

1 **Speech cortical activation and connectivity in typically developing children and those with**
2 **listening difficulties**

3

4 **Abbreviated title:** Cortical speech networks and listening

5

6 Hannah J. Stewart^{1,2,3}, Erin K. Cash², Lisa L. Hunter², Thomas Maloney⁴, Jennifer Vannest^{2,5}
7 and David R. Moore^{2,6,7}

8

- 9 1. Division of Psychology and Language Sciences, University College London, London,
10 UK
11 2. Communication Sciences Research Center, Cincinnati Children's Hospital Medical
12 Center, Cincinnati, Ohio, USA
13 3. Department of Psychology, Lancaster University, Lancaster, UK
14 4. Pediatric Neuroimaging Research Consortium, Cincinnati Children's Hospital Medical
15 Center, Cincinnati, Ohio, USA
16 5. Department of Communication Sciences and Disorders, University of Cincinnati, Ohio,
17 USA
18 6. Department of Otolaryngology, College of Medicine, University of Cincinnati,
19 Cincinnati, Ohio, USA
20 7. Manchester Centre for Audiology and Deafness, University of Manchester, Manchester,
21 M13 9PL, UK

22

23 **Correspondence:** Hannah J. Stewart
24 Department of Psychology
25 Lancaster University
26 Lancaster UK
27 h.stewart5@lancaster.ac.uk

28

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

30

31

32 **Abstract:**

33 Listening difficulties (LiD) in people who have normal audiometry are a widespread but poorly
34 understood form of hearing impairment. Recent research suggests that childhood LiD are cognitive
35 rather than auditory in origin. We examined decoding of sentences using a novel combination of
36 behavioral testing and fMRI with 43 typically developing children and 42 age matched (6-13 years
37 old) children with LiD, categorized by caregiver report (ECLiPS). Both groups had clinically
38 normal hearing. For sentence listening tasks, we found no group differences in fMRI brain cortical
39 activation by increasingly complex speech stimuli that progressed in emphasis from phonology to
40 intelligibility to semantics. Using resting state fMRI, we examined the temporal connectivity of
41 cortical auditory and related speech perception networks. We found significant group differences
42 only in cortical connections engaged when processing more complex speech stimuli. The strength
43 of the affected connections was related to the children's performance on tests of dichotic listening,
44 speech-in-noise, attention, memory and verbal vocabulary. Together, these results support the
45 novel hypothesis that childhood LiD reflects difficulties in language rather than in auditory or
46 phonological processing.

47

48 **Abbreviations:**

49 CANS central auditory nervous system
50 DLD developmental language disorder
51 rs-fMRI resting state-fMRI
52 FOV field-of-view
53 FWE family wise error correction
54 GLM general linear model
55 HUSH Hemodynamics Unrelated to Sounds of Hardware
56 LiD listening difficulties
57 MEG magnetoencephalography
58 MNI Montreal Neurological Institute
59 MTG middle temporal gyrus
60 FDR false detection rate
61 ROI region-of-interest
62 SLI specific language impairment
63 STG superior temporal gyrus
64 STS superior temporal sulcus
65 TD typical developing
66 TE echo time
67 TMS transcranial magnetic stimulation

68 TR repetition time

69 INTRODUCTION

70 Listening is commonly defined as active attention to sound. In human communication, listening
71 also involves memory, language and executive function (Rudner & Signoret, 2016; Schiller et al.,
72 2022). Listening difficulties (LiD) are frequently reported by caregivers of children with clinically
73 normal hearing (Petley et al., 2021; Roeser et al., 2007). However, the mechanisms underlying
74 childhood LiD without hearing loss remain poorly understood. Several studies have shown that
75 such children typically also have a variety of academic, speech, language, attention and other
76 developmental learning problems (Ferguson et al., 2011; Moore et al., 2018; Sharma et al., 2009).
77 Some of these children receive a clinical diagnosis of auditory processing disorder (APD;
78 American Academy of Audiology, 2010; British Society of Audiology, 2018; Dawes & Bishop,
79 2009).

80
81 The heterogeneous nature of these children's difficulties has led to considerable controversy
82 surrounding APD (Iliadou et al., 2019; Iliadou et al., 2018; McFarland & Cacace, 2009; Moore,
83 2018; Neijenhuis et al., 2019; Rees, 1973). At the center of this controversy is the question of
84 whether the children have primarily auditory sensory problems ("bottom-up"), cognitive and
85 language problems ("top-down"), or some combination of the two (Dillon & Cameron, 2021;
86 Moore et al., 2010). In this and other recent studies (Hunter et al., 2021; Petley et al., 2021) we
87 sidestep this controversy by using the umbrella term LiD (Dillon & Cameron, 2021), and the
88 quantitative metrics of a well-validated, reliable and standardized caregiver questionnaire
89 (ECLiPS; Barry & Moore, 2021) to operationalize LiD.

90
91 Speech understanding requires the coalescence of bottom-up processing of acoustic features in the
92 auditory pathway; with top-down processing of linguistic features in speech and language
93 pathways. These top-down processing pathways are thought to include frontal, temporal and
94 parietal cortices involved in semantic representation, memory and attention (Hickok & Poeppel,
95 2007; Pichora-Fuller et al., 2016; Rönnberg et al., 2019; B. G. Shinn-Cunningham, 2008).
96 Impaired interactions between the auditory system and these higher-level cortical speech
97 perception systems could hold the key to further understanding of how LiD presents in children.
98 Poor performance on complex auditory tasks (e.g. speech listening in noise) has been more closely

99 associated with cognitive function than with performance on simpler, non-speech based auditory
100 tasks (Moore et al., 2010).

101
102 Speech morphs into language as it progresses through a hierarchy of processing competencies
103 leading to comprehension. First, auditory information is classified into linguistically meaningful
104 units called phonemes. This stage bridges bottom-up and top-down processing (Brodbeck et al.,
105 2018). Second, the listener links phonemes into words to judge the intelligibility of the auditory
106 information by comparing the words to memory templates. Third, meaning (semantics) is attached
107 to one or more intelligible words. Basic and speech-related acoustic cues and features enable each
108 of these processing competencies. For example, pitch, loudness, temporal order are required to
109 distinguish individual phonemes. Speech-related features (e.g. voicing, vowel quality, direction of
110 formant transition and noise burst cues) aid in linking the phonemes into words and words into
111 sentences.

112
113 Scott et al. (2000) designed a paradigm, subsequently modified by Halai et al. (2015), to investigate
114 cortical activity produced by different levels of speech processing competencies in adults (Figure
115 1A). They used Clear speech and two types of speech-like stimuli: Rotated speech, that spectrally
116 rotated the stimuli around 2 kHz (Blessner, 1972); and Rotated+Vocoded speech, that first rotated
117 and then noise vocoded the stimuli into 6 frequency bands (Shannon et al., 1995). In each of the
118 three stimulus types, the sequential order of phonemes was preserved, thus retaining the prosody
119 (i.e. rhythm) from the Clear speech. In the Rotated speech, the non-spectral features used to
120 identify phonemes remained, while the spectral features were transformed. For example, the
121 formant transition and noise burst cues that distinguish consonants were shifted to new frequencies
122 and so were processed by different regions of the cochlea. This removed the intelligibility of the
123 stimuli while maintaining the prosodic and phonetic features. Finally, in the Rotated+Vocoded
124 speech, the intonation and phonetic features were removed, along with intelligibility.

125
126 To explore which cortical areas were recruited for each of the three speech processing
127 competencies, Scott et al. (2000) created analysis contrasts by subtracting sentence types from one
128 another (Figure 1A). First, phonology was assessed by contrasting Rotated with Rotated+Vocoded
129 sentences. Second, the judgement of intelligibility was assessed by contrasting Clear sentences

130 with Rotated sentences. Finally, the ability to attach meaning to the intelligible sentence was
131 assessed by contrasting Clear sentences with Rotated+Vocoded sentences where the acoustic
132 complexity was matched but all meaning was removed.

133
134 In this study we adopted the methods developed by Scott et al. (2000) to identify cortical areas
135 used in different levels of speech processing. Consistent with a top-down, linguistic model, and
136 previous research in adults (Halai et al., 2015; Scott et al., 2000), we hypothesised that there would
137 be no group difference in primary auditory cortex (i.e., Heschl's gyrus) between the three contrasts
138 of speech processing competencies. However, Intelligibility and Semantic contrasts would show
139 less activation of the speech processing areas (e.g. superior temporal gyrus, Wernicke's) in
140 children with LiD than in TD children. It was expected that the spread of cortical activation for the
141 three contrasts would increase from Phonology to Intelligibility to Semantics, and would include
142 overlap in auditory and speech areas.

143
144 We next used the task-based cortical activation results to create regions-of-interest (ROIs) for
145 functional connectivity analysis of a separate rs-fMRI acquired in the same scanning session. A
146 rs-fMRI scan allows measurement of the temporal correlation of non-stimulus evoked fluctuations
147 in the BOLD signal between anatomically separated brain regions (Biswal, 2012), capturing
148 interactions between regions in functionally associated networks (Damoiseaux et al., 2006). We
149 hypothesised that children with LiD would show diminished functional connectivity in cortical
150 speech processing networks compared with TD children.

