The recognition of human voice in deaf and hearing infants

Arnaud Coëz[1,2,3], Natalie Loundon[1,*], Isabelle Rouillon[1], Marine Parodi[1], Marion Blanchard[1], Sophie Achard[1], Eréa-Noël Garabédian[1], Eric Bizaguet[2], Tabassome Simon[5], Natacha Tessier[4], Françoise Denoyelle[1], Judit Gervain[6,7]

[1]Service d'oto-rhino-laryngologie pédiatrique, Centre de Réference des Surdités Génétiques, Hôpital Necker Enfants Malades, AP-HP, Université de Paris, Paris, France
[2]Laboratoire de correction auditive, Bizaguet, Audition Santé, Paris, France
[3]INSERM ERL "Developmental Trajectories & Psychiatry", Université Paris Saclay, Ecole Normale Supérieure Paris-Saclay, Université de Paris, CNRS, Centre Borelli.
[4]Service d'oto-rhino-laryngologie pédiatrique, Hôpital Robert Debré, AP-HP, Université de Paris, Paris, France
[5]URC du Groupe Hospitalier des Hôpitaux Universitaires de l'Est Parisien (URC-Est) Hôpital Saint-Antoine, AP-HP, Paris, France
[6]Integrative Neuroscience and Cognition Center, CNRS & Université de Paris, France
[7]Department of Developmental and Social Psychology, University of Padua, Italy

*Corresponding author:
Dr Natalie Loundon
Hôpital Necker-Enfants Malades
Département ORL
149 rue de Sèvres
75015 PARIS
France
email : natalie.loundon@aphp.fr

word count: 2690 words (main text, figure legends)

Abstract

Deafness in infancy has long-lasting consequences on brain organization. To achieve the best developmental outcomes in case of profound deafness, cochlear implantation (CI) needs to take place in a critical period during the first year of life, before the cross-modal reorganization of the brain due to auditory deprivation stabilizes, preventing the typical development of the auditory cortices. Despite its importance for implantation outcomes, the nature of this critical period and the underlying neural reorganization have not been fully explored. To fill this knowledge gap, we investigated the cortical responses of 12-month-old profoundly deaf infants and their age-matched controls to sounds produced by a human voice as compared to non-human sounds using functional near-infrared spectroscopy (fNIRS). The deaf infants were tested before undergoing CI surgery and wore hearing aids (median threshold 70 dB HL at 250 Hz, min 45 dB HL max 100 dB HL), allowing them to perceive low frequencies. Human voice stimuli have been shown to trigger brain responses early in development, possibly due to its evolutionary relevance for survival. We found increased brain responses to the human voice in the deaf infants in the bilateral fronto-temporal areas, and their responses correlated with their residual hearing thresholds. These results suggest that even the limited sound stimulation that these deaf infants receive due to their residual hearing allowed the temporal cortices to develop sensitivity to the human voice prior to implantation. The hearing control group showed an inverted hemodynamic response to both voice and non-voice stimuli in the left parietal and right temporal areas, suggesting that by 12 months of age, they habituate rapidly to these very familiar stimuli. While both groups showed an inverted response to the non-voice stimuli, they differed in their responses to voice stimuli. We hypothesize that the responses found in deaf infants to human voice can be consider a good cochlear implant prognosis.

Key words: fNIRS, infants, deafness, cochlear implant, neurodevelopment, brain imaging

Introduction

Functional neuroimaging has made it possible in recent years to better understand the impact of deafness and hearing aids on the organization of cortical maps (Coez et al., 2021). Many of these imaging techniques (e.g. MRI, CT, PET) are not easily tolerable by young infants, yet auditory deprivation impacts brain reorganization early in life and the deprivation-induced neural changes in brain organization are often irreversible after neuroplasticity decreases at the end of the critical / sensitive period (Werker & Hensch, 2015). Such changes have profound effects on deaf and hearing-impaired individuals' development, for instance causing delays in speech and language acquisition, thereby impacting their social integration, academic and professional success. They also have consequences for the outcomes of interventions, e.g. for the success of cochlear implantation. It is, therefore, of utmost importance to better understand the mechanisms of neural plasticity and reorganization caused by auditory deprivation very early in human development.

