



Published in final edited form as:

Trends Cogn Sci. 2014 December ; 18(12): 642–646. doi:10.1016/j.tics.2014.10.001.

Listen up! Speech is for thinking during infancy

Athena Vouloumanos and

Department of Psychology, New York University, 6 Washington Place, New York, NY, 10003-6603, USA

Sandra R. Waxman

Department of Psychology, Northwestern University, 2029 Sheridan Rd, Evanston, IL 60208-2710, USA

Abstract

Infants' exposure to human speech within the first year of life promotes more than speech processing and language acquisition: new developmental evidence suggests that listening to speech shapes infants' fundamental cognitive and social capacities. Speech streamlines infants' learning, promotes the formation of object categories, signals communicative partners, highlights information in social interactions, and offers insight into the minds of others. These results, which challenge the claim that for infants, speech offers no special cognitive advantages, suggests a new synthesis: Far earlier than researchers had imagined, an intimate and powerful connection between human speech and cognition guides infant development, advancing infants' acquisition of fundamental psychological processes.

Speech is not just for language (even for infants)

Infants' rapid progress in speech perception stands as a clarion case of our species' natural proclivity to learn language. Until recently, infant speech perception was considered primarily a foundation upon which to build language. Research focused on the rapidity with which infants tune to the sounds of their native language [1,2] and use these as building blocks for the acquisition of phonology, syntax, and meaning. But infants' natural affinity for processing the speech signal has implications that reach far beyond the acquisition of language. New evidence now shows that from the first months of life, listening to speech is a powerful engine: it promotes the acquisition of fundamental psychological processes including pattern learning, the formation of object categories, the identification of communicative partners, knowledge acquisition within social interactions, and the development of social cognition.

© 2014 Elsevier Ltd. All rights reserved.

Corresponding author: Vouloumanos, A. (athena.vouloumanos@nyu.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Human speech is a privileged signal from birth

From birth, speech is a privileged signal for humans. Newborns prefer the vocalizations of humans and non-human primates (Rhesus macaques: *Macaca mulatta*) to other sounds [3,4]. By 3 months, they tune in specifically to human speech even favoring human speech over other human vocalizations, including emotional (e.g., laughing) and physiological (e.g., sneezing) vocalizations [3,5] (see Box 1). Interestingly, 3-month-olds' preference for speech is broad enough to include native as well as non-native speech sounds. This suggests that infants privilege the speech signal itself—and not simply the familiar sounds of their own native language.

BOX 1

Tuning mechanisms as pervasive developmental processes

The tuning of infants' speech bias between birth and 3 months, from preferring vocalizations including those of other primates to being speech-specific [3], mirrors similar tuning processes at work in face perception, cross-modal speech perception, and phoneme perception [1,2,43–45]. Infants are initially able to recognize faces of individuals from different species but by 9 months and into adulthood show better recognition of human faces compared to the faces of other species [43]. Similarly, infants' ability to discriminate between many different phonemes may initially rely on language-general discrimination abilities, which become language-specific by 6–12 months [1,2]. Tuning mechanisms sharpen initially broad biases into more specific ones across many perceptual domains in infants' first year of life.

These behavioral preferences converge well with neural evidence: At one month of age, human speech and rhesus calls activate similar neural areas, but by 3 months speech and rhesus calls elicit distinctly different neural responses [6,7]. The developmental change in patterns of activation likely reflects neural specialization. Specifically, 1-month-olds' response to human speech is already localized to the left hemisphere; over the next few months, the left hemisphere maintains its activation to speech, but becomes less responsive to non-speech sounds [6]. This developmental pattern suggests that from birth, listening to speech sounds preferentially activates specific areas of the temporal cortex, and that a pruning process underlies further neural specialization for speech in the left hemisphere [8].

Infants' rapid behavioral and neural tuning to the signal of human speech, remarkable in its own right, has powerful developmental consequences that extend beyond their listening preferences alone. Infants' preference for listening to human speech shapes how infants learn.

