



Research Paper

Why does language not emerge until the second year?

Rhodri Cusack^{a, b, *}, Conor J. Wild^b, Leire Zubiaurre-Elorza^{b, c}, Annika C. Linke^{b, d}^a Trinity College Institute of Neuroscience, Trinity College Dublin, Ireland^b Brain and Mind Institute, Western University, London, Canada^c Department of Methods and Experimental Psychology, University of Deusto, Bilbao, Spain^d San Diego State University, San Diego, CA, USA

ARTICLE INFO

Article history:

Received 22 December 2017

Received in revised form

30 April 2018

Accepted 7 May 2018

Available online 9 May 2018

ABSTRACT

From their second year, infants typically begin to show rapid acquisition of receptive and expressive language. Here, we ask why these language skills do not begin to develop earlier. One evolutionary hypothesis is that infants are born when many brain systems are immature and not yet functioning, including those critical to language, because human infants have large heads and their mother's pelvis size is limited, necessitating an early birth. An alternative proposal, inspired by discoveries in machine learning, is that the language systems are mature enough to function but need auditory experience to develop effective representations of speech, before the language functions that manifest in behaviour can emerge. Growing evidence, in particular from neuroimaging, is supporting this latter hypothesis. We have previously shown with magnetic resonance imaging (MRI) that the acoustic radiation, carrying rich information to auditory cortex, is largely mature by 1 month, and using functional MRI (fMRI) that auditory cortex is processing many complex features of natural sounds by 3 months. However, speech perception relies upon a network of regions beyond auditory cortex, and it is not established if this network is mature. Here we measure the maturity of the speech network using functional connectivity with fMRI in infants at 3 months ($N = 6$) and 9 months ($N = 7$), and in an adult comparison group ($N = 15$). We find that functional connectivity in speech networks is mature at 3 months, suggesting that the delay in the onset of language is not due to brain immaturity but rather to the time needed to develop representations through experience. Future avenues for the study of language development are proposed, and the implications for clinical care and infant education are discussed.

© 2018 Published by Elsevier B.V.

1. Introduction

At the time of their first birthday, human infants understand and speak just a few words. Only by their second birthday are they learning new words rapidly, and have typically built a vocabulary of three hundred words (Bloom, 1976; Frank et al., 2017). Here, we ask why rapid word learning does not happen earlier. Given that infants hear a million spoken words per month (Hart and Risley, 1995a), dominated by high frequency tokens (Piantadosi, 2014), what delays rapid language acquisition for a year?

It is possible that late language acquisition might reflect broader initial sluggishness in the life history of human cognitive

development. Human infants are slow to develop in many ways and are helpless for a long time following birth, compared to animals with simpler brains (Jones et al., 2009). Consider motor function, for example: lambs and chickens are walking within a few days, while human infants take 9 months to crawl.

One explanation for this sluggish development is that human infants have large heads but the size of the mother's pelvis (Rosenberg and Trevathan, 2002) or the mother's metabolism (Dunsworth et al., 2012) are limited, and so birth must happen relatively early in gestation, while the infant's brain is still immature. This hypothesis continues to be influential and is supported by elegant modelling of evolutionary pressures that might have driven selection for intelligence (Piantadosi and Kidd, 2016). According to this hypothesis, human infants are helpless and their language delayed, because their brains must develop postnatally.

We propose an alternative explanation inspired by machine learning. Deep (many-layered) neural networks have in recent

* Corresponding author. Trinity College Institute of Neuroscience, Trinity College Dublin, College Green, Dublin 2, Ireland.

E-mail address: rhodri@cusacklab.org (R. Cusack).

years come to dominate artificial intelligence, performing many tasks better than humans, such as visual recognition, playing Go and Chess, and driving cars. Although deep neural networks have been around for decades, they were not useful in practice because they could not be effectively trained. Given the enormous number of degrees of freedom of these networks, they had the tendency to overfit to the initial training examples, and thus to learn to perform tasks in idiosyncratic ways that did not generalize well to new examples. An important breakthrough was the innovation that networks should be *pretrained* (Hinton and Salakhutdinov, 2006). During pretraining, the network learns the statistics of the sensory input ($P(S)$). This is then followed by training which categories are associated with particular sensory input ($p(c|S)$). Pretraining is often used in deep learning, and has been shown to be particularly beneficial for more complex neural networks (Erhan et al., 2010). By analogy, our proposal is that human infants, which have complex brains, benefit from pretraining. In the case of language, we propose that the first year is spent learning the statistics of sound and developing the motoric models that underlie production. Only in the second year are these representations sufficiently mature for rapid language acquisition to begin.

The “immature brain” and “pretraining” hypotheses cannot be distinguished by observing language development in the first year, as they both predict it will be slow to develop. However, they make different predictions at the neural level. The immature brain hypothesis predicts that language systems will be immature for the first year, while the pretraining hypothesis predicts that they will be functioning. These neural measures need to be made in infants, as other animals do not have language, and have simpler brains (and so less need for pretraining).

From previous work, on the one hand, there is evidence to suggest the language system is immature. The auditory cortex is one part of the language network that is important for processing complex sounds (Peelle et al., 2010). Post-mortem anatomical studies have emphasized the immaturity of auditory cortex, finding that in newborns it does not yet have discernible laminar structure, and that it does not receive myelinated projections from the thalamus (Moore and Linthicum, 2007). In an extrapolation from ideas proposed for the motor system (Marin-Padilla and Marin-Padilla, 1982) it has been suggested that the acoustic radiation, which in adults carries rich information from the thalamus to the auditory cortex, is not yet mature, and that auditory input before six months may only be directly through the reticular activating system in the brainstem (Eggermont and Moore, 2012). Structural connectivity in the cortical language network has also been measured using diffusion-weighted magnetic resonance imaging (MRI) and tractography, and some studies have found evidence for immaturity (Perani et al., 2011). These findings support the immaturity hypothesis. Other studies using neuroimaging in infants in the first year have suggested auditory function is more mature. The acoustic radiation was recently traced using tractography in infants through the first year. It was present as early as one month, and its microstructure, as assessed using fractional anisotropy (FA) and other methods, changed only subtly through the first year. This suggests that rich auditory information might, therefore, be delivered to auditory cortex early in the first year (Zubiaurre-Elorza et al., 2018).