Many studies have addressed the question of neuroplasticity and reorganization after cochlear implantation (CI) in older children and adults (Bortfeld, 2019; Kral & Sharma, 2012; Sharma & Campbell, 2011; Sharma & Dorman, 2011). However, very few studies have addressed this issue in the youngest infants, partly due to the practical and methodological difficulties of testing young infants with brain imaging (Saliba et al., 2016; Sevy et al., 2010). The current study aims to fill this gap. With the advent of more infant-friendly brain imaging techniques, such as near-infrared spectroscopy (NIRS) (Gervain et al., 2011; Lloyd-Fox et al., 2009; Yücel et al., 2021), it is now possible to test young infants' neural responses. The current study aims to fill this gap. NIRS is a fully non-invasive, portable, comfortable and low-cost brain imaging modality well suited for testing babies and young children, as it requires no sedation, isotopes, carrier substances or strong magnetic fields. It is silent and has good motion tolerance. Importantly from the perspective of hearing research, it is compatible with CIs as the two devices and signals do not interfere. We, therefore, use near-infrared spectroscopy (NIRS) to explore the neural responses of 12-month-old deaf infants and their age-matched hearing controls to human voice and non-voice sounds.

The human voice is our species' communicative signal and as such has utmost importance for survival. It is thus not surprising that dedicated brain areas are responsible for its processing. Studies carried out over the past 20 years have made it possible to describe these in detail, at least in healthy, hearing humans (Belin et al., 2000). Typical studies have compared the processing of human voice (e.g. speaking, singing, humming, crying, laughter etc.) to other non-human sounds (e.g. cars honking, doors closing, leaves rattling etc.). Both voices and sounds induce bilateral brain activity along the superior temporal sulcus, but the response is greater to the voice than to non-voice stimuli. Indeed, the human voice recognition area has been identified in infants as young as 4-7 months of age (Grossmann et al., 2010). Studies using fNIRS in 4-7-month-old babies found increased responses in left and right superior temporal cortex to human voice when compared to nonvocal sounds, suggesting that voice-sensitive brain systems emerge between 4 and 7 months of age (Grossmann et al., 2010). Even earlier in development, experiments measuring changes in heart rate during the presentation of different voices demonstrate that newborns can discriminate voices and recognize the voices of their parents, an ability even present in fetuses before birth (Kisilevsky et al., 2003). Furthermore, voice recognition has been successfully used to assess the effectiveness of cochlear implant outcomes in adults using PET (Coez et al., 2008).

The current study aims to analyze the effect of deafness and the resulting auditory deprivation on the brain responses of deaf infants to human voice as opposed to non-voice sounds. The deaf infants were tested prior to CI surgery and were using their hearing aids at the time of test so they could rely on their residual hearing.

Materials and Methods

Participants

The clinical group consisted of 16 infants (3 females; median age: 370 days; range: 330 days – 387 days). Of these, 8 were not included in the data analysis due to fussiness and crying, an insufficient number of valid trials due to poor data quality or parental interference. The control group consisted of 13 hearing infants (6 females; median age: 370 days; range: 350 days – 393 days). Of these 4 infants were not included in the data analysis due to fussiness and crying, an insufficient number of valid trials due to poor data quality or parental interference. There were no age difference across group ($p = 0.9$). Rejection due to poor data quality was performed in batch, following the same criteria for all infants (see Data Processing and Analysis below), prior to statistical analysis.

The children were followed up and tested at age 24 months of age, i.e. 12 months after CI surgery, and some after 36 months of age, i.e. 24 months after surgery, as well. However, due to technical difficulties to place the headgear over the implant, these sessions didn't yield exploitable data.

All parents gave informed consent before the experiment. The present experiment was approved by the Saint Louis Hospital ethics committee (approval 2014/58).

Design and Stimuli

The experiment comprised 14 blocks per condition, i.e. 14 voice and 14 non-voice blocks, for a total of 28 blocks. The voice stimuli were those used in the original fMRI study of Belin et al. (2000) and were adapted to the fNIRS paradigm. The voice stimuli consisted of speech (30% of stimuli: words, non words, foreign language) and nonspeech (70% of stimuli: laughs, sighs, various onomatopoeia). The non-voice stimuli consisted of natural sounds (19%), animals (24%), modern human environment (47%) and musical instruments (10%). Stimuli were delivered in free field at a mean 70 dB SPL. Blocks lasted 16 seconds. The order of voice and non-voice blocks was pseudo-randomized and counterbalanced across participants. The blocks were separated by baseline periods of jittered duration between 25 and 35 sec. A movie of non-object-like rays of light moving while changing color was used as a visual attention-getter to keep infants focused and attentive. The experimenter also entertained the baby with silent toys or movements if necessary. The whole experiment lasted about 30 minutes.

