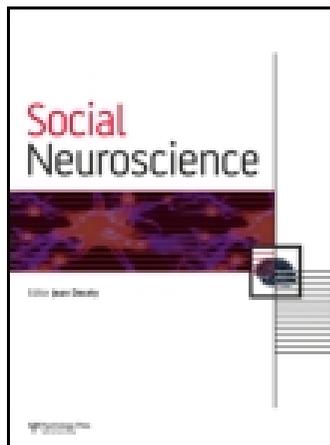


This article was downloaded by: [MPI Max-Planck-Institut fuer Kognitions- und Neurowissenschaften]

On: 19 August 2015, At: 02:00

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London, SW1P 1WG



Social Neuroscience

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/psns20>

When during development do our brains get tuned to the human voice?

Tobias Grossmann^a & Angela D. Friederici^a

^a Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Published online: 24 Oct 2011.

To cite this article: Tobias Grossmann & Angela D. Friederici (2012) When during development do our brains get tuned to the human voice?, *Social Neuroscience*, 7:4, 369-372, DOI: [10.1080/17470919.2011.628758](https://doi.org/10.1080/17470919.2011.628758)

To link to this article: <http://dx.doi.org/10.1080/17470919.2011.628758>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

When during development do our brains get tuned to the human voice?

Tobias Grossmann and Angela D. Friederici

Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

In the previous issue of *Social Neuroscience*, Lloyd-Fox and colleagues (2012) provide evidence that voice-sensitivity in temporal cortex emerges between 4 and 7 months of age. We discuss the implications of these findings and the overall progress that has been made in understanding the development of voice processing in infancy. In this commentary, we also examine important methodological and theoretical issues raised by this new work in the emerging field of developmental social neuroscience.

Keywords: Voice; Temporal cortex; fNIRS; fMRI.

The human voice is one of the most important stimuli in our auditory environment: It carries the acoustical signature of our species and is a rich source of information that is critical for survival and social communication. In adults, human voices are processed in highly specialized auditory brain regions located in the upper bank of the superior temporal sulcus (STS) bilaterally (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000). Other primates for whom vocal communication also plays an important role, such as macaque monkeys, have been shown to possess a voice region in the superior temporal plane that preferentially responds to conspecific vocalizations, suggesting that recognizing the sound that is the vocalization of a species member is an evolutionarily conserved brain function in primates that is independent of language (Petkov et al., 2008; Petkov, Logothetis, & Obleser, 2009). These voice-selective areas in auditory cortex, similar to face-selective areas in visual cortex identified in both humans and monkeys (Kanwisher, McDermott, & Chun, 1997; Tsao, Freiwald, Tootell, & Livingstone, 2006), are thought to serve important functions by binding the processing of crucial, socially relevant information to sensory systems. However, only recently has research begun to examine the ontogenetic origins of this fundamental brain process.

In a first study, by using functional near-infrared spectroscopy (fNIRS) to measure localized brain responses in two groups of infants, Grossmann, Oberecker, Koch, and Friederici (2010) showed that 7-month-olds, but not 4-month-olds, demonstrate increased responses in left and right superior temporal cortex to the human voice compared to non-vocal sounds, suggesting that voice-sensitive brain systems emerge between 4 and 7 months of age. However, this developmental conclusion was challenged by Blasi and colleagues' (2011) recent study of a group of infants ranging from 3 to 7 months of age, using functional magnetic resonance imaging (fMRI), in which increased activation in a right anterior temporal region in response to vocal sounds compared to environmental sounds was reported. Since no effect of age on this voice-sensitive region was observed, the data were taken to indicate that infants' brains are specialized in processing voices earlier than previously thought.

However, in this issue of *Social Neuroscience*, Lloyd-Fox and colleagues (2012), using fNIRS, show that voice sensitivity in the right and left temporal cortex increases from 4 to 7 months of age. These findings are consistent with the earlier findings from Grossmann et al.'s study (2010) and thus support the notion that the brain specialization for processing

Correspondence should be addressed to: Tobias Grossmann, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany. E-mail: grossman@cbs.mpg.de

voices emerges over the first 6 months of age. It is important to note that the similarity in the developmental effect observed in the two fNIRS studies is independent of the particular stimuli used. While Grossmann et al. (2010) used speech and non-speech stimuli in the voice condition of their study, Lloyd-Fox et al. (2012) presented the infants only with non-speech vocalizations such as coughing, yawning, and throat clearing. Thus, there are now two studies (Grossmann et al., 2010; Lloyd-Fox et al., 2012) that report the emergence of voice sensitivity during infant development but one fMRI study that did not find such a developmental effect (Blasi et al., 2011). Let us now consider the factors that might explain these differences.

