Effect of listening difficulties on speech cortical activation and connectivity in children

Abbreviated title: Cortical speech networks and listening

Hannah J. Stewart [1, 2], Erin K. Cash [2], Thomas Maloney [3], Jennifer Vannest [2, 4] and David R. Moore [2, 5, 6]

1. Division of Psychology and Language Sciences, University College London, London, UK
2. Communication Sciences Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA
3. Pediatric Neuroimaging Research Consortium, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA
4. Department of Pediatrics, University of Cincinnati College of Medicine, Ohio, USA
5. Department of Otolaryngology, College of Medicine, University of Cincinnati, Cincinnati, Ohio, USA
6. Manchester Centre for Audiology and Deafness, University of Manchester, Manchester, M13 9PL, UK

Correspondence: Hannah J. Stewart
Division of Psychology and Language Sciences
University College London
London UK
h.stewart@ucl.ac.uk

Keywords: listening difficulties, pediatric, fMRI, resting state, connectivity, speech perception

Abstract:
Listening difficulties (LiD) are a widespread but poorly understood form of hearing impairment in people who have normal audiometry. Recent research suggests that childhood LiD are cognitive rather than auditory in origin. We assessed that hypothesis using behavioral testing and fMRI with 43 typically developing children and 42 children with LiD, categorized by caregiver report. The children with LiD had clinically normal hearing. For sentence listening tasks, we found no group differences in fMRI brain cortical activation by increasingly complex speech, from phonology to intelligibility to semantics. Using unstimulated fMRI connectivity analysis, we examined how cortical auditory and related speech perception networks interact. Group differences were found only in cortical connections engaged by more complex speech processing. Strength of the affected connections was related to the children's performance on tests of dichotic listening, speech-in-noise, attention, memory and verbal vocabulary. Childhood LiD thus reflects cognitive and language rather than auditory deficits.

Abbreviations:

CANS central auditory nervous system
DLD developmental language disorder
rs-fMRI resting state-fMRI
FOV field-of-view
GLM general linear model
HUSH Hemodynamics Unrelated to Sounds of Hardware
LiD listening difficulties
MCC multiple comparison correction
MEG magnetoencephalography
MTG middle temporal gyrus
MVPA multi-voxel pattern analysis
p-FDR false detection rate
ROI region-of-interest
SLI specific language impairment
STG superior temporal gyrus
STS superior temporal sulcus

TD typical developing

TE echo time

TMS transcranial magnetic stimulation

TR repetition time

INTRODUCTION

Hearing loss is the one of most prevalent causes, globally, of years lived with disability (Vos et al., 2016). However, about half of US adults who have difficulty hearing or listening have a normal pure-tone audiogram, the standard clinical test of hearing loss (Edwards, 2020). Their difficulty may be due to a number of causes including 'sub-clinical' ear pathology, and impaired central auditory nervous system (CANS) or higher cortical function (Moore, 2018). Listening difficulties (LiD) are also frequently reported among children with clinically normal hearing, but with academic, language and attention problems (Moore et al., 2018), a condition sometimes referred to as developmental 'auditory processing disorder'.

The mechanisms underlying childhood LiD without hearing loss remain poorly understood. Speech understanding requires coalescence of acoustic features from bottom-up streaming in the auditory pathway with top-down linguistic processes deriving from widespread cortical areas involved with semantic representation and cognitive functions, notably memory and attention (Hickok & Poeppel, 2007; Rönnberg et al., 2019; Shinn-Cunningham, 2008). Whether LiD are primarily bottom-up or top-down in nature is a long-standing debate (Neijenhuis et al., 2019). Interactions between cortical auditory and related speech perception systems, resulting in impaired higher level processing of target sounds, may hold the key since poor performance on complex auditory tasks such as speech listening in noise is more associated with cognitive function than with more simple auditory tasks (Moore et al., 2010). To identify cortical areas used in this process, we adopted fMRI methods developed to examine the hierarchy of cortical speech processing in adults (Halai et al., 2015; Scott, 2000).

RESULTS

LiD - auditory or cognitive?

Child participants (6-13 years old; Table 1) all had sensitive audiograms (0.25 – 16 kHz, Figure 1A), normal ear function (David R. Moore et al., 2020) and typical auditory brainstem and envelope following responses (Blankenship et al., in prep.), suggestive of normal sub-cortical auditory function. Children scoring within the bottom 10th percentile on a caregiver checklist of everyday listening skills, the ECLiPS (Barry et al., 2015; Roebuck & Barry, 2018), and children

with a diagnosis of auditory processing disorder (see Methods), were classified as having LiD (Figure 1B).

We used task-based and resting state fMRI (rs-fMRI) to assess how speech perception networks function in typically developing (TD) children and children with LiD. Between task-based MRI acquisition intervals (Figure 1C), children listened to clear ('human'; Figure 1D) or distorted ('alien') spoken sentences that we asked them to match to a visual cartoon representation. Distorted sentences were speech that was rotated (Figure 1E) or rotated AND vocoded (Figure 1F) to retain prosody while successively eliminating meaning and spectral cues. We were able to explore the neural processing of different speech features by contrasting the different sentence types. By contrasting sentences that had intonation, phonetics and prosody with those having only prosody (Phonology: rotated > rotated+vocoded), we targeted the bridge from bottom-up to top-down processing where auditory information is classified into linguistically meaningful units (Brodbeck et al., 2018). As part of the top-down processes, sentences are judged for intelligibility by the listener. We accessed this judgement by contrasting sentences that had intelligibility, intonation, phonetics and prosody to those retaining intonation, phonetics and prosody (Intelligibility: clear > rotated). Finally, meaning was attached to the intelligible sentence, a process we accessed by contrasting two sentence types, one with all speech attributes and the other lacking all attributes except prosody (Semantics: clear > rotated+vocoded).

All children responded with high accuracy on the sentence recognition task, and did so rapidly, (Figure 1G) suggesting that they maintained attention throughout the task. TD children were more accurate than children with LiD ($F(1, 83) = 8.77$, $p = .004$, $\eta_p^2 = .096$), but there was no difference in accuracy between sentence types ($F(1.78, 148.03) = 2.11$, $p = .13$, $\eta_p^2 = .025$. There was no significant difference in reaction time between the groups ($F(1, 83) = .89$, $p = .35$, $\eta_p^2 = .011$). While there was no difference in accuracy between sentence types ($F(1.78, 148.03) = 2.11$, $p = .13$, $\eta_p^2 = .025$), across all participants, clear sentences were responded to fastest ($F(2, 166) = 11.50$, $p < .001$, $\eta_p^2 = .12$) compared to rotated ($p < .001$, $d = -.41$) and rotated+vocoded ($p < .001$, $d = -.46$) sentences. There were no interactions.

<Figure 1: audiograms, ECLiPS, paradigm, stimuli, task accuracy/RT>

<Table 1: participant details>

Cortical activation

Consistent with a top-down, linguistic model, we predicted that all stimuli would activate the primary auditory cortex (i.e., Heschl's gyrus) equally in all children, but that intelligibility and semantic stimuli would activate the speech processing areas (e.g. superior temporal gyrus, Wernicke's) less in children with LiD than in TD children. However, the two groups did not differ significantly in cortical activation in the Phonology, Intelligibility or Semantics contrasts after correcting for multiple comparisons (Figures 2A-C, threshold = 2.3, MCC .95). Averaged across all children (n = 85), the contrasts (Figure D-F) showed bilateral activation for all stages of speech listening (Phonology, Intelligibility and Semantics). Coordinates for the maximum intensity of the activated regions are shown in Supplementary Material Table 2. The Phonology contrast (rotated > rotated+vocoded; Figure 2D) showed bilateral activation in the middle and superior temporal gyrus, including Heschl's gyrus, and temporal pole. Activation was also found in the left hemisphere in the temporal fusiform cortex, angular gyrus and lateral occipital cortex. The Intelligibility contrast (clear > rotated; Figure 2E) produced similar activation to the Phonology contrast, bilaterally in the middle and superior temporal gyrus (anterior and posterior) and left frontal orbital cortex. Activation extended anteriorly along the left temporal gyrus and into Broca's area. The Semantics contrast (clear > rotated+vocoded; Figure 2F) also showed bilateral activation in the auditory cortices (middle and superior temporal gyrus including Heschl's gyrus and planum temporale) with activation extending along the left temporal fusiform and frontal orbital cortices and right parahippocampal gyrus.

<Figure 2 - fMRI activation>

Resting state fMRI

We predicted that children with LiD would show diminished functional connectivity in speech perception networks compared with TD children. To investigate how those cortical areas work together, we used the cortical activation results to inform a region-of-interest (ROI) based functional connectivity analysis during a separate, rs-fMRI acquired in the same scanning

session. The rs-fMRI scan allowed measurement of the temporal correlation of spontaneous fluctuations in the BOLD signal between anatomically separated brain regions (Biswal, 2012), capturing spontaneous interactions between brain regions in functionally associated networks (Damoiseaux et al., 2006). Networks of ROIs suitable for rs-fMRI connectivity analysis were created by parceling activation produced by each sentence task contrast (Phonology, Intelligibility, and Semantics), combined across groups, using regions from the pediatric ADHD-200 sample (Bellec et al., 2017; Craddock et al., 2012). This data-driven, spatially constrained method divided Phonology-derived cortical activity into 17 ROIs, Intelligibility into 23 ROIs, and Semantics into 31 ROIs (Figure 3; Brodman's area, maximum intensity coordinates and ROI sizes are in Supplementary Material Table 1). Connectivity was compared between groups (TD = 42, LiD = 39), controlling for age, using a general linear model (GLM) for each of the three networks (Figure 3A-C, Supplementary Material Table 1).

