

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

1 **Abstract:**

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

14

15 **Abbreviations:**

16 CANS central auditory nervous system
17 DLD developmental language disorder
18 rs-fMRI resting state-fMRI
19 FOV field-of-view
20 GLM general linear model
21 HUSH Hemodynamics Unrelated to Sounds of Hardware
22 LiD listening difficulties
23 MCC multiple comparison correction
24 MEG magnetoencephalography
25 MTG middle temporal gyrus
26 MVPA multi-voxel pattern analysis
27 p-FDR false detection rate
28 ROI region-of-interest
29 SLI specific language impairment
30 STG superior temporal gyrus
31 STS superior temporal sulcus

- 32 TD typical developing
- 33 TE echo time
- 34 TMS transcranial magnetic stimulation
- 35 TR repetition time

36 INTRODUCTION

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

45
46 The mechanisms underlying childhood LiD without hearing loss remain poorly understood.
47 Speech understanding requires coalescence of acoustic features from bottom-up streaming in the
48 auditory pathway with top-down linguistic processes deriving from widespread cortical areas
49 involved with semantic representation and cognitive functions, notably memory and attention
50 (Hickok & Poeppel, 2007; Rönnberg et al., 2019; Shinn-Cunningham, 2008). Whether LiD are
51 primarily bottom-up or top-down in nature is a long-standing debate (Neijenhuis et al., 2019).
52 Interactions between cortical auditory and related speech perception systems, resulting in
53 impaired higher level processing of target sounds, may hold the key since poor performance on
54 complex auditory tasks such as speech listening in noise is more associated with cognitive
55 function than with more simple auditory tasks (Moore et al., 2010). To identify cortical areas
56 used in this process, we adopted fMRI methods developed to examine the hierarchy of cortical
57 speech processing in adults (Halai et al., 2015; Scott, 2000).

58

59 RESULTS

60 *LiD - auditory or cognitive?*

61 Child participants (6-13 years old; Table 1) all had sensitive audiograms (0.25 – 16 kHz, Figure
62 1A), normal ear function (David R. Moore et al., 2020) and typical auditory brainstem and
63 envelope following responses (Blankenship et al., in prep.), suggestive of normal sub-cortical
64 auditory function. Children scoring within the bottom 10th percentile on a caregiver checklist of
65 everyday listening skills, the ECLiPS (Barry et al., 2015; Roebuck & Barry, 2018), and children

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

68
69 We used task-based and resting state fMRI (rs-fMRI) to assess how speech perception networks
70 function in typically developing (TD) children and children with LiD. Between task-based MRI
71 acquisition intervals (Figure 1C), children listened to clear ('human'; Figure 1D) or distorted
72 ('alien') spoken sentences that we asked them to match to a visual cartoon representation.
73 Distorted sentences were speech that was rotated (Figure 1E) or rotated AND vocoded (Figure
74 1F) to retain prosody while successively eliminating meaning and spectral cues. We were able to
75 explore the neural processing of different speech features by contrasting the different sentence
76 types. By contrasting sentences that had intonation, phonetics and prosody with those having
77 only prosody (Phonology: rotated > rotated+vocoded), we targeted the bridge from bottom-up to
78 top-down processing where auditory information is classified into linguistically meaningful units
79 (Brodbeck et al., 2018). As part of the top-down processes, sentences are judged for
80 intelligibility by the listener. We accessed this judgement by contrasting sentences that had
81 intelligibility, intonation, phonetics and prosody to those retaining intonation, phonetics and
82 prosody (Intelligibility: clear > rotated). Finally, meaning was attached to the intelligible
83 sentence, a process we accessed by contrasting two sentence types, one with all speech attributes
84 and the other lacking all attributes except prosody (Semantics: clear > rotated+vocoded).

85
86 All children responded with high accuracy on the sentence recognition task, and did so rapidly,
87 (Figure 1G) suggesting that they maintained attention throughout the task. TD children were
88 more accurate than children with LiD ($F(1, 83) = 8.77, p = .004, \eta_p^2 = .096$), but there was no
89 difference in accuracy between sentence types ($F(1.78, 148.03) = 2.11, p = .13, \eta_p^2 = .025$). There
90 was no significant difference in reaction time between the groups ($F(1, 83) = .89, p = .35, \eta_p^2 =$
91 $.011$). While there was no difference in accuracy between sentence types ($F(1.78, 148.03) = 2.11,$
92 $p = .13, \eta_p^2 = .025$), across all participants, clear sentences were responded to fastest ($F(2, 166) =$
93 $11.50, p < .001, \eta_p^2 = .12$) compared to rotated ($p < .001, d = -.41$) and rotated+vocoded ($p <$
94 $.001, d = -.46$) sentences. There were no interactions.

95
96

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

98 <Table 1: participant details>

99

100 *Cortical activation*

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

120

121 <Figure 2 - fMRI activation>

122

123 *Resting state fMRI*

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

128 session. The rs-fMRI scan allowed measurement of the temporal correlation of spontaneous
129 fluctuations in the BOLD signal between anatomically separated brain regions (Biswal, 2012),
130 capturing spontaneous interactions between brain regions in functionally associated networks
131 (Damoiseaux et al., 2006). Networks of ROIs suitable for rs-fMRI connectivity analysis were
132 created by parceling activation produced by each sentence task contrast (Phonology,
133 Intelligibility, and Semantics), combined across groups, using regions from the pediatric ADHD-
134 200 sample (Bellec et al., 2017; Craddock et al., 2012). This data-driven, spatially constrained
135 method divided Phonology-derived cortical activity into 17 ROIs, Intelligibility into 23 ROIs,
136 and Semantics into 31 ROIs (Figure 3; Brodman's area, maximum intensity coordinates and ROI
137 sizes are in Supplementary Material Table 1). Connectivity was compared between groups (TD =
138 42, LiD = 39), controlling for age, using a general linear model (GLM) for each of the three
139 networks (Figure 3A-C, Supplementary Material Table 1).

140
141 Figure 3 (D-F) shows resting state functional connectivity in the three speech networks, summed
142 across groups. In the Phonology network, each group had connectivity among regions covering
143 bilateral middle and superior temporal gyri, temporal pole, planum temporale, and left planum
144 polare and supramarginal gyrus (Figure 3D). No significant group differences (p-FDR) were
145 found in the Phonology network (Figure 4A). Each group had connectivity within the
146 Intelligibility network covering bilateral middle/superior temporal gyrus and temporal pole along
147 with left pars opercularis, frontal orbital cortex and supramarginal gyrus (Figure 3E). After false
148 discovery rate (FDR) correction, TD children were found to have a significantly weaker temporal
149 correlation between ROIs in Broca's area (left pars opercularis) and left middle temporal gyrus
150 (posterior) compared to children with LiD (connection 14-23 in Figures 4B, D, Table 3). In the
151 Semantics network, both groups of children had connectivity between bilateral middle and
152 superior temporal gyrus, left Heschl's gyrus, pars triangularis, frontal orbital cortex, planum
153 temporale, temporal fusiform gyrus and right parahippocampal gyrus and planum polare (Figure
154 3F). Group comparisons showed that, compared to children with LiD, the TD children had
155 weaker temporal correlations between ROIs in the right parahippocampal gyrus and left Heschl's
156 gyrus, left middle temporal gyrus, right superior temporal gyrus and right planum temporale
157 (connections 16-9, 16-19, 16-1 and 16-5, respectively, in Figures 4C, D, Table 3). However, they
158 also had stronger temporal correlations between ROIs in the left temporal fusiform cortex and

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

164

165 *<Figure 3 - parcellated ROIs and listening networks>*

166 *<Table 2 - selected regions of interest used in rs-fMRI>*

167 *<Figure 4 - connectivity between groups and effect sizes>*

168 *<Table 3 - significant connectivity stat values>*

169

170 *Relation of brain connectivity to behavioral measures*

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

185

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

187

188

189

190 **DISCUSSION**

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

199
200 Behavioural measures of performance were consistent with these observations. No significant
201 correlations were observed between temporal connectivity and sensory-dominant tasks (e.g,
202 LiSN-S advantage scores and SCAN filtered words). However, significant correlations were
203 observed between temporal connectivity and more cognitive-dominant tasks, especially speech-
204 in-noise and task switching. Overall, the data are consistent with the hypothesis that children
205 with LiD are primarily affected by cognitive deficits.

