Anatomical and behavioral correlates of auditory perception in developmental dyslexia

Ting Qi \textsuperscript{1,2}, Maria Luisa Mandelli \textsuperscript{1,2}, Christa L. Watson Pereira \textsuperscript{1,2}, Emma Wellman \textsuperscript{1,2}, Rian Bogley \textsuperscript{1,2}, Abigail E. Licata \textsuperscript{1,2}, Edward F. Chang \textsuperscript{3}, Yulia Oganian \textsuperscript{3,4,*}, and Maria Luisa Gorno-Tempini \textsuperscript{1,2,*}

Affiliation:

\textsuperscript{1} Department of Neurology, University of California San Francisco, United States

\textsuperscript{2} UCSF Dyslexia Center, University of California San Francisco, United States

\textsuperscript{3} Department of Neurological Surgery, University of California San Francisco, United States

\textsuperscript{4} Center for Integrative Neuroscience, University of Tübingen, Tübingen, Germany

Corresponding author:

Ting Qi

Department of Neurology, University of California San Francisco, CA 94158, United States

Email: ting.t.qi@gmail.com

* Shared senior authorship
Abstract

Developmental dyslexia (DD) is typically associated with difficulties in manipulating speech sounds and, sometimes, in basic auditory processing. However, the neuroanatomical correlates of auditory difficulties in DD and their contribution to individual clinical phenotypes are still unknown. Recent intracranial electrocorticography (ECoG) findings associated processing of sound amplitude rises and speech sounds with posterior and middle superior temporal gyrus (STG), respectively. We hypothesize that regional STG anatomy will relate to specific auditory abilities in DD and that auditory processing abilities will relate to behavioral difficulties. One hundred and ten children (78 DD, 32 typically developing, age 7-15 years) completed amplitude rise time (ART) and speech in noise discrimination (SiN) tasks. They also underwent a battery of cognitive tests. Anatomical MRI scans were used to identify regions in which local cortical gyrification complexity correlated with auditory tasks in DD. Behaviorally, ART but not SiN performance was impaired in DD. Neurally, ART and SiN performance correlated with gyrification in posterior STG and middle STG, respectively. Furthermore, ART significantly contributed to reading impairments in DD, while SiN explained variance in phonological awareness only. Finally, ART and SiN performance was not correlated, and each task was correlated with distinct neuropsychological measures, such that distinct DD subgroups could be identified. Overall, we provide a direct link between the neurodevelopment of the left STG and individual variability in auditory processing abilities in DD. The dissociation between speech and non-speech deficits supports distinct DD phenotypes and implicates different approaches to interventions.

Keywords: developmental dyslexia, auditory processing deficit, local gyrification, superior temporal gyrus, amplitude rise time, speech in noise
Significance statement

The capacity to read is crucial for human development yet challenging for individuals with developmental dyslexia (DD). Individuals with DD show a range of auditory and speech processing deficits. We tested non-speech and speech processing abilities in DD and typically developing children, and combined them with measures of neuroanatomical structure in the human auditory speech cortex on the superior temporal gyrus (STG).

This unique combination revealed a behavioral and neuroanatomical dissociation between the speech and non-speech processing. Each task was uniquely related to a subdivision of the STG, and a distinct set of cognitive abilities. Our findings contribute to the understanding of auditory processing deficits in dyslexia and have clinical implications for individual phenotypes of individuals with DD.
Introduction

Developmental dyslexia (DD) is characterized by difficulties with reading and spelling that persist throughout life and cannot be attributed to general cognitive abilities or poor education opportunities (1). Although DD is diagnosed based on reading performance, it is often accompanied by a range of auditory deficits in processing and manipulating of speech and non-speech sounds (2). In this view, reading, which requires mapping from orthography to speech sounds (phonology), breaks down because of impaired speech sound representations or access to these representations (Phonological deficit theory of dyslexia, 3–5).

Recent theories of DD frame it as a more general auditory processing deficit, suggesting that general auditory impairments drive the inability to develop and access phoneme representations, known as phonological awareness (6, 7). This is in line with the view that to extract speech sounds from auditory streams, the auditory system has to identify a range of complex acoustic cues (8). A key acoustic feature for speech comprehension are amplitude modulations, specifically amplitude rises, which cue speech structure at phrasal and syllabic levels (9). Indeed, a large body of work has found impaired processing of amplitude rises in DD (10–14). Furthermore, DD sometimes show deficits in the perception of speech in noisy backgrounds, which is more challenging than under optimal listening conditions and might require more precise phoneme representations (15, 16). It remains debated, however, whether these auditory deficits characterize all or only subgroups of individuals with DD and how they relate to reading and phonological abilities.

These behavioral deficits are complemented by multiple studies reporting atypical neuroanatomical patterns in auditory pathways in DD (3, 17–21). Altered cortical thickness, myelinated cortical thickness ratio, and surface area lateralization have also been observed in the auditory temporal cortex in DD (22–26) and in individuals with familial risk for DD (27, 28). Moreover, electrophysiological M/EEG studies of DD found reduced neural responses to amplitude modulations in speech and non-speech sounds (29, 30). Interestingly, some but not all functional MRI studies report atypical activation of left hemispheric temporal regions in DD (3, 31–33). This observation has fueled the idea of functional subtypes in DD (34, 35). Yet, it remains unclear how neuroanatomical cortical structure, particularly in auditory cortical areas, might be related to auditory behavioral deficits in DD.

This gap is widened by our limited understanding of the neural computations underlying the processing of speech sounds in human auditory cortices. Recently, advances in intracranial electrophysiology (iEEG) recordings from the superior temporal gyrus (STG), a high-level auditory cortical region specialized for speech processing, have revealed the rapid dynamics of speech sound representations
in this area (36). Most relevant to the behavioral deficits in DD, recent studies established a spatial map for the encoding of amplitude rises at phrasal onsets in posterior STG, and phonemes and syllabic amplitude rises in middle STG (9, 37, 38). This detailed spatial brain map for speech sound processing opens new avenues for mapping auditory processing deficits in DD to underlying neural substrates.

