7-2020

Sound Discrimination and Explicit Mapping of Sounds to Meanings in Preschoolers with and Without Developmental Language Disorder

Carolyn Quam
Portland State University, cquam@pdx.edu

Holly Cardinal
University of Arizona

Celeste Gallegos
University of Arizona

Todd Bodner
Portland State University, tbodner@pdx.edu

Citation Details

Sound Discrimination and Explicit Mapping of Sounds to Meanings in

Preschoolers with and Without Developmental Language Disorder

Carolyn Quam[1]

Department of Speech and Hearing Sciences, Portland State University, USA

Departments of Speech, Language, and Hearing Sciences and Psychology,
University of Arizona, USA

Holly Cardinal

Department of Speech, Language, and Hearing Sciences, University of Arizona,
USA

Celeste Gallegos

Department of Speech, Language, and Hearing Sciences, University of Arizona,
USA

and

Todd Bodner

Department of Psychology, Portland State University, USA

[1] Corresponding author. Portland State University Speech and Hearing Sciences, PO Box 751, Portland, OR 97207-0751, USA. 1-503-725-3558. cquam@pdx.edu.

Structured Abstract

Purpose: To investigate links between sound discrimination and explicit sound-meaning mapping by preschoolers with and without developmental language disorder (DLD).

Method: We tested 26 children with DLD and 26 age- and gender-matched peers with typical language development (TLD). Inclusion was determined via results of standardized assessments of language and cognitive skills and a hearing screening. Children completed two computerized tasks designed to assess pitch and duration discrimination and explicit mapping of pitch- and duration-contrasting sounds to objects.

Result: Children with TLD more successfully mapped pitch categories to meanings than children with DLD. Children with TLD also showed significantly better overall sound discrimination than children with DLD. Sound-discrimination scores were significantly associated with sound-meaning mapping, and this relationship did not differ across groups or acoustic cues. Thus, significant sound-discrimination differences between the groups likely contributed to differences in sound-meaning-mapping accuracy.

Conclusion: Children with DLD had more difficulty mapping sound categories to meanings than TLD peers. We discuss possible explanations for this finding, focusing on the link between sound discrimination and mapping sounds to meaning, as well as implications for theoretical accounts of the etiology of DLD. *(185 words in abstract)*

The present study investigates explicit learning of sound-meaning mappings by preschoolers with and without developmental language disorder (DLD). DLD is a highly prevalent developmental disorder. The prevalence in children of a subtype of DLD (Bishop et al., 2017), specific language impairment (SLI), has been estimated at over 7% (Tomblin et al., 1997). DLD is primarily characterized by delays in learning the grammatical rules of language, but auditory deficits in DLD have also emerged in a variety of tasks. This study tested two main research questions. The first was whether children with DLD, relative to peers with typical language development (TLD), would have difficulty linking speech-sound categories to meanings in an explicit-learning task. The second was whether links would emerge between speech-perception difficulties in DLD and mapping of sounds to meanings.

Two influential accounts of the etiology of DLD are relevant to the present study. The first is the Procedural Deficit Hypothesis (PDH; Ullman & Pierpoint, 2005). The PDH argues that the core deficit in DLD is the procedural-memory system, which supports and underlies a form of implicit learning (Hedenius et al., 2011). The declarative-memory system, which supports and underlies explicit learning, is argued to be relatively spared. Thus, the PDH might predict that children with DLD would not exhibit difficulty learning sound-meaning mappings in the explicit task used here.

The second relevant account of the etiology of DLD, the auditory-deficit account, argues that auditory-processing impairments are central to characterizing the disorder (Tallal et al., 1996; Wright et al., 1997). This account is informed by evidence of auditory deficits in DLD. Neural studies have revealed brainstem impairments for tracking rapid frequency changes (e.g., Basu, Krishnan, & Weber-Fox, 2010). Children with DLD show impairments in discrimination (Elliott & Hammer, 1988) and identification (Schwartz, Scheffler, & Lopez, 2013) of speech sounds, and

in talker discrimination (Dailey, Plante, & Vance, 2013). The auditory-deficit account would predict that children with DLD in the present study *would* exhibit difficulty with explicit sound-meaning mapping due to the auditory components of the task, and that speech-perception impairments would impact sound-meaning-mapping accuracy.