206

207 **Lateralization of speech in 6-12 year old children**

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

217

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

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

226
227 Bilateral activation reported here differs from the report of Scott et al. (2000) who found a left
228 lateralized pathway for speech comprehension using PET. Our results extend those of Halai et al.
229 (2015) who found bilateral activation using continuous MRI. Both these previous studies tested
230 young adults and used similar stimuli to those used here, but in continuous, passive speech
231 presentation, where comprehension was assessed after the scanning session. In contrast, we
232 required children to provide a behavioural response to each short sentence presentation as a
233 check that they were paying attention. In addition, we used a sparse/HUSH protocol compared to
234 the previous continuous scanning protocols. Scanner noise superimposed on speech listening has
235 been shown to increase listening effort (Peelle et al., 2010) and engage different cortical areas at
236 varying intensities. For example, in a meta-analysis of 57 speech comprehension studies, Adank
237 (2012) found that continuous scanning more strongly activated regions of the supplementary
238 motor area and anterior cingulate gyrus, while sparse scanning showed more extensive activation
239 in the STS.

240
241 The study reported here focused on speech listening in children with and without LiD, rather
242 than on differences across the tested age range. Our results support the hypothesis of bilateral
243 processing in speech listening in a 6-12 year old population. Bilateral activation produced by a
244 speech comprehension task has previously been shown in preschoolers (Sroka et al., 2015;
245 Holland et al., 2007), with an increase in lateralization for skills that develop after age 5 (e.g.
246 word-picture matching and syntactic prosody) (Holland et al., 2007). A meta-analysis of 27
247 developmental (ages 4-13) fMRI language comprehension studies suggested that lateralization in
248 the developing brain is associated with increased sensitivity to syntax (Enge et al., 2020).
249 However, using an alternative approach (Plante et al., 2015), cortical activation in young adults
250 became more left-lateralized as they learned an unfamiliar language. Plante's results thus suggest
251 that increasing speech lateralization in older children may occur through language learning rather

252 than through maturation. Recently, transcranial magnetic stimulation (TMS) has provided clear
253 evidence that stimulation of either left or right STG decreases understanding of speech-in-noise
254 in adults (Kennedy-Higgins et al., 2020). Together, these brain imaging and stimulation results
255 suggest that bilateral cortical activation is a life-long, obligate aspect of speech perception. The
256 results reported here are the first in an ongoing longitudinal study where the same cohort is being
257 invited back for behavioural and neuroimaging assessment two and four years after this baseline.
258

259 **Temporal connectivity: Brain and behaviour**

260 Phonology is the system of processing the smallest units of speech sounds and their linguistically
261 appropriate combinations. We found no group difference in the temporal connectivity of this
262 network. As we progressed to the Intelligibility network, we found that the temporal connection
263 between the left inferior frontal gyrus (Intelligibility ROI#14 in Table 2) and left posterior
264 middle temporal gyrus (Intelligibility ROI#23 in Table 2) was stronger in the children with LiD.
265 These areas are well known for language production (Broca's in the inferior frontal gyrus,
266 Hagoort, 2014) and comprehension (left MTG, (Acheson & Hagoort, 2013; Dronkers et al.,
267 2004). Posterior MTG has been associated with lexical and semantic access in a sound-to-
268 meaning network (Hickok & Poeppel, 2000, 2004). It is unknown whether stronger or weaker
269 temporal connectivity is beneficial. However, the relationship found here between connectivity
270 strength and an individual child's behavioural results provides evidence on this key issue.
271 Children with stronger connectivity performed more poorly on cognitive tasks assessing
272 vocabulary, switching attention and episodic memory. Reduced connectivity may thus be
273 indicative of increased processing efficiency and/or suppression of a task-relevant network, in
274 this case language. Note that the behavioural measures used here have been shown to be either
275 little affected by age (audiogram: Hunter et al., 2020) or were standardized across age (ECLiPS,
276 LiSN-S, SCAN and NIH toolbox; Petley et al., 2020).

277
278 Group comparisons in the Semantics network highlight two focal points. First, the right
279 parahippocampal gyrus, associated with memory encoding and retrieval (Luck et al., 2010), had
280 stronger temporal connections with auditory areas (left Heschl's, right planum temporale and
281 bilateral STG) in children with LiD. Second, the left temporal fusiform cortex, associated with
282 word recognition and the recovery of meaning from an impoverished acoustic signal (Davis &

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

294
295 The forebrain focal points identified in this study highlight language production, memory
296 encoding and retrieval, and word recognition as target areas of further research in children with
297 LiD. Combined, these focal points suggest that children with LiD have difficulty matching
298 meaning to understanding words in an auditory environment. The connectivity analysis also
299 highlights the left inferior frontal gyrus that, along with the superior temporal and premotor
300 regions, has been shown to be recruited for statistical learning of language (Davis & Gaskell,
301 2009; Karuza et al., 2013). Another avenue of research could therefore be how children with LiD
302 learn language compared to their peers. Language development is supported by distributed
303 neural networks connecting both cortical and subcortical regions throughout the brain
304 (Lieberman, 2002; Mesulam, 1990). However, such networks have not been closely or
305 systematically examined in children with LiD.

306
307 Assessing where the ‘break’ in speech listening occurs would provide a clearer avenue for
308 research into effective evidence-based treatments. Future studies could utilize paradigms with
309 anomalous and mispronounced words (e.g. Roebuck et al., 2018). Imaging techniques, for
310 example MRI, could be used in parallel to assess whether the cortical networks used in such
311 tasks are affected by LiD. Complementary time-sensitive techniques (EEG, MEG) could
312 investigate whether this difficulty is due to a bottleneck in processing leading to increased
313 listening effort, cognitive effort or fatigue. As LiD may build up over time (Roebuck & Barry,

314 2018), it is important to assess the children's ability throughout the task rather than using
315 summary values (McGarrigle et al., 2020).

316

317 **Relationship to neurodevelopmental disorders and cognitive function**

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

337

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

345
346 The results presented here highlight the importance of non-auditory factors, specifically
347 language, in audiometric testing. There is a growing recognition of the importance of speech
348 perception and, specifically, speech-in-noise (SiN) intelligibility in everyday hearing (Killion et
349 al., 2004; Smits et al., 2013). It has been proposed that such testing could supplement, or even
350 replace pure tone detection as an audiometric gold standard (Hewitt, 2018). However, both SiN
351 test instructions and test items pose a challenge to language and memory as well as auditory
352 function. While those cognitive aspects of auditory testing and learning have been dismissed as
353 procedural issues (Hawkey et al., 2004), they are an intimate component of a SiN test. These
354 results provide insight into mechanisms of how speech perception may be disrupted in LiD, a
355 common form of auditory impairment in both children (Moore et al., 2018) and adults (Edwards
356 et al., 2020). They also add to a growing literature on the role of cognitive function in hearing
357 (David R Moore et al., 2014; Rönnberg et al., 2013; Sharma et al., 2019).