Listening to speech facilitates learning and pattern extraction

Speech is a privileged unit for even the most basic forms of learning, including low-level conditioned responses. From birth, when infants listen to speech, they successfully recognize individual units and their relative positions in the speech sequence [9]. And at 1

month, infants who are conditioned to speech show a stronger response and a steeper learning curve than infants conditioned to either tones or backward speech [10].

By 7 months, speech promotes more sophisticated forms of learning, including the detection of rules and patterns. After hearing only 2 minutes of patterned speech syllable sequences (ABB: la-ga-ga, da-li-li), 7-month-olds extract and generalize rules such as identity and sequential positioning and distinguish ABB (la-ga-ga) from ABA (la-ga-la) [11]. But after 2 minutes of patterned exposure to non-speech sounds (musical tones, animal sounds, timbres), infants do not extract the equivalent ABB or ABA rules. Within the auditory domain, infants can generalize rules to non-speech sounds only if they first hear those rules instantiated in speech [12]. This asymmetry, favoring infants' ability to extract patterns in speech over non-speech sounds, suggests that infants learn better with speech.

Listening to speech promotes categorization

Infants' early preference for speech is powerful. But infants' preferences cannot tell us whether (or when) infants begin to link speech to the objects and events around them. A series of experiments designed to tackle this question focused on object categorization—a building block of cognition [13,14]. In these experiments, infants ranging in age from 3 to 12 months viewed several images from one object-category (e.g., dinosaurs), each accompanied by either a segment of speech or a sequence of sine-wave tones. Next, infants viewed two test images, one from the now-familiar category (a new dinosaur) and one from a novel category (e.g., a fish). If infants formed the object category (here, dinosaurs), they should distinguish between the test images [15]. By 3 months of age, infants listening to speech successfully formed categories; those listening to tones failed to form object categories at any age [14].

Thus, infants are tuned not only to speech but also to a principled and surprisingly early link between speech and the fundamental cognitive process of categorization. Moreover, this link, evident at three months, derives from a broader template that initially encompasses human speech as well as the calls of non-human primates (Madagascar blue-eyed lemurs: *Eulemur macaco flavifrons*). Three- and 4-month-old infants' categorization in the context of hearing lemur calls mirrors precisely their categorization in response to human speech; by 6 months, the link to categorization has become tuned specifically to human vocalizations [13]. This documents a surprisingly early link between human language and core cognitive processes, including object categorization that cannot be attributed to familiarity. Although 3- and 4-month-olds have considerable exposure to speech and none to lemur vocalizations, both signals confer the same cognitive advantage for categorization.

Speech helps identify potential communicative partners

To convey meaning, human communicative partners must integrate, encode, and decode linguistic symbols instantiated in speech, paralinguistic cues (like vocal pitch or intonation), and gestures (see Box 2). The speech signal itself can help identify a potential communicative partner. From the first months of life infants treat people and objects as different kinds of entities: they respond differently to people (with more smiling and emotional sounds) and objects (with more grasping) [16–19] and at 6 months, they also

expect others to also treat people and objects differently [20]. By 5 months, infants use human speech to identify potential conversational partners. When presented with human and monkey faces, 5-month-olds match speech (native or non-native) to human faces and monkey calls to monkey faces, but they do not match other human emotional vocalizations (e.g., laughter) specifically to humans [21].

BOX 2

Are the facilitative effects of speech specific to spoken language?

Might language produced in other modalities, including vision, also confer cognitive advantages in infancy? From birth, infants are prepared to acquire language in either the auditory or visual modality [46]. Both signed language and gesture confer cognitive and social advantages [47,48]. Although there is less work documenting the effects of signed than spoken language in infancy, infants privilege sign language over gestures. At 6 months, naïve hearing infants prefer to look at a person producing sign language, as compared to a person producing gesture [49], and by 7 months infants begin to extract some rules from sequences of sign language [50]. Still, hearing infants' ability to extract rules is less robust when they are presented with sign language than spoken language, which may reflect their experience. Although 9-month-olds already understand gestures such as pointing as being communicative [[NO STYLE for: Krehm]] and a possible precursor to language [52], the communicative function of signed languages might be understood even earlier.