Figure 3 (D-F) shows resting state functional connectivity in the three speech networks, summed across groups. In the Phonology network, each group had connectivity among regions covering bilateral middle and superior temporal gyri, temporal pole, planum temporale, and left planum polare and supramarginal gyrus (Figure 3D). No significant group differences (p-FDR) were found in the Phonology network (Figure 4A). Each group had connectivity within the Intelligibility network covering bilateral middle/superior temporal gyrus and temporal pole along with left pars opercularis, frontal orbital cortex and supramarginal gyrus (Figure 3E). After false discovery rate (FDR) correction, TD children were found to have a significantly weaker temporal correlation between ROIs in Broca's area (left pars opercularis) and left middle temporal gyrus (posterior) compared to children with LiD (connection 14-23 in Figures 4B, D, Table 3). In the Semantics network, both groups of children had connectivity between bilateral middle and superior temporal gyrus, left Heschl's gyrus, pars triangularis, frontal orbital cortex, planum temporale, temporal fusiform gyrus and right parahippocampal gyrus and planum polare (Figure 3F). Group comparisons showed that, compared to children with LiD, the TD children had weaker temporal correlations between ROIs in the right parahippocampal gyrus and left Heschl's gyrus, left middle temporal gyrus, right superior temporal gyrus and right planum temporale (connections 16-9, 16-19, 16-1 and 16-5, respectively, in Figures 4C, D, Table 3). However, they also had stronger temporal correlations between ROIs in the left temporal fusiform cortex and

right superior temporal gyrus (connections 10-8 and 10-27). Figure 4D shows this polarity of hyperconnectivity, among children with LiD, for connections associated with the right parahippocampal gyrus (connections 16-1, 16-5, 16-9, 16-19) and, for TD children, connections associated with the left fusiform cortex (connections 10-8, 10-27). Figure 4D also shows the strong effect sizes of each of these group differences.

<Figure 3 - parcellated ROIs and listening networks>
<Table 2 - selected regions of interest used in rs-fMRI>
<Figure 4 - connectivity between groups and effect sizes>
<Table 3 - significant connectivity stat values>

Relation of brain connectivity to behavioral measures

We finally explored how the cortical connectivity related to behavioral tasks assessing speech in noise ability, dichotic listening and cognition. Interpretation of direction of functional connectivity is ambiguous; an increase in connectivity strength does not necessarily mean improved ability (Parente et al., 2018). Many of the behavioral measures did relate significantly to connectivity (Figure 5, Supplementary Material Table 3 and Figure 1). For Semantics connections involving the left posterior temporal fusiform cortex, positive correlations were found for dichotic listening ability (SCAN: competing words and sentences). In contrast, remaining Intelligibility and Semantics connections had a series of negative correlations for speech in noise and dichotic listening ability (SCAN: auditory figure ground and competing words) and cognition (NIH toolbox: verbal vocabulary, selective attention, task switching and episodic memory). These correlations show that, as the children's behavioural test scores improved, the connection strength within the Intelligibility network decreased, Semantics network connections linked to the right posterior parahippocampal gyrus decreased, and connections linked to the left posterior temporal fusiform cortex increased.

<Figure 5 - heat map of brain connectivity and behavioural measures>

DISCUSSION

Using an active sentence listening task we found that areas of cortical activation extended bilaterally across the superior temporal lobe and into memory and language areas (e.g., right parahippocampal gyrus, Broca's, left temporal fusiform cortex) as the stimuli presented increasingly complex speech elements from phonology to intelligibility to clear speech. These data shed new light on speech processing in the brain. Cortical areas of activation were unaffected by LiD. Functional connectivity between speech activation-defined sites (i) increased in spread with speech complexity, and (ii) differed between TD and children with LiD only during processing of more complex speech.

Behavioural measures of performance were consistent with these observations. No significant correlations were observed between temporal connectivity and sensory-dominant tasks (e.g, LiSN-S advantage scores and SCAN filtered words). However, significant correlations were observed between temporal connectivity and more cognitive-dominant tasks, especially speech-in-noise and task switching. Overall, the data are consistent with the hypothesis that children with LiD are primarily affected by cognitive deficits.

Lateralization of speech in 6-12 year old children

Lateralization of speech processing in adults has been debated across and within imaging modalities (fMRI see Evans & McGettigan, 2017; EEG e.g., Assaneo et al., 2019). For example, Rauschecker & Scott's (2009) unilateral model suggested that the left anterior STG is the hub of successful speech perception, while Hickok and Poeppel's (2000) bilateral model proposed a perceptual pathway in each hemisphere processing speech sounds up to the level of semantics (Hickok & Poeppel, 2007). A middle ground has been proposed by Peelle (2012), with right hemisphere dominated activation for "unconnected" speech (i.e., phonemes to single words) and left hemisphere dominated activation for "connected" speech (i.e., phrases, sentences and narratives).

Using simple but complete sentence stimuli, we consistently found bilateral activation in children for all speech listening contrasts. This activation did not significantly differ between groups, after FDR correction, suggesting that children with LiD use the same cortical areas as

TD children when listening to increasingly complex speech stimuli. However, it is possible that multi-voxel pattern analysis (MVPA) may find finer group differences in the hierarchy for speech processing (e.g., Okada et al., 2010). While the same cortical areas are used by the children with LiD, it is possible that they do so on a different time frame from the TD children. Unfortunately, fMRI does not provide sufficient time resolution to address this possibility.

Bilateral activation reported here differs from the report of Scott et al. (2000) who found a left lateralized pathway for speech comprehension using PET. Our results extend those of Halai et al. (2015) who found bilateral activation using continuous MRI. Both these previous studies tested young adults and used similar stimuli to those used here, but in continuous, passive speech presentation, where comprehension was assessed after the scanning session. In contrast, we required children to provide a behavioural response to each short sentence presentation as a check that they were paying attention. In addition, we used a sparse/HUSH protocol compared to the previous continuous scanning protocols. Scanner noise superimposed on speech listening has been shown to increase listening effort (Peelle et al., 2010) and engage different cortical areas at varying intensities. For example, in a meta-analysis of 57 speech comprehension studies, Adank (2012) found that continuous scanning more strongly activated regions of the supplementary motor area and anterior cingulate gyrus, while sparse scanning showed more extensive activation in the STS.

The study reported here focused on speech listening in children with and without LiD, rather than on differences across the tested age range. Our results support the hypothesis of bilateral processing in speech listening in a 6-12 year old population. Bilateral activation produced by a speech comprehension task has previously been shown in preschoolers (Sroka et al., 2015; Holland et al., 2007), with an increase in lateralization for skills that develop after age 5 (e.g. word-picture matching and syntactic prosody) (Holland et al., 2007). A meta-analysis of 27 developmental (ages 4-13) fMRI language comprehension studies suggested that lateralization in the developing brain is associated with increased sensitivity to syntax (Enge et al., 2020). However, using an alternative approach (Plante et al., 2015), cortical activation in young adults became more left-lateralized as they learned an unfamiliar language. Plante's results thus suggest that increasing speech lateralization in older children may occur through language learning rather

than through maturation. Recently, transcranial magnetic stimulation (TMS) has provided clear evidence that stimulation of either left or right STG decreases understanding of speech-in-noise in adults (Kennedy-Higgins et al., 2020). Together, these brain imaging and stimulation results suggest that bilateral cortical activation is a life-long, obligate aspect of speech perception. The results reported here are the first in an ongoing longitudinal study where the same cohort is being invited back for behavioural and neuroimaging assessment two and four years after this baseline.

Temporal connectivity: Brain and behaviour

Phonology is the system of processing the smallest units of speech sounds and their linguistically appropriate combinations. We found no group difference in the temporal connectivity of this network. As we progressed to the Intelligibility network, we found that the temporal connection between the left inferior frontal gyrus (Intelligibility ROI#14 in Table 2) and left posterior middle temporal gyrus (Intelligibility ROI#23 in Table 2) was stronger in the children with LiD. These areas are well known for language production (Broca's in the inferior frontal gyrus, Hagoort, 2014) and comprehension (left MTG, (Acheson & Hagoort, 2013; Dronkers et al., 2004). Posterior MTG has been associated with lexical and semantic access in a sound-to-meaning network (Hickok & Poeppel, 2000, 2004). It is unknown whether stronger or weaker temporal connectivity is beneficial. However, the relationship found here between connectivity strength and an individual child's behavioural results provides evidence on this key issue. Children with stronger connectivity performed more poorly on cognitive tasks assessing vocabulary, switching attention and episodic memory. Reduced connectivity may thus be indicative of increased processing efficiency and/or suppression of a task-relevant network, in this case language. Note that the behavioural measures used here have been shown to be either little affected by age (audiogram: Hunter et al., 2020) or were standardized across age (ECLiPS, LiSN-S, SCAN and NIH toolbox; Petley et al., 2020).

