient to understand speech, we systematically and unconsciously rely on visual information provided by a speaker's face. Seeing oro-facial gestures of the speaker accelerates core word recognition processes (Fort, Kandel, Chipot, Savariaux, Granjon, & Spinelli, 2012) and enhances intelligibility in noisy environments (Benoît, Mohamadi, & Kandel, 1994). Therefore, most human conversations--except when we

FACES, LANGUAGE, AND NARROWING

4

are on the phone--invoke analysis of facial configurations to locate relevant cues for speech decoding. The integration of audio and facial information is not only crucial to speech perception. There is also strong evidence for interference between facial and vocal information during affect recognition and identity processing (Campanella & Belin, 2007). These observations point to a close link between face and language processing which, we will argue, may reflect a fundamental characteristic of how social communication evolved and how it develops in infants and children. More specifically, we will point to functional links between gestural and oral communication in non-human primates as well as infants, which suggest that social communication is a multimodal system, involving manual and visuo-facial gestures as well as vocalization. We will argue that this multimodal system is gradually tuned during development, with narrowing manifesting itself in all the different modalities of communication.

FACE PROCESSING, LANGUAGE PROCESSING, AND DEVELOPMENT

Human adults are able to recognize familiar faces with ease and are said to be expert at processing faces. Faces form a category of stimuli that are homogenous in terms of the positioning of their internal elements and we have developed a signature way to discriminate them based on configural (i.e., relational) information, such as the particular distance between the eyes, or between lips and chin. Experience likely plays a critical role in the acquisition of this face expertise (Lee, Anzures, Quinn, Pascalis, & Slater, 2011).

Language represents a key tool for social communication since it allows for transmitting cognitively complex information that face expressions cannot. It is a complex cognitive skill requiring recursion and displacement (Chomsky, 1965), yet it is acquired by children swiftly and without instruction, whereas most adults find second language learning challenging. Studies of language acquisition have discovered crucial milestones: Vocalizations are observable at birth,

babbling emerges at around 6-8 months of age, children utter their first words at 10-12 months, and word combinations forming proto-sentences begin at around 20-24 months (Vihman, 1996).

Development within the face and language processing systems has been closely studied and interesting similarities can be observed. Nelson (2001) proposed that face processing develops during the first years of life from a broad non-specific system to a human-tuned face processor. Furthermore, he suggested that faces observed within the infants' visual environment shape and influence the developing face system via a process that has come to be known in the literature as perceptual narrowing: a progression whereby infants maintain ability to discriminate stimuli to which they are exposed, but lose ability to discriminate stimuli to which they are not exposed. This course of responsiveness is similar to that observed for language development. Throughout the first year of life, initial discriminatory ability reflective of a universal sensitivity to the sounds of all human languages 'narrows' as a consequence of predominant exposure to one's native language and a general scarcity of exposure to other languages (Werker & Tees, 1999). During this time window, infants become tuned to their native language and the distribution of phonetic information in the ambient language at the expense of discriminating non-native contrasts. In other words, infants become "experts" at processing frequently experienced faces and native sounds.

Scott, Pascalis, and Nelson (2007) reviewed the literature on narrowing and observed that it cuts across both visual and auditory modalities, possibly reflecting development of a common neural architecture. Lewkowicz and Ghazanfar (2009) suggest that narrowing could be a pan-sensory process, that is, the same phenomenon will be observed in various senses during the same period and will be part of the development of our multisensory representation of the world.

FACES, LANGUAGE, AND NARROWING

6

This line of thinking opens up such questions as: Is perceptual narrowing amodal? Is auditory narrowing linked to visual narrowing?