NEUROIMAGING METHODS

One difference that stands out is the different neuroimaging methods that were used across these infant studies, fNIRS in Grossmann et al. (2010) and Lloyd-Fox et al. (2012) versus fMRI in Blasi et al. (2011). Both methods rely on drawing inferences about cortical activation by measuring a hemodynamic response; however, fNIRS has an inferior spatial resolution when compared to fMRI. Thus, fMRI, due to its better spatial resolution, may have been able to detect voice-sensitive responding in a region that was not measured in the fNIRS study by Grossmann et al. (2010). Nonetheless, these differences in spatial resolution alone are unlikely to explain the differences in terms of the development of voice-sensitive responding described above, because, if anything, they should have produced the opposite pattern of results—that is, the absence of a developmental effect in the fNIRS studies could have been attributed to its inferior spatial resolution, but the findings demonstrate the opposite pattern with the absence of a developmental effect in voice sensitivity in the fMRI study.

STIMULUS MATERIAL

Another important factor to consider is the potential difference in the stimulus material presented to the infants. However, Lloyd-Fox et al. (2012) and Blasi et al. (2011) actually used very similar stimuli, making stimulus selection an unlikely candidate in explaining the presence or absence of developmental effects across studies. Stimulus selection may, however, have affected the location of the voice-sensitive regions identified in the different studies. Specifically, both Lloyd-Fox et al. (2012) and Blasi et al. (2011) used stimuli consisting exclusively of non-speech

vocalizations and found a voice-sensitive region in a more anterior portion of the temporal cortex, whereas Grossmann et al. (2010), who included speech stimuli in their voice condition, found voice-sensitive responses in slightly more posterior regions of the temporal cortex. This raises the interesting hypothesis that in early development, speech (linguistic information) versus non-speech human sounds (non-linguistic information) might at least partly be processed in different regions of temporal cortex. Indeed, it has been argued that, in the adult brain, different aspects of voices, in particular non-linguistic aspects such as speaker identity versus linguistic information derived from speech, may be processed in different but adjacent and functionally associated regions in the temporal cortex (Belin & Zatorre, 2003; Kriegstein & Giraud, 2004; von Kriegstein, Eger, Kleinschmidt, & Giraud, 2003). Future studies should test this hypothesis more directly by contrasting different aspects of vocal information processing in infants.

SAMPLE SIZE

Another methodological factor that might provide an explanation for the differences observed between these studies is sample size. Looking at sample size reveals that both studies that showed a developmental effect had substantially larger samples—16 at each age group, making a total of 32 in Grossmann et al. (2010) and 33 in Lloyd-Fox et al. (2012)—than the study that failed to find such a developmental effect: 21 infants in Blasi et al. (2011). This suggests that a more representative sampling across the ages provides a better way of conducting developmental neuroimaging studies during infancy. Moreover, from a methodological point of view, it has been recommended that measurements should not be combined over more than 1-month intervals when examining development in infant studies (DeBoer, Scott, & Nelson, 2005), precisely because such grouping of the infant data might conceal developmental effects. Therefore, assessing only a fairly small sample of infants grouped across 3 to 7 months of age, as was done in Blasi et al. (2011), is perhaps not the most effective way of assessing infant development and might in fact preclude developmental effects.

STATE OF THE INFANT

Another critical difference between the two fNIRS studies and the fMRI study was that infants were asleep during the fMRI data collection, but not in the fNIRS studies. Sleep has been shown not only to

reduce effects in auditory studies with adults (Czisch et al., 2002) but also to critically alter infant brain responses in auditory experiments measured with fMRI (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002) and event-related brain potentials (Friederici, Friedrich, & Weber, 2002). Indeed, Blasi et al. (2011) acknowledged that the fact that they did not find a modulation of the responses in the voice-sensitive region depending on emotional content of the vocal stimuli, which is in contrast to prior fMRI work with adults (Grandjean et al., 2005) and prior NIRS work with infants (Grossmann et al., 2010), might be due to the infants being asleep during their study. In a similar vein, it is not unlikely that an awake and alert infant might process the contrast between vocal and non-vocal sounds very differently from an infant that is asleep. This overall argument concerning sleep as a potential problem becomes particularly important when one considers the brain specialization for voices as a crucial neural manifestation of infants' developing social communicative and interactive skills. Taken together, the last two factors, namely the state of the infant during the measurement and sample size and the resulting age distribution of the sample are critical methodological factors to consider when assessing current studies and planning future studies on voice specialization in the infant brain.