Group comparisons in the Semantics network highlight two focal points. First, the right parahippocampal gyrus, associated with memory encoding and retrieval (Luck et al., 2010), had stronger temporal connections with auditory areas (left Heschl's, right planum temporale and bilateral STG) in children with LiD. Second, the left temporal fusiform cortex, associated with word recognition and the recovery of meaning from an impoverished acoustic signal (Davis &

Johnsrude, 2003), had stronger temporal connections with auditory areas (right STG), but in TD children. Again the relationship between individuals in connectivity strength and behavioural results shines light on the function of these connections. The first group of connections (right parahippocampal gyrus, Semantics ROI #16 in Table 2) show negative correlations between connectivity strength and dichotic listening, speech in noise and attention (selective and switching). The second group of connections (left temporal fusiform cortex, Semantics ROI #10 in Table 2) show positive correlations between connectivity strength, and dichotic listening and verbal vocabulary. This suggests a major difference between groups is connectivity associated with speech processing. Stronger connections from the right parahippocampal gyrus and weaker connections from the left temporal fusiform cortex were associated with impaired speech listening.

The forebrain focal points identified in this study highlight language production, memory encoding and retrieval, and word recognition as target areas of further research in children with LiD. Combined, these focal points suggest that children with LiD have difficulty matching meaning to understanding words in an auditory environment. The connectivity analysis also highlights the left inferior frontal gyrus that, along with the superior temporal and premotor regions, has been shown to be recruited for statistical learning of language (Davis & Gaskell, 2009; Karuza et al., 2013). Another avenue of research could therefore be how children with LiD learn language compared to their peers. Language development is supported by distributed neural networks connecting both cortical and subcortical regions throughout the brain (Lieberman, 2002; Mesulam, 1990). However, such networks have not been closely or systematically examined in children with LiD.

Assessing where the ‘break’ in speech listening occurs would provide a clearer avenue for research into effective evidence-based treatments. Future studies could utilize paradigms with anomalous and mispronounced words (e.g. Roebuck et al., 2018). Imaging techniques, for example MRI, could be used in parallel to assess whether the cortical networks used in such tasks are affected by LiD. Complementary time-sensitive techniques (EEG, MEG) could investigate whether this difficulty is due to a bottleneck in processing leading to increased listening effort, cognitive effort or fatigue. As LiD may build up over time (Roebuck & Barry,

2018), it is important to assess the children's ability throughout the task rather than using summary values (McGarrigle et al., 2020).

Relationship to neurodevelopmental disorders and cognitive function

Children identified with LiD have difficulty in speech listening compared to TD children (Petley et al., 2020). These groups were distinguished by connectivity differences between cortical areas associated with language production, memory and word recognition, rather than by the activity or connectivity of primary auditory cortical areas. However, at least 50% of children referred/diagnosed with auditory processing disorder (APD), an alternate clinical label that has been used for LiD, have also been diagnosed with specific language impairment (SLI), dyslexia/reading disorders, attention deficit/hyperactivity disorder (ADHD) or more than one of these other neurodevelopmental disorders (Dawes & Bishop, 2010; Ferguson et al., 2011; Gokula et al., 2019; Moore et al., 2018; Sharma et al., 2009). This high level of comorbidity was echoed in this study, where a background caregiver questionnaire showed that half of the children with LiD also reported a diagnosis of ADHD, 9% autism spectrum disorders and 26% had seen a speech language pathologist. This high comorbidity raises the question of whether the group effects reported here were a result of LiD or one of these other disorders. We do not yet know, but we are currently addressing that question using a web-based resource (Neurosynth; Yarkoni et al., 2011) that allows functional connectivity analysis of brain areas defined by a meta-analysis of published fMRI activation coordinates. This analysis is being applied to children with a primary diagnosis of attention or language disorders as well as to the two groups of children reported here. What we do know is that the children identified here as having LiD had specific, atypical neurological characteristics associated with language comprehension.

Further investigation into the neurodevelopmental basis of LiD may also aid in the investigation of language disorders. Our results show typical phonology but impaired non-phonological speech connectivity in children with LiD. This differs from DLD, which presents with both abilities impaired, and dyslexia, which presents with impaired phonological and typical non-phonological abilities in reading (Bishop & Snowling, 2004; Delage & Durrleman, 2018). It is possible that altered cortical language processing leads to LiD. However, it could be that altered cortical language processing may be a consequence of LiD.

The results presented here highlight the importance of non-auditory factors, specifically language, in audiometric testing. There is a growing recognition of the importance of speech perception and, specifically, speech-in-noise (SiN) intelligibility in everyday hearing (Killion et al., 2004; Smits et al., 2013). It has been proposed that such testing could supplement, or even replace pure tone detection as an audiometric gold standard (Hewitt, 2018). However, both SiN test instructions and test items pose a challenge to language and memory as well as auditory function. While those cognitive aspects of auditory testing and learning have been dismissed as procedural issues (Hawkey et al., 2004), they are an intimate component of a SiN test. These results provide insight into mechanisms of how speech perception may be disrupted in LiD, a common form of auditory impairment in both children (Moore et al., 2018) and adults (Edwards et al., 2020). They also add to a growing literature on the role of cognitive function in hearing (David R Moore et al., 2014; Rönnberg et al., 2013; Sharma et al., 2019).

Conclusions

Our results provide the first multifaceted neurological profile for children classified with LiD, based on caregiver report and normal peripheral auditory function (Hunter et al., 2020). Children with LiD recruited the same cortical areas as their peers when listening to increasing complexities of speech. The temporal connections between these areas revealed significant differences between the groups only at the semantic level of speech listening. These differences were related to dichotic listening, speech-in-noise, attention, memory and verbal vocabulary abilities. Overall, the data are consistent with the hypothesis that children with LID are primarily affected by cognitive and language deficits.

METHODS

Participants

Eighty-five participants aged 6-12 years completed the fMRI sentence task and 81 participants completed the rs-fMRI (see Table 1). Ninety percent completed both tasks. All participants had normal audiometric hearing with thresholds < 25 dB HL at all octave-interval frequencies from 0.25 - 16 kHz in both ears (Figure 1A). In this paper we focus on the cortical results and their relationship with behavioural responses from the baseline of our longitudinal 'SICLID' study examining correlates of LiD in children (Hunter et al., 2020; Petley et al., 2020; Blankenship et al., in prep.).

Caregivers of all participants completed a well-validated checklist of everyday listening and related skills (the ECLiPS; Barry & Moore, 2014). Those scoring within the *clinical interest* range on the ECLiPS (< 10th percentile of ECLiPS standardized scores), or had a diagnosis of auditory processing disorder (APD; n=14), were classified as children with LiD (Figure 1B). Participants scoring within the *normal* range on the ECLiPS and with no history of developmental disorders or delays were classified as TD children.

MRI acquisition

MRI was performed via a 3T Philips Ingenia scanner with a 64-channel head coil and Avotec audiovisual system. All participants were awake and non-sedated throughout the scanning. The protocol included a high-resolution T1-weighted anatomical scan, fMRI sentence task (4.9 minutes) and rs-fMRI (5 minutes). The fMRI sentence task was acquired with a sparse scanning protocol ('HUSH', details below); TR/TE = 2000/30 ms, voxel size = 2.5 × 2.5 × 3.5 mm, 39 axial slices. A total of 147 volumes was acquired by alternating scanning for 6 seconds (3 volumes) and not scanning for 6 seconds, 49 times. Cardiac and respiration signals were collected during the fMRI sentence task using the scanner's wireless respirator bellows and Peripheral Pulse Oximeter. The rs-fMRI acquisition was acquired with TR/TE = 2000/30 ms, voxel size = 2.5 × 2.5 × 3.5 mm, 39 axial slices in ascending slice order and 150 volumes. The high-resolution T1-weighted anatomical scan was acquired with TR/TE=8.1/3.7 ms, FOV 25.6 x 25.6 x 16.0 cm, matrix 256 x 256 and slice thickness = 1mm.

fMRI task

With sound levels reaching 118.4 ± 1.3 dB (A) in a 3T MRI system (Price et al., 2001) special considerations must be made when planning an auditory-based MRI study. In order to protect the participant from the loud environment, foam ear plugs and MRI safe circumaural headphones were worn. The scanner noise may also produce masking of the desired stimuli. Therefore, in the fMRI task we used a 'Hemodynamics Unrelated to Sounds of Hardware' (HUSH) scanning protocol (Deshpande et al., 2016; Edmister et al., 1999; Hall et al., 1999; Schmithorst & Holland, 2004) - a sparse temporal sampling protocol where there was no gradient coil noise during presentation of the auditory stimuli. We also used a talker identification task instead of a speech recognition task. Instead of asking the children *what* they heard, we asked them *who* had said it. The children responded with button presses throughout the task so we could ensure they maintained attention to the task.

Sixteen linguistically simple BKB sentences, designed to be familiar to young children, were recorded by a single male North American speaker, mirroring the paradigm used by Scott et al. (2000) and Halai et al. (2015). These were intelligible, clear speech stimuli (e.g. Figure 1D). Rotated speech stimuli (Figure 1E) were created by rotating each sentence spectrally around 2 kHz using the Blesser (1972) technique. Rotated speech was not intelligible, though some phonetic features and some of the original intonation was preserved. Rotated+vocoded speech stimuli (Figure 1F) were created by applying 6 band noise-vocoding to the rotated speech stimuli. While the rotated noise-vocoded speech was completely unintelligible, the character of the envelope and some spectral detail was preserved.