Here, we built on these iEEG findings to hypothesize that the ability to process and manipulate speech and non-speech sounds in DD depends on the neuroanatomical development of the STG. To test this, we behaviorally assessed the ability to discriminate amplitude modulations in sounds and to perceive speech in noise, alongside cognitive, reading, and phonological abilities, in a group of children with a diagnosis of DD and in age-matched typically developing (TD) children. To test how these two auditory tasks map onto the brain’s structure, we determined the local gyrification index (LGI) using anatomical MRI scans in the same cohort. Based on our prior intracranial results, we hypothesized that amplitude modulation processing abilities would be correlated with neuroanatomical structure in the posterior STG (pSTG), whereas we expected speech in noise perception to be associated with the middle STG (mSTG). Further, we hypothesized that the ability to process speech and non-speech sounds might dissociate between two functional subgroups in DD.

Results

Left superior temporal gyrus underlies different auditory processing in children

One hundred and ten participants, including 78 children with DD and 32 TD children, who successfully completed at least one of the auditory tasks (see below for task details) were included in the present study (see detailed demographics and behavioral characteristics in Table 1). We first compared auditory processing, evaluated by a non-speech amplitude rise time (ART) task and a speech in noise (SiN) task, between children with DD and TD children. The ART task evaluated the perceptual threshold for amplitude rise time discrimination using a standard adaptive staircase procedure converging to a 79% just noticeable difference (JND, 42). Children were asked to identify which of two harmonic tones with a triangular amplitude shape had a longer rise time. Tone rise times on subsequent trials were adjusted according to a child’s response and the rise time JND was calculated as the average rise time on the last 8 reversal trials. To evaluate speech in noise perception, during the SiN task, children were asked to repeat the heard one single syllable in background noise at different noise levels (i.e., 6 dB and 12 dB). The percentage of correct responses for each syllable and for individual phonetic features were calculated.

As expected, amplitude rise time discrimination was impaired in DD ($mean_{\text{rawJND}} = 226.57, SD = 155.51$), as evident in significantly elevated thresholds in this group as compared to TD ($mean_{\text{rawJND}} =$
127.81, $SD = 84.28;\ \text{welch'}s-t(70.74) = -3.03, p < 0.01$, Figure 1a). In contrast, groups did not differ in the SiN task (main effect of group: $F_{(1,167)} = 1.97, p = 0.16$; group by noise interaction effect: $F_{(1, 167)} = 0.91, p = 0.34$), with an overall more impaired performance at higher relative noise levels (group average: 6dB: mean = 43.50%, 12 dB: mean = 25.00%, see Figure 1c) in both groups (main effect of noise level: $F_{(1,167)} = 253.10, p < 0.01$). Given the very low accuracy in the 12 dB condition, we focused on the 6 dB condition in all subsequent analyses (results for the 12 dB condition see Figure S1, Table S3, S5, and the Supplementary Results). As previous work (40) showed selective impairments in DD for the perception of certain consonant types, we also analyzed recognition accuracy for single phonetic features. While noise differentially affected phonetic features in both groups (main effect of feature: $F_{(1,268)} = 8.92, p < 0.01$), we observed no differences between groups (main effect of group: $F_{(1,268)} = 0.05, p = 0.83$; group by phonetic features interaction effect: $F_{(1,268)} = 0.24, p = 0.78$). Finally, behavioral performance was not correlated between the two tasks neither in DD ($r = -0.07, p = 0.63$, Table 2), supporting a dissociation between non-speech auditory and speech perception deficits.

Next, we proceeded to investigate whether variation in cortical structure in speech cortical areas might underlie these two aspects of auditory processing. We calculate the local gyrification index (LGI) of the brain, a metric quantifying the amount of cortex buried within the sulcal folds as compared with the amount of cortex on the outer visible cortex (41). Whole-brain LGI and behavior correlation analyses showed that better amplitude rise time discrimination was associated with greater cortical folding in the left pSTG ($p < 0.01$, FWE corrected, Figure 1b) across both groups (n = 78, DD/TD = 56/22, see table S1 for details of participant inclusion in individual analyses). Likewise, SiN task performance was correlated with LGI in the left mid-anterior STG ($p < 0.01$, n = 84, DD/TD = 60/24). Additionally, SiN task performance was also correlated with LGI in clusters in the left insula, precentral ($p < 0.01$) gyrus, and the right pSTG ($p < 0.01$, all FWE corrected, Figure 1d), suggesting that these areas might be part of the network involved in speech comprehension. Of note, similar results were also observed within the DD group for each of the auditory tasks. Whole-brain comparisons between LGI in the DD and TD showed no group differences within the identified clusters, nor at the whole-brain level.

Crucially, we found a high overlap between the two significant left STG clusters and previously identified functional zones in the left STG (9, 37, 38). The significant ART cluster overlapped with the speech onset zone in pSTG, while the SiN cluster overlapped with the phonetic features zone in mSTG (Figure 1e). Overall, these analyses show that cortical folding of different functional subdivisions of the STG is predictive of distinct aspects of auditory processing in children with and without DD.
Figure 1. Auditory processing abilities in children with DD compared to TD and their relationship with cortical folding in the whole brain. a. Children with DD showed decreased amplitude rise time discrimination abilities. b. Amplitude rise time discrimination abilities were associated with the local gyrification index (LGI) of the left posterior superior temporal gyrus (pSTG) \((p < 0.05 \text{ and FWE-corrected})\). c. No significant group difference in speech in noise recognition abilities between DD and TD groups in both noise levels. d. However, speech in noise recognition abilities were correlated with the LGI in the left middle STG (mSTG) \((p < 0.05 \text{ and FWE-corrected})\). e. The left pSTG and mSTG fall in the speech onset (color-coded in green) and phonetic feature zone (color-coded in purple) defined in our ECoG work, respectively [Figure adapted from (9)]. Notes: solid and dashed
gray blobs denote the left pSTG associated with amplitude rise time and the left mSTG associated with speech in noise perception. Red and blue denote DD and TD groups separately. The auditory processing ability scores were z-transformed. \( * p < 0.05. \)

### Table 1. Demographics and behavioral characteristics of the developmental dyslexic (DD) and typically developing (TD) children.

