

Voice-sensitive regions in the infant brain

Early specialization for voice and emotion processing in the infant brain

Anna Blasi^{1*}, Evelyne Mercure^{2*}, Sarah Lloyd-Fox³, Alex Thomson¹, Michael Brammer⁴, Disa Sauter⁵, Quinton Deeley¹, Gareth J Barker⁴, Ville Renvall⁶, Sean Deoni^{7,2}, David Gasston⁴, Steven C.R. Williams^{4,7}, Mark H. Johnson³, Andrew Simmons^{4,8} and Declan G.M. Murphy^{1,8}

¹ King's College London, Institute of Psychiatry, Department of Forensic and Neurodevelopmental Science, London SE5 8AF, United Kingdom.

² University College London, Institute of Cognitive Neuroscience, London WC1N 3AR, United Kingdom.

³ Birkbeck College, Centre for Brain and Cognitive Development, London WC1E 7HX, United Kingdom.

⁴ King's College London, King's Health Partners, Department of Neuroimaging, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK

⁵ Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands

⁶ Aalto University School of Science, Brain Research Unit, Low Temperature Laboratory, Espoo, Finland

⁷ Brown University, Division of Engineering, Advanced Baby Imaging Lab, Providence, RI, USA

⁸ NIHR Biomedical Research Centre for Mental Health at South London and Maudsley NHS Foundation Trust and King's College London Institute of Psychiatry

*These authors contributed equally to the work and their names are ordered alphabetically.

Corresponding authors: Anna Blasi, anna.blasi@kcl.ac.uk
Evelyne Mercure, e.mercure@ucl.ac.uk

Tel: + 44 (0) 2078480942
Fax: +44 (0) 2078480650
Institute of Psychiatry, Box P050
De Crespigny Park, London
SE5 8AF, United Kingdom

Summary

Human voices play a fundamental role in social communication, and areas of the adult ‘social brain’ show specialization for processing voices and their emotional content (superior temporal sulcus - STS, inferior prefrontal cortex, premotor cortical regions, amygdala, insula) [1-8]. However, it is unclear when this specialization develops. Functional magnetic resonance (fMRI) studies suggest the infant temporal cortex does not differentiate speech from music or backward speech [10, 11], but a prior study with functional near infrared spectroscopy revealed preferential activation for human voices in 7-month-olds, in a more posterior location of the temporal cortex than in adults [12]. However, the brain networks involved in processing non-speech human vocalizations in early development are still unknown. To address this issue, in the present fMRI study 3 to 7 month olds were presented with adult non-speech vocalizations (emotionally neutral, emotionally positive and emotionally negative), and non-vocal environmental sounds. Infants displayed significant differential activation in the anterior portion of the temporal cortex, similarly to adults [1]. Moreover, sad vocalizations modulated the activity of brain regions involved in processing affective stimuli such as the orbitofrontal cortex [13] and insula [7, 8]. These results suggest remarkably early functional specialization for processing human voice and negative emotions.

Highlights

- Specialization for the human voice was found along the anterior STS in human infants.
- This voice sensitive area is right lateralized and is in a similar location to adults.
- Orbitofrontal cortex and insula activate when infants process sad vocal emotions.

Voice-sensitive regions in the infant brain

Results

In this study we used functional magnetic resonance imaging (fMRI) to investigate brain activation associated with the perception of adult non-speech vocalizations by infants during natural sleep to address two specific aims. Firstly, to determine whether the temporal cortex of young infants shows specialization for human voices; and secondly, to determine which brain areas are activated when human infants process emotion from non-speech vocalizations. In adults, the human voice [1, 2] and the emotional information conveyed by human vocalizations have been found to modulate the activity of a number of brain areas, including the STS, inferior prefrontal cortex, premotor cortical regions, the amygdala and insula [4-8]. We investigated brain function in 21 infants aged 3-7 months who were presented with 3 types of adult non-speech vocalizations (emotionally neutral, emotionally positive and emotionally negative) and a mixture of non-voice environmental sounds likely to be familiar to infants of that age. Patterns of fMRI activation representing the contrast between voice and non-voice stimuli were studied, as well as the contrast between emotionally neutral, positive and negative vocalizations.