151
152 The aim of the study was to examine the cortical networks involved in different levels of speech
153 processing in children with LiD relative to TD children. Our first research question asked if the
154 two groups of children (LiD and TD) differ in cortical activation during speech listening,
155 specifically in early auditory and speech processing cortical areas. Our second question
156 investigated how the cortical areas activated during these speech processing competencies work
157 together and how they relate to behavioural assessments of speech in noise identification, dichotic
158 listening and cognition.

159

160

161 **RESULTS**

162 We present here data from the first wave of a longitudinal study investigating the audiological and
163 cognitive abilities of children (6-13 years old; Table 1) with LiD and their typically developing
164 peers (6-13 years old; Table 1; Petley et al., 2021). All participants had clinically normal
165 audiometry (≤ 20 dB HL, bilaterally; 0.25 – 8 kHz, Figure 1B), tympanometry (Hunter et al., 2021)
166 and afferent auditory brainstem function (Hunter et al., 2022). Participants who scored at or below
167 the bottom 10th percentile on the total score of a caregiver checklist of everyday listening skills,
168 the ECLiPS (Barry et al., 2015; Barry & Moore, 2021; Roebuck & Barry, 2018), were classified
169 as having LiD (Figure 1C).

170

171 *Task-based cortical activation*

172 During task-based MRI acquisition intervals (Figure 1D), children listened to Clear (“man”) or
173 distorted (“alien”; Rotated and Rotated+Vocoded) spoken sentence types (Figure 1E) that we
174 asked them to match to a visual cartoon representation. All children responded quickly and
175 accurately on the sentence recognition task (Figure 1F), suggesting that they maintained attention
176 throughout the task. Groups did not differ on reaction time ($F(1, 83) = .89, p = .35, \eta^2 = .011$),
177 but TD children were more accurate than children with LiD ($F(1, 83) = 8.77, p = .004, \eta^2 = .096$).
178 Sentence type did not affect response accuracy ($F(1.78, 148.03) = 2.11, p = .13, \eta^2 = .025$) but
179 did influence reaction time ($F(2, 166) = 11.50, p < .001, \eta^2 = .12$) with Clear sentences eliciting
180 quicker responses than Rotated sentences ($p < .001, d = -.41$) and Rotated+Vocoded sentences (p
181 $< .001, d = -.46$). There were no significant interactions.

182

183 The two groups did not differ significantly between contrasts (Phonology, Intelligibility,
184 Semantics; Figure 1A) after correcting for multiple comparisons (Figure 2A-C, clustering
185 threshold = 2.3 voxels, family wise error (FWE) = .95). Further analysis of contrasts was therefore
186 averaged across all children ($n = 85$; Figure 2D-F). The Phonology contrast (Figure 2D) showed

187 bilateral activation in the middle and superior temporal gyrus, including Heschl’s gyrus, and
 188 temporal pole. Activation was also found in the left hemisphere in the temporal fusiform cortex,
 189 angular gyrus and lateral occipital cortex. The Intelligibility contrast (Figure 2E) produced similar
 190 activation to the Phonology contrast, bilaterally in the middle and superior temporal gyrus (anterior
 191 and posterior) and left frontal orbital cortex. Activation extended anteriorly along the left temporal
 192 gyrus and into Broca’s area. The Semantics contrast (Figure 2F) also showed bilateral activation
 193 in the auditory cortices (middle and superior temporal gyrus including Heschl’s gyrus and planum
 194 temporale) with activation extending along the left temporal fusiform and frontal orbital cortices

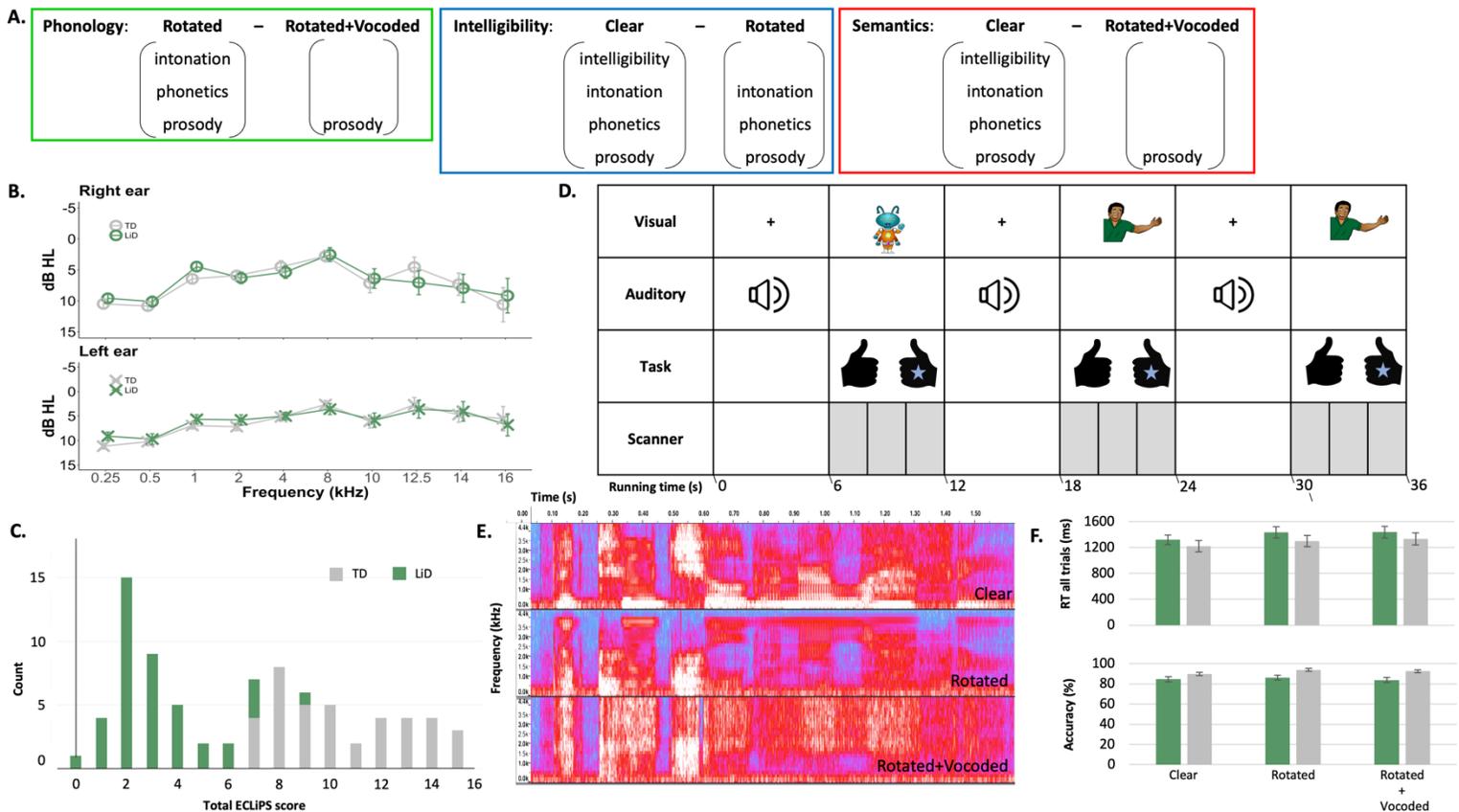


Figure 1: All children had normal tone hearing and were able to perform the speech perception task in the scanner. (A) Stimulus contrasts, illustrating how subtracting one stimulus type from another (Fig. 1E) isolates the specific speech processing competencies of Phonology, Intelligibility and Semantics. (B) Group mean hearing thresholds. (C) Total ECLiPS score. (D) fMRI paradigm. Children were asked if the picture (alien, man) matched who spoke the sentence (stimulus type, Fig. 1E). If it was a match they pressed the right button (with their 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). (E) Stimulus types. 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 (“man”) is intelligible with intonation. Rotated speech (“alien”) is not intelligible, though phonetic features and original intonation are preserved. Rotated+Vocoded speech (“alien”) is completely unintelligible, but preserves the character of the envelope and some spectral detail. (F) Group mean response times and accuracy for each stimulus type. Error bars represent the standard error of the mean. Groups as in Fig. 1C.

195 and right parahippocampal gyrus. Coordinates for the maximum intensity of the activated regions
 196 are shown in Supplementary Material Table 2.

197

| Scan type | Group | N | Age M (SD) | Gender M, F | Maternal education* | Handedness L, both, R | History of tubes | Motion-related artifacts 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) |

198 **Table 1:** Participant details for each scan type – fMRI speech listening task and resting state (RS).

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

200

201 *Cortical functional connectivity*

202 Networks of ROIs suitable for rs-fMRI connectivity analysis were created by parcellating
 203 activation produced by the three contrasts from the task-based MRI (Phonology, Intelligibility,
 204 and Semantics), combined across groups (Figure 2D-F). The pediatric ADHD-200 sample (Bellec
 205 et al., 2017; Craddock et al., 2012) was used as a data-driven, spatially-constrained parcellation
 206 method. ROIs smaller than 4 voxels were removed. Cortical activity from the Phonology contrast
 207 was divided into a network of 16 ROIs, Intelligibility into a network of 20 ROIs, and Semantics
 208 into a network of 24 ROIs (Figure 3A-C; Brodman’s area, maximum intensity coordinates and
 209 ROI sizes are in Supplementary Material Table 3). Connectivity was compared between groups

210 (TD = 42, LiD = 39), controlling for age, using a general linear model (GLM) for each of the three
211 networks (Supplementary Material Table 3).