Assuming that complex auditory information could be delivered, is auditory cortex ready to process it? In adults, structural asymmetries are seen in cortical language regions. These are evident even in preterm newborns (Dubois et al., 2010). To assess cortical processing of stimuli, functional MRI (fMRI) has been used to measure the brain activation to sounds. In the first months, auditory cortex responds to speech (Dehaene-Lambertz et al., 2006; Dehaene-Lambertz et al., 2002; Perani et al., 2011). It encodes not just simple acoustic characteristics (such as frequency

centroid, fundamental frequency, and envelope) but also more complex acoustic features in a similar way to adults (Wild et al., 2017). Hemispheric asymmetries are seen in activation in response to speech but not music in 2 month-old infants (Dehaene-Lambertz et al., 2010), similar to the activation pattern that would be seen in adults. Functional optical imaging has even found cortical responses differ by syllable in preterm infants (Mahmoudzadeh et al., 2013). These studies suggest that auditory cortex is processing the rich acoustic information that is being delivered and that this component of the language system is relatively mature, even in the first months. These results are consistent with the pretraining hypothesis.

A limitation of this evidence for brain maturity is that it focuses on auditory cortex, and there are many other components to the language network, including motor and prefrontal regions. Neuroimaging has shown that components of this system in the prefrontal cortex can be engaged by speech sounds in infants (Dehaene-Lambertz et al., 2006, 2002; Imada et al., 2006). However, it is not clear how mature the network is more broadly. It is not yet possible to identify tasks that will functionally activate each node of the broader network in infants, and so here we take a different approach, of characterising the network's maturity through its connectivity. Functional connectivity analysis has proven highly informative in adults, and is possible in infants (Doria et al., 2010; Fransson et al., 2007; Gao et al., 2011; Perani, 2012; Smyser and Neil, 2015). Specifically, across a broad language network comprising 15 regions we examine the pattern of connectivity – which connections are stronger and which are weaker – in adults and infants. Evidence of language network maturity would further support the pretraining hypothesis.

2. Methods

2.1. Identifying regions of the speech network

The speech network was identified using *neurosynth.org*, an open-source database of thousands of published functional MRI studies (Yarkoni et al., 2011). The keyword “speech” selected 839 studies that contained one or more contrasts of speech against a baseline. As speech production causes head movement, which is a problem for MRI, the vast majority of these contrasts reflected speech perception. These were compared to thousands of other contrasts that reflected other behaviours, to yield a z-score that quantified the likelihood that a speech task was being performed, given activation in a voxel (i.e., reverse inference $p(\text{speech task} | \text{activation})$) – i.e., regions that are selective to some degree for speech function. Fig. 1a shows the resulting map of the speech system, in standard adult (MNI) space.

This map was then parcellated in a semi-automatic way by identifying the lowest z-threshold that would separate each region from its neighbours. The final network of 15 regions is shown in Fig. 1b and comprised the midline supplementary motor area (SMA) and seven regions split into left and right components: the auditory cortex (AUD), thalamus (TH), cerebellum (CER), pallidum (PAL), prefrontal cortex (PFC), insula (INS), and intraparietal sulcus (IPS).

2.2. Participants

Approval for the study was obtained from Western University's Health Sciences Research Ethics Board and informed consent obtained from the adults and infant's caregivers. Sixteen adults (5 male, 11 female, 23 ± 5 years old), twenty-four 3-month old infants and fifteen 9-month old infants were recruited. The following inclusion criteria were applied: (1) no congenital abnormalities; (2)

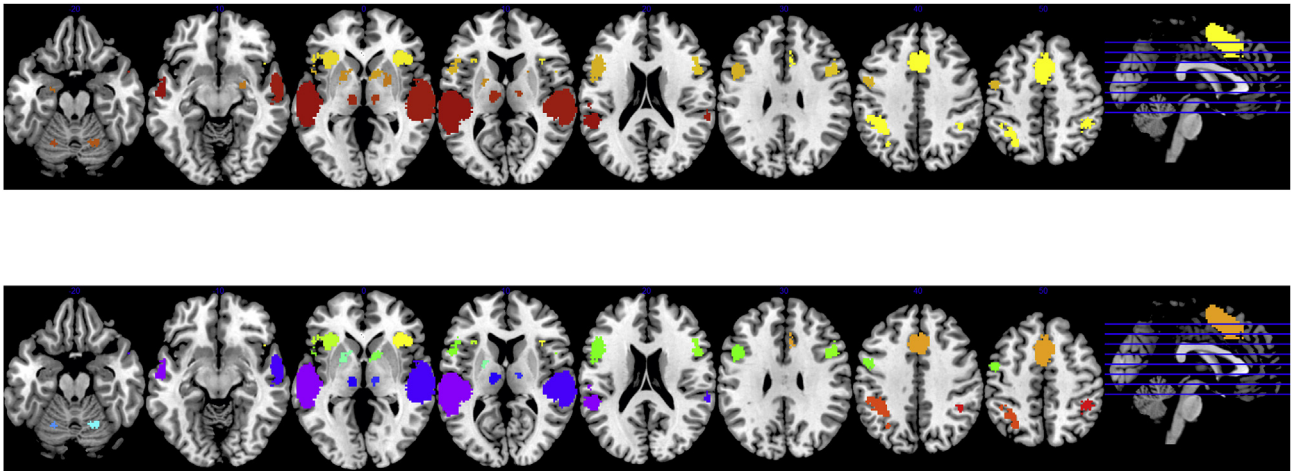


Fig. 1. Voxels activated more by speech tasks than by other tasks, derived from a meta-analysis of functional neuroimaging studies (neurosynth.org, keyword “speech”, reverse inference, threshold FDR $p < 0.01$). The raw z scores (range from 0 to 15) are shown in the top panel (a), and the same map parcellated into regions in the lower panel (b). Slices shown are from MNI $z = -20$ to $+50$ in steps of 10 mm.