Group analyses of all sounds versus rest (Figure 1 and Table S1) revealed significant activation in the middle temporal gyri, right lingual gyrus, medial frontal gyri, right putamen (lentiform nucleus) and right fusiform gyrus. This pattern of activation is

consistent with reports of activation from auditory stimuli in other studies of infants, children and adults [10, 11, 14, 15]. Hence we can confirm that auditory activation was successfully recorded in these naturally sleeping infants.

Neutral vocalizations elicited more activation than *non-voice* stimuli in the right anterior middle and superior temporal gyri (Figure 2a and 2b, Table S2) and, in the medial frontal gyri (Figure 2a, in red). In contrast, the *non-voice* stimuli elicited significantly more activation than *neutral vocalizations* in the left superior temporal gyrus (Figure 2a, in blue). A positive correlation in the Voice > NonVoice contrast was found between age and activation in a small cluster in the left superior temporal gyrus (cluster size = 9 voxels; correlation coefficient = 0.632; p-value = 0.0006, Figure 3, Table S3), suggesting that the voice-selectivity increased with age in this area. No other area showed correlation with age.

No difference was observed between *happy* and *neutral* vocalizations. In contrast, *sad* vocalizations showed stronger activation than *neutral vocalizations* in the insula and gyrus rectus (see Figure 4, Table S4).

Voice-sensitive regions in the infant brain

Discussion

Voice-sensitive activation

Our results demonstrate a strong voice-sensitive area in the temporal cortex of 3-7 month-old infants. The strongest activation was found in the right middle temporal gyrus close to the temporal pole, in a location similar to the anterior portion of the voice-sensitive area reported in adults [1, 2, 16-19]. This area along the anterior superior temporal sulcus (STS) shows more activation to human vocalizations (speech and non speech) than other sounds including non-vocal environmental sounds [1] and animal vocalizations [20]. It also shows adaptation to the speaker's identity in adults [16], and activates more strongly when participants focus on the speaker's voice rather than on the verbal content of spoken sentences [21] whether these voices are familiar or not [22]. Since the stimuli were the same, and only the focus of the participant's attention differed, this result suggests that the activation of this STS area cannot be attributed to low level differences in stimulus categories.. For these reasons, this area of the anterior STS has been described as having an important role in the processing of human voices, especially the identification of speakers [16]. In contrast, Grossmann et al. [12] found voice-sensitivity in a more posterior temporal area in 7 month olds, which has been interpreted as an indication of a developmental difference in the voice-sensitivity of more anterior temporal areas [21]. The current results suggest that this localization discrepancy between prior studies of infants and adults is more likely to reflect differences in brain imaging techniques (for example spatial resolution), than genuine developmental differences.

In our study, the voice-sensitive activation was right lateralized in the temporal cortex, which is congruent with previous fNIRS findings in babies [12], and fMRI findings in adults [20, 23]. The right anterior superior temporal sulcus has been specifically related to the analysis of nonverbal features of speech, showing more activation when focusing on the speaker's voice than when focusing on the semantic content of spoken sentences [21]. Our results suggest that a right hemisphere bias for processing human vocalizations emerges early in development.

Although no correlation was found with age for the main effects in the human voice versus environmental sounds contrast, a small volume in the left superior temporal gyrus did show a positive correlation with age. This small volume was localized in a similar region to findings from an infant fNIRS study that reported age-correlated differences in activation for human voice [12]. Interestingly, while the current findings were left lateralized, the findings from the fNIRS studies suggested bilateral age correlated changes in activation in the temporal cortex. The sleeping state of the infants could have influenced our findings by reducing the BOLD contrast responses as observed in adults [24], or by reducing the differentiation between responses to different stimuli as observed in babies [10]. Also, in fNIRS studies with infants it is necessary to use multimodal stimuli (visual and auditory) to maintain the baby's interest in the task [25]. This multimodal stimuli presentation may have also influenced the responses recorded with fNIRS, accentuating further the differences with our findings.