Do hearing infants initially link visually-produced language to object categories like they do for vocalizations? Although even hearing infants prefer sign to gesture, this preference does not tell us which, if either, they will link to core cognitive capacities. Do they link sign language (but not gesture) to fundamental cognitive and social capacities? How do these links fare over the first year in hearing infants who are not exposed to language in the visual-motor domain?

Infants may thus already expect that humans, but not other animals, are the source of speech (see Box 3). This expectation for human speech (but not emotional vocalizations) suggests that infants are guided by more than their familiarity with the sounds alone. By 6 months, infants are especially attentive to communicative cues including eye gaze and speech produced by their pedagogical partners and use these cues to guide learning [22,23]. Infants appear to use speech to identify the natural class of individuals with whom they can communicate and from whom they can learn.

BOX 3

Can infants use non-linguistic stimuli like they use speech?

A hallmark of speech perception in adults is our ability to perceive distorted or atypical speech as speech. Like adults, infants can also perceive atypical signals as speech, but only under certain circumstances. Nine-month-olds who heard speech-like vocalizations produced by a parrot (which maintain some but not all of the acoustic features of speech) successfully treated the parrot vocalizations similarly to human speech, but only if they

viewed a (static) human face while listening. If they viewed a (static) checkerboard pattern, 9-month-olds treated the parrot vocalizations like non-speech [53]. One question currently under investigation is whether infants would link a parrot's speech-like vocalizations to object categorization or to any other cognitive and social capacities.

Another hallmark of being human is our capacity to infuse communicative status into a host of non-linguistic signals (e.g., Morse code). Infants, too, have this flexibility [54]. Like adults, under certain circumstances, infants will interpret an otherwise inert signal as communicative. Six-month-old infants participated in a categorization task involving sine-wave tone sequences, a signal that fails to promote infant object categorization [14,55,56]. But first, before the categorization task, infants watched a 2-minute videotaped conversation in which one person spoke and the other responded with 'beeps' in sine-wave tones. Embedding the tones within a rich communicative episode convinced infants that the tones had communicative status; tones now supported infants' object categorization. Although infants privilege speech, they can flexibly extend some of its most important communicative and cognitive functions to other initially non-privileged signals.

Speech indexes the transfer of information

When listening to a conversation in a foreign language, even if we cannot understand the meaning of a single word, we nonetheless infer that information is being conveyed. Thus, for adults, understanding the communicative *function* of speech does not require understanding the *contents* of the speech. Infants show a similar understanding. By 6 months, although infants understand only very few words [24], they are already sensitive to the communicative function of speech and appreciate that speech is a powerful conduit through which people share information. When an actor can no longer reach a target object, infants at 6 and 12 months infer that she can still obtain that target object from a second actor by using speech but not coughing and other non-speech vocalization [25,26]. Inferring that speech allows people to transfer information may allow infants to more easily deduce the focus of a person's attention, and to make inferences about what information they intend to share. This early understanding of the communicative function of speech may provide a mechanism for acquiring language and knowledge about the world. Speech is a conduit for moving information between people and a cue that information is being shared.

Speech gives insight into others' minds

Understanding the goals and intentions of others is one of the most complex problems facing infants. How do infants come to gain insight into the minds of others? The foundations of social cognition begin to take shape in the first year of life [27]. By the end of their first year, infants appreciate that people (and other agents) have intentions [28] and they distinguish between agents who can behave intentionally and non-agents, who can't [29–34].

By 12 months, infants use speech to learn about aspects of the world that are beyond their direct perception, including the minds of others [35]. Twelve-month-olds watched as an

actor attempted (but failed) to stack a ring on a funnel. If the actor then spoke to a new actor (who had not observed the failed attempts), infants expected the second actor to stack the ring. But if the actor produced non-speech sounds (e.g., coughs), infants had no such expectation. Infants appreciate that speech (but not non-speech) permits us to share our internal mental states, desires and beliefs. They expect that speech is a powerful vehicle for communicating our intentions and understanding the intentions of others.