212

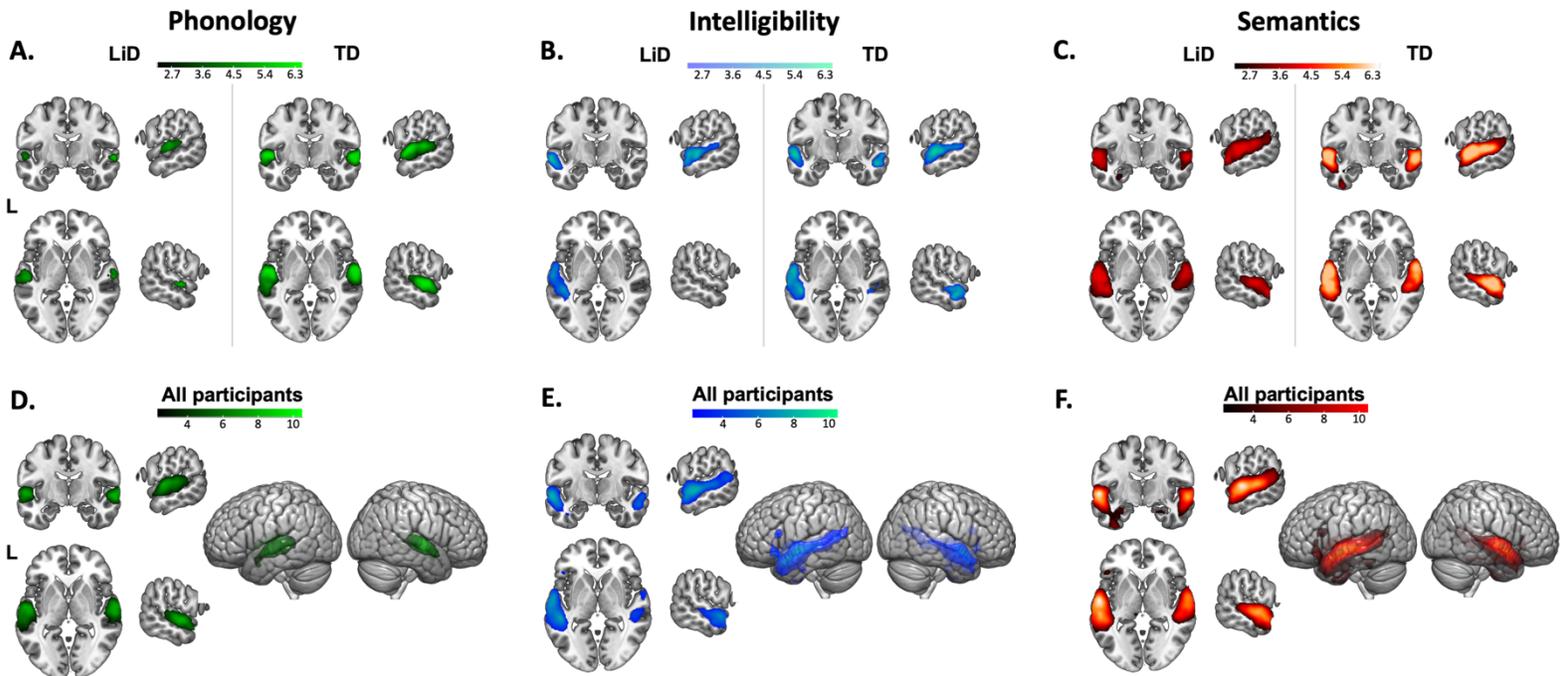


Figure 2: LiDLiD and TD groups showed similar areas of cortical activation in all three contrasts from the fMRI listening task, with no statistical difference between the two groups after correcting for multiple comparisons. Second level GLM analysis for (A-C) groups (clustering threshold = 2.3 voxels, FWE = .95) and (D-F) all participants (clustering threshold = 4.0 voxels, FWE = .95) in the fMRI task, coordinates ($\pm 60, -5, 0$). 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). We took the cortical activation across all participants and created parcellated ROIs (Figure 3) for use in the rs-fMRI analysis (figure 4). Images are in neurological orientation. MRIcroGL was used for visualization.

213 Figure 3 (D-F) shows resting state functional connectivity in the three speech networks, summed
214 across groups. In the Phonology network, each group had connectivity among regions covering
215 bilateral middle and superior temporal gyri, temporal pole, planum temporale, and left planum
216 polare and supramarginal gyrus (Figure 3D). No significant group differences (p-FDR) were found
217 in the Phonology network (Figure 4A, D). Each group had connectivity within the Intelligibility
218 network, covering bilateral middle/superior temporal gyrus and temporal pole along with left pars
219 opercularis, frontal orbital cortex and supramarginal gyrus (Figure 3E). After false discovery rate
220 (FDR) correction, TD children were found to have a significantly stronger temporal correlation
221 between ROIs in left temporal lobe and left middle temporal gyrus (posterior) compared to
222 children with LiD (connection 4-14 in Figure 4B; Table 3). In the Semantics network, both groups

223 of children had connectivity between bilateral middle and superior temporal gyrus, left Heschl's
224 gyrus, pars triangularis, frontal orbital cortex, planum temporale, temporal fusiform gyrus and
225 right parahippocampal gyrus and planum polare (Figure 3F). Group comparisons showed that,
226 compared to children with LiD, the TD children had stronger temporal correlations between ROIs
227 in the right anterior superior temporal gyrus and the left posterior temporal fusiform cortex
228 (connection 7-24 in Figure 4C; Table 3).
229

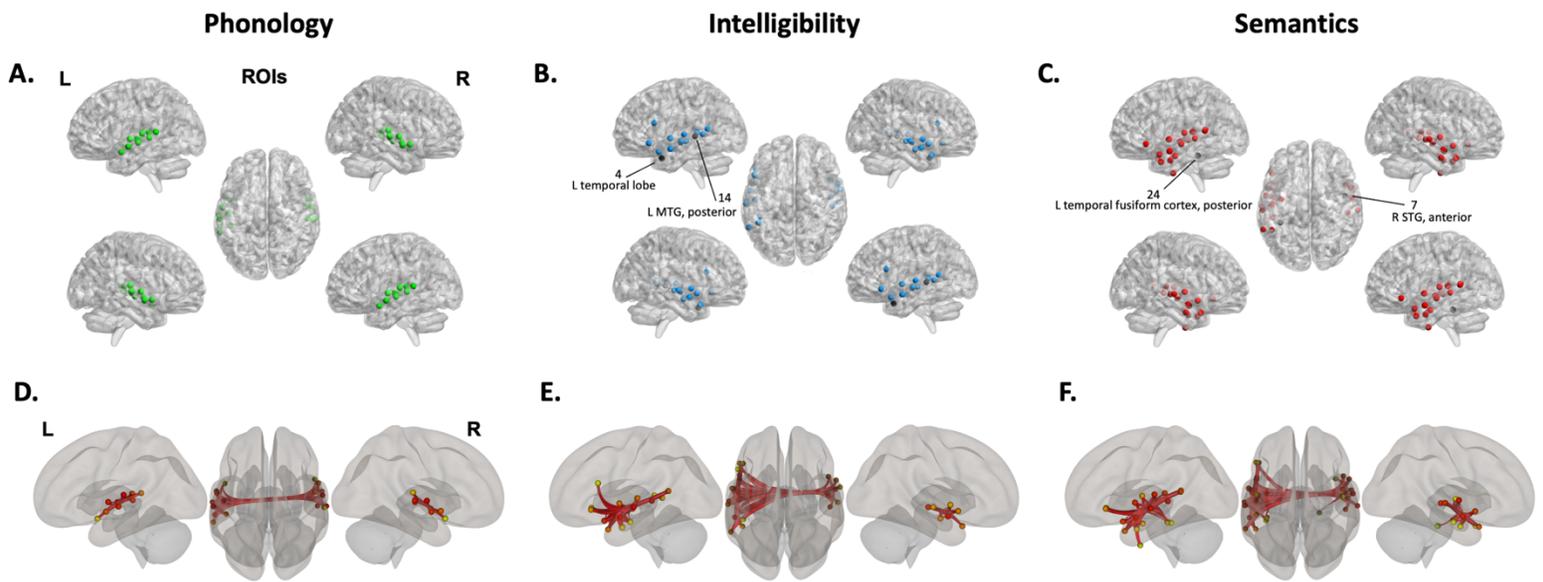


Figure 3: ROIs used for the rs-fMRI connectivity and their networks across all participants. The cortical activity across all participants in the fMRI task (see Figure 2 D-F) covered large areas and so they were parcellated into smaller ROIs (A-C) for the rs-fMRI connectivity analysis by applying data-driven 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 3. 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).

230

231 *Relation of cortical functional connectivity to behavioral measures*

232 We explored how the functional connectivity of anatomically separate cortical areas related to
233 behavioral tasks assessing speech in noise ability, dichotic listening and cognition. As shown by
234 [Petley et al. \(2021\)](#), performance on those behavioural tasks was significantly better in the TD than
235 in the LiD group.

236

237 The Intelligibility connection had a series of significant positive correlations with speech-in-noise
238 ability (SCAN: auditory figure ground), vocabulary and task switching, a form of executive

239 function (NIH Cognition Toolbox). In contrast, the Semantics connection, involving the left
 240 posterior temporal fusiform cortex (7-24; Figure 4C), had significant positive correlations with
 241 dichotic listening ability (SCAN: competing words and sentences) and vocabulary (NIH Cognition
 242 Toolbox). Interpretation of direction of functional connectivity is ambiguous; an increase in
 243 strength of connectivity does not necessarily mean improved ability (Parente et al., 2018).
 244 However, these correlations show that, as the children’s behavioural test scores improved, the
 245 connection strength within the Intelligibility and Semantic networks increased. The remaining
 246 behavioral measures did not relate significantly to connectivity (Figure 4E, Supplementary
 247 Material Table 4 and Figure 1).