Voice-sensitive regions in the infant brain

Emotion-modulated activation

Our second aim was to explore brain areas in which activation was modulated by the emotional content of human vocalizations in young infants. Many studies with adults have found that the emotional information conveyed by human vocalizations modulates the activity of certain brain areas, including the temporal voice-sensitive area, inferior prefrontal cortex, premotor cortical regions, the amygdala and insula [4, 5, 6, 7, 8]. A prior study of infants using fNIRS [12] also found that emotional prosody in spoken words increased the activation of the voice-sensitive area of the right temporal hemisphere in 7 month olds. In the same age group, event-related potentials revealed a positive slow wave over temporal electrodes when words were spoken with a happy or angry prosody but not in a neutral prosody [26]. However, ERPs have a very low spatial resolution and one technical limitation of fNIRS is that it can only be used to measure the activity in superficial areas of the cortex. In addition, the technique is limited by the placement of the fNIRS probes, which in the aforementioned study was over the inferior frontal and temporal cortex [12]. Therefore, a unique contribution of our study was the investigation of emotion within voice over the whole infant brain using an imaging technique of high spatial resolution. Moreover, in contrast to studies on speech, the present study used emotional signals that pre-verbal infants produce themselves from a very early age (crying and laughter), and that are thought to reflect innate behaviors to communicate emotional states [27].

In the present study, the activation elicited by happy vocalizations did not differ from that of emotionally neutral vocalizations, whereas sad vocalizations elicited significantly greater responses than emotionally neutral vocalizations in the insula and orbitofrontal cortex. These results are congruent with findings of increased activation in the insula when adults listen to emotionally salient non-speech vocalizations (especially sad and fearful vocalizations) [7, 8]. The orbitofrontal cortex has been described as one of the ‘least understood regions of the human brain’ [13]. Available data suggest a role of this area in the processing of affective stimuli, although little is known about the maturation processes of the orbitofrontal cortex in childhood and adolescence [13].

The temporal cortex area identified as voice-selective in the present study did not show any modulation of activity based on the emotional content of the stimuli. This is in contrast with others using fNIRS [12] that found an increase of activity with emotional prosody in infants in the voice-sensitive temporal channels, and others using fMRI [3], that found enhanced response of the temporal voice-sensitive area with words spoken with an angry prosody in adults. The relative lack of differences between the emotional (especially happy) and neutral conditions may reflect the fact the participants in the present study were asleep, therefore these responses may have been more difficult to differentiate [10], whereas they were awake in previous studies of emotional vocalizations. Moreover, the ‘neutral’ vocal stimuli in the present study (coughing, sneezing, throat clearing) may be emotionally salient to infants because they indicate the presence of adults. In adult studies, classification of stimuli as *neutral* or *emotional* is based on participants’ rating (eg.: [28]), but an infant’s perception of their emotional

Voice-sensitive regions in the infant brain

content may differ. The terms ‘sad’, ‘happy’ or ‘neutral’ used here should be taken as labels to describe the stimuli as perceived by adults and do not imply that the same emotions are evoked in infants when hearing these stimuli. Furthermore, the finding that activation was evident for sad versus neutral, but not happy versus neutral vocalizations could be an effect of unfamiliarity since young infants are likely to be less frequently exposed to crying than adult laughter and neutral vocalizations. The same activation may not be observed for sad vocalizations when studying the response to the sound of a child crying in a group of infants who have older siblings or who attend childcare settings from an early age. Studying the infants of depressed mothers may also be a way of assessing the role of early experience, since these infants may be exposed to an atypical balance of neutral, happy, sad and angry vocalizations. Nevertheless, our findings indicate that emotional vocalizations are associated with differential activity in emotion processing networks from an early age.