Procedure

Infants were tested with a NIRx NIRScout 1616 machine (source-detector separation: 3 cm; two wavelengths of 760 nm and 850 nm using pulsed LED light; sampling rate: 15.625 Hz) at the maternity ward of the Robert Debré Hospital and of the Necker Enfants Malades, Paris, France. The testing session lasted about 15-20 min. The optical sensors were inserted into a stretchy EEG cap and were placed bilaterally on the infants' head using surface anatomical landmarks (inion, nasion, vertex and the bilateral preauricular points), targeting the classical auditory areas in the bilateral temporal, frontal and parietal cortices (Figure 1).

During testing, infants were seated on a caregiver's lap in a quiet, dimmed room. A Dell Opti-Plex PC380 computer played the sound stimuli through loudspeakers placed in front of the infant bilaterally. The baseline visual stimuli were presented on a screen placed directly in from of the infant as a distance of 50 cm. The experiment was run using E-Prime.

Another DellOpti-Plex computer recorded the NIRS signal. If infants lost attention during the baseline periods, an experimenter hidden behind the computer screen presented silent toys by raising them above the screen and moving them to redirect infants' attention to the screen.

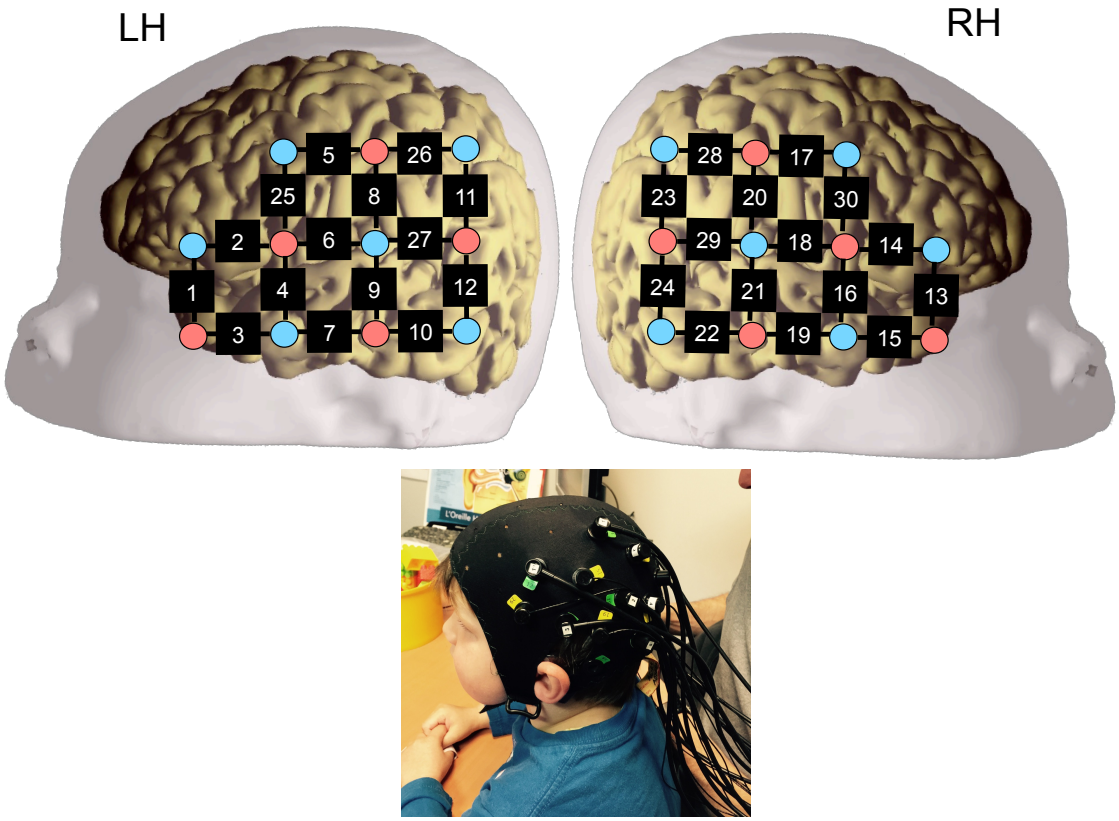

Figure 1. The localization of the NIRS optodes

Data Processing and Analysis

The NIRS machine measured the intensity of the transmitted of red and near-infrared light, from which relative changes in the concentration of oxygenated hemoglobin (oxyHb) and deoxygenated hemoglobin (deoxyHb) were calculated as indicators of neural activity. To eliminate high-frequency noises (e.g., heartbeat) and overall trends, the data were band pass-filtered between 0.01 and 0.7Hz. Movement artifacts, defined as concentration changes greater than 0.1 mmol*mm over 2 samples, were removed by rejecting block-channel pairs in which artifacts occurred. For the non-rejected, valid blocks, a baseline was linearly fitted between the means of the 5 s preceding the onset of the block and the 5 s starting 15 s after offset of the block. Infants were included in the analysis if they had at least 20% valid data.

All statistical analyses were carried out over both oxyHb and deoxyHb. However, this latter didn't yield significant results, as is common in infant populations (Gervain et al., 2011; Lloyd-Fox et al., 2009) and will thus not be reported in detail. We conducted two types of analyses: = as well as the two conditions between them. We conducted paired samples t-tests comparing infants' responses between the voice and non-voice conditions, as well as one sample t-tests comparing the responses in each condition to a zero baseline in each channel. Furthermore, for hearing-impaired infants, we have correlated their responses to the voice condition (in the channels that showed a significant response to this condition) with their