One argument for the link between development of face and language processing comes from neuroanatomy. The superior temporal sulcus (STS) is associated with face processing and auditory representation of speech components (Démonet, Thierry, & Cardebat, 2005; Haxby, Hoffman, & Gobbini, 2000). Belin et al. (2011) further suggest that the posterior part of the STS may be considered as an amodal ‘convergence’ zone which plays a key role in integrating face and voice information. These findings provide evidence for similar, interacting, and common brain circuits for face and speech processing

We argue that the way narrowing has been described so far in the literature fails to take into account the evolution and historical timing of when face and language processing emerged. Our point is that what drives or motivates the development of both face and language processing is the urge to communicate. To anticipate the remainder of the article, we first describe several functions of human communication, and then explain how each of those functions are modified by perceptual narrowing, concluding that narrowing is a common characteristic of all social communication.

GESTURAL AND ORAL COMMUNICATION

Human language is described as unique even if some form of communication does exist in other species. Understanding the emergence of language during evolution is a challenge, as fossil evidence does not provide much insight into oral language. Two communication means are seen as potential precursors to human language: vocal calls and gestures. It is still a highly debated issue whether language originated in manual gestures or evolved exclusively in the vocal domain. The former hypothesis considers pointing as the initial means to communicate which

FACES, LANGUAGE, AND NARROWING

7

later developed into a gestural “language”. Corballis (2003) claims that language evolved from manual gestures, gradually incorporating vocal elements, implying that language involves reciprocity in the actions of partners. The mechanism could be supported by mirror neurons, located in Broca’s area in humans (Buccino et al., 2001). This area is involved with vocalization as well as manual action and could first have been used as a neural substrate for interspecific communication, and subsequently for speech processing.

In addition, gestures, and more specifically pointing, appear to be tightly associated with language development (Kita, 2003). Ocular pointing (or deictic gaze, at 6 - 9 months) and later index finger pointing (deictic gesture, at 9 - 11 months) have been shown to be two key stages in cognitive development that are correlated with stages in speech development. Finger pointing is associated with lexicon construction and when accompanied with word production (16-20 months), fosters morphosyntax emergence. At later stages, children start using prosodic focus, i.e., vocal pointing (Ménard, Løevenbruck, & Savariaux, 2006), or constructions involving a deictic pronoun (Diessel & Tomasello, 2000). It has also been suggested that the different pointing modalities may share a common cerebral network: Results of an fMRI study of multimodal pointing show a common left parietal activation in ocular, digital, and prosodic pointing (Loevenbruck, Dohen, & Vilain, 2009). These findings reinforce the argument of a link between gesture and language.

The referential and combinatorial properties of primate vocal communication suggest, however, that language is also rooted in vocalization (Arnold & Zuberbuhler, 2008): Chimpanzees produce and understand functionally referential calls, and monkeys can combine existing calls into higher-order meaningful sequences. Furthermore, MacNeilage (1998) suggests that syllables may derive from cycles of mandibular oscillation involved in chewing, sucking,

FACES, LANGUAGE, AND NARROWING

8

and licking, which take on communicative significance as lip smacks, tongue smacks, and teeth chatters. It may therefore be more parsimonious to propose a direct evolutionary trajectory from primate vocalizations to human speech rather than a complex route requiring an intermediate stage of gestural communication.

Our view is that evidence of functional links between gestural and oral communication, observed in non-human primates as well as infants, suggests that communication is a multimodal system, involving manual and visuo-facial gestures as well as vocalization. Human communication may have switched to oral-dominant language for several reasons, including accessibility without seeing the other person (during the night or if far away) and accessibility while doing something else with the forelimbs such as carrying or using tools (Corballis, 2003). The oro-facial region would have become gradually more used than the hand in human communication.

It seems clear that different kinds of communication existed before oral language. Vocalizations, facial expressions, and visuo-facial gestures would have been part of it.

NARROWING ACROSS DOMAINS THAT INVOLVE SOCIAL COMMUNICATION**Faces**

Whereas 6-month-olds show recognition memory for different races of human faces as well as for different monkey faces, 9- to 10-month-old infants reliably discriminate only own-species faces and own-race faces (for a review, see Lee et al., 2011). Successful social communication relies on our ability to process indexical information (i.e., all the elements that allow for person identification) about the individuals we interact with, such as their identity, age, and sex.