It is a contentious question to what degree auditory-processing difficulties cascade to contribute to children's language impairments. Tuomainen, Stuart, and van der Lely (2015) tested cue weighting and sound identification in adolescents with DLD. While they found deficits in language-impaired adolescents in both cue weights and sound identification, they found no significant correlations between performance on these measures and scores on standardized and non-standardized tests of receptive and expressive vocabulary and grammar. However, no systematic attempts were made to ensure that the assessments of language skills relied directly on the speech-perception abilities being probed. Testing 6- to 8-year-olds with DLD, Evans, Viele, Kass, and Tang (2002) found impairments to sound discrimination. However, when they probed for links between sound discrimination and use of grammatical morphology, they also failed to find significant correlations between speech-sound processing and grammatical skills, even when discrimination of fricatives was correlated with use of fricatives, and non-fricatives with non-fricatives. The fact that auditory difficulties have not been convincingly correlated with the severity of productive language difficulties has led to arguments that auditory difficulties are not a primary contributor to the disorder (e.g., Leonard, Eyer, Bedore, & Grela, 1997).

The current study design may be more sensitive to links between speech perception and mapping of sounds to meanings than these prior studies, for two reasons. First, we tested younger children than in similar previous studies. Testing preschoolers, whose speech-perception abilities (e.g., Creel & Jimenez, 2012; Nittrouer, 1996), and language skills (e.g., Beyer & Hudson Kam,

2009) are both still significantly developing, could increase sensitivity to links between speech-perception deficits and sound-meaning mapping. Second, as in many prior studies, Tuomainen et al. and Evans et al. both evaluated knowledge of English. Differences between groups (e.g., children with vs. without DLD) in static, existing knowledge of English reflect the accumulation of many contributing factors over the course of years. By contrast, our study links sound discrimination to learning of sound-object mappings within the same experimental session, for sound dimensions that are not used contrastively in English. Zeroing in on learning capacity, instead of existing knowledge, may boost sensitivity to links between speech perception and sound-meaning mapping.

We used a child-friendly task to tap explicit learning of sound-meaning mapping. To define sound categories, we chose two sound dimensions, pitch and duration, that were comparable in the following ways. Both are used contrastively in languages other than English (pitch in tone languages like Mandarin; duration in vowel-length contrasting languages like Dutch). Neither dimension is used contrastively in English, though they both play other roles, including serving as secondary cues to voicing.[2] Assigning different pitch or duration values to different meanings (two toys) therefore requires learning to attend to a dimension that is not typically used as a primary dimension for contrasting meanings in English. We were interested in whether children with DLD could identify and use the new dimension of contrast when it was taught explicitly.

Method

Participants

We included 52 4- and 5-year-olds in the study: 26 in the DLD group and 26 in the TLD

---

[2] Voiced sounds like /b/ tend to be lower in pitch than voiceless sounds like /p/; vowels preceding voiced codas tend to be longer than those preceding voiceless codas.

group. All children were native English speakers for whom English had been their primary language since birth. Most children were recruited and tested at public and private preschools and kindergartens in the greater Tucson area (one child was tested in Portland, Oregon[3]). However, two children were recruited via a participant database and tested in the laboratory and a private home, respectively. Five children were recruited and tested via a summer language intervention camp run in the university clinic. Despite efforts to recruit SES-balanced groups, maternal education was significantly higher in the TLD group than the DLD group (see **Table 1** for means and p-value). Individual children in the TLD group were gender- and age-matched (within 6 months) to individual children in the TLD group. Demographic and age information are provided in **Table 1**.