residual hearing threshold at 250Hz, i.e. the frequency close to the fundamental frequency of the human voice.

Given the non-absolute nature of the hemodynamic responses measured with fNIRS, between-group comparisons are not readily possible.

Results

The hemodynamic responses obtained for the clinical (hearing-impaired) and control (hearing) groups are shown in Figures 2 and 3, respectively.

Clinical (hearing-impaired) group

Voice vs. non-voice. The permutation test over oxyHb concentration changes directly comparing the two conditions yielded a spatial cluster involving channel 7 ($p=0.091$) in the LH and channels 13 & 16 ($p=0.055$) in the RH with marginally greater activation to the voice as compared to the non-voice stimuli.

Voice vs. baseline. The permutation test over oxyHb concentration changes comparing the voice condition to a zero baseline yielded a spatial cluster involving channels 7 & 10 ($p<0.001$) in the LH with greater activation to the voice stimuli than to baseline.

Non-voice vs. baseline. The permutation test over oxyHb concentration changes comparing the non-voice condition to a zero baseline yielded a spatial cluster involving channel 5 ($p=0.001$) in the LH and channels 14, 17 & 30 ($p=0.004$) in the RH with greater deactivation, i.e. a more negative, inverted response, to the non-voice stimuli than to baseline.

Control (hearing) group

Voice vs. non-voice. The permutation test over oxyHb concentration changes directly comparing the two conditions yielded no spatial clusters with significant differences between the two conditions.

Voice vs. baseline. The permutation test over oxyHb concentration changes comparing the voice condition to a zero baseline yielded a spatial cluster involving channel 8 ($p=0.02$) in the LH with greater deactivation, i.e. a more negative, inverted response, to the voice stimuli than to baseline.

Non-voice vs. baseline. The permutation test over oxyHb concentration changes comparing the non-voice condition to a zero baseline yielded a spatial cluster involving channel 8 ($p=0.004$) in the LH and channel 22 ($p<0.001$) in the RH with greater deactivation, i.e. a more negative, inverted response, to the non-voice stimuli than to baseline.

Correlations between hemodynamic responses and residual hearing

As deaf infants showed increased brain responses to human voice in channels 7 and 10, we correlated their responses in these channels with their residual hearing threshold (median: 70 dB HL at 250 Hz, min. 45 dB HL, max. 100 dB HL). The correlation was not significant in channel 7 ($r = -0.31$, ns.) but it reached significance in channel 10 ($r = -0.769$, $p = 0.025$).