Specialization for own-race faces improves indexical information extraction. Regarding voice recognition, Johnson, Westrek, Nazzi, and Cutler (2011) found that 7-month-olds only detected

FACES, LANGUAGE, AND NARROWING

9

voice changes if the language was in the native tongue. This outcome suggests that voice recognition develops in pace with increasing competence in language processing. However, younger infants' ability has yet not been reported and we cannot therefore conclude that narrowing has occurred in this domain.

In addition to face recognition, infants also develop facial expression recognition skills, which further feed into their social communication abilities (Quinn et al., 2011). Perceptual narrowing has been found for emotion recognition in 9-month-old infants, but only for own-race faces (Vogel, Monesson & Scott, 2012), suggesting that perceptual narrowing affects stimuli that are relevant to communication with our relatives.

Audiovisual speech

Lewkowicz and collaborators (Lewkowicz & Ghazanfar, 2009; Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009) have reported a decline in responsiveness to non-native audiovisual inputs by the end of the first year of life for other species sound-face matching and non-native language. Weikum and colleagues (2007) used silent video clips of a bilingual speaker telling a story in two languages and found that monolingual 4- and 6-month-olds were able to visually discriminate between the two languages, whereas monolingual 8-month-olds were not. The link between face and language processing is further illustrated by Lewkowicz and Hansen-Tift's (2012) study investigating infants' face scanning while infants watched and listened to a female speaking their native or a non-native language. The investigators reported developmental differences in the pattern of face scanning between 4- and 8-months of age regardless of the language spoken, but 12-month-olds behaved differently for the native and non-native language. The older infants looked the same amount of time at the eyes and mouth when native speech was spoken, whereas they looked longer at the mouth for the non-native language. The same pattern

FACES, LANGUAGE, AND NARROWING

10

of results was found by Kubicek et al. (2013) who showed a decrease of looking toward the mouth when the face spoke the native language and an increase when the non-native language was spoken.

Music-rhythm

Music is important for communication and may be involved in functions such as comforting, courtship, movement coordination, and social cohesion (Brown, 2003). It requires social skills such as vocal/gestural imitation, and involves cultural transmission. It may even be considered as an early form of oral communication that emerged before language (Fitch, 2006). If narrowing happens for any form of communication, it should also be found for music. Indeed, it has been found that 6-month-olds are able to discriminate rhythms that are specific to their culture and those unfamiliar to them; however, 12-month-olds could only do so with a rhythm specific to their own culture (Hannon & Trehub, 2005). Furthermore, early and active exposure to culture-specific music rhythms and tonalities may accelerate perceptual narrowing in music (Trainor, Marie, Gerry, Whiskin, & Unrau, 2012).

Auditory speech

Narrowing of speech perception is also well documented. Infants' speech perception becomes tuned toward their native language at around 10-12 months of age. Young infants discriminate fine phonetic differences, such as differences in voice onset time, between consonants such as /pa/ vs /ba/ (Eimas et al., 1971). Infants are also able to discriminate vowels (e.g., /a/ vs /i/ or /i/ vs /u/, Trehub, 1973). Crucially, infants younger than 6-8 months of age can not only categorically discriminate native phonetic contrasts, they can also discriminate those that fall outside their native language. English-learning 6- to 8-month-olds, for instance, can discriminate

the non-native dental/retroflex contrasts such as the Hindi /ʈa/ versus /ta/ (Werker & Tees, 1999). However, a decline in cross-language consonant perception is observed at 10-12 months of age.

Sign language

Narrowing has additionally been observed for perception of sign language (Palmer, Fais, Golinkoff, & Werker, 2012). Hearing infants are able to discriminate ASL signs at 4- but not at 14-months of age, whereas ASL-learning infants are still able to discriminate signs at the later age. This result supports the conclusion that narrowing happens for language regardless of the support of language: gesture versus oral.