Let us now return to the crucial research question: When in human development does the brain specialization for processing voices emerge? There is now corroborating evidence from two fNIRS studies with awake infants that the infant brain becomes sensitive to voices during the first 6 months of life (Grossmann et al., 2010; Lloyd-Fox et al., 2012). Given this developmental pattern, another important issue to be addressed is, how specific is this emerging voice sensitivity in infants? Grossmann and colleagues (2010) contrasted human vocal sounds to a variety of control sounds (including animal vocalizations) and observed increased temporal responses to human vocal sounds at 7 but not at 4 months of age. This suggests that the infant brain becomes tuned to processing species-specific vocal information between 4 and 7 months of age. This conclusion is in line with a recent fNIRS study which found that 4-month-old infants showed similarly strong responses in temporal cortex to monkey calls as to human speech, indicating that at this young age infants are still sensitive to vocalizations from both species (Minagawa-Kawai et al., 2011). These findings together with the other work on voice processing in infancy support the view that voice processing may be gradually narrowed to human vocal sounds during the first year of life. Such narrowing effects have also been reported for

infants' development of face processing with a narrowing toward human faces (de Haan, Pascalis, & Johnson, 2002; Pascalis, de Haan, & Nelson, 2002). This perceptual narrowing consistently observed in numerous studies during infancy has been argued to be an important sensory phenomenon (Lewkowicz & Ghazanfar, 2009; Scott, Pascalis, & Nelson, 2007) that is a critical prerequisite for later developmental achievements (Jansson-Verkasalo et al., 2010). The reviewed evidence from current infant neuroimaging studies provides the first insights into the neural processes that may underlie perceptual narrowing to the human voice during early development.

Original manuscript received 13 September 2011

Revised manuscript accepted 19 September 2011

First published online 25 October 2011

REFERENCES

- Belin, P., Zatorre, R., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*, 309–312.
- Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*, *14*, 2105–2109.
- Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., et al. (2011). Early specialization for voice and emotion processing in the infant brain. *Current Biology*, *21*, 1220–1224.
- Czisch, M., Wetter, T. C., Kaufmann, C., Pollmächer, T., Holsboer, F., & Auer, D. P. (2002). Altered processing of acoustic stimuli during sleep: Reduced auditory activation and visual deactivation detected by a combined fMRI/EEG study. *NeuroImage*, *16*, 251–258.
- DeBoer, T., Scott, L. S., & Nelson, C. A. (2005). ERPs in developmental populations. In T. C. Handy (Ed.), *Event-related potentials* (pp. 263–297). Cambridge, MA: MIT Press.
- de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, *14*, 199–209.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional imaging of speech perception in infants. *Science*, *298*, 2013–2015.
- Friederici, A. D., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *NeuroReport*, *13*, 1251–1254.
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, *8*, 145–146.
- Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. *Neuron*, *65*(6), 852–858.
- Jansson-Verkasalo, E., Ruusuvirta, T., Huotilainen, M., Alku, P., Kushnerenko, E., Suominen, K., et al. (2010). Atypical perceptual narrowing in prematurely born

- infants is associated with compromised language acquisition at 2 years of age. *BMC Neuroscience*, *11*, 88. doi:10.1186/1471-2202-11-88
- Kanwisher, N., McDermott, J., & Chun, M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for the perception of faces. *Journal of Neuroscience*, *17*, 4303–4211.
- Kriegstein, K. V., & Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *NeuroImage*, *22*, 948–955.
- Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences*, *13*, 470–478.
- Lloyd-Fox, S., Blasi, A., Mercure, E., Elwell, C. E., & Johnson, M. H. (2012). The emergence of cerebral specialization for the human voice over the first months of life. *Social Neuroscience*, *7*, 317–330.
- Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cerebral Cortex*, *21*, 254–261.
- Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, *296*, 1321–1323.
- Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., & Logothetis, N. K. (2008). A voice region in the monkey brain. *Nature Neuroscience*, *11*, 367–374.
- Petkov, C. I., Logothetis, N. K., & Obleser, J. (2009). Where are the human speech and voice regions, and do other animals have anything like them? *Neuroscientist*, *15*, 419–429.
- 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.
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*, 670–674.
- von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, *17*, 48–55.