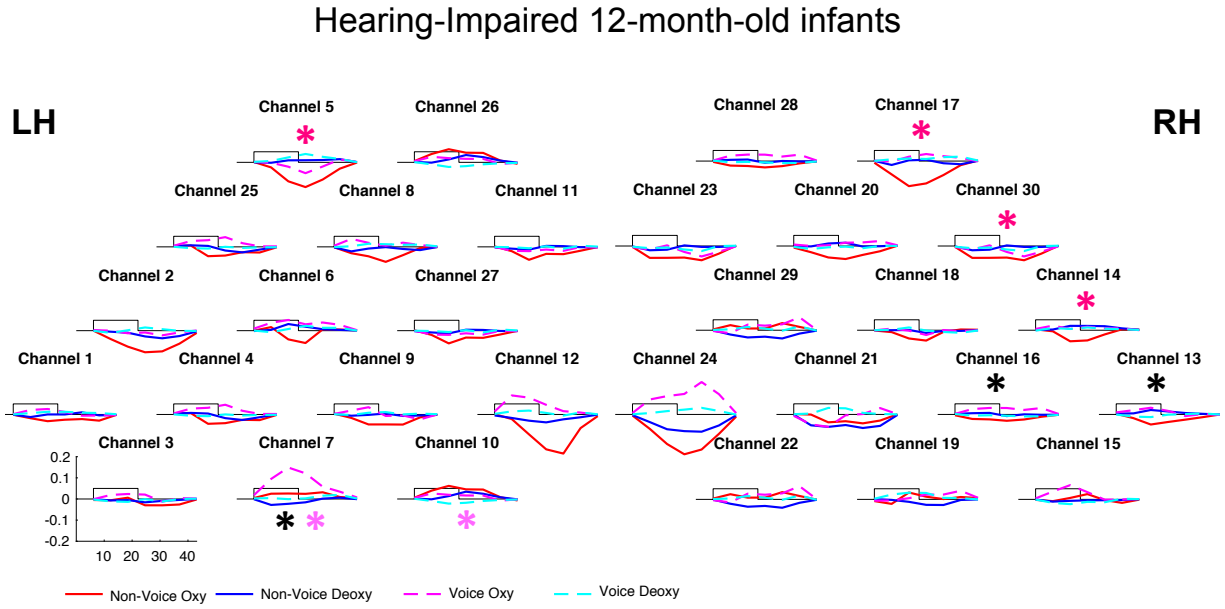

Figure 2. Hearing impaired 12-month-old infants' responses to voice and non-voice stimuli. Channels are plotted following the probe placement indicated in Figure 1. In each subplot, the x-axis represents time in sec. The y-axis shows concentration change in mmol × mm. The rectangle along the x-axis indicates time of stimulation (in seconds). The curves indicate grand average responses for the voice (oxyHb: pink line, deoxyHb: turquoise line) and non-voice (oxyHb: red line, deoxyHb: blue line) conditions. Asterisks indicate significant spatial clusters obtained in the permutation tests: black for the voice vs. non-voice comparison, pink for the voice vs. baseline comparison and red for the non-voice vs. baseline comparison.