<table>
<thead>
<tr>
<th>Demographics</th>
<th>DD (n = DD/TD)</th>
<th>Mean±SD</th>
<th>Range</th>
<th>TD (n = DD/TD)</th>
<th>Mean±SD</th>
<th>Range</th>
<th>p-values</th>
<th>t-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years - 78/32)</td>
<td>10.67±1.89</td>
<td>7.40-14.60</td>
<td>11.35±2.11</td>
<td>8.20-15.00</td>
<td>0.12</td>
<td>1.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex (F/M - 78/32)</td>
<td>32/46</td>
<td>16/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Single word/nonword reading (%ile)

**TOWRE-2**

- Sight word efficiency (78/32)* | 15.51±18.07 | 0.10-70.00 | 65.63±22.59 | 16.00-99.00 | <0.01 | 10.38 |
- Phonemic decoding efficiency (78/32)* | 15.59±18.07 | 0.30-60.00 | 66.09±24.03 | 21.00-99.00 | <0.01 | 10.06 |

**WJ-IV**

- Word identification (ID) (77/n.a.)*,1 | 24.03±22.66 | 0.10-95.00 | n.a. | n.a. | <0.01 | -14.78 |
- Word attack (78/n.a.)*,1 | 36.53±25.14 | 1.00-97.00 | n.a. | n.a. | <0.01 | -16.24 |

### Phonological awareness (%ile)

- Segmentation (78/n.a.)*,1,+ | 61.99±21.72 | 14.00-95.00 | n.a. | n.a. | <0.01 | 4.87 |
- Sound blending (78/n.a.)*,1 | 55.36±26.51 | 5.00-98.00 | n.a. | n.a. | 0.08 | 1.79 |
- Sound awareness (78/n.a.)*,1 | 39.56±25.94 | 4.00-95.00 | n.a. | n.a. | <0.01 | -3.55 |

### Cognitive and language (%ile)

- Matrix reasoning (78/29)*,2 | 66.29±21.51 | 21.00-97.00 | 73.24±14.96 | 50.00-98.00 | 0.06 | 1.88 |
- Digit span forward (77/28)* | 29.79±24.40 | 1.00-84.00 | 55.21±29.48 | 2.00-95.00 | <0.01 | 3.90 |
- Digit span backward (77/28)* | 35.90±23.53 | 5.00-91.00 | 54.43±26.62 | 5.00-98.00 | <0.01 | 3.40 |
- ROWPVT vocabulary (76/n.a.)*,+ | 70.66±22.32 | 14.00-99.80 | n.a. | n.a. | <0.01 | 8.07 |
- Rapid picture naming (78/20) | 31.02±22.72 | 0.30-96.00 | 53.50±20.86 | 18.00-90.00 | 0.88 | -0.15 |

**Notes:** All scores reflect percentiles (%ile) relative to age-matched population data. Welch’s two-sample t-tests and Fisher’s Exact tests were conducted to examine the differences between the DD and TD groups for continuous behavioral measures and sex. \(^1\)For tests completed by the DD group only, we reported one-sample t-tests against the 50\(^{th}\) percentile. \(^2\)Matrix reasoning showed significant differences between TD and DD for the whole cohort, however for the sub-samples that completed each of the auditory tasks it showed only trend-level differences \((p = 0.16 \text{ and } 0.07 \text{ for the amplitude rise time task and speech in noise task, respectively})\). Numbers in brackets denote the number of children in the DD and TD groups, respectively. SD: standard deviation. \( * p < 0.05. + \) scores where the DD group performed above the 50\(^{th}\) percentile of age-matched peers.

**Amplitude rise time is associated with reading and phonological abilities in DD**

Next, we aimed to understand how the two auditory tasks are related to the main deficits in reading and phonology in DD. Children with DD underwent a comprehensive battery of neuropsychological and academic testing. This battery consisted of several standardized reading tests, along with matrix reasoning, digit spans, receptive one-word picture (ROWPVT) vocabulary, and rapid picture naming tests. Due to the small sample size and subset of behavioral data available in the TD group, we conducted all behavioral analyses within the DD group. Pairwise Pearson’s correlations between auditory tasks and reading measures
showed a complex pattern of relationships between the two auditory tasks and different measures of general cognitive abilities and literacy skills (Table 2). Specifically, amplitude rise time discrimination, untimed reading accuracies (WJ Word ID and Word attack), and phonological awareness were all positively correlated (Figure 2a), with similar, but weaker and non-significant, associations with timed reading tasks. Notably, effects for amplitude rise time discrimination were stronger for words than for nonwords. In addition, we found a complex pattern of correlations between amplitude rise time discrimination, reading, and phonological awareness and the rest of the tasks. In particular, amplitude rise time discrimination abilities were correlated with digit span backward, but not correlated with matrix reasoning and age ($r = 0.11$, $p = 0.42$, Table 2 and Figure 2a).