248

| Contrast | Max. intensity MNI coordinates | | | | Brain Regions (Harvard-Oxford atlas) | Cluster size (voxels) | |
|--|-----------------------------------|----|-----|-----|---|----------------------------------|-----|
| | BA | x | y | z | | | |
| Intelligibility Clear > Rotated | 4 | 20 | -46 | 12 | -34 | L temporal pole | 82 |
| | 14 | 21 | -50 | -34 | -6 | L middle temporal gyrus, post | 635 |
| Semantics Clear > Rotated+Vocoded | 7 | 20 | 44 | 0 | -26 | R STG, ant | 108 |
| | 24 | 37 | -38 | -40 | -22 | L temporal fusiform cortex, post | 4 |

249 **Table 2:** Selected regions of interest (ROIs, from the fMRI task; Fig. 3B and C) used in the rs-fMRI ROI-to-ROI
 250 analysis. Clustering threshold = 4.0 voxels. Family-wise error correction = .95. The full table of all ROIs can be
 251 found in Supplementary Material Table 3.

252

253

254

255

| Contrast | ROI # – ROI # (Fig. 3) | t (78) | p-FDR |
|-----------------|---------------------------|--------|---------|
| Intelligibility | 4 – 14 | 3.11 | 0.026 |
| Semantics | 7 - 24 | 3.54 | < 0.001 |

256 **Table 3:** Network temporal connectivity. Statistical comparison between group differences (TD vs. LiD, corrected
 257 for age) for the Intelligibility and Semantics networks (Fig. 4D). No significant group difference was found in the
 258 Phonology network. False discovery rate (FDR) was used for multiple comparison correction.

259

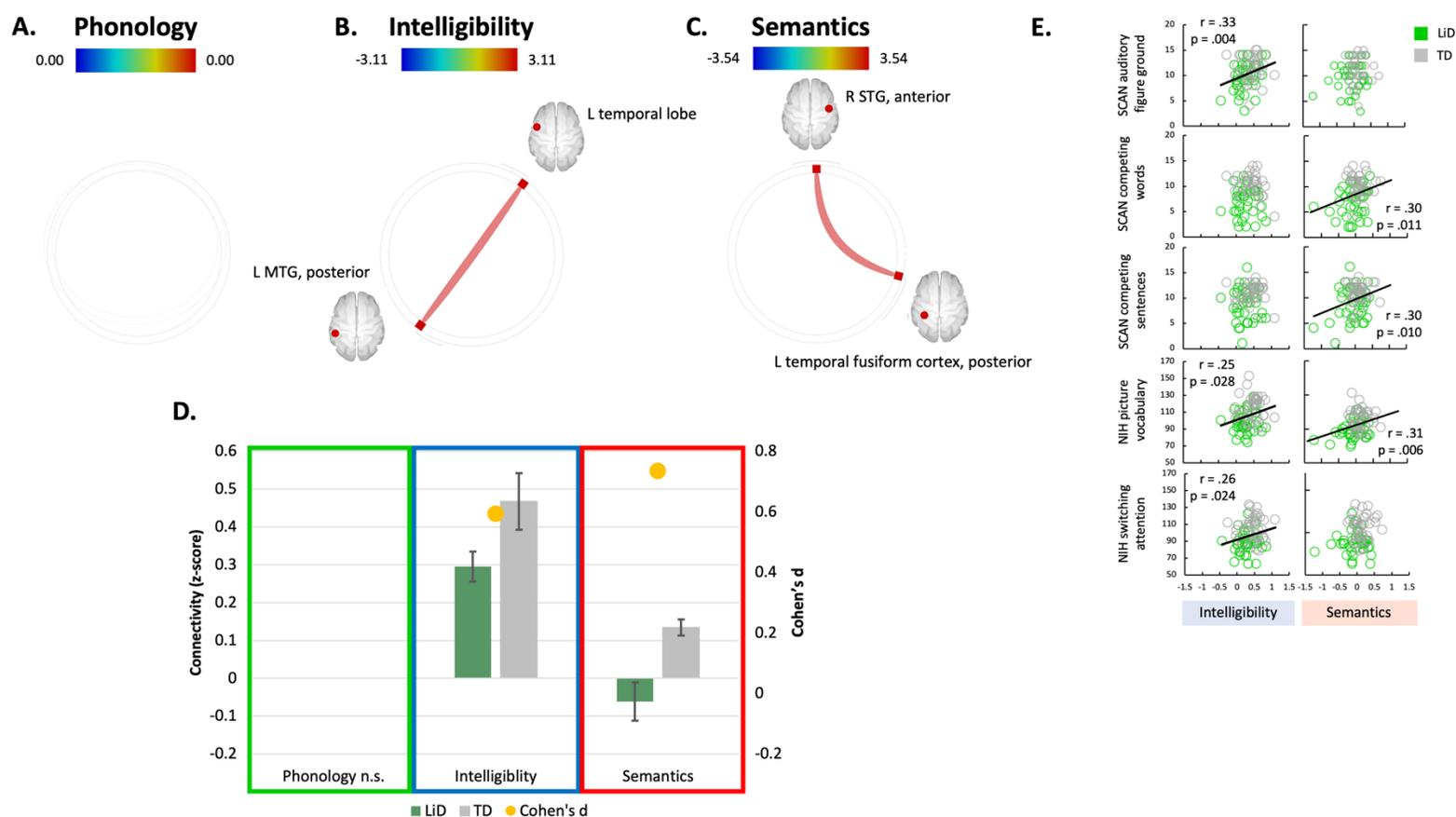


Figure 4: ROI-to-ROI resting state connectivity: the difference between the groups' listening networks grew from no statistical differences in the phonology network to a minor difference in the intelligibility network to more widespread differences in the semantics network. (A-C) The group comparisons without the effect of age for (A) Phonology, (B) Intelligibility and (C) Semantics 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 marker, 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 show the TD group as having significantly stronger connectivity than the LiD group. (E) Scatter plots of brain and behavioral scores demonstrate the correlated patterns of connectomic and behavioural features. Coloured dots in each panel indicate the participant group (LiD: green, TD: grey). The solid black line marks significant correlations (see Supplementary Table 4 for correlation details).

260 **DISCUSSION**

261 This study compared TD children and children with LiD on fMRI measures of cortical responses
262 to speech stimuli that assessed processing at the levels of phonology, intelligibility and semantics.
263 The participant groups showed similar cortical activation, with all three sentence contrasts eliciting
264 bilateral activation across the auditory cortex (e.g. Heschl's gyrus, superior temporal lobe).
265 Activation extended bilaterally into areas thought to be involved in memory and language (e.g.,
266 Broca's area, right parahippocampal gyrus, left temporal fusiform cortex) as the sentence contrasts
267 progressed from emphasizing phonology to intelligibility to semantics. These findings provide
268 support for the hypothesis that primary auditory function does not differ between the two groups
269 when listening to speech. The findings did not support our prediction that children with LiD would
270 show less activation in the speech processing areas than those in the TD group. However,
271 activation was examined using BOLD responses which do not capture the excitatory and inhibitory
272 balance or temporal relationships between cortical areas.

273
274 Analysis of the separate rs-fMRI data showed that the TD and LiD groups had different rsMRI
275 responses for the three sentence contrasts. The children with LiD showed less functional
276 connectivity as the sentence contrasts progressed from emphasizing phonology to intelligibility to
277 semantics. This finding supported our functional connectivity hypothesis, but only in networks
278 required for processing speech intelligibility and semantics. It also supports a novel interpretation
279 of childhood LiD as being less of a phonological disorder and more of a language disorder affecting
280 intelligibility and semantics.

281
282 Behavioural measures of performance were consistent with the MRI observations in two areas.
283 First was the absence of a significant correlation between functional connectivity for intelligibility
284 and semantic processing in the cortex and the behavioural measure of spatial advantage (the ability
285 to use differences in speaker location to enhance speech intelligibility in noise; LiSN-S), thought
286 to assess auditory processing independently of language and cognition. Second was the significant
287 correlations between functional connectivity for intelligibility and semantic processing with the
288 behavioural measures auditory figure-ground (the ability to hear words-in-noise; SCAN filtered
289 words), dichotic listening (SCAN competing words and SCAN competing sentences), vocabulary
290 (NIH picture vocabulary) and switching attention (NIH switching attention), all thought to assess

291 more cognitively-dominant processes that rely on language and executive functions (Petley et al.,
292 2021). Overall, these data were consistent with the hypothesis that children with LiD have
293 primarily cognitive and language processing deficits rather than auditory processing deficits.

294

295 **Bilateralization of speech in 6-12 year old children**

296 Using simple but complete sentence stimuli, we consistently found bilateral activation in children
297 for all speech listening contrasts. After FDR correction, this activation did not significantly differ
298 between groups. These findings suggest that children with LiD use the same cortical areas as TD
299 children for increasingly complex speech processing competencies. However, it is possible that
300 multi-voxel pattern analysis may find finer group differences in the hierarchy for speech
301 processing (e.g., Okada et al., 2010). While the same cortical areas are used by the children with
302 LiD, it is possible that they do so on a different time frame from the TD children. Unfortunately,
303 fMRI does not provide sufficient time resolution to address this possibility.