Participants were told that they would be completing a matching game where they would hear a sentence and then see a picture (of a man or an alien). If the picture matched who said the sentence (man - clear speech, alien - rotated or rotated+vocoded speech), the participant pressed a button with their right thumb, if the picture did not match, they pressed a second button with their left thumb. A sticker was placed on the participants' right hand to provide a reminder as to which hand was correct for matching voice and picture.

Before scanning, each participant was familiarized with the sentence task and completed three practice trials with verbal feedback from the tester. If a trial was completed incorrectly, the stimuli and instructions were reintroduced until the participant showed understanding. During scanning, each participant completed 48 matching trials, 16 of each sentence type, with no feedback. To maintain scanner timings the behavioural task continued regardless of whether the participant responded. However, if the child did not press a response button on three trials in a row the tester provided reminders/encouragement over the scanner intercom between stimuli presentations.

fMRI data analysis

First-level fMRI data were processed using FSL (FMRIB Software Library, https://fsl.fmrib.ox.ac.uk/fsl/). The T1 brain data were extracted using BET and normalized and resampled to the 2 mm isotropic MNI ICBM 152 non-linear 6th generation template using FLIRT.

For the sparse HUSH acquisition, the volumes were separated and combined into three files according to the volume's order during the scanner-on period. Each of the three files was pre-processed separately and first-level statistics computed. The three statistical images were then averaged together using a one sample t-test. This was done to account for the difference in intensity among the volumes due to T2* relaxation effects. The pre-processing steps included the following. FSL's BET was used for brain extraction of the functional data. Outlying functional volumes were detected with 'fsl_motion_outliers' using the RMS intensity difference metric. AFNI's '3dretroicor' was used to regress out the cardiac and respiration signals using a RETROICOR approach (Glover et al., 2000). Motion correction was carried out by MCFLIRT. A GLM was used to regress motion-related artifacts from the data using 6 regressors for the motion parameters and an additional regressor for each outlying volume. The amount of motion during the scans (the number of outlying volumes for each participant) did not differ between groups, $p = .62$ (Table 1). The data were spatially smoothed using a Gaussian kernel with a sigma of 3 mm and temporally filtered with a high pass filter with a sigma of 30 seconds. The results were interpolated to a 2 mm isotropic voxel size and aligned to the Montreal Neurological Institute (MNI) template by first co-registering it with the participant's T1 using FSL's FLIRT.

Second-level analysis was also conducted using FSL. A GLM approach was used to create group activation maps based on contrasts between conditions for all participants (i.e. regardless of LiD/TD status) with age as a covariate. Group composite images were thresholded using a family-wise error correction ($p < 0.001$) and clustering threshold of $k = 4$ voxels. Three BOLD activation contrasts were used to search for brain loci responding to different aspects of listening to language (Halai et al., 2015; modified from Scott, 2000). First, a 'Semantics' activation map, whereby the signal with intelligibility, intonation, phonetics, prosody and sound was contrasted with one lacking all of these attributes except prosody and sound (clear > rotated+vocoded). Second, an 'Intelligibility' activation map contrasted the signal with all speech attributes to one retaining intonation, phonetics, prosody and sound (clear > rotated). Third, a 'Phonetics' activation map contrasted a signal with intonation, phonetics, prosody and sound with one having only prosody and sound (rotated > rotated+vocoded).

For the rs-fMRI scan, pre-processing and analysis was performed in the CONN toolbox using standard spatial and temporal pipelines (Whitfield-Gabrieli & Nieto-Castanon, 2012). For spatial smoothing a FWHM of 8mm was used. The Artifact Detection Tool (ART, https://www.nitrc.org/projects/artifact_detect) within CONN was used to regress out framewise motion. The number of frames regressed out was compared between groups with no significant group differences, $p = .44$ (Table 1).

Behavioural responses from the fMRI task were assessed using a 2 (group: TD, LiD) $\times$ 3 (sentence type: clear, rotated, rotated+vocoded) repeated measures analysis of variance (ANOVA) for accuracy and again for RT (Figure 1G). Where the assumption of sphericity was violated, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

Resting state fMRI

We maintained visual fixation during this period with a cross in the center of a black screen. During this second scan the participant was asked to lie still, keep eyes open and let their mind wander; they were not performing an exogenous task. Eyes were monitored by the tester through CCTV and no child fell asleep during the task.

Resting state ROI-to-ROI analysis

The areas of activation from the fMRI task's three contrasts were used to define the ROIs of advancing speech listening networks (Phonology, Intelligibility and Semantics). However, as these areas of activation were large, we applied the parcellation from the pediatric ADHD-200 sample (Bellec et al., 2017) to parcellate each network. This created smaller and more appropriate ROIs for connectivity analysis of each network (Figure 3D-F, Supplementary Material Table 1).

A ROI-to-ROI analysis was used in Conn to test the functional relationship between each pair of ROIs identified in the fMRI sentence listening task. The mean time course of all voxels within each ROI was used to calculate individual pairwise Pearson correlations. The *r* values were normalized to *z* values via Fisher's z-transformation. We then used these *z* values to explore the relationship between the three listening networks and behavioural measures. Statistical thresholds were set to $p < .05$ (corrected) at the single voxel level, and the resulting connections were thresholded at seed-level by intensity FDR ($p < .05$).

Caregiver questionnaire

Everyday listening skills - ECLiPS (Barry & Moore, 2014)

The ECLiPS is a standardized parental-report measure to screen for listening and communication difficulties. Caregivers rated 38 simple statements on a five-point scale, ranging from strongly disagree to strongly agree. Children who score below the 10th percentile of the standardized composite scores (M = 10; SD = 3) are identified as experiencing clinically significant difficulties in listening and communication.

Behavioral measures

Resting state temporal connections with significant group differences were correlated with behavioural measures. Study data were collected and managed using REDCap electronic data capture tools hosted at Cincinnati Children's Hospital (Harris et al., 2009, 2019). REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture; 2)

audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources.

Listening in Spatialized Noise-Sentences (Cameron & Dillon, 2007, 2008, 2009)

LiSN-S (US version) is a standardized test assessing speech in noise ability. Binaural target (T) sentences were presented through headphones along with two other distracting sentences (D1, D2). The children were asked to repeat the sentences of the target voice only. Distracting sentences remained constant at 55 db SPL. After each correct trial the target voice descended in level (4 dB), but if the child incorrectly repeated back over 50% of the sentence the level increased (by 2 dB).

Four listening conditions are made by manipulating D1 and D2 with respect to T (same voice, different voices; same direction, 0°, different direction, ± 90° azimuth). Three difference scores are calculated in order to control for language and cognitive demands: Talker advantage (different voices - same voice); Spatial advantage (different directions – same direction); and Total advantage (different voices and directions – same voices and directions). The LISN-S software calculated the difference scores for each participant.

Auditory processing disorder tests - SCAN-3:C (Keith, 2009; 2000)

The SCAN-3 is a US-standardized test battery often used by audiologists to diagnose APD in children (Emanuel et al., 2011). Subtests used in our battery were Auditory Figure Ground - assessing the ability to repeat words presented against background multi-talker speech at signal/noise ratio of +8 dB; Competing words - a dichotic listening task where the child repeats different words presented simultaneously to each ear, but repeating that from a designated ear first; Filtered words - assessing ability to identify words that are low pass filtered at 750 Hz; and Competing Sentences - a dichotic listening task where different sentences are presented simultaneously to each ear, and the child is asked to repeat the sentence from a designated ear. Both subtest and a standardized composite score are calculated.

Cognition - NIH toolbox (Weintraub et al., 2013)

The NIH toolbox - Cognition Battery is a collection of US-standardized tests from which we used measures of selective attention (Flanker Inhibitory Control and Attention Test), episodic memory (Picture Sequence Memory Test), executive functioning (Dimensional Change Card Sort Test) and picture vocabulary. Each visually administered test took 5 - 15 minutes to complete on an iPad. Age-corrected subtest and an overall 'early childhood composite' scores were calculated for each participant.

DATA AVAILABILITY

The dataset generated during and analysed during the current study are available from the corresponding author. ROIs used in the rs-fMRI analysis are available at GitHub: https://github.com/stewarthannahj/ROIs_SICLiD.git

ACKNOWLEDGEMENTS

Thanks to Dr Kim Leikin and Prof. Scott K Holland for aspects of study design and to Prof. Stuart Rosen and Dr Peter Chiu for the scripts to rotate and vocode the stimuli for the fMRI sentence task. Thanks also to Audrey Perdew, Nicholette Sloat and Dr Ronan McGarrigle for assisting in MRI acquisition. Many thanks to the MRI techs for many hours of scanning.

FUNDING

This research was supported by NIH R01DC014078, NIH 2UL1TR001425 and by Cincinnati Children's Research Foundation. DRM is supported in part by the NIHR Manchester Biomedical Research Centre.

COMPETING INTERESTS

The authors report no competing interests.