Given this complex pattern of dependencies, we thus continued to test whether amplitude rise time discrimination scores are predictive of reading accuracy (WJ Word ID) after accounting for shared variance with other tasks. We conducted a hierarchical regression analysis with reading as the dependent variable, age, sex, matrix reasoning, digit spans, rapid picture naming and ROWPVT vocabulary as variables of no covariates interest, and phonological awareness and amplitude rise time as independent variables of interest. Modeling results showed a significant unique contribution of rise time ($\Delta R^2 = 0.06$, $t = 2.28$, $p = 0.03$, $\beta = 0.28$, Table S4 for hierarchical regressions of other reading scores in DD) to word reading accuracy even after accounting for phonological awareness and all other neuropsychological scores.

In a final step, we asked how amplitude rise time discrimination abilities and phonological awareness relate in their contribution to word reading accuracy (WJ Word ID). A mediation analysis tested whether auditory processing contributes to reading directly or via an effect on phonological awareness. It showed a major direct contribution of amplitude rise time discrimination to word reading accuracy, as well as a marginal indirect impact via phonology ($p_{direct} = 0.02$, $p_{indirect} = 0.06$, $p_{prod\text{-mediated}} = 0.06$, Figure 2b). Taken together, our behavioral analyses show that DD are impaired in their perception of amplitude rise time, independent of general cognitive ability, vocabulary, and age. Importantly, they confirm that amplitude rise time discrimination is a major contributing factor to DD’s impaired reading abilities.

![Figure 2. Amplitude rise time (ART) is predictive of reading in children with DD. a. Amplitude rise time abilities were correlated with word reading accuracy (WJ Word ID), and phonological awareness (sound awareness), but not with matrix reasoning in DD.](image-url)
b. Amplitude rise time influences word reading accuracy partially through phonological awareness. $c$ denotes total effect and $c'$ denotes direct effect. * $p < 0.05$, , $p < 0.10$.

Table 2. Pairwise Pearson’s correlation between amplitude rise time (ART), speech in noise (SiN), and reading, phonology, and cognitive measures.

<table>
<thead>
<tr>
<th>r (p, n)</th>
<th>ART</th>
<th>SiN</th>
<th>TOWRE SWE</th>
<th>TOWRE PDE</th>
<th>WJ Word ID</th>
<th>Word attack</th>
<th>Sound awareness</th>
<th>Digit span forward</th>
<th>Digit span backward</th>
<th>Rapid picture naming</th>
<th>Matrix reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiN</td>
<td>-0.07</td>
<td>(0.63, 51)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOWRE SWE</td>
<td>0.24</td>
<td>(0.07, 58)</td>
<td>0.15</td>
<td>(0.23, 66)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOWRE PDE</td>
<td>0.12</td>
<td>(0.36, 58)</td>
<td>0.08</td>
<td>(0.54, 66)</td>
<td>0.65</td>
<td>(&lt;0.01, 78)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WJ Word ID</td>
<td>0.41</td>
<td>(&lt;0.01, 57)</td>
<td>0.16</td>
<td>(0.20, 65)</td>
<td>0.64</td>
<td>(&lt;0.01, 77)</td>
<td>&lt;0.01, 77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WJ Word attack</td>
<td>0.28</td>
<td>(0.04, 58)</td>
<td>0.05</td>
<td>(0.66, 66)</td>
<td>0.37</td>
<td>(&lt;0.01, 78)</td>
<td>(&lt;0.01, 78)</td>
<td>&lt;0.01, 77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sound awareness</td>
<td>0.28</td>
<td>(0.03, 58)</td>
<td>0.24</td>
<td>(0.05, 66)</td>
<td>0.33</td>
<td>(&lt;0.01, 78)</td>
<td>(&lt;0.01, 78)</td>
<td>&lt;0.01, 77)</td>
<td>0.57</td>
<td>(&lt;0.01, 78)</td>
<td></td>
</tr>
<tr>
<td>Digit span forward</td>
<td>0.01</td>
<td>(0.98, 58)</td>
<td>0.26</td>
<td>(0.04, 65)</td>
<td>0.08</td>
<td>(0.49, 77)</td>
<td>0.20</td>
<td>(0.08, 77)</td>
<td>0.26</td>
<td>(0.02, 76)</td>
<td>0.22</td>
</tr>
<tr>
<td>Digit span backward</td>
<td>0.29</td>
<td>(0.03, 58)</td>
<td>-0.14</td>
<td>(0.27, 65)</td>
<td>0.32</td>
<td>(&lt;0.01, 77)</td>
<td>(&lt;0.01, 77)</td>
<td>0.05, 76)</td>
<td>0.22</td>
<td>(0.07, 77)</td>
<td>0.19</td>
</tr>
<tr>
<td>Rapid picture naming</td>
<td>-0.17</td>
<td>(0.22, 58)</td>
<td>0.28</td>
<td>(0.02, 66)</td>
<td>0.24</td>
<td>(0.03, 78)</td>
<td>0.26</td>
<td>(0.02, 78)</td>
<td>0.14</td>
<td>(0.22, 77)</td>
<td>0.01</td>
</tr>
<tr>
<td>Matrix reasoning</td>
<td>0.21</td>
<td>(0.12, 58)</td>
<td>0.08</td>
<td>(0.51, 66)</td>
<td>-0.05</td>
<td>(0.66, 78)</td>
<td>-0.01</td>
<td>(0.90, 78)</td>
<td>0.21</td>
<td>(0.07, 76)</td>
<td>0.21</td>
</tr>
<tr>
<td>ROWPVT vocabulary</td>
<td>0.08</td>
<td>(0.56, 57)</td>
<td>0.33</td>
<td>(0.01, 64)</td>
<td>0.13</td>
<td>(0.28, 76)</td>
<td>0.13</td>
<td>(0.43, 76)</td>
<td>0.09</td>
<td>(0.11, 75)</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Notes: Age and sex were controlled for in the analysis. Numbers in brackets denote the uncorrected $p$-values. The auditory processing ability scores were $z$-transformed. Bold texts denote $p < 0.05$, italic texts denote $p < 0.10$.