358
359 **Conclusions**

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

368

369 **METHODS**

370 **Participants**

371 Eighty-five participants aged 6-12 years completed the fMRI sentence task and 81 participants
372 completed the rs-fMRI (see Table 1). Ninety percent completed both tasks. All participants had
373 normal audiometric hearing with thresholds < 25 dB HL at all octave-interval frequencies from
374 0.25 - 16 kHz in both ears (Figure 1A). In this paper we focus on the cortical results and their
375 relationship with behavioural responses from the baseline of our longitudinal ‘SICLID’ study
376 examining correlates of LiD in children (Hunter et al., 2020; Petley et al., 2020; Blankenship et
377 al., in prep.).

378
379 Caregivers of all participants completed a well-validated checklist of everyday listening and
380 related skills (the ECLiPS; Barry & Moore, 2014). Those scoring within the *clinical interest*
381 range on the ECLiPS (< 10th percentile of ECLiPS standardized scores), or had a diagnosis of
382 auditory processing disorder (APD; n=14), were classified as children with LiD (Figure 1B).
383 Participants scoring within the *normal* range on the ECLiPS and with no history of
384 developmental disorders or delays were classified as TD children.

385

386 **MRI acquisition**

387 MRI was performed via a 3T Philips Ingenia scanner with a 64-channel head coil and Avotec
388 audiovisual system. All participants were awake and non-sedated throughout the scanning. The
389 protocol included a high-resolution T1-weighted anatomical scan, fMRI sentence task (4.9
390 minutes) and rs-fMRI (5 minutes). The fMRI sentence task was acquired with a sparse scanning
391 protocol (‘HUSH’, details below); TR/TE = 2000/30 ms, voxel size = $2.5 \times 2.5 \times 3.5$ mm, 39
392 axial slices. A total of 147 volumes was acquired by alternating scanning for 6 seconds (3
393 volumes) and not scanning for 6 seconds, 49 times. Cardiac and respiration signals were
394 collected during the fMRI sentence task using the scanner’s wireless respirator bellows and
395 Peripheral Pulse Oximeter. The rs-fMRI acquisition was acquired with TR/TE = 2000/30 ms,
396 voxel size = $2.5 \times 2.5 \times 3.5$ mm, 39 axial slices in ascending slice order and 150 volumes. The
397 high-resolution T1-weighted anatomical scan was acquired with TR/TE=8.1/3.7 ms, FOV 25.6 x
398 25.6 x 16.0 cm, matrix 256 x 256 and slice thickness = 1mm.

399

400 **fMRI task**

401 With sound levels reaching 118.4 ± 1.3 dB (A) in a 3T MRI system (Price et al., 2001) special
402 considerations must be made when planning an auditory-based MRI study. In order to protect the
403 participant from the loud environment, foam ear plugs and MRI safe circumaural headphones
404 were worn. The scanner noise may also produce masking of the desired stimuli. Therefore, in the
405 fMRI task we used a ‘Hemodynamics Unrelated to Sounds of Hardware’ (HUSH) scanning
406 protocol (Deshpande et al., 2016; Edmister et al., 1999; Hall et al., 1999; Schmithorst & Holland,
407 2004) - a sparse temporal sampling protocol where there was no gradient coil noise during
408 presentation of the auditory stimuli. We also used a talker identification task instead of a speech
409 recognition task. Instead of asking the children *what* they heard, we asked them *who* had said it.
410 The children responded with button presses throughout the task so we could ensure they
411 maintained attention to the task.

412

413 Sixteen linguistically simple BKB sentences, designed to be familiar to young children, were
414 recorded by a single male North American speaker, mirroring the paradigm used by Scott et al.
415 (2000) and Halai et al. (2015). These were intelligible, clear speech stimuli (e.g. Figure 1D).
416 Rotated speech stimuli (Figure 1E) were created by rotating each sentence spectrally around 2
417 kHz using the Blesser (1972) technique. Rotated speech was not intelligible, though some
418 phonetic features and some of the original intonation was preserved. Rotated+vocoded speech
419 stimuli (Figure 1F) were created by applying 6 band noise-vocoding to the rotated speech
420 stimuli. While the rotated noise-vocoded speech was completely unintelligible, the character of
421 the envelope and some spectral detail was preserved.

422

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

429

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

438

439 **fMRI data analysis**

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

444

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

461
462 Second-level analysis was also conducted using FSL. A GLM approach was used to create group
463 activation maps based on contrasts between conditions for all participants (i.e. regardless of
464 LiD/TD status) with age as a covariate. Group composite images were thresholded using a
465 family-wise error correction ($p < 0.001$) and clustering threshold of $k = 4$ voxels. Three BOLD
466 activation contrasts were used to search for brain loci responding to different aspects of listening
467 to language (Halai et al., 2015; modified from Scott, 2000). First, a ‘Semantics’ activation map,
468 whereby the signal with intelligibility, intonation, phonetics, prosody and sound was contrasted
469 with one lacking all of these attributes except prosody and sound (clear > rotated+vocoded).
470 Second, an ‘Intelligibility’ activation map contrasted the signal with all speech attributes to one
471 retaining intonation, phonetics, prosody and sound (clear > rotated). Third, a ‘Phonetics’
472 activation map contrasted a signal with intonation, phonetics, prosody and sound with one
473 having only prosody and sound (rotated > rotated+vocoded).

474
475 For the rs-fMRI scan, pre-processing and analysis was performed in the CONN toolbox using
476 standard spatial and temporal pipelines (Whitfield-Gabrieli & Nieto-Castanon, 2012). For spatial
477 smoothing a FWHM of 8mm was used. The Artifact Detection Tool (ART,
478 https://www.nitrc.org/projects/artifact_detect) within CONN was used to regress out framewise
479 motion. The number of frames regressed out was compared between groups with no significant
480 group differences, $p = .44$ (Table 1).

481
482 Behavioural responses from the fMRI task were assessed using a 2 (group: TD, LiD) \times 3
483 (sentence type: clear, rotated, rotated+vocoded) repeated measures analysis of variance
484 (ANOVA) for accuracy and again for RT (Figure 1G). Where the assumption of sphericity was
485 violated, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

486 487 **Resting state fMRI**

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

492

493 **Resting state ROI-to-ROI analysis**

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

500

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

508

509 **Caregiver questionnaire**

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

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

516

517 **Behavioral measures**

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

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

526

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

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

534

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

541

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

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

552

553

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

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

561

562 **DATA AVAILABILITY**

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

566

567 **ACKNOWLEDGEMENTS**

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

572

573 **FUNDING**

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

577

578 **COMPETING INTERESTS**

579 The authors report no competing interests.

580

581 **CONTRIBUTIONS**

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

584 EC - behavioural and rs-fMRI analysis, figures, assisted with manuscript preparation and
585 revisions
586 TM - fMRI preprocessing, assisted with manuscript preparation
587 JV - study design, analysis and presentation, manuscript preparation and revisions
588 DRM - study design and presentation, funding, manuscript preparation and revisions
589
590