For inclusion, children had to: (1) pass a binaural hearing screening for 1000, 2000, and 4000 Hz tones at 25 dB;[4] and (2) attain a composite standard score of at least 75 on the Kaufman Assessment Battery for Children – Second Edition (KABC-II) non-verbal subtests (Kaufman & Kaufman, 2004) to rule out a cognitive deficit. The threshold of 75 on the KABC-II has been used extensively in work in this area using the term SLI (e.g., Dailey, Plante, & Vance, 2013). It is based on the DSM-5 threshold for intellectual disability, IQ of 70 + 1 SEM (Spaulding, Plante, & Vance, 2008). However, it should be noted that nonverbal IQ criteria have varied across studies, with other SLI studies using a more stringent nonverbal IQ criterion. A standard score below 87 on the Structured Photographic Expressive Language Test—Preschool: 2nd Edition (SPELT-P2; Dawson et al., 2005) was required for inclusion in the DLD group, and at or above 87 for the typical control group. A cutoff score of 87 has been previously demonstrated to provide the highest

---

[3] This child was in the TLD group. Procedures were identical and relied on the same internal laboratory protocols and assessment procedures.

[4] An Ear Scan 3 portable pure tone air conduction audiometer manufactured by Micro Audiometrics Corporation was used for hearing screenings.

sensitivity and specificity for children in the Tucson, Arizona area (Greenslade, Plante, & Vance, 2009).

Table 1: Demographic information and standardized test scores for preschoolers with typical language development (TLD) and developmental language disorder (DLD).

Variable	DLD (N = 26)			TLD (N = 26)			
Race (N)							
White	15			20			
African American	3			2			
Hawaiian/Pacific Islander	0			1			
American Indian/Alaska Native	0			1			
Asian	0			1			
Multiple Races	4			1			
Unspecified	4			0			
Ethnicity (N)							
Hispanic	10			10			
Non-Hispanic	14			15			
Unspecified	2			1			
Gender (N)							
M	19			19			
F	7			7			
	Mean	*SD*	*Range*	*Mean*	*SD*	*Range*	*T-test*
Age (years; months)	4;11	4 mo.	4;2-5;7	4;9	4 mo.	4;0-5;10	$t(50) = 1.40, p = .17$
Mother's education level (years)	14.38	2.04	11-18	16.00	1.90	12-18	$t(50) = 2.96, p = .005$
KABC-II[a]	101.23	14.37	79-130	111.88	9.38	87-130	$t(50) = 3.17, p = .003$
PPVT-4[a]	96.92	12.98	72-118	113.81	12.58	86-132	$t(50) = 4.76, p < .001$
SPELT-P2[a]	73.27	10.05	41-86	112.54	8.16	100-132	$t(50) = 15.47, p < .001$

Note. KABC-II = Kaufman Assessment Battery for Children – Second Edition; PPVT-4 = Peabody Picture Vocabulary Test, Fourth Edition; SPELT-P2 = Structured Photographic Expressive Language Test-Preschool 2. [a] Standard scores with a mean of 100 and a standard deviation of 15. If a child's SPELT-P2 score fell below the cutoff of 87, we verified that this was still the case even when potentially articulation-based errors were counted as correct.

Children were not enrolled in the study if they had a previously suspected or diagnosed cognitive or sensory impairment or speech delays preventing collection of a valid expressive language score (see next paragraph). Children who were not previously diagnosed with attention

deficit (hyperactivity) disorder, but who exhibited some attentional difficulties or hyperactivity, were only excluded if they were unable to complete the tasks. Two children were excluded from the TLD group because the research speech-language pathologist was concerned that they did not have typical language development, based on language samples in combination with borderline SPELT-P2 scores. Two children were excluded from the TLD group because the PsychoPy program froze or crashed and experimental data did not save.

Scores on the KABC-II, SPELT-P2 and Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007) are reported in **Table 1**. Due to time constraints, the Goldman-Fristoe Test of Articulation—2nd Edition (GFTA-2; Goldman & Fristoe, 2000) was only administered when teachers, school SLPs, or parents reported concerns or history of articulation difficulties, or study staff had concerns. Standard scores for the DLD group (N=17) ranged from 50-105. Standard scores for the TLD group (N=3) ranged from 75-101. If a child's SPELT-P2 score fell below the cutoff of 87, we verified that this was still the case even when potentially articulation-based errors were counted as correct. The groups significantly differed on the SPELT-P2, as expected, but also on the PPVT-4 and KABC-II. While DLD primarily impacts grammatical skills, vocabulary and non-verbal IQ differences are also often reported (e.g., Dailey, Plante, & Vance, 2013).