Speech in noise is associated with phonological, but not reading abilities in DD

For the SiN task, despite the absence of group differences, pairwise correlations (Table 2) showed that speech in noise recognition abilities were positively correlated with sound awareness in the DD group ($n = 66$, Figure 3), but not with any reading measures. Further, speech in noise performance was also significantly correlated with digit span forward, rapid picture naming, ROWPVT vocabulary, and age ($r = 0.32$, $p < 0.01$), but not with matrix reasoning (Figure 3 and Table 2). We thus focused further analyses on the contribution of speech in noise perception to phonological awareness. Hierarchical regression modeling showed that the best model to explain phonological awareness ($R^2 = 0.11$, $F_{(3,60)} = 3.58$, $p = 0.02$) included speech in noise recognition, sex and matrix reasoning. Critically, speech in noise recognition contributed significantly to this model ($t = 2.21$, $p = 0.03$, $\beta = 0.26$), after accounting for variance by all other factors.
Even if all other tasks were included, speech in noise perception still accounted for unique variance in phonological awareness ($\Delta R^2 = 0.05$, $t = 2.00$, $p < 0.05$, $\beta = 0.29$, Table S5), even though the full model was not significant (full model $R^2 = 0.04$, $F_{(8,55)} = 1.36$, $p = 0.23$). Collectively, these results support a relationship between speech in noise recognition and phonological awareness in neurotypical and DD populations, but not directly with single word reading.

![associations with speech in noise (SiN) in DD](image)

**Figure 3.** Speech in noise (SiN) recognition abilities were associated with phonological awareness (sound awareness) in children with DD, but not with word reading accuracy (WJ Word ID) and matrix reasoning. * $p < 0.05$.

**ART and SiN tasks tap into different linguistic and cognitive abilities**

In a final set of analyses, we aimed to directly characterize subgroups of the DD cohort with specific impairments in one of the two auditory processing tasks. To this end, we subset the entire cohort to two subgroups, with particularly low (bottom 25th percentile) or high (upper 25th percentile) performance on each task. As expected, ART-based subgroups ($n = 21$ and 16 for high and low performance) differed in all measures that were found to be correlated with amplitude rise time discrimination abilities in the entire DD group (Table 1), including WJ Word ID, all three phonological awareness measures, and digit span backward scores (Table 3). Likewise, SiN-based subgroups ($n = 18$ and 17 for high and low performance) differed in tasks that were correlated with speech in noise abilities in the entire DD cohort, including ROWPVT vocabulary and rapid picture naming, with marginal differences in phonological sound awareness and digit span forward (Table 3). Next, direct comparisons of correlations across tasks showed significantly different correlations for digit span backward ($p = 0.02$, $z = 2.37$) and rapid picture naming ($p = 0.01$, $z = 2.49$), whereas all other correlations did not differ significantly across tasks. Overall, this analysis suggests that speech in noise recognition and amplitude rise time discrimination are associated with different domains of verbal memory. That is, verbal working memory is associated with non-speech auditory processing, whereas processing speed is associated with speech perception abilities. In short, this subgrouping analysis replicates the correlational patterns within the DD group and suggests that the two auditory tasks tap into (partially) distinct neurocognitive abilities.
Table 3. Group differences between two subgroups of DD formulated according to the performance in amplitude rise time (ART) and speech in noise (SiN) tasks.

<table>
<thead>
<tr>
<th>Art DD subgroups</th>
<th>Art DD subgroups</th>
<th>SiN DD subgroups</th>
<th>SiN DD subgroups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean±SD</td>
<td>Mean±SD</td>
<td>Mean±SD</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>11.27±1.88</td>
<td>10.71±1.72</td>
<td>0.35</td>
</tr>
<tr>
<td>Sex (M/F)</td>
<td>8/13</td>
<td>5/11</td>
<td>0.89</td>
</tr>
<tr>
<td><strong>Experimental Auditory task</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ART</td>
<td>65.95±27.17</td>
<td>437.42±51.72</td>
<td>&lt;0.01*</td>
</tr>
<tr>
<td>SiN (%)</td>
<td>42.59±8.27</td>
<td>42.77±5.61</td>
<td>0.74</td>
</tr>
<tr>
<td><strong>Single word/nonword reading (%ile)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOWRE-2</td>
<td>24.46±22.94</td>
<td>11.26±12.68</td>
<td>0.06</td>
</tr>
<tr>
<td>TOWRE PDE</td>
<td>22.22±21.85</td>
<td>15.06±17.46</td>
<td>0.32</td>
</tr>
<tr>
<td>WJ-IV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word ID</td>
<td>40.20±27.46</td>
<td>17.41±16.42</td>
<td>&lt;0.01*</td>
</tr>
<tr>
<td>Word attack</td>
<td>49.76±23.74</td>
<td>33.63±25.23</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>Phonological awareness (%ile)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segmentation</td>
<td>70.43±19.26</td>
<td>51.00±22.37</td>
<td>&lt;0.01*</td>
</tr>
<tr>
<td>Sound blending</td>
<td>69.00±22.17</td>
<td>50.38±22.74</td>
<td>&lt;0.01*</td>
</tr>
<tr>
<td>Sound awareness</td>
<td>48.57±26.82</td>
<td>30.75±26.09</td>
<td>0.04*</td>
</tr>
<tr>
<td><strong>Cognitive and language (%ile)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matrix reasoning</td>
<td>70.67±21.76</td>
<td>63.56±19.52</td>
<td>0.07</td>
</tr>
<tr>
<td>Digit span forward</td>
<td>32.90±28.60</td>
<td>31.38±21.57</td>
<td>0.59</td>
</tr>
<tr>
<td>Digit span backward</td>
<td>47.10±24.30</td>
<td>31.25±19.88</td>
<td>0.03*</td>
</tr>
<tr>
<td>ROWPVT vocabulary</td>
<td>69.40±24.57</td>
<td>66.45±24.98</td>
<td>0.66</td>
</tr>
<tr>
<td>Rapid picture naming</td>
<td>25.62±22.44</td>
<td>34.38±22.04</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Notes: Welch’s two-sample t-test was used to compare two DD subgroups. Fisher’s exact test was conducted for the group difference regarding sex. Age and sex were controlled for in the group comparison. * and . denotes p < 0.05 and p < 0.10, respectively. SD: standard deviation.