591 **REFERENCES**

- 592 Acheson, D. J., & Hagoort, P. (2013). Stimulating the brain's language network: Syntactic
593 ambiguity resolution after TMS to the inferior frontal gyrus and middle temporal gyrus.
594 *Journal of Cognitive Neuroscience*, 25(10), 1664–1677.
- 595 Adank, P. (2012). The neural bases of difficult speech comprehension and speech production:
596 Two Activation Likelihood Estimation (ALE) meta-analyses. *Brain and Language*, 122(1),
597 42–54.
- 598 Assaneo, M. F., Orpella, J., Ripolles, P., Diego-Balaguer, D., Poeppel, D., & others. (2019). The
599 lateralization of speech-brain coupling is differentially modulated by intrinsic auditory and
600 top-down mechanisms. *Frontiers in Integrative Neuroscience*, 13, 28.
- 601 Barry, J. G., & Moore, D. R. (2014). *Evaluation of children's listening and processing skills*
602 *(ECLiPS)*.
- 603 Barry, J. G., Tomlin, D., Moore, D. R., & Dillon, H. (2015). *Use of Questionnaire-Based*
604 *Measures in the Assessment of Listening Difficulties in School-Aged Children*. 1–14.
- 605 Bellec, P., Chu, C., Chouinard-Decorte, F., Benhajali, Y., Margulies, D. S., & Craddock, R. C.
606 (2017). The neuro bureau adhd-200 preprocessed repository. *Neuroimage*, 144, 275–
607 286.
- 608 Bishop, D. V., & Snowling, M. J. (2004). Developmental dyslexia and specific language
609 impairment: Same or different? *Psychological Bulletin*, 130(6), 858.
- 610 Biswal, B. B. (2012). Resting state fMRI: a personal history. *Neuroimage*, 62(2), 938–944.
- 611 Blesser, B. (1972). Speech perception under conditions of spectral transformation: I. Phonetic
612 characteristics. *Journal of Speech and Hearing Research*, 15(1), 5–41.
- 613 Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid Transformation from Auditory to
614 Linguistic Representations of Continuous Speech. *Current Biology*, 28(24), 3976-
615 3983.e5. <https://doi.org/10.1016/j.cub.2018.10.042>
- 616 Cameron, S., & Dillon, H. (2007). Development of the listening in spatialized noise-sentences

- 617 test (LISN-S). *Ear and Hearing*, 28(2), 196–211.
- 618 <https://doi.org/10.1097/AUD.0b013e318031267f>
- 619 Cameron, S., & Dillon, H. (2008). The listening in spatialized noise-sentences test (LISN-S):
620 Comparison to the prototype LISN and results from children with either a suspected
621 (central) auditory processing disorder or a confirmed language disorder. *Journal of the*
622 *American Academy of Audiology*, 19(5), 377–391. <https://doi.org/10.3766/jaaa.19.5.2>
- 623 Cameron, S., & Dillon, H. (2009). *Listening in Spatialized Noise–Sentences test (LiSN-S)*.
- 624 Craddock, R. C., James, G. A., Holtzheimer III, P. E., Hu, X. P., & Mayberg, H. S. (2012). A
625 whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human*
626 *Brain Mapping*, 33(8), 1914–1928.
- 627 Damoiseaux, J. S., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., &
628 Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects.
629 *Proceedings of the National Academy of Sciences*, 103(37), 13848–13853.
- 630 Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning:
631 Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B:*
632 *Biological Sciences*, 364(1536), 3773–3800.
- 633 Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language
634 comprehension. *The Journal of Neuroscience : The Official Journal of the Society for*
635 *Neuroscience*, 23(8), 3423–3431. <https://doi.org/23/8/3423> [pii]
- 636 Dawes, P., & Bishop, D. V. M. (2010). Psychometric profile of children with auditory processing
637 disorder and children with dyslexia. *Archives of Disease in Childhood*, 95(6), 432–436.
638 <https://doi.org/10.1136/adc.2009.170118>
- 639 Delage, H., & Durrleman, S. (2018). Developmental dyslexia and specific language impairment:
640 Distinct syntactic profiles? *Clinical Linguistics & Phonetics*, 32(8), 758–785.
- 641 Deshpande, A. K., Tan, L., Lu, L. J., Altaye, M., & Holland, S. K. (2016). fMRI as a preimplant
642 objective tool to predict postimplant oral language outcomes in children with cochlear

- 643 implants. *Ear and Hearing*, 37(4), e263–e272.
- 644 Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion
645 analysis of the brain areas involved in language comprehension. *Cognition*, 92(1–2),
646 145–177. <https://doi.org/10.1016/j.cognition.2003.11.002>
- 647 Edmister, W. B., Talavage, T. M., Ledden, P. J., & Weisskoff, R. M. (1999). Improved auditory
648 cortex imaging using clustered volume acquisitions. *Human Brain Mapping*, 7(2), 89–97.
- 649 Edwards, B. (2020). *Emerging technologies, market segments, and MarkeTrak 10 insights in*
650 *hearing health technology*.
- 651 Emanuel, D. C., Ficca, K. N., & Korczak, P. (2011). Survey of the diagnosis and management of
652 auditory processing disorder. *American Journal of Audiology*.
- 653 Enge, A., Friederici, A. D., & Skeide, M. A. (2020). A meta-analysis of fMRI studies of language
654 comprehension in children. *NeuroImage*, 116858.
- 655 Evans, S., & McGettigan, C. (2017). Comprehending auditory speech: Previous and potential
656 contributions of functional MRI. *Language, Cognition and Neuroscience*, 32(7), 829–846.
- 657 Ferguson, M. a, Hall, R. L., Riley, A., & Moore, D. R. (2011). Communication, listening, cognitive
658 and speech perception skills in children with auditory processing disorder (APD) or
659 Specific Language Impairment (SLI). *Journal of Speech, Language, and Hearing*
660 *Research : JSLHR*, 54(1), 211–227. [https://doi.org/10.1044/1092-4388\(2010/09-0167\)](https://doi.org/10.1044/1092-4388(2010/09-0167))
- 661 Glover, G. H., Li, T.-Q., & Ress, D. (2000). Image-based method for retrospective correction of
662 physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine:*
663 *An Official Journal of the International Society for Magnetic Resonance in Medicine*,
664 44(1), 162–167.
- 665 Gokula, R., Sharma, M., Cupples, L., & Valderrama Valenzuela, J. T. (2019). Comorbidity of
666 auditory processing, attention, and memory in children with word reading difficulties.
667 *Frontiers in Psychology*, 10, 2383.
- 668 Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region

- 669 and beyond. *Current Opinion in Neurobiology*, 28, 136–141.
- 670 Halai, A. D., Parkes, L. M., & Welbourne, S. R. (2015). Dual-echo fMRI can detect activations in
671 inferior temporal lobe during intelligible speech comprehension. *NeuroImage*, 122, 214–
672 221. <https://doi.org/10.1016/j.neuroimage.2015.05.067>
- 673 Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R.,
674 Gurney, E. M., & Bowtell, R. W. (1999). “Sparse” temporal sampling in auditory fMRI.
675 *Human Brain Mapping*, 7(3), 213–223.
- 676 Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O’Neal, L., McLeod, L.,
677 Delacqua, G., Delacqua, F., Kirby, J., & others. (2019). The REDCap consortium:
678 Building an international community of software platform partners. *Journal of Biomedical*
679 *Informatics*, 95, 103208.
- 680 Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., Conde, J. G., & others. (2009). A
681 metadata-driven methodology and workflow process for providing translational research
682 informatics support. *J Biomed Inform*, 42(2), 377–81.
- 683 Hawkey, D. J., Amitay, S., & Moore, D. R. (2004). Early and rapid perceptual learning. *Nature*
684 *Neuroscience*, 7(10), 1055–1056.
- 685 Hewitt, D. (2018). Audiometry and Its Discontents. *Hearing Review*, 25(1), 20–23.
- 686 Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception.
687 *Trends in Cognitive Sciences*, 4(4), 131–138.
- 688 Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding
689 aspects of the functional anatomy of language. *Cognition*, 92(1–2), 67–99.
- 690 Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature*
691 *Reviews Neuroscience*, 8(5), 393–402.
- 692 Hunter, L. L., Blankenship, C. M., Lin, L., Sloat, N. T., Perdew, A., Stewart, H., & Moore, D. R.
693 (2020). Peripheral auditory involvement in childhood listening difficulty. *Ear and Hearing*.
- 694 Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013).