no contraindications for MRI; (3) an age at MRI of 3 or 9 postnatal months or older than 18 years. One adult was rejected following the detection of an abnormal finding by the MR technologist ($n = 15$). Infants were scanned during natural sleep, without the use of sedation. A neonatal/perinatal nurse was present during the duration of the scan to monitor the infant’s well-being and assist in the case of a medical emergency. Not all infants could be persuaded to sleep, one 9-month old had poor coverage in inferior regions in the speech network, and following stringent exclusion for movement, which is particularly important when studying functional connectivity (Power et al., 2011), eighteen 3-month olds (yielding $n = 6$) and eight 9-month olds (yielding $n = 7$) were excluded from further analyses. The details of the participants included in the final analysis are shown in Table 1. Some infants were recruited through community advertising, and some via referral from a neonatologist at the neonatal intensive care unit (NICU) of London Health Sciences Centre. No brain injury was apparent in the MRI of any of the infants included in this study, although 4/12 were diagnosed with potential neurological challenges during their period at the NICU.

2.3. MRI acquisition

A Siemens Prisma 3T MRI scanner with a 12-channel head coil (Siemens, Erlangen Germany) was used to acquire two sessions (each 7.5 min long) of fMRI. Multiband acceleration (Feinberg et al., 2010; Xu et al., 2013; Nunes et al., 2006) was used to increase sampling rate and reduce sensitivity to movement (36 slices of

64×64 matrix size with 3×3 mm in-plane resolution, and slice thickness 3 mm, multiband factor 4). The echo-time was adjusted for the 3-month group ($TE = 40$ ms) to reflect the longer $T2^*$ relaxation due to increased water content in brain tissue (Rivkin et al., 2004), but a typical value was used in the 9-month and adult group ($TE = 30$ ms). There was slight variation in protocol between subjects, reflecting ongoing optimization (3 month: slice gap = $0-0.3$ mm, $TR = 776-861$ ms. Adults and 9 month: slice gap = $0-0.3$ mm; $TR = 686-861$ ms). Both T1 and $T2^*$ weighted structural images were acquired (36 oblique slices of 3 mm thickness, 64×64 matrix, voxel size $3 \times 3 \times 3$ mm³, $TR = 4000$ ms, $TE = 120$ ms) at the beginning of the MRI testing.

To detect atypical development, at the scanning visits we also acquired a standardised measure of language, the Receptive Expressive Emergent Language Test (Version 3, REEL-3), and a standardised measure of broader development, the Vineland Adaptive Behaviour Scales (Version 2, VABS-2).

2.4. Analysis

Data were analyzed with automatic analysis (aa) version 4.2 (Cusack et al., 2015) and SPM 8 in Matlab (Mathworks, Natick MA). The EPs were motion corrected, and de-noising was performed using a nine-parameter regression model that has been shown to be effective for functional connectivity analyses (Ciric et al., 2017a). The nine parameters comprised six motion parameters (three translations, three rotations) and three regional time series (mean

Table 1
Details of the infant participants.

ID	Sex	GA at Birth (weeks)	Age at Scan, Corrected for Prematurity (days)	Observations
3_1	M	29	92	
3_2	F	28	92	
3_3	F	41	89	
3_4	M	41	91	Hypoxia, stroke, seizures
3_5	F	41	92	Stroke, white matter hemorrhages
3_6	F	40	92	
9_1	M	40	273	
9_2	M	40	274	
9_3	M	27	273	
9_4	M	27	273	IVH grade II
9_5	M	29	273	
9_6	M	25	334	IVH grade II

white matter, mean cerebrospinal fluid, global mean). To remove low-frequency drift the time series were then high-pass filtered with a 128 s cut-off (0.008 Hz).

As part of another study that has been reported elsewhere (Wild et al., 2017), sounds were presented during fMRI in a block design with 26-s cycles (15-s on, 11-s off). However, the current manuscript examines brain connectivity rather than sound-evoked activation. Functional connectivity has been previously found to be similar between resting-state and tasks including those in which sounds were presented (Shah et al., 2016). To investigate any residual effect of sound on connectivity, we also repeated all analyses with the effect of sound regressed out. For this, as the hemodynamic response in infants can be different to adults (Arichi et al., 2012) we used a flexible finite impulse response (FIR) model, with 26 1-s bins modelling the 26 s cycles. This was highly effective in removing sound-evoked peaks in the frequency spectrum. In the main manuscript, we present results following this modelling, but for comparison in the supplementary materials we include the results without this regression, which are very similar, suggesting that functional connectivity of the language network is dominated by spontaneous rather than sound-evoked fluctuations.

To calculate the mapping between an individual adult's brain and the adult MNI template space, we used non-linear warping as implemented in SPM 8's "segment and normalise". The mapping between the infant and adult brain space was calculated with a two-stage transformation. First, the same algorithm was used to derive the non-linear mapping from an individual infant's brain to the one-year-old UNC template (Shi et al., 2011). Inspection by eye showed this template gave good normalization for all infants in this study. Next, the transformations from each of the UNC template to the adult MNI space was calculated, again using non-linear warping with SPM. Using the inverse of these two transformations in sequence, the ROIs derived from the adult meta-analysis in MNI space could then be projected back into the space of individual infants. This allowed the time-course of BOLD activity for each infant in each ROI to be extracted.

The functional connectivity between each pair of ROIs in the speech network was calculated using Pearson correlation between the de-noised and filtered time-series. For R ROIs, there are $\frac{R(R-1)}{2}$ unique pairwise comparisons. MRI sensitivity differs through development, due to the difference in brain size, the magnitude of the hemodynamic response (Arichi et al., 2012), and changes in the composition of brain tissue that affect MRI relaxation times (Rivkin et al., 2004). These will have affected the signal-to-noise ratio of the different groups, and so to remove this overall difference, we focussed upon the relative strength of connectivity between different regions-of-interest within an individual, by z-scoring the set of pairwise correlations for each individual. A necessary limitation of this approach is that it will not highlight changes in connectivity that add or scale the entire network equally. However, the interactive specialization framework proposes that tuning within brain regions is dependent on the balance of inputs from different regions (Johnson, 2011) and accordingly overall changes in connectivity would not be expected to contribute to maturation.