NARROWING AS A CATEGORIZATION PROCESS SERVING SOCIAL NEEDS

Narrowing is thus observed for different cognitive abilities commonly involved in communication, even though not all empirical evidence uniformly shows that narrowing occurs simultaneously across different domains (see, for instance, Hayden, Bhatt, Kangas, Zieber, & Joseph, 2012, for evidence of own-race specialization several months before language narrowing). The underlying mechanism might then not be specific to one cognitive ability, but common to all communicative tools. In terms of evolution, it will have emerged first for face and facial expression processing. It should therefore have been part of rhythm and gesture primitive language before becoming part of oral language.

An unresolved question is that the concomitance of occurrence in multiple modalities does not explain why there is narrowing. To provide informed speculation on this question, we would observe that infants are born into a social group which has developed a culture of communication that is unique, opaque (i.e., association between an oral/gestural sign and a referent may be arbitrary), and subject to evolution. The best way to integrate within the group may be to adapt rapidly to the group's social habits and communication traditions. During the

FACES, LANGUAGE, AND NARROWING

12

first 12 months, when infants mainly interact with the mother/caregiver, they will have to rapidly learn the appropriate way of communicating when interacting within the social group. The mother/caregiver thus rapidly transmits the basic aspects of communication that are crucial to be part of the community: smiling, language characteristics, and recognition of specific faces.

The child then calibrates its various communication systems using various learning abilities including imitation. If the child is exposed to several individuals, he/she will use convergence mechanisms to calibrate the system and will end up with finely tuned representations of the faces present in the environment as well as detailed representations of the phonemes and prosodic patterns in the ambient language(s). By this account, narrowing is a categorization process that serves social needs. In the language domain, infants build a broad category inclusive of the non-native contrasts that are lost and retain tightly tuned categories for native contrasts. In the same way, in the face domain, infants build a large category for other-race faces inclusive of multiple other-race face categories (i.e., for infants exposed mainly to Caucasian faces, this category would include Asian and African faces) and build tightly tuned categories organized around subordinate-level identity information for same-race faces (i.e., Olivier vs. Paul vs. Helene). Narrowing can therefore be conceived of as a system that allows the infant to become more efficient or specialized for the social stimuli at hand in the close environment.

CONCLUSION

In the present article, we have argued that perceptual narrowing should be observed for all forms of social communication. During evolution, our social communication has used different perceptual and cognitive systems--face, facial expression, gesture, vocalization, sound, and oral language--which have emerged at different times. These systems are interactive in adults and

FACES, LANGUAGE, AND NARROWING

13

their neural mechanisms are linked to some extent. Their development presents similarities as infants adjust to their native social group.

We suggest that the adaptation is accomplished via a specific mechanism dedicated to social cognition, which encompasses the different modalities of communication, including manual and visuo-facial gesture processing as well as vocalization processing abilities, although we are uncommitted to whether such a mechanism is present at birth or is a product of development. Future behavioral and neuroimaging studies could bring further evidence for the intertwining of the development of these social abilities. Our suggestion also has predictions in the field of neurological or developmental disorders. We predict that deficits in either the development of manual gesture processing, facial gesture processing, or vocalization processing should result in disorders of social communication. This prediction is supported by work on autism spectrum disorders suggesting that social communication strongly relies on the healthy development of these different abilities (Adolphs, Sears, & Piven, 2001; Baron-Cohen, 1989).

REFERENCES

1. Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism *Journal of Cognitive Neuroscience*, *13*, 232-240
2. Arnold, K., & Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. *Current Biology*, *18*, R202-R203.
3. Baron-Cohen, S. (1989). Perceptual role taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology*, *7*, 113-127
4. Belin, P., Bestelmeyer, P., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, *102*, 711-725.
5. Benoît, C., Mohamadi, T., & Kandel, S. (1994). Effects of phonetic context on audio-visual intelligibility of French. *Journal of Speech, Language and Hearing Research*, *37*, 1195-1203.
6. Brown, S. (2003). Biomusicology, and three biological paradoxes about music. *Bulletin of Psychology and the Arts*, *4*, 15-17.
7. Buccino, G., Binkofski, F., Fink, G., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R., Zilles, K., Rizzolatti, G., & Freund, H. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study *European Journal of Neuroscience*, *13*, 400-404
8. Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*, 535-543.
9. Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
10. Corballis, M. C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences*, *26*, 199-260.