- 695 The neural correlates of statistical learning in a word segmentation task: An fMRI study.
696 *Brain and Language*, 127(1), 46–54.
- 697 Keith, R. W. (2000). Development and Standardization of SCAN-C Test for Auditory Processing
698 Disorders in Children. *J Am Acad Audiol*, 11(8), 438–445.
- 699 Keith, R. W. (2009). *SCAN-3 for Children. Tests of Auditory Processing Disorders*. Pearson.
- 700 Kennedy-Higgins, D., Devlin, J. T., Nuttall, H. E., & Adank, P. (2020). The causal role of left and
701 right superior temporal gyri in speech perception in noise: A transcranial magnetic
702 stimulation study. *Journal of Cognitive Neuroscience*, 32(6), 1092–1103.
- 703 Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004).
704 Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in
705 normal-hearing and hearing-impaired listeners. *Journal of Acoustical Society of America*,
706 116, 2395–2405.
- 707 Lieberman, P. (2002). On the nature and evolution of the neural bases of human language.
708 *American Journal of Physical Anthropology: The Official Publication of the American*
709 *Association of Physical Anthropologists*, 119(S35), 36–62.
- 710 Luck, D., Danion, J.-M., Marrer, C., Pham, B.-T., Gounot, D., & Foucher, J. (2010). The right
711 parahippocampal gyrus contributes to the formation and maintenance of bound
712 information in working memory. *Brain and Cognition*, 72(2), 255–263.
- 713 McGarrigle, R., Rakusen, L., & Mattys, S. (2020). *Effortful listening under the microscope:*
714 *Examining relations between pupillometric and subjective markers of effort and tiredness*
715 *from listening*.
- 716 Mesulam, M.-M. (1990). Large-scale neurocognitive networks and distributed processing for
717 attention, language, and memory. *Annals of Neurology: Official Journal of the American*
718 *Neurological Association and the Child Neurology Society*, 28(5), 597–613.
- 719 Moore, D R. (2018). Auditory processing disorder (APD). *Ear and Hearing*, 39(4), 617.
- 720 Moore, D R, Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., & Riley, A. (2010). Nature of

- 721 Auditory Processing Disorder in Children. *Pediatrics*, 126(2), E382–E390.
722 <https://doi.org/10.1542/peds.2009-2826>
- 723 Moore, D R, Sieswerda, S. L., Grainger, M. M., Bowling, A., Smith, N., Perdew, A., Eichert, S.,
724 Alston, S., Hilbert, L. W., Summers, L., & others. (2018). Referral and diagnosis of
725 developmental auditory processing disorder in a large, United States hospital-based
726 audiology service. *Journal of the American Academy of Audiology*, 29(5), 364–377.
- 727 Moore, David R, Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R.
728 H., & Munro, K. J. (2014). Relation between speech-in-noise threshold, hearing loss and
729 cognition from 40–69 years of age. *PloS One*, 9(9), e107720.
- 730 Moore, David R., Hugdahl, K., Stewart, H. J., Vannest, J., Perdew, A. J., Sloat, N. T., Cash, E.
731 K., & Hunter, L. L. (2020). Listening Difficulties in Children: Behavior and Brain Activation
732 Produced by Dichotic Listening of CV Syllables. *Frontiers in Psychology*, 11.
733 <https://doi.org/10.3389/fpsyg.2020.00675>
- 734 Neijenhuis, K., Campbell, N., Cromb, M., Luinge, M., Moore, D. R., Rosen, S., & de Wit, E.
735 (2019). An Evidence-based Perspective on ‘Misconceptions’ Regarding Pediatric
736 Auditory Processing Disorder. *Frontiers in Neurology*, 10, 287.
- 737 Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I. H., Saberi, K., Serences, J. T., &
738 Hickok, G. (2010). Hierarchical organization of human auditory cortex: Evidence from
739 acoustic invariance in the response to intelligible speech. *Cerebral Cortex*, 20(10),
740 2486–2495. <https://doi.org/10.1093/cercor/bhp318>
- 741 Parente, F., Frascarelli, M., Mirigliani, A., Di Fabio, F., Biondi, M., & Colosimo, A. (2018).
742 Negative functional brain networks. *Brain Imaging and Behavior*, 12(2), 467–476.
- 743 Peelle, J. E. (2012). The hemispheric lateralization of speech processing depends on what
744 “speech” is: A hierarchical perspective. *Frontiers in Human Neuroscience*, 6, 309.
- 745 Peelle, J. E., Eason, R. J., Schmitter, S., Schwarzbauer, C., & Davis, M. H. (2010). Evaluating
746 an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory

- 747 processing. *Neuroimage*, 52(4), 1410–1419.
- 748 Petley, L., Hunter, L. L., Motlagh Zadeh, L., Sloat, N., Perdew, A., & Moore, D. R. (2020).
749 Listening Difficulties in Children with Normal Audiograms: Relation to Hearing and
750 Cognition. *MedRxiv*. <https://doi.org/10.1101/2020.10.05.20205468>
- 751 Plante, E., Almryde, K., Patterson, D. K., Vance, C. J., & Asbjørnsen, A. E. (2015). Language
752 lateralization shifts with learning by adults. *Laterality: Asymmetries of Body, Brain and*
753 *Cognition*, 20(3), 306–325.
- 754 Price, D. L., De Wilde, J. P., Papadaki, A. M., Curran, J. S., & Kitney, R. I. (2001). Investigation
755 of acoustic noise on 15 MRI scanners from 0.2 T to 3 T. *Journal of Magnetic Resonance*
756 *Imaging: An Official Journal of the International Society for Magnetic Resonance in*
757 *Medicine*, 13(2), 288–293.
- 758 Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman
759 primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718–724.
- 760 Roebuck, H., & Barry, J. G. (2018). Parental perception of listening difficulties: An interaction
761 between weaknesses in language processing and ability to sustain attention. *Scientific*
762 *Reports*, 8(1), 1–10.
- 763 Roebuck, H., Sindberg, H., & Weismer, S. E. (2018). The role of language in nonlinguistic
764 stimuli: Comparing inhibition in children with language impairment. *Journal of Speech*
765 *Language and Hearing Research*, 61(5), 1216. [https://doi.org/10.1044/2018_JSLHR-L-](https://doi.org/10.1044/2018_JSLHR-L-17-0294)
766 17-0294
- 767 Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language
768 understanding. *International Journal of Audiology*, 58(5), 247–261.
- 769 Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö.,
770 Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & others. (2013). The Ease of Language
771 Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in*
772 *Systems Neuroscience*, 7, 31.

- 773 Schmithorst, V. J., & Holland, S. K. (2004). Event-related fMRI technique for auditory processing
774 with hemodynamics unrelated to acoustic gradient noise. *Magnetic Resonance in*
775 *Medicine: An Official Journal of the International Society for Magnetic Resonance in*
776 *Medicine*, 51(2), 399–402.
- 777 Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for
778 intelligible speech in the left temporal lobe. *Brain*, 123(12), 2400–2406.
779 <https://doi.org/10.1093/brain/123.12.2400>
- 780 Sharma, M., Purdy, S. C., & Humburg, P. (2019). Cluster analyses reveals subgroups of
781 children with suspected auditory processing disorders. *Frontiers in Psychology*, 10,
782 2481.
- 783 Sharma, M., Purdy, S. C., & Kelly, A. S. (2009). Comorbidity of Auditory Processing, Language,
784 and Reading Disorders. *Journal of Speech Language and Hearing Research*, 52(3), 706.
785 [https://doi.org/10.1044/1092-4388\(2008/07-0226\)](https://doi.org/10.1044/1092-4388(2008/07-0226))
- 786 Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in*
787 *Cognitive Sciences*, 12(5), 182–186.
- 788 Smits, C., Theo Goverts, S., & Festen, J. M. (2013). The digits-in-noise test: Assessing auditory
789 speech recognition abilities in noise. *The Journal of the Acoustical Society of America*,
790 133(3), 1693–1706.
- 791 Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., Carter, A., Casey, D. C.,
792 Charlson, F. J., Chen, A. Z., & others. (2016). Global, regional, and national incidence,
793 prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A
794 systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*,
795 388(10053), 1545–1602.
- 796 Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi,
797 N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., & others. (2013). Cognition assessment
798 using the NIH Toolbox. *Neurology*, 80(11 Supplement 3), S54–S64.