CONTRIBUTIONS

HJS - study design, MRI acquisition, MRI preprocessing and analysis, behavioural analysis, figures and tables, manuscript preparation and revisions

EC - behavioural and rs-fMRI analysis, figures, assisted with manuscript preparation and revisions

TM - fMRI preprocessing, assisted with manuscript preparation

JV - study design, analysis and presentation, manuscript preparation and revisions

DRM - study design and presentation, funding, manuscript preparation and revisions

REFERENCES

Acheson, D. J., & Hagoort, P. (2013). Stimulating the brain's language network: Syntactic ambiguity resolution after TMS to the inferior frontal gyrus and middle temporal gyrus. *Journal of Cognitive Neuroscience*, *25*(10), 1664–1677.

Adank, P. (2012). The neural bases of difficult speech comprehension and speech production: Two Activation Likelihood Estimation (ALE) meta-analyses. *Brain and Language*, *122*(1), 42–54.

Assaneo, M. F., Orpella, J., Ripolles, P., Diego-Balaguer, D., Poeppel, D., & others. (2019). The lateralization of speech-brain coupling is differentially modulated by intrinsic auditory and top-down mechanisms. *Frontiers in Integrative Neuroscience*, *13*, 28.

Barry, J. G., & Moore, D. R. (2014). *Evaluation of children's listening and processing skills (ECLiPS)*.

Barry, J. G., Tomlin, D., Moore, D. R., & Dillon, H. (2015). *Use of Questionnaire-Based Measures in the Assessment of Listening Difficulties in School-Aged Children*. 1–14.

Bellec, P., Chu, C., Chouinard-Decorte, F., Benhajali, Y., Margulies, D. S., & Craddock, R. C. (2017). The neuro bureau adhd-200 preprocessed repository. *Neuroimage*, *144*, 275–286.

Bishop, D. V., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin*, *130*(6), 858.

Biswal, B. B. (2012). Resting state fMRI: a personal history. *Neuroimage*, *62*(2), 938–944.

Blesser, B. (1972). Speech perception under conditions of spectral transformation: I. Phonetic characteristics. *Journal of Speech and Hearing Research*, *15*(1), 5–41.

Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. *Current Biology*, *28*(24), 3976-3983.e5. https://doi.org/10.1016/j.cub.2018.10.042

Cameron, S., & Dillon, H. (2007). Development of the listening in spatialized noise-sentences

test (LISN-S). *Ear and Hearing*, *28*(2), 196–211. https://doi.org/10.1097/AUD.0b013e318031267f

Cameron, S., & Dillon, H. (2008). The listening in spatialized noise-sentences test (LISN-S): Comparison to the prototype LISN and results from children with either a suspected (central) auditory processing disorder or a confirmed language disorder. *Journal of the American Academy of Audiology*, *19*(5), 377–391. https://doi.org/10.3766/jaaa.19.5.2

Cameron, S., & Dillon, H. (2009). *Listening in Spatialized Noise–Sentences test (LiSN-S)*.

Craddock, R. C., James, G. A., Holtzheimer III, P. E., Hu, X. P., & Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, *33*(8), 1914–1928.

Damoiseaux, J. S., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences*, *103*(37), 13848–13853.

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3773–3800.

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(8), 3423–3431. https://doi.org/23/8/3423 [pii]

Dawes, P., & Bishop, D. V. M. (2010). Psychometric profile of children with auditory processing disorder and children with dyslexia. *Archives of Disease in Childhood*, *95*(6), 432–436. https://doi.org/10.1136/adc.2009.170118

Delage, H., & Durrleman, S. (2018). Developmental dyslexia and specific language impairment: Distinct syntactic profiles? *Clinical Linguistics & Phonetics*, *32*(8), 758–785.

Deshpande, A. K., Tan, L., Lu, L. J., Altaye, M., & Holland, S. K. (2016). FMRI as a preimplant objective tool to predict postimplant oral language outcomes in children with cochlear

implants. *Ear and Hearing*, *37*(4), e263–e272.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*(1–2), 145–177. https://doi.org/10.1016/j.cognition.2003.11.002

Edmister, W. B., Talavage, T. M., Ledden, P. J., & Weisskoff, R. M. (1999). Improved auditory cortex imaging using clustered volume acquisitions. *Human Brain Mapping*, *7*(2), 89–97.

Edwards, B. (2020). *Emerging technologies, market segments, and MarkeTrak 10 insights in hearing health technology*.

Emanuel, D. C., Ficca, K. N., & Korczak, P. (2011). Survey of the diagnosis and management of auditory processing disorder. *American Journal of Audiology*.

Enge, A., Friederici, A. D., & Skeide, M. A. (2020). A meta-analysis of fMRI studies of language comprehension in children. *NeuroImage*, 116858.

Evans, S., & McGettigan, C. (2017). Comprehending auditory speech: Previous and potential contributions of functional MRI. *Language, Cognition and Neuroscience*, *32*(7), 829–846.

Ferguson, M. a, Hall, R. L., Riley, A., & Moore, D. R. (2011). Communication, listening, cognitive and speech perception skills in children with auditory processing disorder (APD) or Specific Language Impairment (SLI). *Journal of Speech, Language, and Hearing Research : JSLHR*, *54*(1), 211–227. https://doi.org/10.1044/1092-4388(2010/09-0167)

Glover, G. H., Li, T.-Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *44*(1), 162–167.

Gokula, R., Sharma, M., Cupples, L., & Valderrama Valenzuela, J. T. (2019). Comorbidity of auditory processing, attention, and memory in children with word reading difficulties. *Frontiers in Psychology*, *10*, 2383.

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region

and beyond. *Current Opinion in Neurobiology*, *28*, 136–141.

Halai, A. D., Parkes, L. M., & Welbourne, S. R. (2015). Dual-echo fMRI can detect activations in inferior temporal lobe during intelligible speech comprehension. *NeuroImage*, *122*, 214–221. https://doi.org/10.1016/j.neuroimage.2015.05.067

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., Gurney, E. M., & Bowtell, R. W. (1999). “Sparse” temporal sampling in auditory fMRI. *Human Brain Mapping*, *7*(3), 213–223.

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O’Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & others. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., Conde, J. G., & others. (2009). A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*, *42*(2), 377–81.

Hawkey, D. J., Amitay, S., & Moore, D. R. (2004). Early and rapid perceptual learning. *Nature Neuroscience*, *7*(10), 1055–1056.

Hewitt, D. (2018). Audiometry and Its Discontents. *Hearing Review*, *25*(1), 20–23.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, *4*(4), 131–138.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1–2), 67–99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402.

Hunter, L. L., Blankenship, C. M., Lin, L., Sloat, N. T., Perdew, A., Stewart, H., & Moore, D. R. (2020). Peripheral auditory involvement in childhood listening difficulty. *Ear and Hearing*.

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013).

The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, *127*(1), 46–54.

Keith, R. W. (2000). Development and Standardization of SCAN-C Test for Auditory Processing Disorders in Children. *J Am Acad Audiol*, *11*(8), 438–445.

Keith, R. W. (2009). *SCAN-3 for Children. Tests of Auditory Processing Disorders.* Pearson.

Kennedy-Higgins, D., Devlin, J. T., Nuttall, H. E., & Adank, P. (2020). The causal role of left and right superior temporal gyri in speech perception in noise: A transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, *32*(6), 1092–1103.

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *Journal of Acoustical Society of America*, *116*, 2395–2405.

Lieberman, P. (2002). On the nature and evolution of the neural bases of human language. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *119*(S35), 36–62.

Luck, D., Danion, J.-M., Marrer, C., Pham, B.-T., Gounot, D., & Foucher, J. (2010). The right parahippocampal gyrus contributes to the formation and maintenance of bound information in working memory. *Brain and Cognition*, *72*(2), 255–263.

McGarrigle, R., Rakusen, L., & Mattys, S. (2020). *Effortful listening under the microscope: Examining relations between pupillometric and subjective markers of effort and tiredness from listening.*

Mesulam, M.-M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *28*(5), 597–613.

Moore, D R. (2018). Auditory processing disorder (APD). *Ear and Hearing*, *39*(4), 617.

Moore, D R, Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., & Riley, A. (2010). Nature of

Auditory Processing Disorder in Children. *Pediatrics*, *126*(2), E382–E390. https://doi.org/10.1542/peds.2009-2826

Moore, D R, Sieswerda, S. L., Grainger, M. M., Bowling, A., Smith, N., Perdew, A., Eichert, S., Alston, S., Hilbert, L. W., Summers, L., & others. (2018). Referral and diagnosis of developmental auditory processing disorder in a large, United States hospital-based audiology service. *Journal of the American Academy of Audiology*, *29*(5), 364–377.

Moore, David R, Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R. H., & Munro, K. J. (2014). Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. *PloS One*, *9*(9), e107720.

Moore, David R., Hugdahl, K., Stewart, H. J., Vannest, J., Perdew, A. J., Sloat, N. T., Cash, E. K., & Hunter, L. L. (2020). Listening Difficulties in Children: Behavior and Brain Activation Produced by Dichotic Listening of CV Syllables. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00675

Neijenhuis, K., Campbell, N., Cromb, M., Luinge, M., Moore, D. R., Rosen, S., & de Wit, E. (2019). An Evidence-based Perspective on ‘Misconceptions’ Regarding Pediatric Auditory Processing Disorder. *Frontiers in Neurology*, *10*, 287.

Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I. H., Saberi, K., Serences, J. T., & Hickok, G. (2010). Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cerebral Cortex*, *20*(10), 2486–2495. https://doi.org/10.1093/cercor/bhp318

Parente, F., Frascarelli, M., Mirigliani, A., Di Fabio, F., Biondi, M., & Colosimo, A. (2018). Negative functional brain networks. *Brain Imaging and Behavior*, *12*(2), 467–476.

Peelle, J. E. (2012). The hemispheric lateralization of speech processing depends on what “speech” is: A hierarchical perspective. *Frontiers in Human Neuroscience*, *6*, 309.

Peelle, J. E., Eason, R. J., Schmitter, S., Schwarzbauer, C., & Davis, M. H. (2010). Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory

processing. *Neuroimage*, *52*(4), 1410–1419.

Petley, L., Hunter, L. L., Motlagh Zadeh, L., Sloat, N., Perdew, A., & Moore, D. R. (2020). Listening Difficulties in Children with Normal Audiograms: Relation to Hearing and Cognition. *MedRxiv*. https://doi.org/10.1101/2020.10.05.20205468

Plante, E., Almryde, K., Patterson, D. K., Vance, C. J., & Asbjørnsen, A. E. (2015). Language lateralization shifts with learning by adults. *Laterality: Asymmetries of Body, Brain and Cognition*, *20*(3), 306–325.

Price, D. L., De Wilde, J. P., Papadaki, A. M., Curran, J. S., & Kitney, R. I. (2001). Investigation of acoustic noise on 15 MRI scanners from 0.2 T to 3 T. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *13*(2), 288–293.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*(6), 718–724.

Roebuck, H., & Barry, J. G. (2018). Parental perception of listening difficulties: An interaction between weaknesses in language processing and ability to sustain attention. *Scientific Reports*, *8*(1), 1–10.

Roebuck, H., Sindberg, H., & Weismer, S. E. (2018). The role of language in nonlinguistic stimuli: Comparing inhibition in children with language impairment. *Journal of Speech Language and Hearing Research*, *61*(5), 1216. https://doi.org/10.1044/2018_JSLHR-L-17-0294

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology*, *58*(5), 247–261.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & others. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*, 31.

Schmithorst, V. J., & Holland, S. K. (2004). Event-related fMRI technique for auditory processing with hemodynamics unrelated to acoustic gradient noise. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *51*(2), 399–402.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400–2406. https://doi.org/10.1093/brain/123.12.2400

Sharma, M., Purdy, S. C., & Humburg, P. (2019). Cluster analyses reveals subgroups of children with suspected auditory processing disorders. *Frontiers in Psychology*, *10*, 2481.

Sharma, M., Purdy, S. C., & Kelly, A. S. (2009). Comorbidity of Auditory Processing, Language, and Reading Disorders. *Journal of Speech Language and Hearing Research*, *52*(3), 706. https://doi.org/10.1044/1092-4388(2008/07-0226)

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(5), 182–186.

Smits, C., Theo Goverts, S., & Festen, J. M. (2013). The digits-in-noise test: Assessing auditory speech recognition abilities in noise. *The Journal of the Acoustical Society of America*, *133*(3), 1693–1706.

Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., & others. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1545–1602.

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., & others. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S54–S64.

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). *Conn*: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity*, *2*(3), 125–141. https://doi.org/10.1089/brain.2012.0073

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. *Nature Methods*, *8*(8), 665–670.

Table 1: Participant details broken down by scan type - fMRI speech listening task and the resting state (RS).

Scan type	Group	N	Age M (SD)	Gender M, F	Maternal education*	Handedness L, both, R	History of tubes	Artifacts removed M (SD)
fMRI	LiD	43	10.02 (2.13)	30, 13	6, 27	3, 5, 35	12	12.16 (4.63)
	TD	42	9.78 (1.93)	25, 17	0, 42	2, 1, 39	15	11.67 (4.30)
RS	LiD	42	10.06 (2.09)	29, 13	8, 34	3, 6, 33	12	28.64 (32.17)
	TD	39	9.75 (1.93)	23, 16	0, 39	2, 1, 36	15	23.00 (29.20)

*High school graduate or less, some college or more

Table 2: Selected regions of interest (ROIS, from the fMRI task) used in the rs-fMRI ROI-to-ROI analysis. Threshold = 4.0, MCC .95. The full table of all ROIs can be found in Supplementary Material Table 1.

Contrast		Max. intensity MNI coordinates				Brain Regions (Harvard-Oxford atlas)	Cluster size (voxels)
		BA	x	y	z		
Intelligibility normal > rotated	14	48	-50	18	18	L inferior frontal gyrus, pars opercularis (Broca's)	6
	23	20	-50	-14	-20	L MTG, post	1
Semantics normal > rotated+vocoded	1	21	68	-20	-10	R MTG, post	301
	5	21	66	-4	-6	R superior temporal gyrus, ant.	420
	8	20	44	0	-26	R STG, ant	108
	9	20	-46	-8	-22	L MTG, post	593
	10	37	-38	-40	-22	L temporal fusiform cortex, post	4
	16	30	22	-30	-24	R parahippocampal gyrus, post	1
	19	21	-50	-34	-8	L MTG, post	899
	27	20	52	-18	-16	R middle temporal gyrus, post	625

Table 3: No significant group difference (TD > LiD, corrected for age) was found in the temporal connectivity of the Phonology network. Group differences were found in the more advanced speech listening networks (Intelligibility and Semantics). See Figure 4.

Contrast	ROI # – ROI #	t (78)	p-FDR
Intelligibility	14 – 23	-3.57	0.016
Semantics	16 – 1	-3.79	0.009
	16 – 5	-3.05	0.033
	16 – 19	-3.04	0.033
	16 – 9	-2.92	0.034
	10 – 27	3.18	0.041
	10 – 8	3.1	0.041