**Discussion**

We provide a neuroanatomical and behavioral dissociation between non-speech auditory processing of sound amplitudes and speech recognition abilities in TD children and children with DD. Behaviorally, DD were impaired in non-speech amplitude rise time discrimination, but not in speech in noise recognition. Neurally, cortical LGI showed a dissociation between neural substrates related to the two tasks in DD. Across groups, amplitude rise time discrimination was positively correlated with LGI of left pSTG, whereas speech in noise perception was positively correlated with LGI of left mSTG. This
dissociation was further manifested in distinct association patterns for reading, phonology, and other cognitive abilities with the two tasks.

The observed dissociation between non-speech amplitude rise time and speech in noise perception is in line with recently discovered distinct response profiles in posterior and middle STG in intracranial EEG recordings (9, 37, 38). Prior work in DD using MRI found atypical STG activation in a range of auditory and phonetic tasks in DD (17, 18, 20, 32). Others reported atypical cortical morphometry in temporal cortex (23, 24, 26). Our results extend that literature and suggest that neuroanatomically dissociated aspects of the STG support phonological and non-speech auditory processing in DD. Furthermore, this dissociation highlights the distinct roles of posterior and middle STG in speech sound processing in developmental populations. Moreover, the correlation between gyrification and behavioral performance in our cohort is in line with prior studies that argue in favor of a functional gradient for processing speech sounds along the posterior-to-anterior axis of the STG (18).

Interestingly, brain to behavior correlations were consistent across groups in our cohort, with no overall LGI differences between DD and TD children. This stands in contrast to previously reported structural alteration in the temporal cortex in DD (22, 24, 26). However, although several studies showed sulcal and folding-related alterations in DD, no direct evidence of atypical STG LGI has been ever reported in DD (42–44). In our cohort, this might be due to a large variability in LGI development in children and the small size of our TD group. However, the overall similar correlational pattern in TD and DD might also indicate that underlying deficits in DD lead to overall reduced LGI but do not alter the role of the STG in sound processing.

As expected, we found that children with DD were impaired in amplitude rise time discrimination. In fact, rise time discrimination deficits in children and adults with DD are well-documented and among the most robust auditory deficits in DD (6, 13, 14). Prior studies also found that rise time deficits were related to reading and phonological awareness (12, 13, 45, 46). Our results are in line with those studies, showing that rise time deficits are direct predictors of reading abilities, beyond variance that can be attributed to phonological abilities. Interestingly, amplitude rise time discrimination abilities were not correlated with age, supporting a precursor role of amplitude rise time for phonological and reading skills in DD, as has been proposed in the past (10, 47–50). Finally, we want to note that recent studies showed genetic correlations between left posterior superior temporal cortex, where onset rise time is represented, and reading and language measures in cohorts including young children as well as young adults (51), which might further point to the heritable and innate nature of rise time discrimination.

In contrast, speech in noise perception was not impaired in our cohort of children with DD and was predictive of phonological awareness but not of reading skills. Indeed, prior findings on speech perception
in DD have been mixed [impaired: (10, 15, 52, 53); intact, particularly in adults: (54–56)]. It has been previously suggested that speech in noise abilities improve with reading instruction and phonological awareness and might not persist into adulthood (10, 57, 58). Notably, our cohort of DD received targeted interventions with phonological awareness training. Together with the correlation between phonological awareness and speech in noise perception, this suggests that our participants overcompensated for their speech in noise deficits. This account is supported by the improved speech in noise perception in older participants - those who had more reading and phonological training. It is further supported by neuroimaging studies that found a compensatory role of right STG in speech in noise perception in adults with DD (59). This is in line with higher LGI of the right STG with better speech in noise recognition in our cohort. Overall, our results suggest that speech perception deficits are related to phonological processing skills in DD but, unlike deficits with amplitude rise time perception, are not directly related to reading abilities.

Behavioral comparisons between amplitude rise time and speech in noise perception showed that the two tasks were not directly related but rather were most closely related to distinct sets of cognitive skills, in line with past results (10, 52, 54). Specifically, while amplitude rise time perception was associated with verbal working memory skills, speech in noise perception was related to general linguistic skills and verbal short-term memory. Thus both tasks are constrained by distinct aspects of verbal memory (60–62). Within and across large numbers of experimental trials, in addition to sensory sensitivity, memory and attention are required to accomplish auditory tasks. In particular, the rise time task requires a child to decide from two tones, which may load upon memory differently than a task that presents each stimulus only once (62). The association between tone discrimination and verbal working memory might support a general difficulty in the automatic extraction of stimulus regularities from auditory inputs in DD (7, 63, 64). In contrast, speech in noise abilities were associated with receptive vocabulary and lexical retrieval speed, and marginally with verbal short-term memory, favoring the hypothesis of potential linguistic compensation for speech in noise deficits during development (54, 55). Furthermore, while both tasks were related to phonological awareness, only rise time contributed unique predictive power to reading skills. Taken together with the neural dissociation between the tasks, our results emphasize that the specific expression of auditory deficits differs between individuals, with different auditory impairments affecting phoneme awareness and reading. The distinct neural correlates of our two specific tasks highlight that these differences also stem from distinct neural impairments. Overall, these results support the existence of distinct DD phenotypes, which is critical for clinical applications, and provide a framework for the design and evaluation of differential approaches to interventions for DD (65). In the future, longitudinal studies in young children would be warranted to elucidate the potential causal relationships between different auditory abilities and their relation to neurodevelopment along the STG in neurotypical and impaired populations.
Several limitations should be considered when interpreting our results. First, our study included a relatively small cohort of TD children. As discussed above, null results at the group level in behavior and neural analyses warrant future studies with a larger sample size of TD. Next, although LGI is considered to be one of the most sensitive neural measures to distinguish DD and TD (66), it is largely under-investigated compared to other cortical geometric properties. This highlights the need for investigations of the underlying biological mechanism from different perspectives (67, 68). In addition, according to previous literature (57, 58), we argue that some auditory processing abilities could be ameliorated in later childhood. However, this needs to be confirmed by future longitudinal studies.