799 Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). *Conn*: A Functional Connectivity Toolbox for
800 Correlated and Anticorrelated Brain Networks. *Brain Connectivity*, 2(3), 125–141.

801 <https://doi.org/10.1089/brain.2012.0073>

802 Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-
803 scale automated synthesis of human functional neuroimaging data. *Nature Methods*,
804 8(8), 665–670.

805

806

807

808

809

810 **Table 1:** Participant details broken down by scan type - fMRI speech listening task and the
 811 resting state (RS).
 812

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)

813 *High school graduate or less, some college or more

814

815

816

817 **Table 2:** Selected regions of interest (ROIs, from the fMRI task) used in the rs-fMRI ROI-to-
 818 ROI analysis. Threshold = 4.0, MCC .95. The full table of all ROIs can be found in
 819 Supplementary Material Table 1.
 820

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	

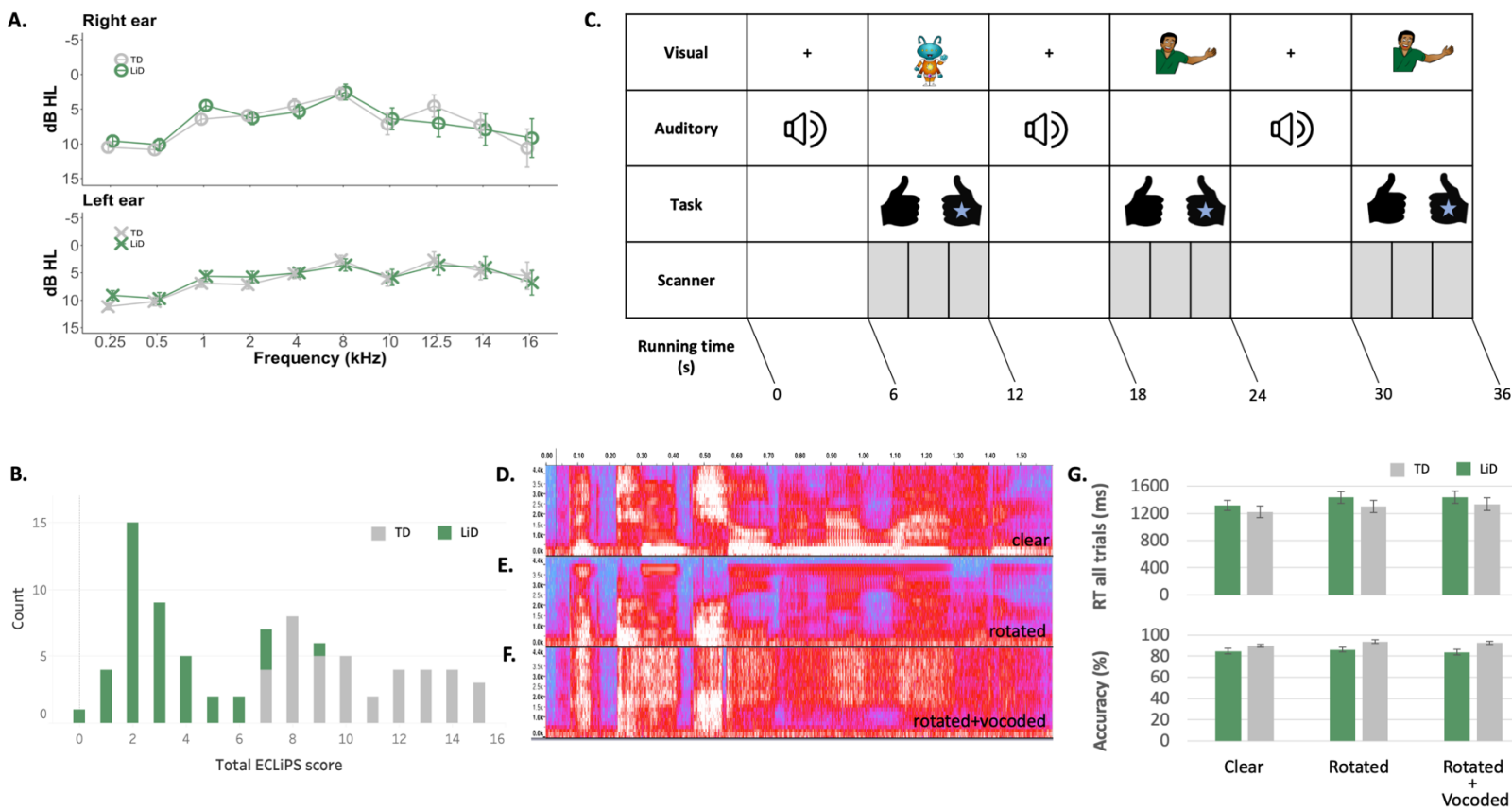
821

822

823 **Table 3:** No significant group difference (TD > LiD, corrected for age) was found in the
824 temporal connectivity of the Phonology network. Group differences were found in the more
825 advanced speech listening networks (Intelligibility and Semantics). See Figure 4.
826