At this age, infants also begin to forge more precise expectations about the functions of human language: They discover that different kinds of words refer to objects, events, and categories [36]. This more precise set of expectation permits infants to make more precise inferences about speakers' intentions. The advantage that speech has on categorization in 3–6 month-olds becomes far more precise: by 12 months, infants expect words that are presented in naming phrases (“Look at the *blick*”) to refer to objects and object categories, but have no such expectation for words presented alone (“Wow”) or for speech that does not involve naming (“Oh, look!”) [37,38]. Moreover, they expect novel nouns to refer to objects and object categories but not to surface properties (e.g., color or pattern) [39]. And by 14 months, infants expect that novel words also refer to actions and events. Although infants at this age tend not to imitate an adult experimenter's unconventional action (e.g., using her forehead—rather than her hand—to turn on a light; [40], if the unconventional action is named (“I'm going to *blick* the light!”), infants imitate it spontaneously [41]. As infants' expectations about the different functions of language become more precise, so too do the ways in which listening to speech comes to shape cognition.

Intriguing new evidence suggests that individual differences in infants' preferences for speech may even be linked to differences in their acquisition of fundamental social cognitive capacities. Infants who exhibit reduced preferences for human speech at 12 months display more autistic-like behaviors at 18 months [42]. Inasmuch as autistic traits include social communicative deficits beyond simple language difficulties (DSM-5), this suggests a potent link between simple speech biases and complex social communicative behaviors.

Conclusions

Before infants begin talking, they are listening. We have proposed that even before infants can understand the meaning of the speech that surrounds them, listening to speech transforms infants' acquisition of core cognitive capacities. This transformation is unlikely to be explained by appealing to low-level perceptual effects or issues of stimulus familiarity. Instead, what begins as a natural preference for listening to speech actually provides infants with a powerful natural mechanism for learning rapidly about the objects, events, and people that populate their world.

Acknowledgments

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under Award Numbers R01HD072018 (AV) and R01HD30410 (SRW), and National Science Foundation BCS 0950376 (SRW).