304

305 Our results support the theory of bilateral processing of speech listening in 6-12 year old children.
306 Lateralization of speech processing in adults has been debated across and within imaging
307 modalities (fMRI see Evans & McGettigan, 2017; EEG e.g., Assaneo et al., 2019). For example,
308 Albouy et al. (2020) discussed how acoustic structure leads to a left bias for fast modulation
309 (speech) and a right bias for slow modulation (music) and Rauschecker & Scott's (2009) unilateral
310 model suggested that the left anterior STG is the hub of successful speech perception. In contrast,
311 Hickok and Poeppel's (2000) bilateral model proposed a perceptual pathway in each hemisphere
312 processing speech sounds up to the level of semantics (Hickok & Poeppel, 2007). A middle ground
313 has been proposed by Peelle (2012), with right hemisphere dominated activation for
314 "unconnected" speech (i.e., phonemes, syllables, single words) and left hemisphere dominated
315 activation for "connected" speech (i.e., phrases, sentences and narratives).

316

317 Bilateral activation reported here differs from the findings of Scott et al. (2000) who reported a
318 left lateralized pathway for speech comprehension using PET, a silent scanning technique. Our
319 results extend those of Halai et al. (2015) who found bilateral activation during continuous fMRI
320 scanning. Both these studies tested young adults and used similar stimuli to those used here.
321 However, speech presentation was continuous, and passive in the sense that the participants did

322 not perform any task during scanning. Rather, comprehension was assessed after scanning and
323 outside the scanning room. In contrast, we required children to provide a behavioural response to
324 each short sentence presentation to encourage attention and as a metric of attention. We
325 additionally used a silent, “sparse” acquisition protocol (Hall et al., 1999; Vannest et al., 2009), in
326 contrast to the previous (Halai et al., 2015) continuous fMRI scanning protocol. Scanner noise
327 superimposed on speech listening has been shown to increase listening effort (Peelle et al., 2010)
328 and to engage additional and or different cortical areas compared with silent acquisition protocols.
329 For example, in a meta-analysis of 57 speech comprehension studies, Adank (2012) showed that
330 continuous scanning more strongly activated regions of the supplementary motor area and anterior
331 cingulate gyrus, while sparse scanning showed more extensive activation in the STS.

332

333 **Functional connectivity: Brain and behaviour**

334 We used temporal correlations in rs-fMRI to compare how cortical areas associated with speech
335 processing competencies worked together in children with and without LiD. Phonology is the
336 system of processing the smallest units of speech sounds and their linguistically appropriate
337 combinations. We found no group difference in the connectivity of this network. As we progressed
338 up the speech processing competency hierarchy to the Intelligibility network, we found that
339 temporal correlation between left temporal lobe and left posterior middle temporal gyrus activity
340 was stronger in the TD children. These areas are well known for processing auditory information
341 (temporal lobe) and speech comprehension (left MTG, Acheson & Hagoort, 2013; Dronkers et al.,
342 2004). Posterior MTG has been associated with lexical and semantic access in a sound-to-meaning
343 network (Hickok & Poeppel, 2000, 2004). It is unknown whether stronger or weaker temporal
344 correlations between brain regions is beneficial (Parente et al., 2018). However, relationships
345 found here between the variability of brain connectivity strength and of behaviour provides
346 evidence on this key issue. Children with stronger connectivity performed better on cognitive tasks
347 assessing vocabulary and switching attention. Increased connectivity may thus be indicative of
348 increased processing efficiency and/or suppression of a task-relevant network, in this case
349 language. Note that the behavioural measures used here have been shown to be either little affected
350 by age (audiogram: Hunter et al., 2021) or were standardized across age (ECLiPS, LiSN-S, SCAN
351 and NIH toolbox (Petley et al., 2021).

352

353 Group comparisons in the Semantics network highlighted that, compared to children with LiD, the
354 TD children had stronger temporal correlations between auditory areas (right STG) and the left
355 temporal fusiform cortex, associated with word recognition and the recovery of meaning from an
356 impoverished acoustic signal (Davis & Johnsruide, 2003). Brain-behaviour correlations suggest
357 that decreased connectivity between auditory and word recognition cortical areas was associated
358 with impaired speech listening (dichotic listening and picture vocabulary; Table 2). Recently,
359 brain-behaviour correlations have been called into debate due to extremely large datasets being
360 required to reach even very small correlations (Marek et al., 2022). However, we consistently
361 found mild to moderate sized correlations with a peak of $|r| = 0.33$ (mean $|r| = 0.29$, first percentile
362 $|r| = 0.26$; Supplementary Material Table 4).

363
364 The cortical areas highlighted in this study (Figure 4 B and C) draw focus to language production,
365 memory encoding and retrieval, and word recognition (Acheson & Hagoort, 2013; Davis &
366 Johnsruide, 2003; Dronkers et al., 2004; Hickok & Poeppel, 2000, 2004) as target areas of further
367 research in children with LiD. Group differences were found in networks of latter stages of speech
368 processing competencies, in the intelligibility and semantics networks. Specifically, between the
369 MTG/STG and the temporal lobe. The reduction in connectivity in these networks suggest that
370 childhood LiD is a semantic and intelligibility disorder. However, no group differences were found
371 in the phonological networks, an early stage of speech processing competencies, suggesting that
372 childhood LiD is not a phonological disorder.

373
374 Further assessment where the ‘break’ in speech listening occurs in LiD would provide a clearer
375 avenue for research into effective evidence-based treatments. Future studies could utilize
376 paradigms with anomalous and mispronounced words, with imaging techniques used in parallel to
377 assess if the type of speech processing errors during such tasks are affected by LiD. For example,
378 was it detected when presented (i.e. when judging it’s intelligibility) or did the listener reach the
379 end of the sentence and have to ‘go back’ when they are unable to connect meaning to the sentence.
380 Complementary, time-sensitive techniques (EEG, MEG) could also investigate whether this
381 difficulty is due to a bottleneck in processing leading to increased listening effort, cognitive effort
382 or fatigue (McGarrigle et al., 2014). As LiD may build up during a listening event (i.e., over several
383 minutes; Roebuck & Barry, 2018), it is important to assess the children's ability throughout the

384 task with cross-sectional time points rather than using summary values of the complete listening
385 event (McGarrigle et al., 2020).

386

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

388 Children identified with LiD have difficulty in speech listening compared to TD children (Petley,
389 Hunter, Motlagh Zadeh, et al., 2021). These groups were further distinguished here by functional
390 connectivity differences between cortical areas associated with language production, memory and
391 word recognition, rather than by the activity or connectivity of primary auditory cortical areas.
392 However, at least 50% of children referred/diagnosed with auditory processing disorder (APD;
393 British Society of Audiology, 2018; Dillon & Cameron, 2021; Jerger & Musiek, 2000), a clinical
394 label closely related to LiD (Dillon & Cameron, 2021), have also been diagnosed with
395 developmental language disorder (DLD), dyslexia/reading disorders, attention
396 deficit/hyperactivity disorder (ADHD) or more than one of these other neurodevelopmental
397 disorders (Dawes & Bishop, 2010; Ferguson et al., 2011; Gokula et al., 2019; Moore et al., 2018;
398 Sharma et al., 2009). This high level of comorbidity was echoed in the study reported here. While
399 LiD was the primary referral for the study, as with other neurodevelopmental disorders it does not
400 stand in isolation. A background caregiver questionnaire showed that half of the children with LiD
401 also reported a diagnosis of ADHD, 9% autism spectrum disorders and 26% had seen a speech
402 language pathologist. However, these data were not entered into the analysis as we did not conduct
403 gold standard assessments for these other developmental disorders. We are currently using a web-
404 based resource (Neurosynth; Yarkoni et al., 2011) that allows functional connectivity analysis of
405 brain areas defined by a meta-analysis of published fMRI activation coordinates to explore whether
406 there are shared audio-based neurological patterns in children with primary diagnoses of other
407 developmental disorders, such as ASD and ADHD, with LiD.

408

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

415 cortical language processing may be a consequence of LiD. Further longitudinal analysis of these
416 children is designed in part to address these issues.

417
418 The neuroimaging results presented here highlight the importance of non-auditory factors,
419 specifically language, in audiological testing at a cortical level. There is a growing recognition of
420 the importance of language and, specifically, speech-in-noise (SiN) intelligibility in everyday
421 hearing (Dillon & Cameron, 2021; Golestani et al., 2009; Killion et al., 2004; Magimairaj et al.,
422 2021; Smits et al., 2013). It has been proposed that such testing could supplement, or even replace
423 pure tone detection as an audiometric gold standard (Hewitt, 2018). However, both SiN test
424 instructions and test items pose a challenge to language and memory as well as auditory function.
425 While those cognitive aspects of auditory testing and learning have previously been dismissed as
426 procedural issues (Hawkey et al., 2004), they are an intimate component of a SiN test. These
427 neuroimaging results provide insight into mechanisms of how the advanced stages of speech
428 processing where auditory information is translated into language during speech perception may
429 be disrupted in LiD, a common form of auditory impairment in both children (Moore et al., 2018)
430 and adults (Edwards, 2020). They also add to a growing literature on the role of cognitive function
431 in hearing (Moore et al., 2014; Rönnberg et al., 2013; Sharma et al., 2019; B. Shinn-Cunningham,
432 2017).