11. Coulon, M., Guellai, B., & Streri, A. (2011). Recognition of unfamiliar talking faces at birth. *International Journal of Behavioral Development, 35*, 282-287.
12. Démonet, J., Thierry, G., & Cardebat, D. (2005). Renewal of the neurophysiology of language: functional neuroimaging. *Physiological Reviews, 85*, 49-95.
13. Diessel, H., & Tomasello, M. (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguistics, 11*, 131-152.
14. Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303-306.
15. Fitch, W. (2006). The biology and evolution of music: A comparative perspective. *Cognition, 100*, 173-215.
16. Fort, M., Kandel, S., Chipot, J., Savariaux, C., Granjon, L., & Spinelli, E. (2012). Seeing the initial articulatory gestures of a word triggers lexical access. *Language and Cognitive Processes, 28*, 1207-1223.
17. Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences (USA), 102*, 12639-12643.
18. Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233.
19. Hayden, A., Bhatt, R. S., Kangas, A., Zieber, N., & Joseph, J. E. (2012). Race-based perceptual asymmetry in face processing is evident early in life. *Infancy, 17*, 578-590.
20. Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14*, 1002-1011.

21. Kita, S. (Ed.) (1983). *Pointing: Where language, culture and cognition meet*. Mahwah, NJ: Lawrence Erlbaum Associates.
22. Kubicek, C., Hillairet de Boisferon, A., Dupierrix, E., Løevenbruck, H., Gervain, J., & Schwarzer, G. (2013). Face-scanning behavior to silently talking faces in 12-month-old infants: The role of pre-exposed auditory speech. *International Journal of Behavioral Development, 37*, 77-78.
23. Lee, K., Anzures, G., Quinn, P. C., Pascalis, O., & Slater, A. (2011). Development of face processing expertise. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 753-778). New York: Oxford University Press.
24. Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences, 13*, 470-478.
25. Lewkowicz, D. J., & Hansen-Tift, A. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences (USA), 109*, 1431-1436.
26. Løevenbruck H., Dohen M., & Vilain C. (2009). Pointing is 'special'. In S. Fuchs, H. Løevenbruck, D. Pape, & P. Perrier (Eds.), *Some aspects of speech and the brain* (pp. 211-258). Peter Lang: Berlin.
27. MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences, 21*, 499-511.
28. Ménard, L., Løevenbruck, H., & Savariaux, C. (2006). Articulatory and acoustic correlates of contrastive focus in French: a developmental study. In J. Harrington & M.

Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* (pp. 227-251). Psychology Press: New York.

29. Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development, 10*, 3-18.
30. Palmer, S. B., Fais, L., Golinkoff, R. M., & Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the first year of life. *Child Development, 83*, 543-553
31. 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 USA, 106*, 10598-10602.
32. Quinn, P. C., Anzures, G., Izard, C. E., Lee, K., Pascalis, O., Slater, A. M., & Tanaka, J. W. (2011). Looking across domains to understand infant representation of emotion. *Emotion Review, 3*, 197-206.
33. Sai, F. Z. (2005). The role of the mother's voice in developing mother's face preference: evidence for intermodal perception at birth. *Infant and Child Development, 14*, 29-50
34. Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science, 16*, 197-201.
35. Trainor, L. J., Marie, C., Gerry, D., Whiskin, E., & Unrau, A. (2012). Becoming musically enculturated: effects of music classes for infants on brain and behavior. *Annals of the New York Academy of Sciences, 1252*, 129-138.
36. Trehub, S. E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology, 9*, 91-96.

37. Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Oxford: Blackwell.
38. Vogel, M., Monesson, A., & Scott, L. S. (2012). Building biases in infancy: the influence of race on face and voice emotion matching. *Developmental Science, 15*, 359-372.
39. Weikum, W., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science, 316*, 1159.
40. Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology, 50*, 509-535.