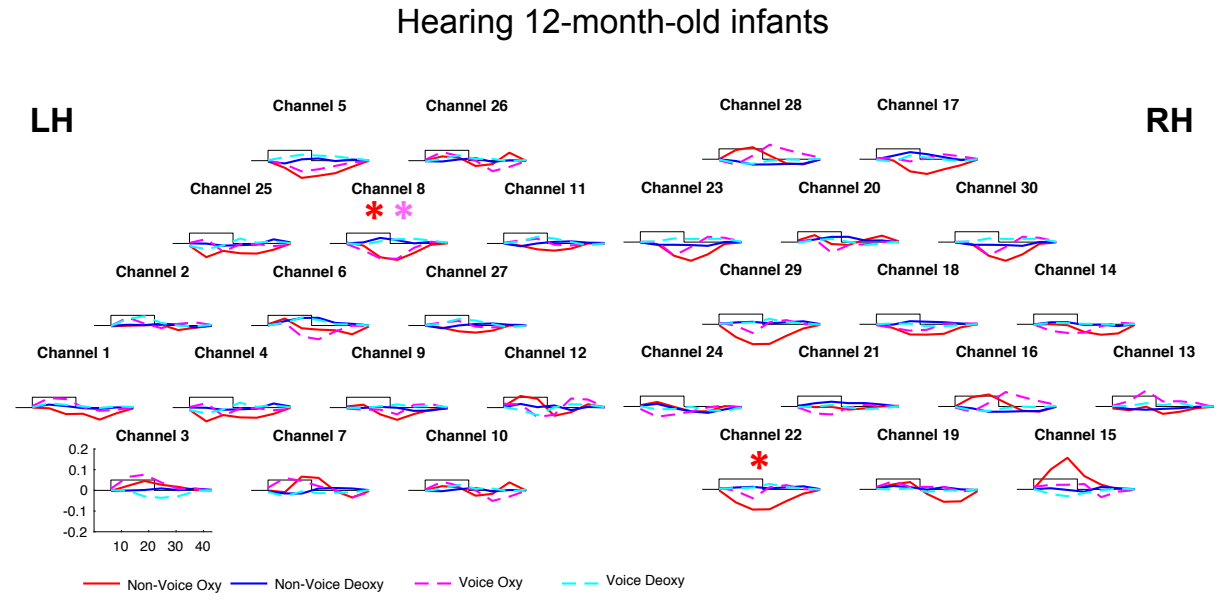

Figure 3. Hearing 12-month-old infants' responses to voice and non-voice stimuli. The plotting conventions are the same as in Figure 2.

Discussion

In this study, we compared 12-month-old hearing-impaired and hearing infants' responses to human voices and non-voice stimuli using near-infrared spectroscopy.

The most important finding of the study is that hearing-impaired infants showed a differential response in the bilateral temporal areas bilaterally to the voice stimuli as compared to non-voice stimuli. Moreover, in the left temporal areas, they showed an increased, positive response to the voice stimuli as compared to baseline, while they showed a negative, inverted response to non-voice stimuli in the bilateral fronto-parietal areas. These results suggest that the temporal voice recognition area, as described by Belin (2000), is robust even in the case of radical auditory deprivation. We hypothesize that the area emerges very early in development and requires no or little sensory input to function. The most relevant perceptual input for this ability to emerge is the perception of low frequencies, as the fundamental frequency of the human voice lies in the 100-300Hz range. Such low frequency stimulation is available to infants already prior to birth, as low frequency sounds may be perceived in utero both through auditory and vibro-tactile transmission (Gervain, 2015, 2018). Given the importance of caregiving and thus the recognition of other humans for the survival of human infants (Grossmann et al., 2010; Grossmann & Friederici, 2012), the robustness and precocious development of voice recognition may be evolutionarily adaptive. We thus speculate that 12-month-old hearing-impaired infants' responses to human voices represent a developmentally typical, but delayed response, for which strongly reduced low frequency stimulation available through residual hearing through hearing aids is sufficient. The correlation we observed between infants' responses and their hearing thresholds provide empirical support to this idea. Furthermore, adults have been found to activate the temporal voice area even for low-pass filtered voice stimuli(Belin et al., 2000).

Hearing infants, i.e. the control group, by contrast, showed an inverted response to both voice and non-voice stimuli in the left parietal and right temporal areas, and no difference between the two conditions. Inverted hemodynamic responses or deactivations are often found in the infant NIRS literature, for instance as the neural signature of habituation, i.e. decreased processing effort, upon repeated stimulus presentation (Issard & Gervain, 2018). This interpretation explains well our findings in the control group, as we find inverted responses to both stimuli in 12-month-old hearing infants. Arguably, these infants have ample experience with both human voices and non-voice sounds, the processing of these sounds does not, therefore, require considerable metabolic effort. By contrast, hearing-impaired infants may need greater effort to process these sounds, given their more limited experience and degraded input, similarly to the positive hemodynamic responses found in 4-7-month-old infants (Grossmann et al., 2010). This further supports the hypothesis that hearing impaired infants' responses may be typical, but delayed.

As a limitation of our study, sample sizes were relatively small. While this is not unusual for studies with clinical populations, in particular with infants, there are important methodological lessons to be learned specific to hearing impaired populations. Movement causes artifacts and distortions in the brain signal, leading to rejection of artifacted data and reduction of sample sizes and statistical power. These movement artifacts were stronger in our hearing impaired sample. This was obvious both in the NIRS signal, leading to more rejected data, and when observing infants' behavior. Since deaf and hearing impaired infants are deprived of auditory perception to a large extent, they rely on vision to gain information about their environment. This translates into frequent head turns and other motions to allow the infant to scan the environment (e.g. look for parents for social referencing etc.), leading to more artifacted signals. Future testing in this population should aim to reduce these behaviors, e.g. by actively engaging infants' attention or by placing a parent in front of the infant.