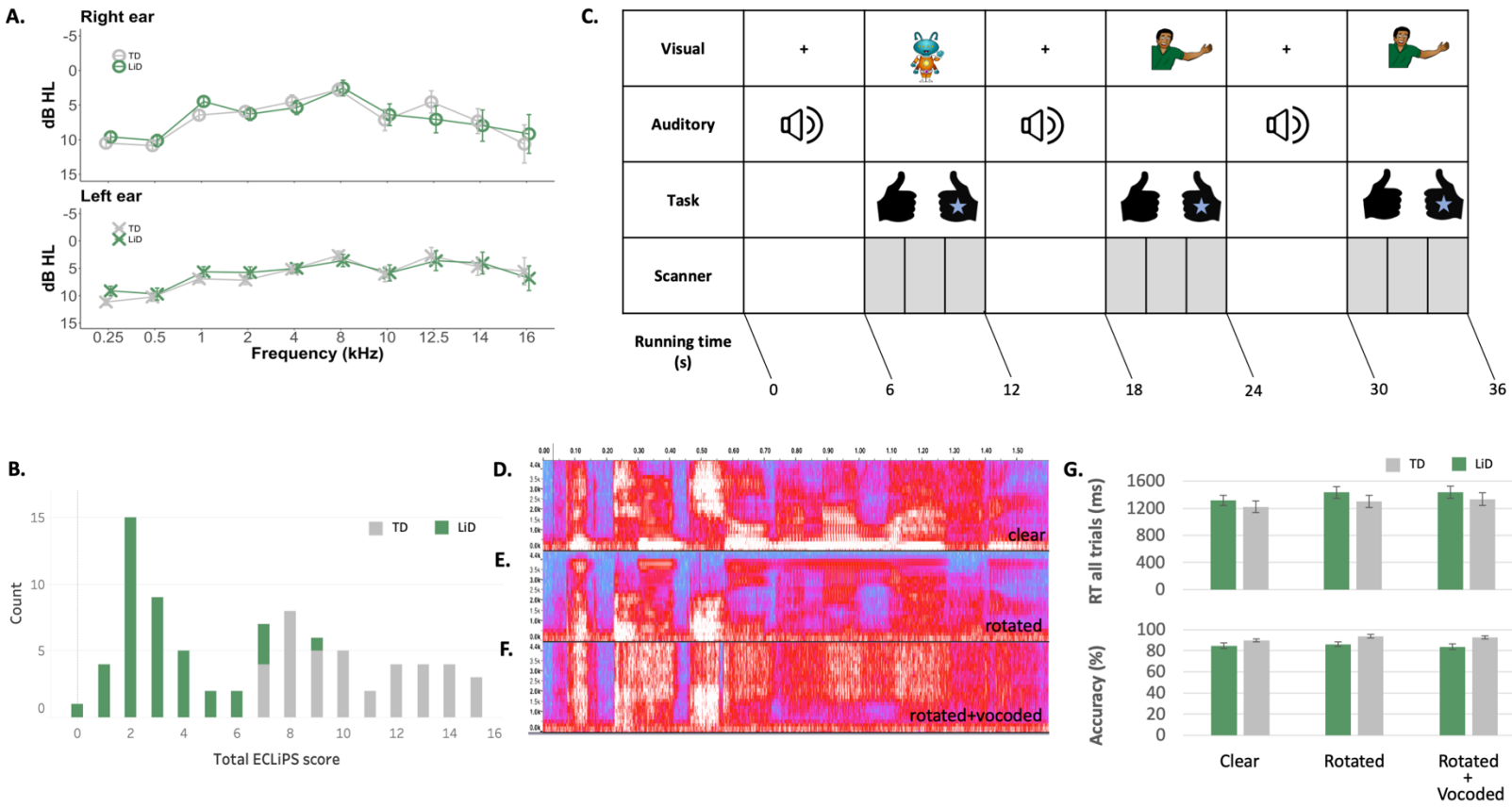

Figure 1: All children had normal tone hearing and were able to perform the speech perception task in the scanner. (A) Group mean hearing thresholds. (B) Total ECLiPS score. (C) fMRI paradigm, the children were asked if the picture matched who spoke the sentence. If it was a match they pressed the right button (the stickered hand), if it was not a match they pressed the left button. Data acquisition (grey shading) was turned off/on for the presentation of the auditory stimuli (HUSH/sparse scanning). (D-F) Spectrograms of 'The two children are laughing'. Time is represented on the x-axis (0.0–1.60 s) and frequency on the y-axis (0.0–4.4 kHz). The shading of the trace in each time/frequency region is controlled by the amount of energy in the signal at that particular frequency and time (red = more energy, blue = less energy). Clear speech is intelligible with intonation. Rotated speech is not intelligible, though some phonetic features and some of the original intonation are preserved. Rotated noise-vocoded speech is completely unintelligible, but preserves the character of the envelope and some spectral detail. (G) Group mean response times and accuracy for clear sentences, rotated sentences and rotated+vocoded

sentences. Error bars represent the standard error of the mean. TD children are in grey and LiD in green.

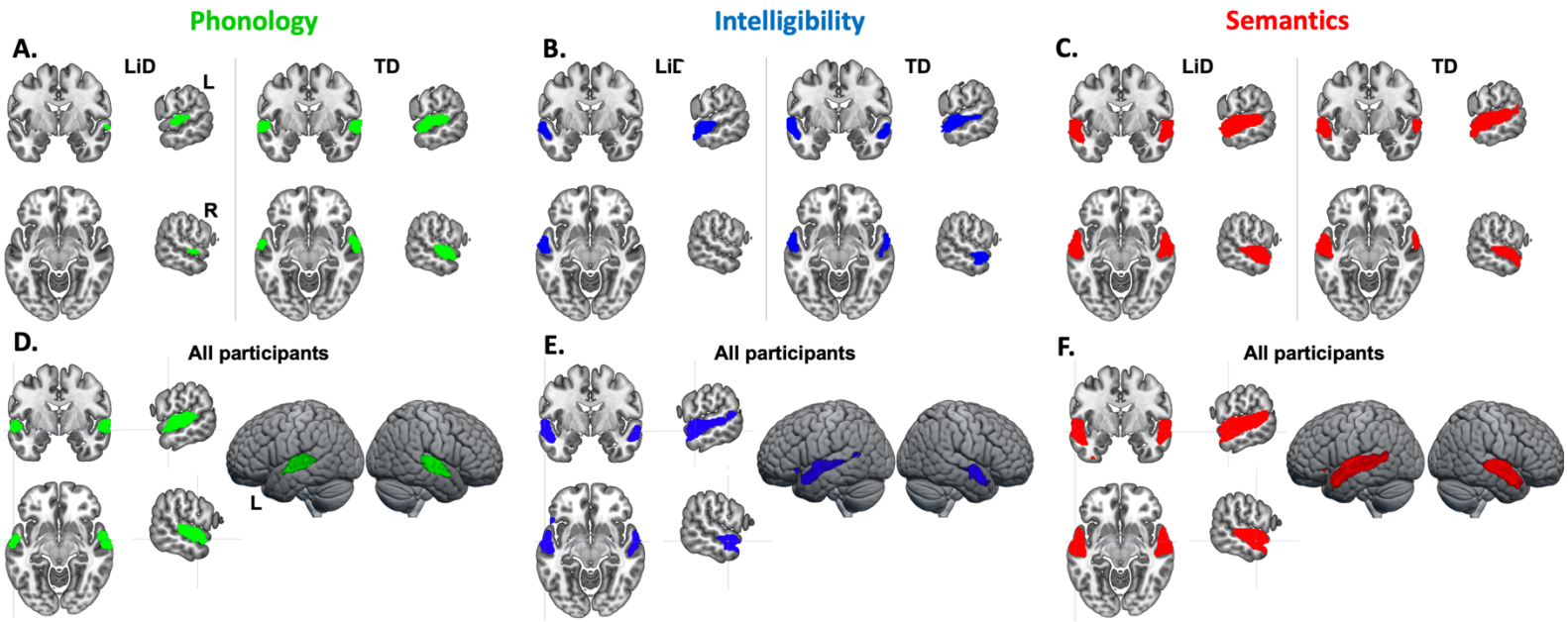

Figure 2: LiD and TD groups showed similar areas of cortical activation in all three contrasts from the fMRI listening task, with no statistical difference between the two groups after correcting for multiple comparisons. We took the cortical activation across all participants and created parcellated ROIs (Figure 3) for use in the rs-fMRI analysis (figure 4). Second level GLM analysis for (A-C) groups (threshold = 2.3, MCC .95) and (D-F) all participants (threshold = 4.0, MCC .95) in the fMRI task, coordinates (±60, -5, -10). Coordinates for maximum intensity voxels for (D-F) can be found in Supplementary Material Table 2. Contrasts are: (A, D) Phonology (green: rotated > rotated+vocoded), (B, E) Intelligibility (blue: clear > rotated) and (C, F) Semantics (red: clear > rotated+vocoded). Images are in neurological orientation. MRIcroGL was used for visualization.

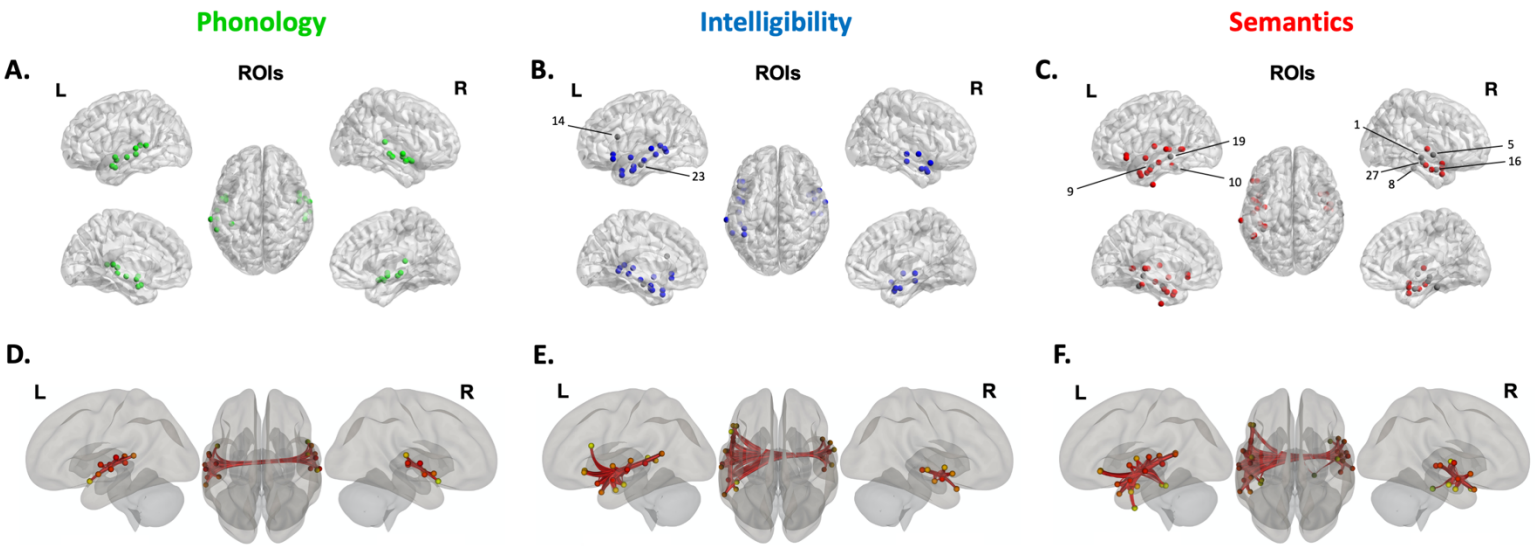

Figure 3: ROIs used for the rs-functional connectivity and their networks across all participants. The cortical activity across all participants in the fMRI task (see Figure 2 D-F) covered large areas and so they were parcellated into smaller ROIs (A-C) for the rs-functional connectivity analysis by applying data-driven spatially constrained parcellation to the areas of activation from the fMRI sentence recognition task using the pediatric ADHD-200 sample. The red lines (D-F) indicate the ROI-to-ROI connections analyzed in each network. Maximum intensity coordinates can be found in Supplementary Material Table 1. Networks are: (A, D) Phonology (green: rotated > rotated+vocoded), (B, E) Intelligibility (blue: clear > rotated) and (C, F) Semantics (red: clear > rotated+vocoded). Images are in neurological orientation. For visualization, BrainNet software was used to display foci in (A-C) (Xia, Wang & He, 2013).