In conclusion, voice-sensitive activation was observed along the anterior STS in 3-7 month olds in a location very similar to that described in the adult brain. This suggests that the infant temporal cortex shows more refined functional specialization than previously reported. This early functional specialization for processing the human voice parallels infants’ surprisingly early ability to extract subtle information from human vocalizations. Indeed, newborns prefer to listen to their mother’s voice [28] and their mother tongue [29], while young infants can also discriminate emotional vocalizations [31] and can differentiate male/female and child/adult vocalizations [32]. Moreover, the emotional content of vocalizations, especially sadness, may modulate the activity of areas known to be involved in processing affective stimuli in the adult brain, such as the

orbitofrontal cortex and insula. These results represent a rare demonstration that cortical specialization exists very early in development, a fundamental advance in our understanding of infant development. The findings of this study also highlight the potential of fMRI as a tool for studying the development of specific brain responses to auditory stimuli in infancy.

Experimental procedures

Participants

Data from 21 naturally sleeping infants was included in the analysis (age range 91-203 days, mean age: 152 ± 30 days, gestation-corrected). All parents gave their written informed consent. The study was approved by the Institute of Psychiatry and South London and Maudsley research ethics committee.

Stimuli

Infants were presented with 4 categories of auditory stimuli: neutral vocalizations (natural non-speech vocalizations with no strong emotional content such as coughing, sneezing, yawning and throat clearing), happy voices (laughing), sad voices (crying) and

Voice-sensitive regions in the infant brain

non-voice sounds (environmental sounds that are not human or animal produced, but are likely to be familiar to infants of that age, such as toy sounds and water sounds). All voice sounds were adult vocalizations. Each condition was presented for 21 seconds, with an interval of 9 seconds of rest, organised in a block design.

Data analysis

The MRI data was analysed with XBAM software (www.brainmap.co.uk/xbam.htm) using a data-driven approach based on the standard general linear model (GLM). Data were normalised to Talairach space using an infant template previously described by Dehaene-Lambertz et al [10]. Statistical analysis was performed using a random effects model, and only clusters of at least 3 voxels in size are reported.

Supplemental information:

Supplemental information includes an additional figure, a detailed description of the study participants, testing procedure and data analysis with supplemental references to complement these.

Acknowledgements:

Funding was provided by the Medical Research Council (MRC, UK) and the MRC UK AIMS network (G0400061/69344, DGMM, principle investigator); the NIHR Biomedical Research Centre for Mental Health at King's College London, Institute of Psychiatry and South London & Maudsley NHS Foundation Trust. MJ and SL-F were supported by MRC G0701484. We would like to thank Dr. Ghislaine Dehaene-Lambertz for her valuable input on infant MRI scanning.

Voice-sensitive regions in the infant brain

References

1. Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403, 309-312.
2. Belin, P., Zatorre, R.J. and Ahad, P. (2002) Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, 13, 17–26.
3. Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M.L., Scherer, K.L., and Villeumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, 8(2), 145-146.
4. Fecteau, S, Fecteau, Armony, J.L., Joanette, Y., and Belin, P. (2005) Sensitivity to voice in human prefrontal cortex. *J Neurophysiol*, 94, 2251–2254.
5. Warren, J.E., Sauter, D.A., Eisner, F., Wiland, J., Dresner, M.A., Wise, R. J.S., et al. (2006). Positive emotions preferentially engage an auditory-motor "mirror" system. *Journal of Neuroscience*, 26(50), 13067-13075.
6. Fecteau, S., Belin, P., Joanette, Y., and Armony, J.L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage*, 36, 480-487.
7. Morris, J.S., Scott, S.K., Dolan, R.J. (1999). Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia*, 37, 1155-1163.
8. Sander, K., Scheich H. (2001). Auditory perception of laughing and crying activated human amygdala regardless of attentional state. *Cognitive Brain Research*, 12, 181-198.
9. Courchesne, E., Pierce, K., Schumann, C.M., Redcay, E., Buckwalter, J.A., Kennedy, D.P., and Morgan J. (2007). Mapping Early Brain Development in Autism. *Neuron* 56, 399-413.
10. Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional Neuroimaging of Speech Perception in Infants. *Science*, 298, 2013-2015.
11. Dehaene-Lambertz, G., Montavont, A., Jobert, A., Allirol, L., Dubois, J., Hertz-Pannier, L., and Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain & Language*, 114(2), 53-65.