433
434 **Conclusions**
435 Our results provide the first multifaceted neurological profile for children classified with LiD,
436 based on caregiver report and normal peripheral auditory function. Children with LiD recruited
437 the same cortical areas as their peers when processing increasing complexities of speech. However,
438 how these cortical areas work together does differ between the two groups. Differences in
439 functional connectivity were found at the more advanced stages of speech listening where the
440 intelligibility and semantics of the speech are processed, specifically in the left temporal lobe,
441 posterior middle temporal gyrus, posterior temporal fusiform cortex and right superior temporal
442 gyrus. These highlighted cortical connections related to the children's behavioural abilities in
443 dichotic listening, speech-in-noise, attention, memory and verbal vocabulary abilities. Overall, the
444 data provide support for the hypothesis that children with LID are primarily affected by cognitive
445 and, particularly, language processing deficits.

446 **METHODS**

447 **Participants**

448 Eighty-five participants aged 6-12 years completed the fMRI sentence task and 81 participants
449 completed the rs-fMRI (see Table 1). Seventy-six of participants who completed the fMRI
450 sentence task also completed the rs-fMRI. All participants had normal audiometric hearing with
451 thresholds < 25 dB HL at all octave-interval frequencies from 0.25 - 8 kHz in both ears (Figure
452 1A). In this paper we focus on the fMRI and its relationship with behavioural responses from the
453 baseline of our longitudinal ‘SICLID’ study examining correlates of LiD in children. Extensive
454 analysis of ear function and behavioural responses of these children is reported elsewhere (Hunter
455 et al., 2021, 2022; Petley et al., 2021).

456
457 Caregivers of all participants completed a standardized, validated and reliable checklist of
458 everyday listening and related skills (ECLiPS; Barry & Moore, 2021). Children recruited through
459 advertisement all scored < 10th percentile of ECLiPS standardized scores. Additional children
460 with an audiological diagnosis of auditory processing disorder (APD; n=14) were placed in the
461 LiD group. All but one of these also scored < 10th percentile on the ECLiPS (Figure 1C).
462 Participants scoring within the upper 90th percentile on the ECLiPS and with no history of
463 developmental disorders or delays were classified as TD (further details in Petley et al., 2021).

464 Eligibility for both groups included English as the child’s native language, and a reported absence
465 of otologic, neurologic or psychiatric disease, or of intellectual insufficiency that would prevent
466 or restrict their ability to complete testing procedures. TD participants were additionally required
467 to have no known history of developmental delay, or attention or language disorder. While half of
468 the children with LiD also reported a diagnosis of ADHD, 9% autism spectrum disorders and 26%
469 had seen a speech language pathologist. Eligibility was determined based on caregiver responses
470 on a medical and educational history ‘Background’ questionnaire (Supplementary Material Table
471 1).

472 This study was approved by the Institutional Review Board of Cincinnati Children’s Hospital
473 (CCH) Research Foundation. Prior to completion of study-related imaging and behavioural testing,
474 caregivers reviewed the informed consent form with a study staff member. Children aged 11 and

475 above were also assented using a child-friendly version of the consent document, per institutional
476 policy. All participants received financial compensation for their participation.

477 **MRI acquisition**

478 MRI was performed via a 3T Philips Ingenia scanner with a 64-channel head coil and Avotec
479 audiovisual system. All participants were awake throughout the scanning. The protocol was
480 modified from a previous large, cross-sectional examination of brain function (Holland & Vannest,
481 2015) and included a high-resolution T1-weighted anatomical scan, fMRI sentence task (4.9
482 minutes) and rs-fMRI (5 minutes). The fMRI sentence task was acquired with a sparse scanning
483 protocol ('HUSH', details below); TR/TE = 2000/30 ms, voxel size = $2.5 \times 2.5 \times 3.5$ mm, 39 axial
484 slices. A total of 147 volumes was acquired by alternating scanning for 6 seconds (3 volumes) and
485 not scanning for 6 seconds, 49 times. Cardiac and respiration signals were collected during the
486 fMRI sentence task using the scanner's wireless respirator bellows and Peripheral Pulse Oximeter.
487 The rs-fMRI acquisition was acquired with TR/TE = 2000/30 ms, voxel size = $2.5 \times 2.5 \times 3.5$ mm,
488 39 axial slices in ascending slice order and 150 volumes. The high-resolution T1-weighted
489 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
490 256 and slice thickness = 1mm.

491

492 **fMRI task**

493 With sound levels reaching 118.4 ± 1.3 dB (A) in a 3T MRI system (Price et al., 2001) special
494 considerations must be made when planning an auditory-based MRI study. In order to protect the
495 participant from the loud environment, foam ear plugs and MRI safe circumaural headphones were
496 worn. The scanner noise may also produce masking of the desired stimuli. Therefore, in the fMRI
497 task we used a 'Hemodynamics Unrelated to Sounds of Hardware' (HUSH) scanning protocol
498 (Deshpande et al., 2016; Edmister et al., 1999; Hall et al., 1999; Schmithorst & Holland, 2004) - a
499 sparse temporal sampling protocol where there was no gradient coil noise during presentation of
500 the auditory stimuli. Stimulus levels were elevated to produce adequate signal/noise ratios for
501 accurate responding. We also used a talker identification task instead of a speech recognition task.
502 Instead of asking the children *what* they heard, we asked them *who* had said it (Fig. 1). The children
503 responded with button presses throughout the task so we could ensure they maintained attention
504 to the task.

505

506 Sixteen linguistically simple BKB sentences (Bench et al., 1979), designed to be familiar to young
507 children, were recorded by a single male North American speaker, mirroring the paradigm used
508 by Scott et al. (2000) and Halai et al. (2015). These were the Clear speech stimuli (e.g. Figure 1E).
509 Rotated speech stimuli were created by rotating each sentence spectrally around 2 kHz using the
510 Blesser (1972) technique. Rotated speech was not intelligible, though some phonetic features and
511 some of the original intonation was preserved. Rotated+Vocoded speech stimuli were created by
512 applying 6 band noise-vocoding (Shannon et al., 1995) to the Rotated speech stimuli. While the
513 Rotated+Vocoded speech was completely unintelligible, the character of the envelope and some
514 spectral detail was preserved.

515
516 Participants were told that they would be completing a matching game where they would hear a
517 sentence and then see a picture (of a “man” or an “alien”). If the picture matched who said the
518 sentence (man - Clear speech, alien - Rotated or Rotated+Vocoded speech), the participant pressed
519 a button with their right thumb, if the picture did not match, they pressed a second button with
520 their left thumb. The participants were asked to respond as fast and as accurately as possible. A
521 sticker was placed on the participants’ right hand to provide a reminder as to which hand was
522 correct for matching voice and picture.

523
524 Before scanning, each participant was familiarized with the sentence task and completed three
525 practice trials with verbal feedback from the tester. If a trial was completed incorrectly, the stimuli
526 and instructions were reintroduced until the participant showed understanding. During scanning,
527 each participant completed 48 matching trials, 16 of each sentence type, with no feedback. To
528 maintain scanner timings the behavioural task continued regardless of whether the participant
529 responded. However, if the child did not press a response button on three trials in a row the tester
530 provided reminders/encouragement over the scanner intercom between stimulus presentations.

531 532 **fMRI data analysis**

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

536

537 For the sparse HUSH acquisition, the volumes were separated and combined into three files
538 according to the volume's order during the scanner-on period. Each of the three files was pre-
539 processed separately and first-level statistics computed. The three statistical images were then
540 averaged together using a one sample t-test. This was done to account for the difference in
541 intensity among the volumes due to T2* relaxation effects. The pre-processing steps included the
542 following. FSL's BET was used for brain extraction of the functional data. Outlying functional
543 volumes were detected with 'fsl_motion_outliers' using the RMS intensity difference metric.
544 AFNI's '3dretroicor' was used to regress out the cardiac and respiration signals using a
545 RETROICOR approach (Glover et al., 2000). Motion correction was carried out by MCFLIRT. A
546 GLM was used to regress motion-related artifacts from the data using 6 regressors for the motion
547 parameters and an additional regressor for each outlying volume. The amount of motion during
548 the scans (the number of outlying volumes for each participant) did not differ between groups, p
549 = .62 (Table 1). The data were spatially smoothed using a Gaussian kernel with a sigma of 3 mm
550 and temporally filtered with a high pass filter with a sigma of 30 seconds. The results were
551 interpolated to a 2 mm isotropic voxel size and aligned to the Montreal Neurological Institute
552 (MNI) template by first co-registering it with the participant's T1 using FSL's FLIRT.

553
554 Second-level analysis was also conducted using FSL. A GLM approach was used to create group
555 activation maps based on contrasts between conditions for all participants (i.e. regardless of
556 LiD/TD status) with age as a covariate. Group composite images were thresholded using a family-
557 wise error correction ($p < 0.05$) and clustering threshold of $k = 4$ voxels. Three BOLD activation
558 contrasts were used to search for brain loci responding to different aspects of listening to language
559 (Halai et al., 2015; modified from Scott, 2000). First, a 'Semantics' activation map, whereby the
560 signal with intelligibility, intonation, phonetics, and prosody was contrasted with one lacking all
561 of these attributes except prosody (Clear > Rotated+Vocoded). Second, an 'Intelligibility'
562 activation map contrasted the signal with all speech attributes to one retaining intonation, phonetics
563 and prosody (Clear > Rotated). Third, a 'Phonetics' activation map contrasted a signal with
564 intonation, phonetics and prosody with one having only prosody (Rotated > Rotated+Vocoded).