To quantify the similarity of the overall connectivity in the broader speech network between each pair of individuals, we used a second-order correlation between the $\frac{R(R-1)}{2}$ correlation values in one individual's connectivity matrix and the corresponding values for the other individual. As correlation is invariant to additive or multiplicative changes, the same result would have been obtained on the raw r values rather than the z-scored correlation values. To test for within-group consistency of the pattern, the resulting second-order correlations were then tested against zero with a one-sample t -test with $\frac{N(N-1)}{2} - 1$ degrees of freedom, where N is the number of subjects in the group. Under the null hypothesis that

the expected second-order correlation across subjects is zero, this set of correlations is independent (i.e., the covariance of the pairwise correlations is zero) ensuring the validity of the one-sample t -test (Nel, 1985; Wilson et al., 2008). We furthermore compared the similarity of connectivity in each infant to the mean of the adults, again using second-order correlation. The resulting r values across infants in each of the groups were tested for significance using a one-sample t -test. To test if any similarity between the infant groups and adults was driven by the distance between the regions (e.g., regions further apart having weaker connections), we repeated this analysis using partial correlation and inter-region distance as a nuisance variable.

3. Results

Fig. 2c and f shows the functional connectivity in the broader speech network in adults. Pairwise comparison between adults showed that the pattern of connectivity within the speech network was highly consistent within the adult cohort ($r = 0.65 \pm 0.01$, $t(119) = 62.77$, $p < 0.001$). Inter-hemispheric connectivity between homologous regions was notable for all of the cortical regions and the thalamus. The prefrontal cortices, insulae and SMA were also tightly interconnected, but the cerebellum was more weakly connected to the parietal and frontal regions, and to the pallidum.

Fig. 2a, b, d and e show the strength of pairwise connectivity between nodes in three- and nine-month olds. There was good consistency across subjects in the connectivity within the three-month group ($r = 0.37 \pm 0.04$, $t(14) = 9.06$, $p < 0.001$) and the nine-month group ($r = 0.43 \pm 0.03$, $t(20) = 12.89$, $p < 0.001$). Contrary to the prediction of immature connectivity prior to the development of function, both groups show a pattern of pairwise functional connectivity that is strikingly similar to the adults. Each infant's connectivity pattern was strongly similar to the adult mean, for both the three-month group ($r = 0.57 \pm 0.03$, $t(5) = 18.09$, $p < 0.001$) and the nine-month group ($r = 0.59 \pm 0.07$, $t(6) = 8.95$, $p < 0.001$).

In recent years, awareness has increased of the effect of movement on functional connectivity fMRI (Power et al., 2011; Van Dijk et al., 2011). In particular, longer-range connections are disrupted more by movement than shorter-range connections (Ciric et al., 2017b). It was important to establish that the similarity we observe between infants and adults does not merely reflect a movement-induced effect of distance between the nodes. We therefore repeated the correlation analyses, but using partial correlation with inter-node distance as a confounding variable to be removed. The results were similar within the 3 month group ($r = 0.38 \pm 0.04$, $t(119) = 66.17$, $p < 0.001$), the 9 month group ($r = 0.43 \pm 0.03$, $t(20) = 12.89$, $p < 0.001$), and between the adults and the 3 month group ($r = 0.57 \pm 0.03$, $t(5) = 18.80$, $p < 0.001$), and the adults and the 9 month group ($r = 0.59 \pm 0.06$, $t(6) = 9.10$, $p < 0.001$).

The REEL-3 was obtained in all of the infants, and the VABS-2 in all but one of the infants. These standardised scores are designed to have a mean of 100 and standard deviation of 15. None of the infants were below the cutoff that indicates impairment (< 70) on either test (3 months: REEL 98 ± 11 , VABS 100 ± 15 ; 9 months: REEL: 94 ± 19 , VABS 106 ± 13).

Finally, although the pairwise connectivity patterns were similar for the infants and adults, perhaps there was a higher-order difference in network structure. To investigate this, we used hierarchical clustering to group together nodes that are more connected to each other, for each of the three age groups. The results (Fig. 3) showed that the higher-order structure of each of the infant groups and the adults was also strikingly similar.