12. Grossman, T., Oberecker, R., Koch, S.P., and Friederici A.D. (2010) The Developmental Origins of Voice Processing in the Human Brain. *Neuron*, 65, 1-7.
13. Kringelbach, M.L. (2005) The human orbitofrontal cortex: linking reward to hedonic experience. *Nature Reviews Neuroscience*, 6, 691-702.
14. Altman, N.R., and Bernal, B. (2001). Brain Activation in Sedated Children: Auditory and Visual Functional MR Imaging. *Radiology*, 221, 56–63.
15. Redcay, E, Kennedy, D.P. and Courchesne, E. (2007). fMRI during natural sleep as a method to study brain function during early childhood. *NeuroImage* 38, 696–707.
16. Belin, P. and Zatorre, R.J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport* 14(16), 2105-2109.
17. Belin, P., Fecteau, S. and Beddard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences* 8 (3), 129-135.
18. Latinus, M. and Belin, P. (2011). Human voice perception. *Current Biology* 21 (4), R143-R145.
19. Scott, S. (2008). Voice processing in monkey and human brains. *Trends in Cognitive Sciences* 12(9), 323-325.
20. Fecteau, S. Armony, J.L., Joanette, Y., Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *NeuroImage*, 23(3), 840-848.
21. Von Kreigstein, 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(1), 48-55.
22. Von Kreigstein, K, Giraud, AL. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *NeuroImage*, 22 (2), 948-955.
23. Belin, P., and Grosbras M.-H. (2010) Before Speech: Cerebral Voice Processing in Infants. *Neuron*, 65, 733-735.
24. Czisch, M., Wetter, T.C., Kaufmann, C., Pollmächer, T., Holsboer, F., and Auer, D.P. (2002). Altered processing of acoustic stimuli during sleep: reduced

Voice-sensitive regions in the infant brain

- auditory activation and visual deactivation detected by a combined fMRI/EEG study. *NeuroImage*, 16, 251-258.
25. Lloyd-Fox, S., Blasi, A., and Elwell, C.E. (2010). Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy. *Neuroscience and Biobehavioral Reviews* 34, 269–284
 26. Grossmann, T., Striano, T., and Friederici, A.D. (2005). Infants' electric brain responses to emotional prosody. *Neuroreport*, 16(16), 1825-1828.
 27. Barr, R.G., Hopkins, B., and Green, J.A. (2000). *Crying as a sign, a symptom, & a signal*. New York: Cambridge University Press.
 28. Belin, P., Fillion-Bilodeau, S., and Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research* 40 (2), 531-539.
 29. DeCasper, A.J., Fifer, WP. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, 208(4448), 1174-1176.
 30. Moon, C., Cooper, R.P., Fifer, WP. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16(4), 495-500.
 31. Walker-Andrews, A.S., and Lennon, E.M. (1991). Infants' discrimination of vocal expressions: Contributions of auditory and visual information. *Infant Behavior and Development*, 14, 131-142.
 32. Bahrick, L.E., Netto, D., and Hernandez-Reif, M., (1998). Intermodal Perception of Adult and Child Faces and Voices by Infants. *Child Development*, 69(5), 1263-1275.

Figure Legends

Figure 1. All sounds v rest condition. 3D representation of activation to all sounds compared with rest. The activations are significant at $p \leq 0.005$. Related to Table S1.