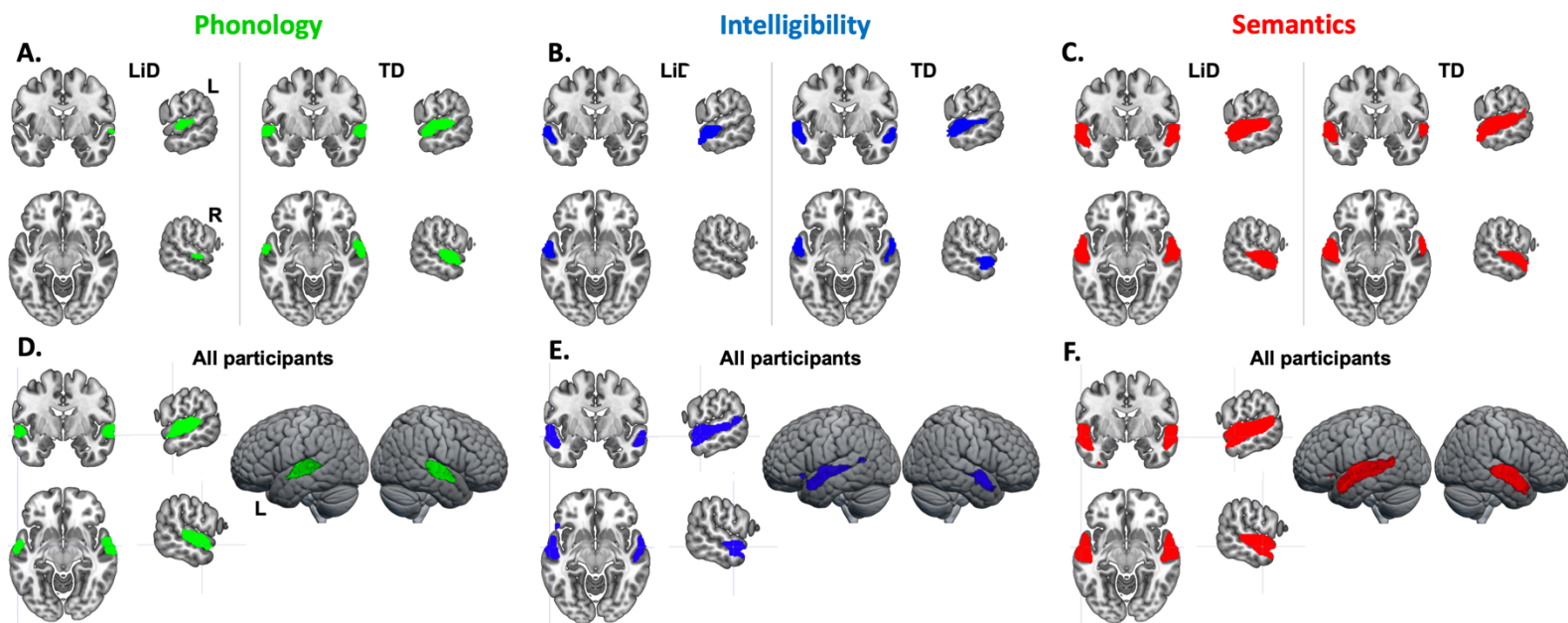
Contrast	ROI # – ROI #	t (78)	p-FDR
Intelligibility	14 – 23	-3.57	0.016
Semantics	16 – 1	-3.79	0.009
	16 – 5	-3.05	0.033
	16 – 19	-3.04	0.033
	16 – 9	-2.92	0.034
	10 – 27	3.18	0.041
	10 – 8	3.1	0.041

827
828
829
830
831
832



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

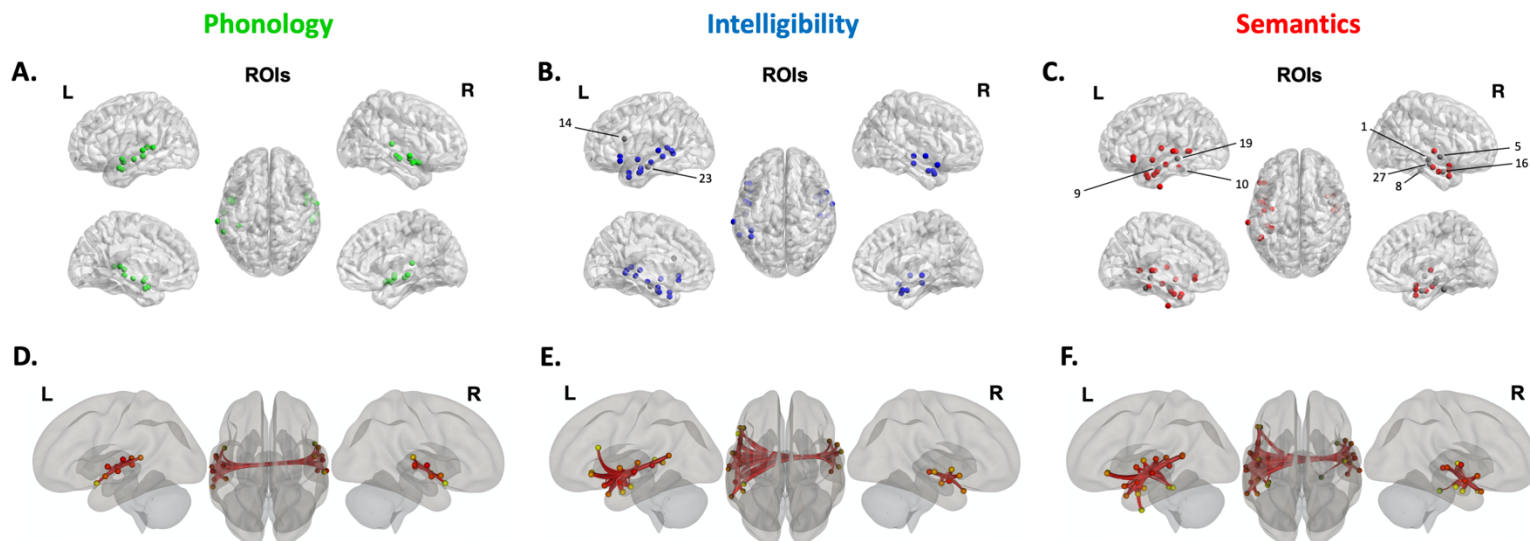
846 sentences. Error bars represent the standard error of the mean. TD children are in grey and LiD
847 in green.
848
849



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

860

861



862 **Figure 3:** ROIs used for the rs-functional connectivity and their networks across all participants.

863 The cortical activity across all participants in the fMRI task (see Figure 2 D-F) covered large

864 areas and so they were parcellated into smaller ROIs (A-C) for the rs-functional connectivity

865 analysis by applying data-driven spatially constrained parcellation to the areas of activation from

866 the fMRI sentence recognition task using the pediatric ADHD-200 sample. The red lines (D-F)

867 indicate the ROI-to-ROI connections analyzed in each network. Maximum intensity coordinates

868 can be found in Supplementary Material Table 1. Networks are: (A, D) Phonology (green:

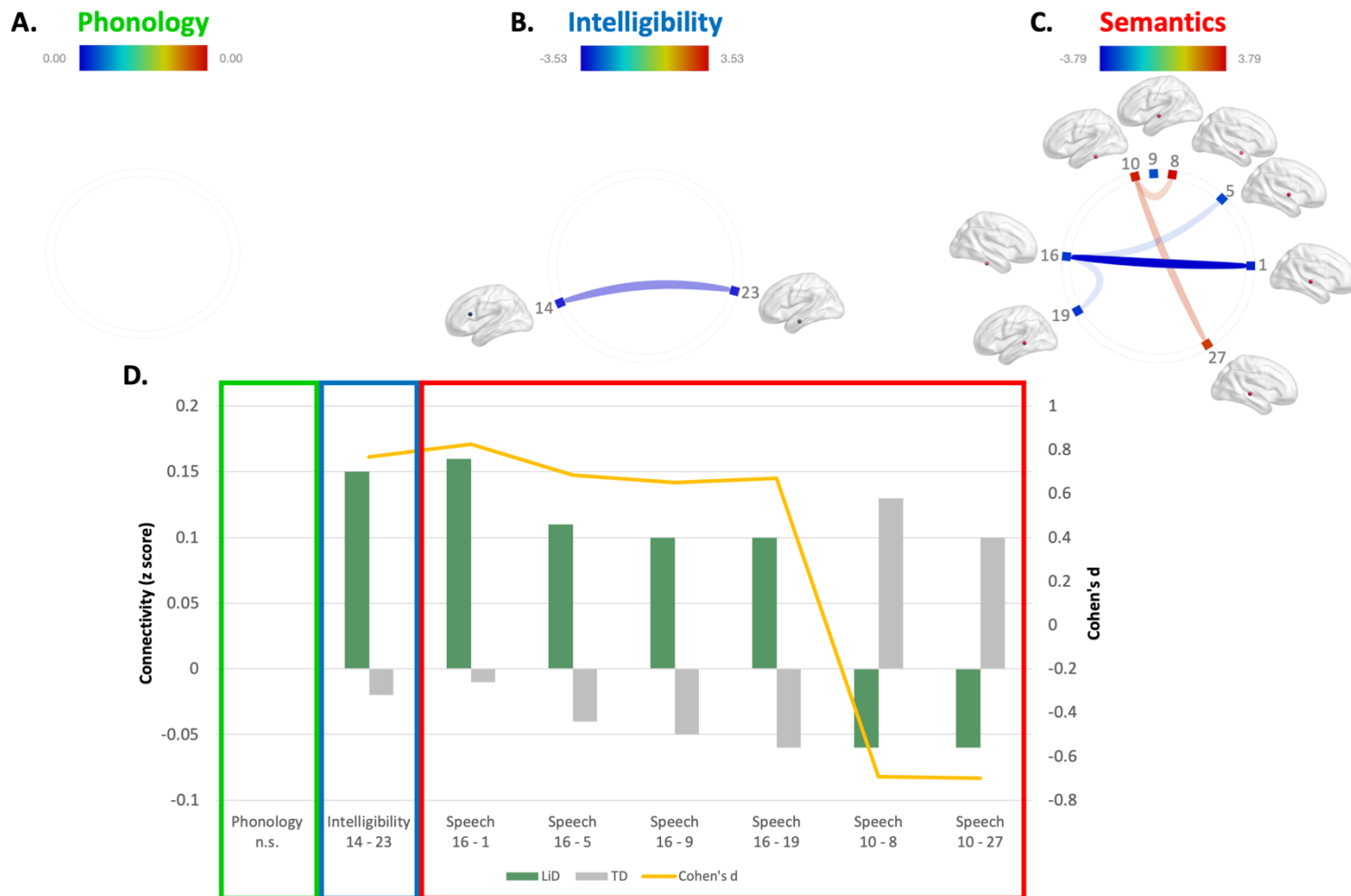
869 rotated > rotated+vocoded), (B, E) Intelligibility (blue: clear > rotated) and (C, F) Semantics