References

1. Kuhl PK, et al. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*. 1992; 255:606–608. [PubMed: 1736364]
2. Werker JF, Tees RC. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behav Dev*. 1984; 7:49–63.
3. Vouloumanos A, et al. The tuning of human neonates' preference for speech. *Child Dev*. 2010; 81:517–527. [PubMed: 20438457]
4. Vouloumanos A, Werker JF. Listening to language at birth: evidence for a bias for speech in neonates. *Dev Sci*. 2007; 10:159–164. [PubMed: 17286838]
5. Shultz S, Vouloumanos A. Three-month olds prefer speech to other naturally occurring signals. *Lang Learn Dev*. 2010; 6:241–257.
6. Shultz S, et al. Neural specialization for speech in the first months of life. *Dev Sci*. 2014
7. Minagawa-Kawai Y, et al. Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb Cortex*. 2011; 21:254–261. [PubMed: 20497946]
8. Huttenlocher PR. Dendritic and synaptic development in human cerebral cortex: time course and critical periods. *Dev Neuropsychol*. 1999; 16:347–349.
9. Gervain J, et al. Binding at birth: the newborn brain detects identity relations and sequential position in speech. *J Cogn Neurosci*. 2012; 24:564–574. [PubMed: 22066581]
10. Reeb-Sutherland BC, et al. One-month-old human infants learn about the social world while they sleep. *Dev Sci*. 2011; 14:1134–1141. [PubMed: 21884328]
11. Marcus GF, et al. Rule learning by seven-month-old infants. *Science*. 1999; 283:77–80. [PubMed: 9872745]
12. Marcus GF, et al. Infant rule learning facilitated by speech. *Psychol Sci*. 2007; 18:387–391. [PubMed: 17576276]
13. Ferry AL, et al. Nonhuman primate vocalizations support categorization in very young human infants. *Proc Natl Acad Sci USA*. 2013; 110:15231–15235. [PubMed: 24003164]
14. Ferry AL, et al. Categorization in 3-and 4-month-old infants: an advantage of words over tones. *Child Dev*. 2010; 81:472–479. [PubMed: 20438453]
15. Aslin RN. What's in a look? *Dev Sci*. 2007; 10:48–53. [PubMed: 17181699]
16. Rönqvist L, von Hofsten C. Neonatal finger and arm movements as determined by a social and an object context. *Early Dev Parent*. 1994; 3:81–94.
17. Legerstee M. Patterns of 4-month-old infant responses to hidden silent and sounding people and objects. *Early Dev Parent*. 1994; 3
18. Legerstee M. Changes in the quality of infant sounds as a function of social and nonsocial stimulation. *First Lang*. 1991; 11:327–343.
19. Legerstee M. The role of person and object in eliciting early imitation. *J Exp Child Psychol*. 1991; 51:423–433. [PubMed: 2072085]
20. Legerstee M, et al. Precursors to the development of intention at 6 months: understanding people and their actions. *Dev Psychol*. 2000; 36:627–634. [PubMed: 10976602]
21. Vouloumanos A, et al. Five-month-old infants' identification of the sources of vocalizations. *Proc Natl Acad Sci USA*. 2009; 106:18867–18872. [PubMed: 19846770]
22. Csibra G, Gergely G. Natural pedagogy as evolutionary adaptation. *Philos Trans R Soc Lond B Biol Sci*. 2011; 366:1149–1157. [PubMed: 21357237]
23. Csibra G, Gergely G. Natural pedagogy. *Trends Cogn Sci*. 2009; 13:148–153. [PubMed: 19285912]
24. Bergelson E, Swingley D. At 6–9 months, human infants know the meanings of many common nouns. *Proc Natl Acad Sci USA*. 2012; 109:3253–3258. [PubMed: 22331874]
25. Vouloumanos A, et al. Do 6-month-olds understand that speech can communicate? *Dev Sci*. 2014
26. Martin A, et al. Understanding the abstract role of speech in communication at 12 months. *Cognition*. 2012; 123:50–60. [PubMed: 22209584]