Figure 2. Neutral v non voice contrast. Representation on a gray matter infant template of the results for the contrast between *neutral* and *non voice* conditions. L = left; R = right. (A) Neutral vocalization > non voice (in red), neutral vocalization < non voice (in blue). The Talairach z-coordinates mark the position of the voxel with maximum activation. (B) 3D render of the contrast neutral vocalization > non voice, showing the right anterior temporal activation. Related to table S2.

Figure 3. Correlation with age. Representation of the cluster in the left superior temporal gyrus (BA 22) showing positive correlation with age for the contrast neutral vocalization – non voice. Related to table S3.

Figure 4. Sad versus neutral vocalizations. The clusters represent increased activation with sad compared with neutral vocalizations. The corresponding Talairach z-coordinate are given above each slice. L = left; R = right. Related to table S3.

Voice-sensitive regions in the infant brain

Figure 1

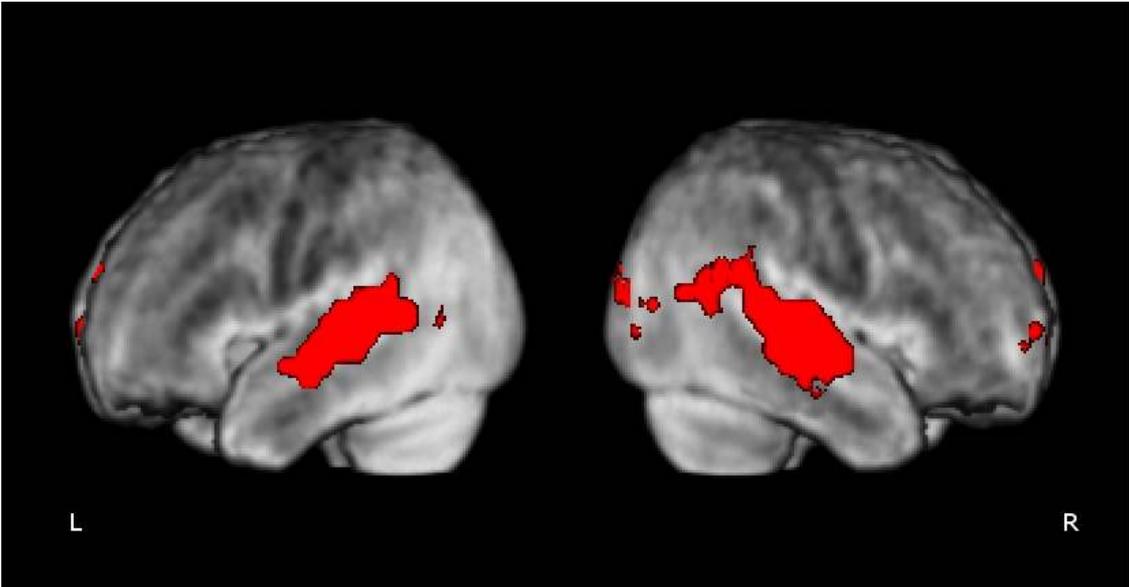
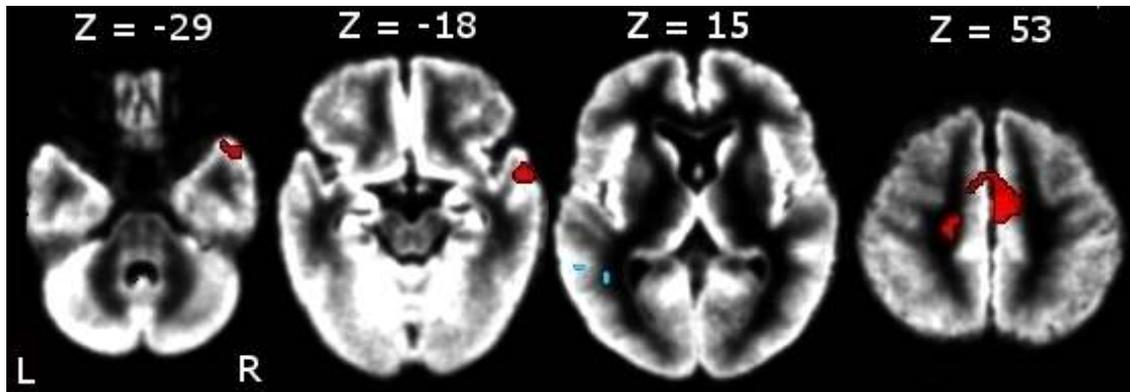
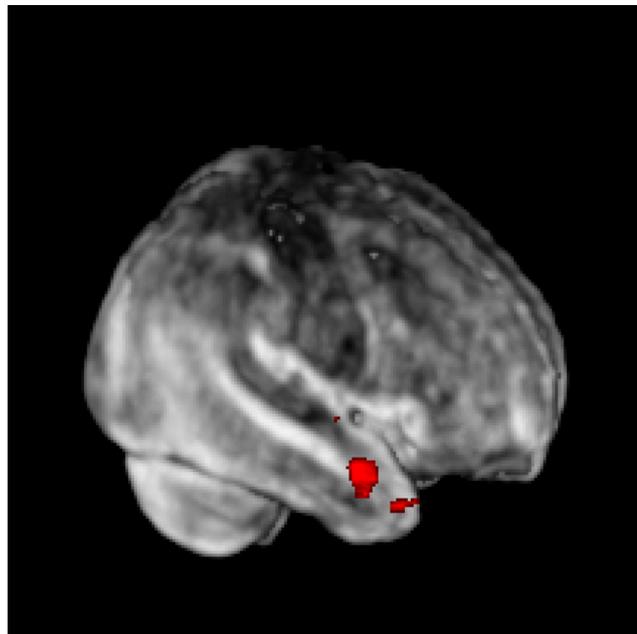