870 (red: clear > rotated+vocoded). Images are in neurological orientation. For visualization,

871 BrainNet software was used to display foci in (A-C) (Xia, Wang & He, 2013).

872

873



874

875 **Figure 4:** ROI-to-ROI resting state connectivity: the difference between the groups' listening

876 networks grew from no statistical differences in the phonology network to a minor difference in

877 the intelligibility network to more widespread differences in the semantics network. (A-C) The

878 group comparisons without the effect of age for (A) Phonology, (B) Intelligibility and (C)

879 Speech networks. Thicker, more saturated color lines represent stronger connections between

880 cortical areas. Note that the colored bar connectivity z score scales vary slightly between

881 connectivity wheels. (D) Details of the ROI-to-ROI connectivity values (left axis) for each group

882 (LiD green, TD grey) and effect sizes (yellow line, right axis) of the group comparisons without

883 the effect of age. The connections plotted are the ones highlighted as having a significant group

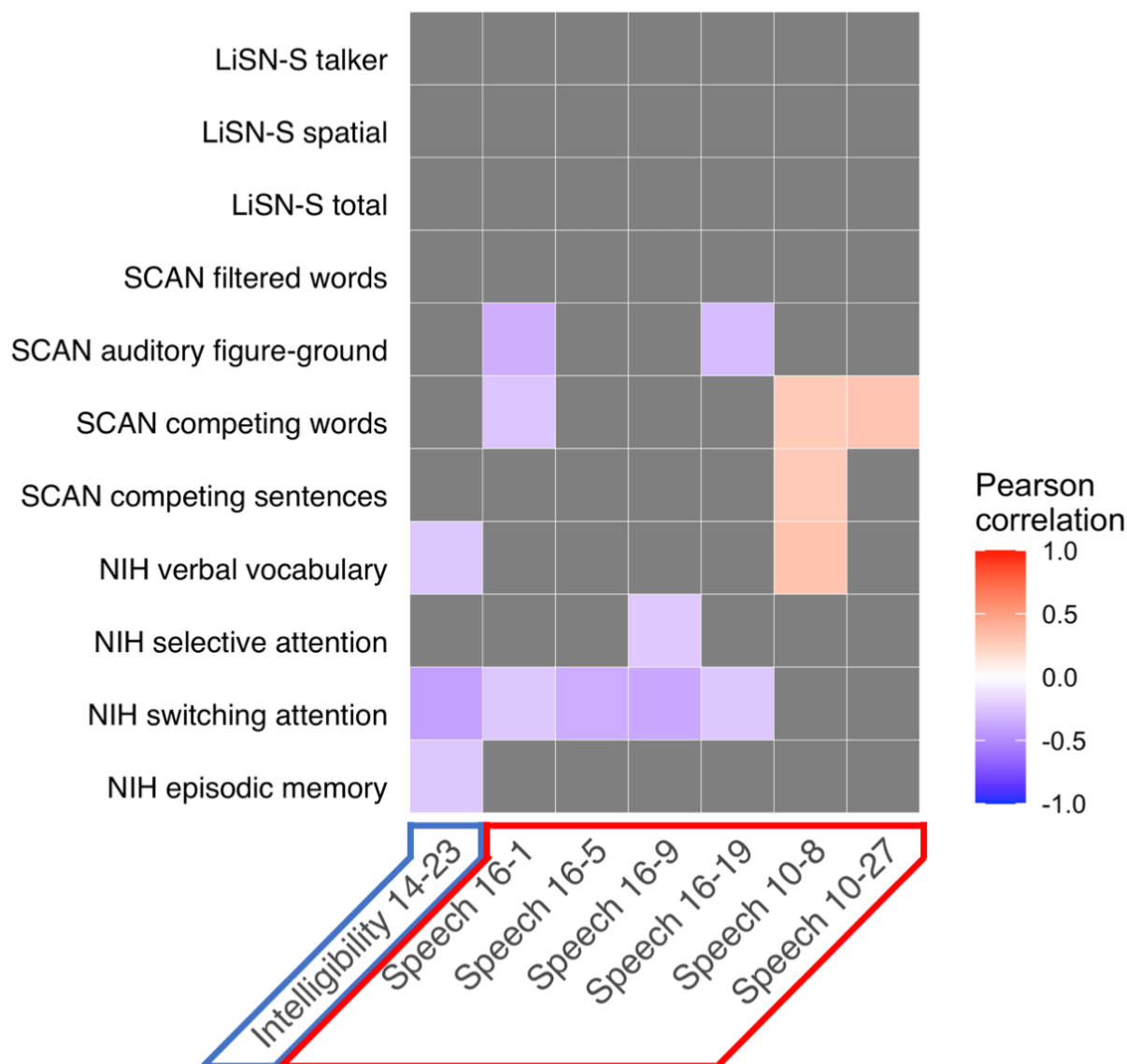
884 difference in the GLM comparing groups without the effect of age (B and C). Connections

885 Intelligibility 14-23 and Speech 16-1,16-5, 16-9 and 16-19 all show the TD group as having

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

888

889



890

891

892 **Figure 5:** Heat map of brain and behavioral score correlations. Colored squares indicate
 893 significant low-moderate Pearson correlations ($p < .05$) between the connections in Figure 3D
 894 and speech in noise, auditory processing skills and cognition. Correlations associated with the
 895 Broca and right posterior parahippocampal gyrus connections are negative while correlations
 896 associated with the left posterior temporal fusiform cortex connections are positive. See
 897 Supplementary Materials Table 3 for correlation details. Note that the direction of the
 898 correlations flip for the same connections as in the group comparisons in Figure 3 B-D.

899 Intelligibility 14-23 and Speech 16-1,16-5, 16-9 and 16-19 showed that the TD group had weaker

900 connectivity compared to the LiD group and these connections show negative correlations with
901 behavioral scores. While Speech 10-8 and 10-27 showed that the TD group had stronger
902 connectivity compared to the LiD group and these connections show positive correlations with
903 behavioral scores.
904