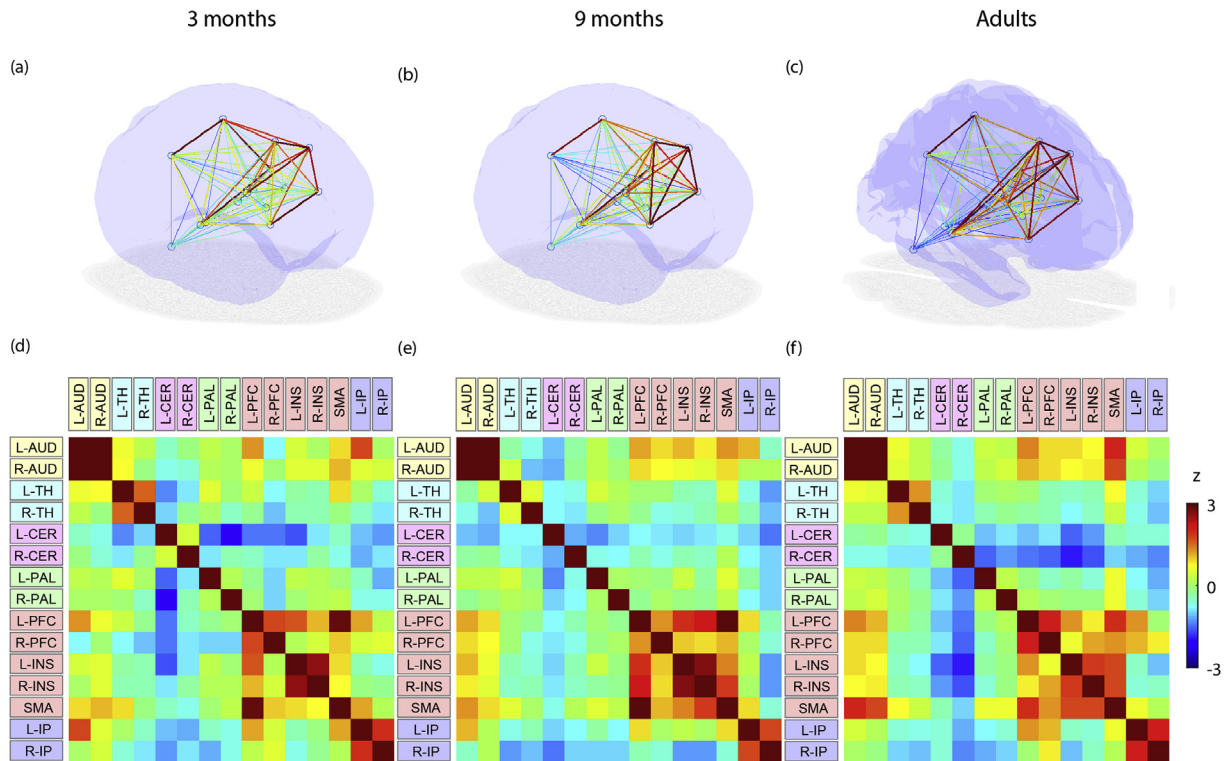


Fig. 2. Functional connectivity within speech network at 3 months (a,d), 9 months (b,e) and in adults (c,f). AUD = Auditory cortex, TH = thalamus; CER = cerebellum; PAL = pallidum; PFC = prefrontal cortex; INS = insula; SMA = supplementary motor area; IP = inferior parietal.

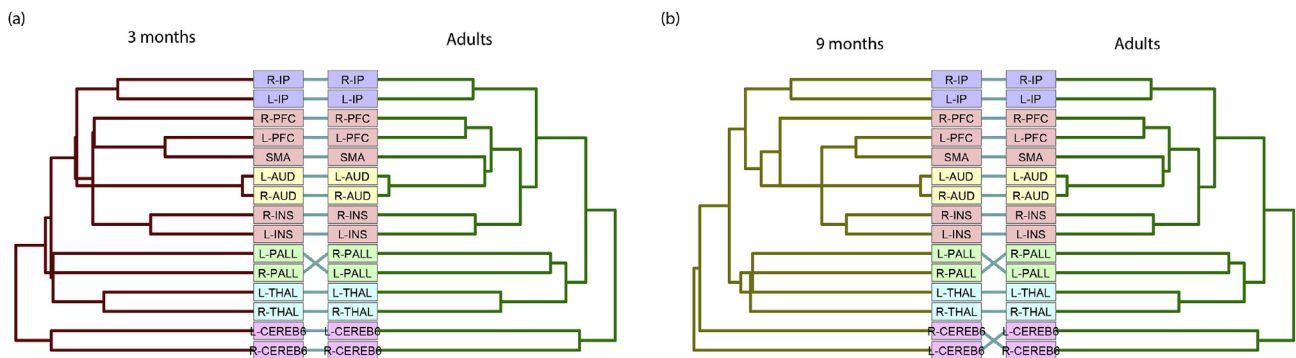


Fig. 3. Hierarchical clustering of the regions by similarity of connectivity, for the adults and the two infant groups.

4. Discussion

We found that at 3 and 9-months old the connectivity of the speech network, as measured with fMRI, was similar to adults. This was true both for the simple pairwise connectivity and the higher-order structure of connectivity, as assessed with hierarchical clustering. These results resonate with the gathering evidence of network maturity from diffusion tractography (Dubois et al., 2009, 2010; Zubiaurre-Elorza et al., 2018) and fMRI (Wild et al., 2017; Dehaene-Lambertz et al., 2002; G Dehaene-Lambertz et al., 2010; Dehaene-Lambertz et al., 2006), supporting the hypothesis that prior to the outwards manifestation of speech acquisition, there is a period of experience-dependent learning that shapes neural representations within these networks.

If the cortical speech network in infants is functioning, and developing representations of sound and language during their first year, it reemphasizes the importance of ensuring infants are

exposed to rich acoustic stimulation through their first year, and the importance of addressing socioeconomic differences in language exposure (Hart and Risley, 1995b; Noble et al., 2005). For premature babies, a rapidly growing body of evidence suggests that excessively quiet neonatal intensive care units (NICUs) might hinder neurodevelopment, by depriving infants of natural sounds during the critical period. Language development is sometimes delayed in premature infants (Vohr, 2013), particularly those from NICUs that were quieter (Pineda et al., 2014; Stromswold and Sheffield, 2004) or that had less talking (Caskey et al., 2014). Furthermore, structural neuroimaging has shown that the environment of the NICU affects the development of auditory brain regions (Pineda et al., 2014; Webb et al., 2015). The consequent language impairment has a tremendous impact on health, achievement and well-being (Barnett and Escobar, 1989; Beitchman et al., 2001; C. J. Johnson et al., 1999). It is therefore important to carefully design the acoustic environment of NICUs, to

encourage parents to talk to their children (<http://thirtymillionwords.org>, <https://stories.clintonfoundation.org/closing-the-word-gap-c1e73c068914>), and to investigate other auditory interventions in infancy (Gerry et al., 2012). Functional connectivity measures, of the kind evaluated in this study, could be informative in early detection of impairments, in tracking the effect of interventions, or in guiding the design of new interventions. In another study of a cohort of infants from the NICU we found that functional connectivity measured at term-equivalent age is predictive of motor skills at 4 and 8 months (Linke et al., n.d.). The current cohort was too small for the study of individual differences, but given the current findings it would be interesting to conduct a future study of the consequences of disruption to the speech network.

Although the early maturation of the speech network resonates with this growing literature on the importance of early language exposure, it also suggests a quite different perspective – which is that to some degree the network may be hard-wired. Future large-scale studies employing genotyping, rich environmental measurements, and parent/child imaging will be necessary to disentangle these contributions.