Figure 2

A



B



Voice-sensitive regions in the infant brain

Figure 3

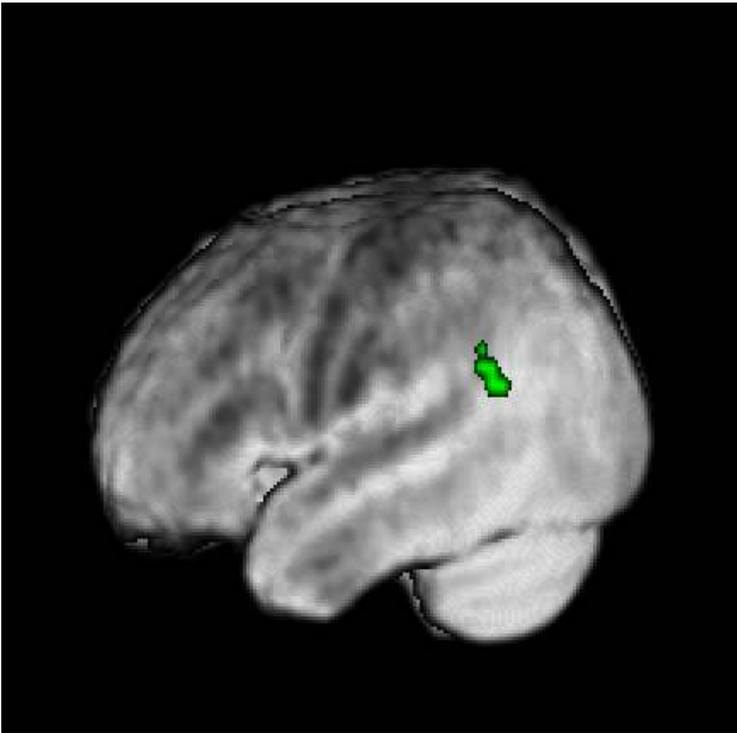


Figure 4