27. Baillargeon, R., et al. Psychological and Sociomoral Reasoning in Infancy. In: Borgida, E.; Bargh, J., editors. *APA Handbook of Personality and Social Psychology: Vol.1. Attitudes and Social Cognition*. Washington DC: APA;
28. Woodward AL. Infants selectively encode the goal object of an actor's reach. *Cognition*. 1998; 69:1–34. [PubMed: 9871370]
29. Király I, et al. The early origins of goal attribution in infancy. *Conscious Cogn*. 2003; 12:752–769. [PubMed: 14656515]
30. Luo Y, Baillargeon R. Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychol Sci*. 2005; 16:601–608. [PubMed: 16102062]
31. Johnson S, et al. Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Dev Sci*. 1998; 1:233–238.
32. Gergely G, et al. Taking the intentional stance at 12 months of age. *Cognition*. 1995; 56:165–193. [PubMed: 7554793]
33. Csibra G. Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*. 2008; 107:705–717. [PubMed: 17869235]
34. Bíró S, Leslie AM. Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Dev Sci*. 2007; 10:379–398. [PubMed: 17444978]
35. Vouloumanos A, et al. Twelve-month-old infants recognize that speech can communicate unobservable intentions. *Proc Natl Acad Sci USA*. 2012; 109:12933–12937. [PubMed: 22826217]
36. Waxman SR, Gelman SA. Early word-learning entails reference, not merely associations. *Trends Cogn Sci*. 2009; 13:258–263. [PubMed: 19447670]
37. Fennell CT, Waxman SR. What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Dev*. 2010; 81:1376–1383. [PubMed: 20840228]
38. Waxman SR, Markow DB. Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cogn Psychol*. 1995; 29:257–302. [PubMed: 8556847]
39. Waxman SR. Specifying the scope of 13-month-olds' expectations for novel words. *Cognition*. 1999; 70:B35–B50. [PubMed: 10384739]
40. Gergely G, et al. Rational imitation in preverbal infants. *Nature*. 2002; 415:755. [PubMed: 11845198]
41. Chen ML, Waxman SR. "Shall we blick?": novel words highlight actors' underlying intentions for 14-month-old infants. *Dev Psychol*. 2013; 49:426–431. [PubMed: 22822935]
42. Curtin S, Vouloumanos A. Speech Preference is Associated with Autistic-Like Behavior in 18-Months-Olds at Risk for Autism Spectrum Disorder. *J Autism Dev Disord*. 2013; 43:2114–2120. [PubMed: 23334808]
43. Pascalis O, et al. Is face processing species-specific during the first year of life? *Science*. 2002; 296:1321–1323. [PubMed: 12016317]
44. Lewkowicz DJ, Ghazanfar AA. The decline of cross-species intersensory perception in human infants. *Proc Natl Acad Sci USA*. 2006; 103:6771–6774. [PubMed: 16618919]
45. Pons F, et al. Narrowing of intersensory speech perception in infancy. *Proc Natl Acad Sci USA*. 2009; 106:10598–10602. [PubMed: 19541648]
46. Senghas A, et al. Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science*. 2004; 305:1779–1782. [PubMed: 15375269]
47. Goldin-Meadow S. Talking and thinking with our hands. *Curr Dir Psychol Sci*. 2006; 15:34–39.
48. Pyers J, Senghas A. Language Promotes False-Belief Understanding: Evidence From Learners of a New Sign Language. *Psychol Sci*. 2009
49. Krentz UC, Corina DP. Preference for language in early infancy: the human language bias is not speech specific. *Dev Sci*. 2008; 11:1–9. [PubMed: 18171360]
50. Rabagliati H, et al. Infant rule learning: advantage language, or advantage speech? *PLoS One*. 2012; 7:e40517. [PubMed: 22815756]
51. Krehm M, et al. I See Your Point: Infants Under 12 Months Understand That Pointing Is Communicative. *J Cogn Dev*. 2014; 15:527–538.
52. Goldin-Meadow S. Pointing sets the stage for learning language--and creating language. *Child Dev*. 2007; 78:741–745. [PubMed: 17517001]

53. Vouloumanos A, Gelfand HM. Infant perception of atypical speech signals. *Dev Psychol.* 2012; 49:815–824. [PubMed: 22709131]
54. Ferguson, B.; Waxman, S. Communication and categorization: New insights into the relation between speech, labels, and concepts for infants. In: Knauff, M.; Pauen, M.; Sebanz, N.; Wachsmuth, I., editors. *Proceedings of the 35th Annual Conference of the Cognitive Science Society.* 2014.
55. Fulkerson AL, Waxman SR. Words (but not tones) facilitate object categorization: evidence from 6- and 12-month-olds. *Cognition.* 2007; 105:218–228. [PubMed: 17064677]
56. Balaban MT, Waxman SR. Do words facilitate object categorization in 9-month-old infants? *J Exp Child Psychol.* 1997; 64:3–26. [PubMed: 9126625]

BOX 4**Outstanding questions**

- What is the range of fundamental cognitive and social processes that are facilitated by speech? Are there processes that are not facilitated by speech?
- What is the range of signals that promote infant cognitive and social development?
 - Does sign language, like spoken language, facilitate infant cognitive and social development?
 - Can atypical speech signals facilitate infant cognitive development?
- What are the mechanisms underlying the cognitive and social advantages conferred by speech?
- How might this new evidence from typically developing infants help design interventions for infants and young children experiencing delays and disorders in language, cognitive, and social development?

Highlights

- Until recently, infant speech perception was considered primarily a foundation for building language.
- New developmental evidence suggests that listening to speech shapes infants' fundamental cognitive and social capacities including learning, categorization, communication, and understanding of other minds.
- There is an early intimate and powerful connection between human speech and cognition that guides infants' acquisition of fundamental psychological processes.