In the literature on early vocabulary development, there has been discussion on whether the “spurt” of rapid acquisition of words from around 18 months is due to the emergence of a new strategy for word learning (Mayor and Plunkett, 2010; Nazzi and Bertoncini, 2003), or whether it merely reflects the consequence of word frequency distributions within a single learning mechanism (McMurray, 2007). An analogous discussion might be held for the earlier stage of vocabulary development considered here, the start of word learning. Does it reflect a continuous learning process that could begin even before birth, passing some threshold that leads to outwards manifestation of language? In support of this, there is some evidence that the meaning of a small number of words is reliably understood much earlier than previously thought, by infants from around 6 months old (Bergelson and Swingle, 2012; Bergelson and Aslin, 2017). Alternatively, there may be a change in the cognitive processes engaged during pretraining and training, in the same way as there are qualitative changes in learning algorithm in machine learning.

There are limitations of our work that highlight the need for future studies. Although our results show that there is a remarkable similarity between the infant and adult language systems, it is likely that with a larger sample size, or with finer spatial resolution, differences will emerge. Forthcoming large-scale studies, such as the Development Human Connectome Project (dHCP) (<http://www.developingconnectome.org>) and Baby Connectome Project (<http://babyconnectomeproject.org>) will be helpful in this regard. Larger sample sizes will reduce the potential for a type II error, and consistent differences in connectivity may emerge, perhaps at a finer scale within the networks. Greater sensitivity will also become possible with longitudinal imaging (Cusack et al., 2018). Complementary methods, such as electroencephalography (EEG) will also provide valuable information, although as with MRI, care must be taken to distinguish developmental differences in the signal reflect a change in the underlying neural processing, and what changes reflect changes in the coupling between the neural processing and the measured signal. Furthermore, future studies could assess developing representations of language, perhaps using multivoxel-pattern analysis with fMRI, and directly test the hypothesis that while language network circuitry is mature, speech representations within the network are not and only emerge with auditory training.

In summary, we find that the functional connectivity of the speech network, as measured with fMRI, is largely mature in young infants. This, in combination with a host of other neuroimaging,

behavioural and clinical measures, suggests that the delay in the emergence of language is not due to brain network immaturity, but rather due to a pretraining period of “quiet learning.”

Acknowledgements

This research was supported by support from the Canada Excellence Research Chairs, Government of Canada in Cognitive Neuroimaging, a Natural Sciences and Engineering Research Council of Canada Discovery Grant [NSERC 418293DG-2012], CIHR/NSERC Collaborative Health Research Project [CHRP 201110CPG] and the Children's Health Research Institute (to RC) from the Children's Heart Foundation. We give special thanks to the families and infants who participated in this study.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.heares.2018.05.004>.

References

- Arichi, T., Fagiolo, G., Varela, M., Melendez-Calderon, A., Allievi, A., Merchant, N., et al., 2012. Development of BOLD signal hemodynamic responses in the human brain. *NeuroImage* 63 (2), 663–673. <https://doi.org/10.1016/j.neuroimage.2012.06.054>.
- Barnett, W.S., Escobar, C.M., 1989. Research on the cost effectiveness of early educational intervention: implications for research and policy. *Am. J. Community Psychol.* 17 (6), 677–704. Retrieved from. <http://link.springer.com/10.1007/BF00922734>.
- Beitchman, J.H., Wilson, B., Johnson, C.J., Atkinson, L., Young, A., Adlaf, E., et al., 2001. Fourteen-year follow-up of speech/language-impaired and control children: psychiatric outcome. *J. Am. Acad. Child Adolesc. Psychiatry* 40 (1), 75–82. Retrieved from. <http://www.jaacap.com/article/S0890856709608182/fulltext>.
- Bergelson, E., Aslin, R.N., 2017. Nature and origins of the lexicon in 6-mo-olds. *Proc. Natl. Acad. Sci. U. S. A.* 114 (49), 12916–12921. <https://doi.org/10.1073/pnas.1712966114>.
- Bergelson, E., Swingle, D., 2012. At 6–9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci.* 109 (9), 3253–3258. <https://doi.org/10.1073/pnas.1113380109>.
- Bloom, L., 1976. *One Word at a Time: the Use of Single Word Utterances before Syntax*. Mouton.
- Caskey, M., Stephens, B., Tucker, R., Vohr, B., 2014. Adult talk in the NICU with preterm infants and developmental outcomes. *Pediatrics* 133 (3), e578–e584. <https://doi.org/10.1542/peds.2013-0104>.
- Ciric, R., Wolf, D.H., Power, J.D., Roalf, D.R., Baum, G.L., Ruparel, K., et al., 2017a. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *NeuroImage* 154, 174–187. <https://doi.org/10.1016/j.neuroimage.2017.03.020>.
- Ciric, R., Wolf, D.H., Power, J.D., Roalf, D.R., Baum, G.L., Ruparel, K., et al., 2017b. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *NeuroImage* 154, 174–187. <https://doi.org/10.1016/j.neuroimage.2017.03.020>.
- Cusack, R., McCuaig, O., Linke, A.C.A.C., 2018. Methodological challenges in the comparison of infant fMRI across age groups. *Dev. Cogn. Neurosci.* Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S1878929317300658>.
- Cusack, R., Wild, C., Linke, A.C., Arichi, T., Lee, D.S.C., Han, V.K., 2015. Optimizing stimulation and analysis protocols for neonatal fMRI. *PLoS One*. <https://doi.org/10.1371/journal.pone.0120202>.
- Dehaene-Lambertz, G., Dehaene, S., Hertz-Pannier, L., 2002. Functional neuroimaging of speech perception in infants. *Sci. (New York, N.Y.)* 298 (5600), 2013–2015. <https://doi.org/10.1126/science.1077066>.
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., Dehaene, S., 2006. Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc. Natl. Acad. Sci. U. S. A.* 103 (38), 14240–14245. <https://doi.org/10.1073/pnas.0606302103>.
- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, L., Dubois, J., Hertz-Pannier, L., Dehaene, S., 2010. Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain Lang.* 114 (2), 53–65. <https://doi.org/10.1016/j.bandl.2009.09.003>.
- Doria, V., Beckmann, C.F., Arichi, T., Merchant, N., Groppo, M., Turkheimer, F.E., et al., 2010. Emergence of resting state networks in the preterm human brain. *Proc. Natl. Acad. Sci. U. S. A.* 107 (46), 20015–20020. <https://doi.org/10.1073/pnas.1007921107>.
- Dubois, J., Benders, M., Lazeyras, F., Borradori-Tolsa, C., Leuchter, R.H.-V., Mangin, J.F., Hüppi, P.S., 2010. Structural asymmetries of perisylvian regions in the preterm newborn. *NeuroImage* 52 (1), 32–42. <https://doi.org/10.1016/j.neuroimage.2010.03.054>.

- Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J.F., Le Bihan, D., Dehaene-Lambertz, G., 2009. Structural asymmetries in the infant language and sensorimotor networks. *Cereb. Cortex* (New York, N.Y. 1991) 19 (2), 414–423. <https://doi.org/10.1093/cercor/bhn097>.
- Dunsworth, H.M., Warrener, A.G., Deacon, T., Ellison, P.T., Pontzer, H., 2012. Metabolic hypothesis for human altriciality. *Proc. Natl. Acad. Sci.* 109 (38), 15212–15216. <https://doi.org/10.1073/pnas.1205282109>.
- Eggermont, J.J., Moore, J.K., 2012. Morphological and functional development of the auditory nervous system. *Hum. Audit. Dev.*
- Erhan, D., Courville, A., Vincent, P., 2010. Why does unsupervised pre-training help deep Learning? *J. Mach. Learn. Res.* 11 (March 2014), 625–660. <https://doi.org/10.1145/1756006.1756025>.
- Feinberg, David A., Moeller, Steen, Smith, Stephen M., Auerbach, Edward, Ramanna, Sudhir, Gunther, Matthias, Glasser, Matt F., Miller, Karla L., Ugurbil, Kamil, Yacoub, Essa, 2010. Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. *PLoS One* 5 (12), e15710.
- Frank, M.C., Braginsky, M., Yurovsky, D., Marchman, V.A., 2017. Wordbank: an open repository for developmental vocabulary data. *J. Child Lang.* 44 (3), 677–694. <https://doi.org/10.1017/S0305000916000209>.
- Fransson, P., Skögl, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., Aden, U., 2007. Resting-state networks in the infant brain. *Proc. Natl. Acad. Sci. U. S. A.* 104 (39), 15531–15536. <https://doi.org/10.1073/pnas.0704380104>.
- Gao, W., Gilmore, J.H., Giovanello, K.S., Smith, J.K., Shen, D., Zhu, H., Lin, W., 2011. Temporal and spatial evolution of brain network topology during the first two years of life. *PLoS ONE* 6 (9). <https://doi.org/10.1371/journal.pone.0025278>.
- Gerry, D., Unrau, A., Trainor, L.J., 2012. Active music classes in infancy enhance musical, communicative and social development. *Dev. Sci.* 15 (3), 398–407. <https://doi.org/10.1111/j.1467-7687.2012.01142.x>.
- Hart, B., Risley, T.R., 1995a. Meaningful Differences in the Everyday Experience of Young American Children. Paul H Brookes Publishing, Baltimore, MD. Retrieved from. <http://www.amazon.ca/Meaningful-Differences-Everyday-Experience-American/dp/1557661979>.
- Hart, B., Risley, T.R., 1995b. Meaningful Differences in the Everyday Experience of Young American Children. Brookes Publishing, Baltimore, MD. Retrieved from. <http://www.amazon.ca/Meaningful-Differences-Everyday-Experience-American/dp/1557661979>.
- Hinton, G., Salakhutdinov, R., 2006. Reducing the dimensionality of data with neural networks. *Science* 313 (July), 504–507. Retrieved from. <http://www.sciencemag.org/content/313/5786/504.short>.
- Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., Kuhl, P.K., 2006. Infant speech perception activates Broca's area: a developmental magnetoencephalography study. *Learning* 17 (10).
- Johnson, C.J., Beitchman, J.H., Young, A., Escobar, M., Atkinson, L., Wilson, B., et al., 1999. Fourteen-year follow-up of children with and without speech/language impairments. *J. Speech Lang. Hear. Res.* 42 (3), 744. Retrieved from. <http://jslhr.pubs.asha.org/article.aspx?articleid=1780992>.
- Johnson, M.H., 2011. Interactive Specialization: a domain-general framework for human functional brain development? *Dev. Cogn. Neurosci.* 1 (1), 7–21. <https://doi.org/10.1016/j.dcn.2010.07.003>.
- Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O'Dell, J., Orme, C.D.L., et al., 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90 (9), 2648. <https://doi.org/10.1890/08-1494.1>.
- Linke, A.C., Wild, C.J., Zubiaurre-Elorza, L., Herzmann, C., Duffy, H., Han, V.K., et al. (n.d.). Resting State Functional Connectivity Networks in Neonates with Perinatal Brain Injury.
- Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., et al., 2013. Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proc. Natl. Acad. Sci.* 110 (12), 1–6. <https://doi.org/10.1073/pnas.1212220110>.
- Marin-Padilla, M., Marin-Padilla, T.M., 1982. Anatomy and Embryology Origin, Prenatal Development and Structural Organization of Layer I of the Human Cerebral (Motor) Cortex, pp. 161–206.
- Mayor, J., Plunkett, K., 2010. Vocabulary Spurt: Are Infants Full of Zipf?, i, pp. 836–841. Retrieved from. <http://csjarchive.cogsci.rpi.edu/Proceedings/2010/papers/0285/paper0285.pdf>.
- McMurray, B., 2007. Moo-cow! Mummy! More! How do children learn so many words? *Significance* 4 (4), 159–163. <https://doi.org/10.1111/j.1740-9713.2007.00258.x>.
- Moore, J.K., Linthicum, F.H., 2007. The human auditory system: a timeline of development. *Int. J. Audiology* 46 (9), 460–478. <https://doi.org/10.1080/14992020701383019>.
- Nazzi, T., Bertoncini, J., 2003. Before and after the vocabulary spurt: two modes of word acquisition? *Dev. Sci.* 6 (2), 136–142. <https://doi.org/10.1111/1467-7687.00263>.