565
566 Behavioural responses from the fMRI task were assessed using a 2 (group: TD, LiD) \times 3 (sentence
567 type: Clear, Rotated, Rotated+Vocoded) repeated measures analysis of variance (ANOVA) for

568 accuracy and again for RT (Figure 1F). Where the assumption of sphericity was violated, degrees
569 of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

570

571 **Resting state fMRI**

572 During this second scan the participant was asked to lie still, keep their eyes open and look at the
573 (central) white cross on the black screen. During the 5 minute period they were not performing an
574 exogenous task. Eyes were monitored by the tester through CCTV and no child fell asleep during
575 the task.

576

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

583

584 **Resting state ROI-to-ROI analysis**

585 The group activation maps from the fMRI task's three contrasts (cluster threshold = 4 voxels; p -
586 FWE = .95) were used to define the ROIs of advancing speech processing competency networks
587 (Phonology, Intelligibility and Semantics). However, as these areas of activation were large (e.g.
588 an area of 4351 voxels in the left frontal lobe as part of the Semantics network, Supplementary
589 Material Table 2), we applied the parcellation from the pediatric ADHD-200 sample (Bellec et al.,
590 2017) to each network. This created smaller and more appropriate ROIs for connectivity analysis
591 of each network (Figure 3D-F, Supplementary Material Table 3). ROIs smaller than 4 voxels were
592 not included in the analysis.

593

594 Conn was used to test the functional relationship between each pair of ROIs identified in the fMRI
595 sentence listening task. The mean time course of all voxels within each ROI was used to calculate
596 individual pairwise Pearson correlations. The r values were normalized to z values via Fisher's z -
597 transformation. We then used these z values to explore the relationship between the three listening
598 networks and behavioural measures. Statistical thresholds were set to $p < .05$ (corrected) at the

599 single voxel level, and the resulting connections were thresholded at seed-level by intensity with
600 FDR correction ($p < .05$).

601

602 **Caregiver questionnaire**

603 *Everyday listening skills – ECLiPS* (Barry & Moore, 2021)

604 The ECLiPS is a standardized caregiver-report measure of listening and communication
605 difficulties. Caregivers rated 38 simple statements about their child on a five-point scale, ranging
606 from strongly disagree to strongly agree.

607

608 **Behavioral measures**

609 Resting state temporal connections with significant group differences were correlated with
610 behavioural measures described briefly below (see Petley et al., 2021 for further detail and data).

611 Study data were collected and managed using REDCap electronic data capture tools hosted at
612 Cincinnati Children’s Hospital (Harris et al., 2009, 2019). REDCap (Research Electronic Data
613 Capture) is a secure, web-based software platform designed to support data capture for research
614 studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking
615 data manipulation and export procedures; 3) automated export procedures for seamless data
616 downloads to common statistical packages; and 4) procedures for data integration and
617 interoperability with external sources.

618

619 *Listening in Spatialized Noise-Sentences* (Brown et al., 2010; Cameron & Dillon, 2007, 2009)

620 LiSN-S (US version) is a standardized test assessing speech in noise ability. Binaural target (T)
621 sentences were presented through headphones along with two other distracting sentences (D1, D2).

622 The children were asked to repeat the sentences of the target voice only. Distracting sentences
623 remained constant at 55 db SPL. After each correct trial the target voice descended in level (4 dB),
624 but if the child incorrectly repeated back over 50% of the sentence the level increased (by 2 dB).

625

626 Four listening conditions are made by manipulating D1 and D2 with respect to T (same voice,
627 different voices; same direction, 0° , different direction, $\pm 90^\circ$ azimuth). Three difference
628 (“Advantage”) scores are calculated to reduce the influence of language and cognitive demands of
629 each condition (Dillon et al., 2014): Talker Advantage (different voices - same voice); Spatial

630 Advantage (different directions – same direction); and Total Advantage (different voices and
631 directions – same voices and directions). LISN-S software calculated these difference scores for
632 each participant.

633

634 *SCAN-3:C* (Keith, 2009; 2000)

635 SCAN-3:C is a US-standardized test battery often used by audiologists to diagnose APD in
636 children (Emanuel et al., 2011). We used the task to obtain comparison measures relative to other
637 studies of auditory processing ability (e.g. Kelley & Littenberg, 2019), but not to group children
638 as LiD or TD. Subtests used in our battery were Auditory Figure Ground - assessing the ability to
639 repeat words presented against background multi-talker speech; Competing words - a dichotic
640 listening task where the child repeats different words presented simultaneously to each ear,
641 repeating that from a designated ear first; Filtered words - assessing ability to identify words that
642 are low pass filtered at 750 Hz; and Competing Sentences - a dichotic listening task where different
643 sentences are presented simultaneously to each ear, and the child is asked to repeat the sentence
644 from a designated ear. Both subtest and a standardized SCAN composite score are calculated.

645

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

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

653

654 **DATA AVAILABILITY**

655 In accordance with ethics requirements, the dataset generated and analysed during the current
656 study are available from the corresponding author. ROIs used in the rs-fMRI analysis are available
657 at GitHub (https://github.com/stewarthannahj/ROIs_SICLiD.git).

658

659

660

661 **ACKNOWLEDGEMENTS**

662 Thanks to Dr Kim Leikin and Prof. Scott K Holland for aspects of study design and to Prof. Stuart
663 Rosen and Dr Peter Chiu for the scripts to rotate and vocode the stimuli for the fMRI sentence
664 task. Thanks also to Audrey Perdew, Nicholette Sloat and Dr Ronan McGarrigle for assisting in
665 MRI acquisition. Many thanks to the CCH MRI techs (Matt Lanier, Lacey Haas, Kaley
666 Bridgewater, Kelsey Murphy, Byrnn Williams and Marty Jones) for many hours of scanning.

667
668 **FUNDING**

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

672
673 **COMPETING INTERESTS**

674 The authors report no competing interests.

675
676 **CONTRIBUTIONS**

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

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

680 LLH - study design, manuscript preparation and revisions

681 TM - fMRI preprocessing, assisted with manuscript preparation

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

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

684

685

686 **REFERENCES**

- 687 Acheson, D. J., & Hagoort, P. (2013). Stimulating the brain's language network: Syntactic
688 ambiguity resolution after TMS to the inferior frontal gyrus and middle temporal gyrus.
689 *Journal of Cognitive Neuroscience*, 25(10), 1664–1677.
- 690 Albouy, P., Benjamin, L., Morillon, B., & Zatorre, R. J. (2020). Distinct sensitivity to
691 spectrotemporal modulation supports brain asymmetry for speech and melody. *Science*,
692 367(6481), 1043–1047.
- 693 American Academy of Audiology. (2010). *Diagnosis , Treatment and Management of Children*
694 *and Adults with Central Auditory Processing Disorder American Academy of Audiology*
695 *Clinical Practice Guidelines*. August.
- 696 Barry, J. G., & Moore, D. R. (2021). *ECLIPS: Evaluation of children's listening and processing*
697 *skills* (2nd ed.). Cincinnati Children's Hospital.
- 698 Barry, J. G., Tomlin, D., Moore, D. R., & Dillon, H. (2015). *Use of Questionnaire-Based*
699 *Measures in the Assessment of Listening Difficulties in School-Aged Children*. 1–14.
- 700 Blesser, B. (1972). Speech perception under conditions of spectral transformation: I. Phonetic
701 characteristics. *Journal of Speech and Hearing Research*, 15(1), 5–41.
- 702 British Society of Audiology. (2018). *Position Statement and Practice Guidance—Auditory*
703 *Processing Disorder (APD)*. [http://www.thebsa.org.uk/wp-](http://www.thebsa.org.uk/wp-content/uploads/2018/02/Position-Statement-and-Practice-Guidance-APD-2018.pdf)
704 [content/uploads/2018/02/Position-Statement-and-Practice-Guidance-APD-2018.pdf](http://www.thebsa.org.uk/wp-content/uploads/2018/02/Position-Statement-and-Practice-Guidance-APD-2018.pdf)
- 705 Brown, D. K., Cameron, S., Martin, J. S., Watson, C., & Dillon, H. (2010). The North American
706 Listening in Spatialized Noise—Sentences Test (NA LiSN-S): Normative data and test-
707 retest reliability studies for adolescents and young adults. *Journal of the American*
708 *Academy of Audiology*, 21(10), 629–641.

- 709 Cameron, S., & Dillon, H. (2007). Development of the listening in spatialized noise-sentences
710 test (LISN-S). *Ear and Hearing*, 28(2), 196–211.
711 <https://doi.org/10.1097/AUD.0b013e318031267f>
- 712 Cameron, S., & Dillon, H. (2009). *Listening in Spatialized Noise–Sentences test (LiSN-S)*.
- 713 Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language
714 comprehension. *The Journal of Neuroscience : The Official Journal of the Society for*
715 *Neuroscience*, 23(8), 3423–3431. <https://doi.org/23/8/3423> [pii]
- 716 Dawes, P., & Bishop, D. (2009). Auditory processing disorder in relation to developmental
717 disorders of language, communication and attention: A review and critique Research
718 Report. *International Journal of Language and Communication Disorders*, 44(4), 440–
719 465. <https://doi.org/10.1080/13682820902929073>
- 720 Dillon, H., & Cameron, S. (2021). Separating the causes of listening difficulties in children. *Ear*
721 *and Hearing*.
- 722 Dillon, H., Cameron, S., Tomlin, D., & Glyde, H. (2014). Comments on “Factors influencing
723 tests of auditory processing: A perspective on current issues and relevant concerns” by
724 Tony Cacace and Dennis McFarland. *Journal of the American Academy of Audiology*,
725 25(07), 699–703.
- 726 Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion
727 analysis of the brain areas involved in language comprehension. *Cognition*, 92(1–2),
728 145–177. <https://doi.org/10.1016/j.cognition.2003.11.002>
- 729 Edwards, B. (2020). *Emerging technologies, market segments, and MarkeTrak 10 insights in*
730 *hearing health technology*.