Overall, we provide the first evidence for distinct contributions of posterior and middle STG to different auditory processing deficits in DD. Our study enhances the understanding of auditory processing deficits in DD by characterizing how distinct auditory tasks are related to reading, phonology, and cortical neuroanatomy. Our results show that auditory and phonological processing difficulties may arise through multiple underlying mechanisms, which vary across individuals. The clinical implications of this pattern call for future studies on the inter-individual variability in DD phenotypes and their response to interventions.

Methods

Participants

This study includes 78 children with DD and 32 TD children who successfully completed at least one of the auditory tasks. A subgroup of 102 (76 DD, 26 TD) completed the MRI session. To maximize sample sizes, each of the following analyses included the maximal subset of children that completed the relevant tasks (see Table S1 for initial sample sizes and Table 1 for included sample sizes). All DD children were selected from the recruitment base at the UCSF Dyslexia Center, a multidisciplinary research program that performs neurological, psychiatric, cognitive, linguistic, and neuroimaging evaluations of children with language-based neurodevelopmental disorders. Of note, the center partners with several schools for young individuals with language-based learning differences, where children participate in Orton-Gillingham-based intervention programs. TD children were recruited through local schools and parent networks. Reading and language abilities were assessed using a battery of standardized reading tests. General cognitive abilities were assessed using the Matrix reasoning test (WASI, 69). All DD children were native speakers of English, aged between 7 and 15 years, and underwent a detailed clinical interview and neurological examination. They had prior formal diagnoses of DD and at least one current reading score.
falling below the 25\textsuperscript{th} percentile of same-aged peers on a standardized reading test despite extensive school-based reading intervention and general cognitive abilities within the normal range (16\textsuperscript{th} percentile) of same-aged peers. Exclusion criteria for both groups included acquired brain injury, neurological disorders such as perinatal injuries, seizures, and severe migraine. TD were excluded if a single score on any of the reading or general cognitive ability tests fell below the normal range (16\textsuperscript{th} percentile) of same-aged peers (see Table 1 below). Further exclusion criteria for the TD group were a history of academic difficulties, prior diagnoses of DD, or other developmental, neurological, or psychiatric disorders. Behavioral assessments were typically completed within 6 months from MRI scans, with a mean interval of 0.09 (SD = 0.15, in years). Written informed consent was obtained from the legal guardian or parent of the children. Additionally, children provided verbal consent for participation before the experiments. The study was approved by the UCSF Committee on Human Research and complied with the declaration of Helsinki.

**Neuropsychological and academic assessment**

Children with DD underwent a comprehensive battery of neuropsychological and academic testing. Neuropsychological testing consisted of matrix reasoning for general cognitive abilities (WASI Matrix Reasoning, 68), digit span forward (DSF) and backward (DSB) (WISC-IV Integrated Digit Span, 69) for verbal short-term memory and verbal working memory respectively, receptive one-word picture vocabulary Test-4 for general vocabulary skills (ROWPVT, 70), and rapid picture naming for lexical processing (retrieval) speed (Woodcock-Johnson IV, 71).

To evaluate their reading abilities, all DD children completed two sets of standardized single word reading and literacy tests:

1) Woodcock-Johnson IV (WJ-IV) subtests: untimed word identification (reading accuracy) and word attack (non-word reading accuracy), as well as tests for three different aspects of phonological awareness: sound blending, segmentation, and sound awareness. We found that as a result of school-specific intervention protocols, our DD cohort performed above 50\% of age-matched peers on two of these tests. We thus only included the third subtest (sound awareness) as a measure of phonological awareness in further analyses.

2) The Timed Test of One-Word Reading Efficiency, version 2 (TOWRE-2, 72), which has two subtests, one measuring sight word recognition efficiency based on timed single-word reading efficiency (SWE), and one measuring phonemic decoding efficiency (PDE) based on timed non-word reading fluency.
Of note, most DD children also completed the Gray Oral Reading Test, version 5 (GORT, 73), which assesses oral reading fluency and comprehension based on passages and stories reading. Because this test assessed complex reading comprehension rather than phonological decoding or single word reading, we do not include it in any of our main analyses. However, scores are reported in supplements for completeness (Table S2).

Due to limitations of time, protocol updates, or subject fatigue, not all DD children were able to complete all of the tasks (see Table 1 for sample size details of each test). TD children participated in an abbreviated study protocol and completed only matrix reasoning, digit span forward and backward, and TOWRE-2 tests.