- Nel, D.G., 1985. A matrix derivation of the asymptotic covariance matrix of sample correlation coefficients. *Linear Algebra Its Appl.* 67 (C), 137–145. [https://doi.org/10.1016/0024-3795\(85\)90191-0](https://doi.org/10.1016/0024-3795(85)90191-0).
- Noble, K.G., Norman, M.F., Farah, M.J., 2005. Neurocognitive correlates of socioeconomic status in kindergarten children. *Dev. Sci.* 8 (1), 74–87. <https://doi.org/10.1111/j.1467-7687.2005.00394.x>.
- Nunes, R.G., Hajnal, J.V., Golay, X., Larkman, D.J., 2006. Simultaneous slice excitation and reconstruction for single shot EPI. *Proc. Intl. Soc. Mag. Reson. Med.* 14.
- Peelle, J.E., Johnsrude, I.S., Davis, M.H., 2010. Hierarchical processing for speech in human auditory cortex and beyond. *Front. Hum. Neurosci.* 4 (51). <https://doi.org/10.3389/fnhum.2010.00051>.
- Perani, D., 2012. Functional and Structural Connectivity for Language and Music Processing at Birth. New York. <https://doi.org/10.1007/s12210-012-0168-2>.
- Perani, D., Saccuman, M.C., Scifo, P., Awander, A., Spada, D., Baldoli, C., et al., 2011. Neural language networks at birth. *Proc. Natl. Acad. Sci.* 108 (38), 1–6. <https://doi.org/10.1073/pnas.1102991108>.
- Piantadosi, S.T., 2014. Zipf's word frequency law in natural language: a critical review and future directions. *Psychon. Bull. Rev.* 21 (5), 1112–1130. <https://doi.org/10.3758/s13423-014-0585-6>.
- Piantadosi, S.T., Kidd, C., 2016. Extraordinary intelligence and the care of infants. *Proc. Natl. Acad. Sci.* 113 (25), 6874–6879. <https://doi.org/10.1073/pnas.1506752113>.
- Pineda, R.G., Neil, J., Dierker, D., Smyser, C.D., Wallendorf, M., Kidokoro, H., et al., 2014. Alterations in brain structure and neurodevelopmental outcome in pre-term infants hospitalized in different neonatal intensive care unit environments. *J. Pediatr.* 164 (1), 52–60.e2. <https://doi.org/10.1016/j.jpeds.2013.08.047>.
- Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2011. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* 59 (3), 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>.
- Rivkin, M.J., Wolraich, D., Als, H., McAnulty, G., Butler, S., Conneman, N., et al., 2004. Prolonged T2 values in newborn versus adult brain: implications for fMRI studies of newborns. *Magnetic Reson. Med.* 51 (6), 1287–1291. <https://doi.org/10.1002/mrm.20098>.
- Rosenberg, K., Trevathan, W., 2002. Birth, obstetrics and human evolution. *BJOG Int. J. Obstetrics Gynaecol.* 109 (11), 1199–1206. <https://doi.org/10.1046/j.1471-0528.2002.00010.x>.
- Shah, L.M., Cramer, J.A., Ferguson, M.A., Birn, R.M., Anderson, J.S., 2016. Reliability and reproducibility of individual differences in functional connectivity acquired during task and resting state. *Brain Behav.* 6 (5), e00456. <https://doi.org/10.1002/brb3.456>.
- Shi, F., Yap, P.-T., Wu, G., Jia, H., Gilmore, J.H., Lin, W., Shen, D., 2011. Infant brain atlases from neonates to 1- and 2-year-olds. *PLoS One* 6 (4), e18746. <https://doi.org/10.1371/journal.pone.0018746>.
- Smyser, C.D., Neil, J.J., 2015. Use of resting-state functional MRI to study brain development and injury in neonates. *Seminars Perinatol* 39 (2), 130–140. <https://doi.org/10.1053/j.semperi.2015.01.006>.
- Stromswold, K., Sheffield, E., 2004. Neonatal Intensive Care Unit Noise & Language Development. Rutgers University Center for Retrieved from. http://rucss.rutgers.edu/tech_rpt/NICU_Noise_TR76.pdf.
- Van Dijk, K.R.A., Sabuncu, M.R., Buckner, R.L., 2011. The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage* 59 (1), 431–438. <https://doi.org/10.1016/j.neuroimage.2011.07.044>.
- Vohr, B., 2013. Speech and language outcomes of very preterm infants. *Seminars Fetal & Neonatal Med.* Retrieved from <http://www.sciencedirect.com/science/article/pii/S1744165X13000991>.
- Webb, A.R., Heller, H.T., Benson, C.B., Lahav, A., 2015. Mother's voice and heartbeat sounds elicit auditory plasticity in the human brain before full gestation. *Proc. Natl. Acad. Sci. U. S. A.* 1–6. <https://doi.org/10.1073/pnas.1414924112>.
- Wild, C., Linke, A.C., Zubiaurre-Elorza, L., Duffy, H., Herzmann, C., Han, V.K., et al., 2017. Processing of Complex Sounds in the Auditory Cortex in the First Year.
- Wilson, S.M., Molnar-Szakacs, I., Iacoboni, M., 2008. Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cereb. Cortex* (New York, N.Y. 1991) 18 (1), 230–242. <https://doi.org/10.1093/cercor/bhm049>.
- Xu, Junqian, Moeller, Steen, Auerbach, Edward J., Strupp, John, Smith, Stephen M., Feinberg, David A., Yacoub, Essa, Ugurbil, Kamil, 2013. Evaluation of slice accelerations using multiband echo planar imaging at 3 T. *NeuroImage* 83 (December), 991–1001.
- Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D., 2011. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* 8 (8), 665–670. <https://doi.org/10.1038/nmeth.1635>.
- Zubiaurre-Elorza, L., Linke, A.C., Herzmann, C., Wild, C.J., Duffy, H., Lee, D.S., et al., 2018. Auditory structural connectivity in preterm and healthy term infants during the first postnatal year. *Dev. Psychobiol.* <https://doi.org/10.1002/dev.21610>.