Conclusion

Our study has shown that the human voice recognition area is functional in 1-year-old infants with hearing loss. Responses are different from those recorded in age-matched normal hearing babies, and resemble more closely those found in younger infants.

Acknowledgements

This study has funded by MeDEL (https://www.medel.com) as well as by the ERC Consolidator Grant "BabyRhythm" nr. 773202 to JG. We thank 'URC Est' for its help with the logistics of the study. The Assistance Publique - Hôpitaux de Paris served as the official promotor of the study.

Declaration of interest statement

The company MeDEL partially funded the study, but had no influence over the methods, results and the publication.

References

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312.

Bortfeld, H. (2019). Functional near-infrared spectroscopy as a tool for assessing speech and spoken language processing in pediatric and adult cochlear implant users. *Developmental Psychobiology*, *61*(3), 430–443.

Coez, A., Fillon, L., Saitovitch, A., Rutten, C., Marlin, S., Boisgontier, J., Vinçon-Leite, A., Lemaitre, H., Grévent, D., Roux, C.-J., Dangouloff-Ros, V., Levy, R., Bizaguet, E., Rouillon, I., Garabédian, E. N., Denoyelle, F., Zilbovicius, M., Loundon, N., & Boddaert, N. (2021). Arterial spin labeling brain MRI study to evaluate the impact of deafness on cerebral perfusion in 79 children before cochlear implantation. *NeuroImage. Clinical*, *29*, 102510.

Coez, A., Zilbovicius, M., Ferrary, E., Bouccara, D., Mosnier, I., Ambert-Dahan, E., Bizaguet, E., Syrota, A., Samson, Y., & Sterkers, O. (2008). Cochlear implant benefits in deafness rehabilitation: PET study of temporal voice activations. *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, *49*(1), 60–67.

Gervain, J. (2015). Plasticity in early language acquisition: The effects of prenatal and early childhood experience. *Current Opinion in Neurobiology*, *35*, 13–20.

Gervain, J. (2018). The role of prenatal experience in language development. *Current Opinion in Behavioral Sciences*, *21*, 62–67.

Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., Shukla, M., & Aslin, R. N. (2011). Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Developmental Cognitive Neuroscience*, *1*(1), 22–46.

Grossmann, T., & Friederici, A. D. (2012). When during development do our brains get tuned to the human voice? *Social Neuroscience*, *7*(4), 369–372.

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.

Issard, C., & Gervain, J. (2018). Variability of the hemodynamic response in infants: Influence of experimental design and stimulus complexity. *Developmental Cognitive Neuroscience*, *33*, 182–193.

Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Huang, H., Ye, H. H., Zhang, K., & Wang,

Z. (2003). Effects of experience on fetal voice recognition. *Psychological Science*, *14*(3), 220–224.

Kral, A., & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in Neurosciences*, *35*(2), 111–122.

Lloyd-Fox, S., Blasi, A., & Elwell, C. E. (2009). Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy. *Neuroscience and Biobehavioral Reviews*.

Saliba, J., Bortfeld, H., Levitin, D. J., & Oghalai, J. S. (2016). Functional near-infrared spectroscopy for neuroimaging in cochlear implant recipients. *Hearing Research*, *338*, 64–75.

Sevy, A. B., Bortfeld, H., Huppert, T. J., Beauchamp, M. S., Tonini, R. E., & Oghalai, J. S. (2010). Neuroimaging with near-infrared spectroscopy demonstrates speech-evoked activity in the auditory cortex of deaf children following cochlear implantation. *Hearing Research*, *270*(1–2), 39–47.

Sharma, A., & Campbell, J. (2011). A sensitive period for cochlear implantation in deaf children. *The Journal of Maternal-Fetal & Neonatal Medicine*, *24*(sup1), 151–153.

Sharma, A., & Dorman, M. (2011). Central auditory system development and plasticity after cochlear implantation. *Auditory Prostheses*, 233–255.

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, *66*, 173–196.

Yücel, M. A., Lühmann, A. v, Scholkmann, F., Gervain, J., Dan, I., Ayaz, H., Boas, D., Cooper, R. J., Culver, J., & Elwell, C. E. (2021). Best practices for fNIRS publications. *Neurophotonics*, *8*(1), 012101.