Auditory processing tasks

*Non-speech amplitude rise time task (ART)*: We evaluated the perceptual threshold for amplitude rise time with a standard adaptive staircase procedure (39), using a 3-steps-down 1-step-up procedure converging to a 79% just noticeable difference (JND). In each trial, participants heard two harmonic tones with a triangular amplitude shape and were asked to identify which of the two tones had a longer rise time (softer onset). Tone rise times on subsequent trials were adjusted according to a child’s response: it was increased following an incorrect response and decreased after a series of three consecutive correct responses. The standard tone rise time was fixed to 15 ms, whereas the test tone had an initial rise time of 300 ms, varying between 15 and 500 ms. The inter-stimulus interval was fixed to 350 ms. The task terminated after eight response reversals (i.e., switches between correct and incorrect responses) or the maximum possible 80 trials (7). To account for worse overall performance in children, we defined successful completion of the task as performance above an accuracy criterion of 65 % (75, 76) (see Table S1 in the Supplementary Methods). The rise time JND was then calculated as the average rise time on the last 8 reversal trials. A lower raw rise time JND indicates better amplitude rise time performance. For further analyses, raw JNDs were z-scored and inverted such that higher z-scores indicate better performance.

*Speech in noise task (SiN)*: Speech in noise perception accuracy was tested using the single syllable in background noise. On each trial, children heard a single syllable and were asked to repeat what they heard. The examiner recorded the responses. Syllables were consonant-vowel combinations, namely 12 consonants covering three phonetic features (voicing, place, and manner) in English and ending with the vowel /a/. Each syllable was repeated 5 times and presented in two noise conditions, at -6 and -12 dB relative to the noise level. Noise conditions were blocked, with the 6 dB condition administered first. Before the noise conditions, all syllables were presented once in quiet in a practice block. We calculated the percentage of correct responses for each syllable at each noise level. In addition, we examined confusion patterns on
error trials (77), to evaluate the percentage of transmitted information for the place, manner, and voicing of articulation.

Image acquisition and processing

Neuroimaging data were acquired with a 3.0 Tesla Siemens Prisma MRI scanner. T1-weighted (T1w) three-dimensional sagittal Magnetization Prepared Rapid Acquisition Gradient Echo (MPRAGE) images were acquired with the following parameters: TR = 2300 ms, TE = 2.98 ms, TI = 900 ms, flip angle = 9°, field of view (FOV) = 256 × 240 × 160 mm³, spatial resolution = 1 × 1 × 1 mm³, parallel imaging acceleration factor (iPAT) = 2.

T1w images were preprocessed using the FreeSurfer toolbox (version 6.0.0) for cortical reconstruction and volumetric segmentation. Once surfaces were reconstructed, an array of anatomical measures, including cortical thickness, surface area, and local gyrification index (LGI, 75), were then automatically calculated at each vertex of the cortex. The LGI was investigated in the current study, as a metric that quantifies the amount of cortex buried within the sulcal folds as compared with the amount of cortex on the outer visible cortex. A large gyrification index indicates a cortex with extensive folding and a small gyrification index indicates a cortex with limited folding. The vertex-wise maps of individuals were aligned to the FreeSurfer fsaverage surface-based template and smoothed using a 5 mm FWHM Gaussian kernel (78) for group analysis (see detailed image processing in the Supplementary Methods).

Behavioral analyses of auditory and language tasks

First, we tested for group-level differences between DD and TD groups’ performance on the ART and SiN tasks using Welch’s two-sample t-test, which accounts for unequal sample sizes between groups.

Pairwise correlations between reading measures and auditory tasks reflect Pearson’s correlations. As auditory processing abilities were not only correlated with reading scores but also with a range of other tasks, we used hierarchical regression analyses to evaluate the unique contribution of auditory processing abilities to reading. Models included reading scores as the dependent variable; age and sex, matrix reasoning, digit span forward, digit span backward, rapid picture naming, and ROWPVT vocabulary as covariates of no interest. Finally, phonological awareness and auditory processing scores were included in the model as independent variables of interest. The full model thus was: reading ~ age + sex + matrix reasoning + digit span backward + digit span forward + rapid picture naming + ROWPVT vocabulary (Step1) + phonological awareness (Step 2) + auditory ability (Step 3). This analysis was performed separately for each of the auditory processing tasks. Finally, to test whether auditory processing contributes to reading directly or via an effect on phonological awareness, we performed a mediation analysis using the mediation (79) package.
in R (v4.0.5). Notably, hierarchical regression and mediation analyses were conducted for each reading measure, but we only reported the results for the most significant reading measure in the main text. Due to the small sample size and subset of behavioral data available in the TD group, we performed the above analyses within the DD group.

Age and sex were controlled for in all analyses. Matrix reasoning was additionally controlled for in the group comparisons between DD and TD to rule out that the observed effects were driven by trend-level differences in this task. If not specified, all \( p \)-values are two-tailed with a threshold of 0.05.

**Brain-behavioral correlations**

We tested whether performance on auditory processing tasks was correlated with local cortical gyration (LGI). This analysis was performed at the whole-brain level for the DD group and for the whole cohort. Age, sex, and total brain volume were included as covariates of no interest in all analyses. All surface-based results were thresholded with a cluster-forming threshold of \( p \)-value < 0.005 at a cluster level of \( p \)-value < 0.05 corrected for multiple comparisons based on random field theory (80). All brain-behavior correlations were performed using the *surfstat* toolbox implemented in MATLAB.

**Dyslexia phenotyping and subgrouping**

Because of the evident individual variability in auditory processing deficits reported in previous studies (2, 14, 62), we tested the differences within the DD cohort by formulating two subgroups according to the performance in each auditory task, i.e., one high (> 75\(^{\text{th}}\) percentile) and one low-performing (< 25\(^{\text{th}}\) percentile) DD subgroup. Subgroup differences were tested using Welch’s two-sample \( t \)-test and controlled for age and sex. Finally, to test the differences in cognitive abilities associated with ART and SiN tasks, we compared the correlations across auditory tasks using the packages *psych* in R (v4.0.5).

**Acknowledgments**

This work was supported by UCSF Dyslexia Center; the Schwab Dyslexia and Cognitive Diversity Center Innovation Fund; and NIH grants (R01-DC012379, NIDCD K24 DC015544).
References