- 731 Golestani, N., Rosen, S., & Scott, S. K. (2009). Native-language benefit for understanding
732 speech-in-noise: The contribution of semantics. *Bilingualism: Language and Cognition*,
733 *12*(3), 385–392.
- 734 Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception.
735 *Trends in Cognitive Sciences*, *4*(4), 131–138.
- 736 Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding
737 aspects of the functional anatomy of language. *Cognition*, *92*(1–2), 67–99.
- 738 Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature*
739 *Reviews Neuroscience*, *8*(5), 393–402.
- 740 Holland, S., & Vannest, J. (2015). *Cincinnati MR Imaging of Neurodevelopment (C-MIND)*.
741 https://nda.nih.gov/edit_collection.html?id=2329
- 742 Hunter, L. L., Blankenship, C. M., Lin, L., Sloat, N. T., Perdew, A., Stewart, H., & Moore, D. R.
743 (2021). Peripheral auditory involvement in childhood listening difficulty. *Ear and*
744 *Hearing*, *42*(1), 29–41.
- 745 Hunter, L. L., Blankenship, C. M., Shinn-Cunningham, B., Hood, L., Motlagh-Zadeh, L., &
746 Moore, D. R. (2022). Brainstem auditory physiology in children with listening
747 difficulties. *MedRxiv*.
- 748 Iliadou, V., Chermak, G. D., Bamiou, D.-E., Rawool, V. W., Ptok, M., Purdy, S., Jutras, B.,
749 Moncrieff, D., Mattsson, T. S., Ferre, J. M., & others. (2018). Letter to the editor: An
750 affront to scientific inquiry re: Moore, DR (2018) editorial: Auditory processing disorder,
751 *Ear Hear*, *39*, 617–620. *Ear and Hearing*, *39*(6), 1236–1242.
- 752 Iliadou, V., Kiese-Himmel, C., Bamiou, D.-E., Grech, H., Ptok, M., Chermak, G. D., Thai-Van,
753 H., Stokkerei Mattsson, T., & Musiek, F. E. (2019). Clinical expertise is core to an

- 754 evidence-based approach to auditory processing disorder: A reply to Neijenhuis et al.
755 2019. *Frontiers in Neurology*, *10*, 1096.
- 756 Jerger, J., & Musiek, F. E. (2000). Report of the Consensus Conference on the Diagnosis of
757 Auditory Processing Disorders in School-Aged Children. *Journal of the American*
758 *Academy of Audiology*, *11*(9), 467–474.
- 759 Kelley, K. S., & Littenberg, B. (2019). Comparison of Children’s Double Dichotic Digits and
760 SCAN-3 Competing Words Free Recall Scores. *Journal of the American Academy of*
761 *Audiology*, *30*(05), 363–369.
- 762 Killion, M. C., Niquette, Patricia, A(Etymotic reseach Inc., 61 Martin lane, Elk Grove Village, I.
763 60007), Gudmundsen, Gail, L., Revit, L. J., & Banerjee, S. (2004). Development of a
764 quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and
765 hearing impaired listeners. *J Acoust. Soc. Am*, *116*(4), 2395–2405.
766 <https://doi.org/10.1121/1.1784440>
- 767 Magimairaj, B. M., Nagaraj, N. K., Champlin, C. A., Thibodeau, L. K., Loeb, D. F., & Gillam,
768 R. B. (2021). Speech Perception in Noise Predicts Oral Narrative Comprehension in
769 Children With Developmental Language Disorder. *Frontiers in Psychology*, 4587.
- 770 Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S.,
771 Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., & others. (2022).
772 Reproducible brain-wide association studies require thousands of individuals. *Nature*,
773 *603*(7902), 654–660.
- 774 McFarland, D. J., & Cacace, A. T. (2009). Modality Specificity and Auditory Processing
775 Disorders. In D. J. Mcfarland & A. T. Cacace (Eds.), *Controversies in Central Auditory*
776 *Processing Disorder* (pp. 199–216). Plural Publishing, Inc.

- 777 McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S.
778 (2014). Listening effort and fatigue: What exactly are we measuring? A British Society of
779 Audiology Cognition in Hearing Special Interest Group ‘white paper’. *International*
780 *Journal of Audiology*, 53(7), 433–445.
- 781 Moore, D. R. (2018). Auditory processing disorder (APD). *Ear and Hearing*, 39(4), 617.
- 782 Moore, D. R., Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R. H.,
783 & Munro, K. J. (2014). Relation between speech-in-noise threshold, hearing loss and
784 cognition from 40–69 years of age. *PloS One*, 9(9), e107720.
- 785 Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., & Riley, A. (2010). Nature
786 of Auditory Processing Disorder in Children. *Pediatrics*, 126(2), E382–E390.
787 <https://doi.org/10.1542/peds.2009-2826>
- 788 Moore, D. R., Sieswerda, S. L., Grainger, M. M., Bowling, A., Smith, N., Perdew, A., Eichert,
789 S., Alston, S., Hilbert, L. W., Summers, L., & others. (2018). Referral and diagnosis of
790 developmental auditory processing disorder in a large, United States hospital-based
791 audiology service. *Journal of the American Academy of Audiology*, 29(5), 364–377.
- 792 Neijenhuis, K., Campbell, N., Cromb, M., Luinge, M., Moore, D. R., Rosen, S., & de Wit, E.
793 (2019). An Evidence-based Perspective on ‘Misconceptions’ Regarding Pediatric
794 Auditory Processing Disorder. *Frontiers in Neurology*, 10, 287.
- 795 Parente, F., Frascarelli, M., Mirigliani, A., Di Fabio, F., Biondi, M., & Colosimo, A. (2018).
796 Negative functional brain networks. *Brain Imaging and Behavior*, 12(2), 467–476.
- 797 Petley, L., Hunter, L. L., Zadeh, L. M., Stewart, H. J., Sloat, N. T., Perdew, A., Lin, L., &
798 Moore, D. R. (2021). Listening Difficulties in Children with Normal Audiograms:
799 Relation to Hearing and Cognition. *Ear and Hearing*.

- 800 Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E.,
801 Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., & others. (2016). Hearing
802 impairment and cognitive energy: The framework for understanding effortful listening
803 (FUEL). *Ear and Hearing*, 37, 5S-27S.
- 804 Rees, N. S. (1973). Auditory processing factors in language disorders: The view from
805 Procrustes' bed. *Journal of Speech and Hearing Disorders*, 38(3), 304–315.
- 806 Roebuck, H., & Barry, J. G. (2018). Parental perception of listening difficulties: An interaction
807 between weaknesses in language processing and ability to sustain attention. *Scientific*
808 *Reports*, 8(1), 1–10.
- 809 Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language
810 understanding. *International Journal of Audiology*, 58(5), 247–261.
- 811 Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö.,
812 Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & others. (2013). The Ease of Language
813 Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in*
814 *Systems Neuroscience*, 7, 31.
- 815 Rudner, M., & Signoret, C. (2016). The role of working memory and executive function in
816 communication under adverse conditions. *Frontiers in Psychology*, 7, 148.
- 817 Schiller, I. S., Remacle, A., Durieux, N., & Morsomme, D. (2022). Effects of Noise and a
818 Speaker's Impaired Voice Quality on Spoken Language Processing in School-Aged
819 Children: A Systematic Review and Meta-Analysis. *Journal of Speech, Language, and*
820 *Hearing Research*, 65(1), 169–199.
- 821 Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech
822 recognition with primarily temporal cues. *Science*, 270(5234), 303–304.

- 823 Sharma, M., Purdy, S. C., & Humburg, P. (2019). Cluster analyses reveals subgroups of children
824 with suspected auditory processing disorders. *Frontiers in Psychology, 10*, 2481.
- 825 Shinn-Cunningham, B. (2017). Cortical and Sensory Causes of Individual Differences in
826 Selective Attention Ability Among Listeners With Normal Hearing Thresholds. *Journal*
827 *of Speech Language and Hearing Research, 60*(10), 2976.
828 https://doi.org/10.1044/2017_JSLHR-H-17-0080
- 829 Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in*
830 *Cognitive Sciences, 12*(5), 182–186.
- 831 Smits, C., Goverts, S. T., & Festen, J. M. (2013). The digits-in-noise test: Assessing auditory
832 speech recognition abilities in noise. *The Journal of the Acoustical Society of America,*
833 *133*(3), 1693–1706.
- 834 Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: A Functional Connectivity Toolbox
835 for Correlated and Anticorrelated Brain Networks. *Brain Connectivity, 2*(3), 125–141.
836 <https://doi.org/10.1089/brain.2012.0073>
- 837 Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-
838 scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8*(8),
839 665–670.
- 840