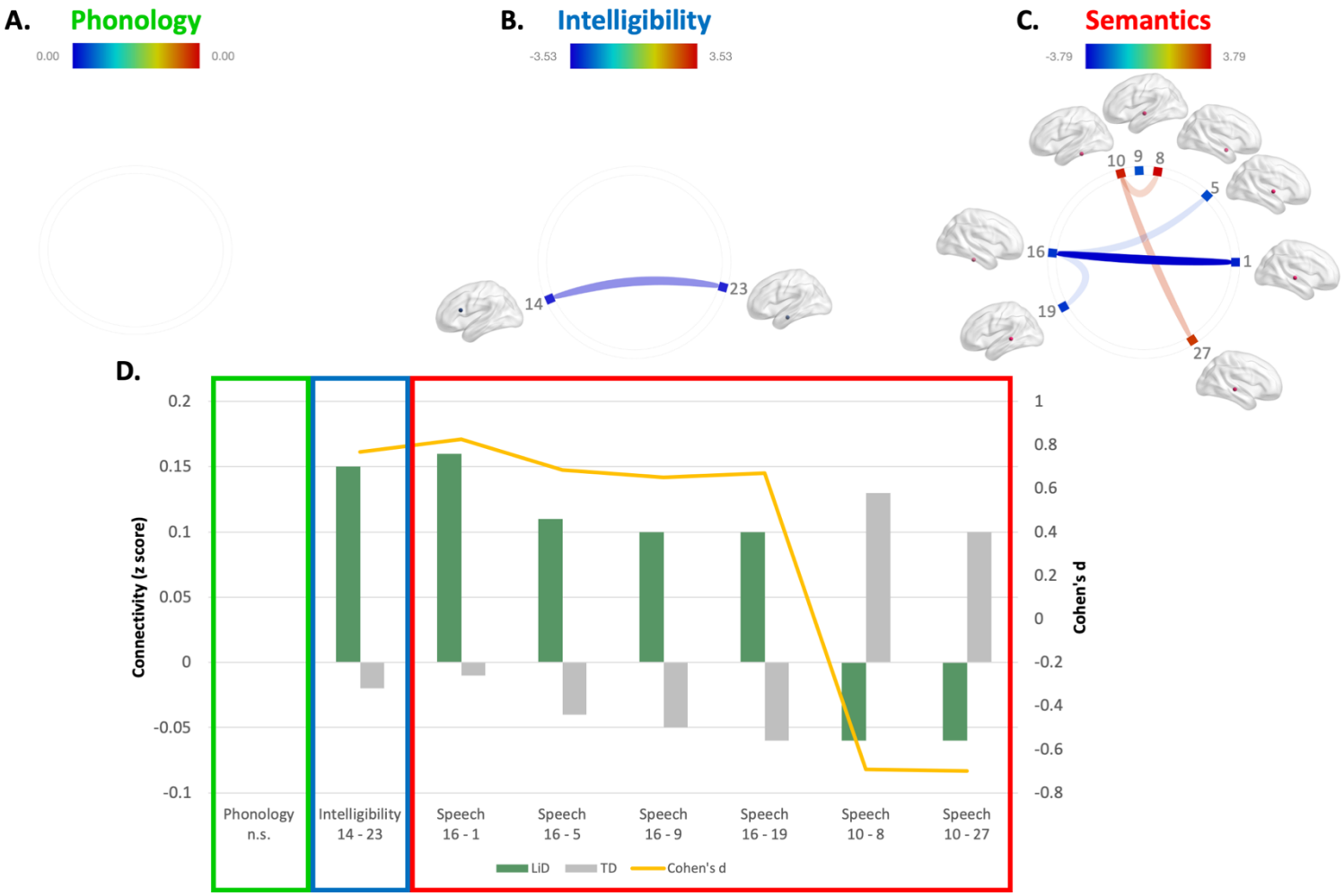

Figure 4: ROI-to-ROI resting state connectivity: the difference between the groups' listening networks grew from no statistical differences in the phonology network to a minor difference in the intelligibility network to more widespread differences in the semantics network. (A-C) The group comparisons without the effect of age for (A) Phonology, (B) Intelligibility and (C) Speech networks. Thicker, more saturated color lines represent stronger connections between cortical areas. Note that the colored bar connectivity z score scales vary slightly between connectivity wheels. (D) Details of the ROI-to-ROI connectivity values (left axis) for each group (LiD green, TD grey) and effect sizes (yellow line, right axis) of the group comparisons without the effect of age. The connections plotted are the ones highlighted as having a significant group difference in the GLM comparing groups without the effect of age (B and C). Connections Intelligibility 14-23 and Speech 16-1,16-5, 16-9 and 16-19 all show the TD group as having

weaker connectivity than the LiD group. While connections Speech 10-8 and 10-27 show the TD group as having stronger connectivity than the LiD group.

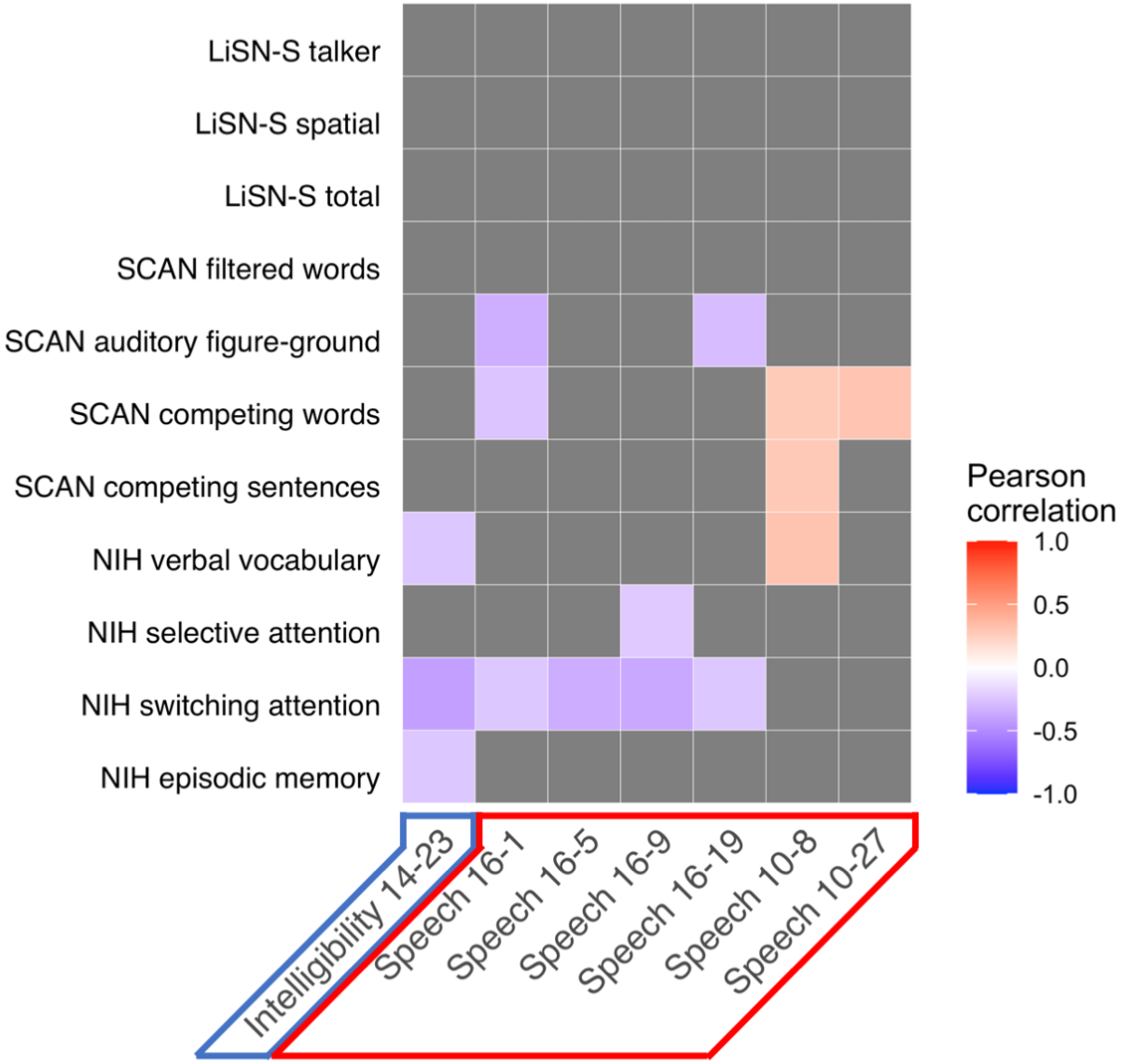

Figure 5: Heat map of brain and behavioral score correlations. Colored squares indicate significant low-moderate Pearson correlations ($p < .05$) between the connections in Figure 3D and speech in noise, auditory processing skills and cognition. Correlations associated with the Broca and right posterior parahippocampal gyrus connections are negative while correlations associated with the left posterior temporal fusiform cortex connections are positive. See Supplementary Materials Table 3 for correlation details. Note that the direction of the correlations flip for the same connections as in the group comparisons in Figure 3 B-D. Intelligibility 14-23 and Speech 16-1,16-5, 16-9 and 16-19 showed that the TD group had weaker

connectivity compared to the LiD group and these connections show negative correlations with behavioral scores. While Speech 10-8 and 10-27 showed that the TD group had stronger connectivity compared to the LiD group and these connections show positive correlations with behavioral scores.