Stimuli

Sounds were isolated vowels synthesized using the KlattGrid speech synthesizer (Klatt & Klatt, 1990; Weenink, 2009) in the Praat program (Boersma & Weenink, 2008). Speech synthesis enabled tight control over acoustic parameters, which was important for structuring sound categories. Sounds averaged 0.6 seconds in duration and 70 decibels (dB). A final pitch decrement to 75% of the original pitch value made stimuli sound less robotic. The last 10 milliseconds of

each soundfile were amplitude-ramped in Matlab so that the soundfile did not end abruptly.

Sound categories were defined by either pitch or duration variation. Two categories were defined along each dimension: low pitch vs. high pitch, or short duration vs. long duration. Each category contained only 3 distinct sounds, making 6 sounds total per dimension. **Figure 1** displays pitch and duration values for all stimuli. Pitch stimuli were spaced according to the Bark scale (which incorporates logarithmic compression to simulate high-frequency compression in the human auditory system; Zwicker, 1961). Pilot testing indicated that children were more sensitive to pitch than duration. We therefore compressed differences among duration stimuli within-category and increased the distance between categories, in an attempt to make the duration categories more differentiable. The second-formant frequency was set to mimic an /u/ vowel for the pitch-differentiated sounds (1254 Hertz, or 8.9 Barks) and an /i/ vowel for the duration-differentiated sounds (2988 Hertz, or 13.85 Barks). For all sounds, the first, third, fourth, and fifth formant frequencies were set to 448, 2722, 4019, and 4898 Hertz, respectively (corresponding to 4.14, 13.3, 15.6, and 16.8 Barks).
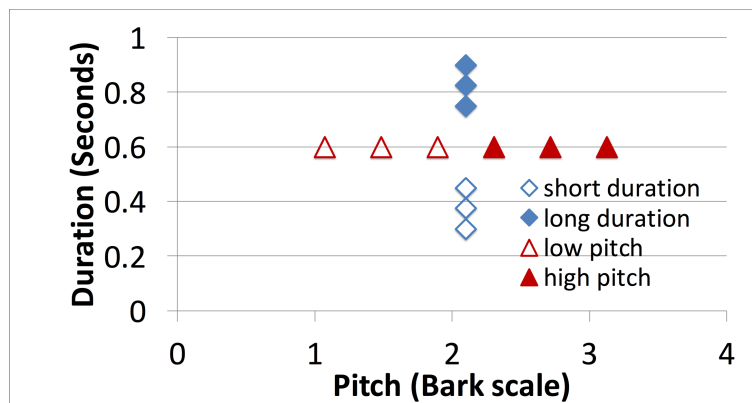

Figure 1: Duration (in seconds) and pitch (in Barks) for the 12 stimuli used across experiments. Two sound dimensions were used: pitch and duration. Two categories were defined along each dimension and each category contained 3 sounds (for pitch, 3 low-pitch sounds and 3 high-pitched sounds; for duration, 3 short sounds and 3 long sounds).

Apparatus and Procedure

The experiments were created and administered via the PsychoPy program (Peirce, 2007) on a Mac Mini computer with an attached Dell monitor, Apple keyboard, Apple mouse, and KidzGear headphones. In each of two experimental sessions, children first completed a test of discrimination of either pitch- or duration-contrasting sounds and then completed a separate task designed to promote either explicit or implicit learning of mappings of the same set of sounds to objects. Only one of these learning tasks, the test of explicit mapping of sounds to objects, is reported here (the test of implicit mapping of sounds to objects will be reported elsewhere; Quam, Cardinal, & Gallegos, in prep.).

Sound-Discrimination Task

At the start of each of the two experimental sessions, children first completed a test of sound discrimination to determine children's baseline sensitivity to either pitch or duration. Roughly half of the children in each group (DLD or TLD) had also participated on a previous test date in an implicit-learning experiment, to be reported elsewhere (teaching whichever sound categories the particular child did *not* learn in the explicit task, either pitch or duration; Quam, Cardinal, & Gallegos, in prep.). All children were included in sound-discrimination analyses for both pitch and duration, whether the particular discrimination test preceded the explicit task or the implicit task not reported here.

Across 12 trials, children heard pairs of sounds sampled from a set of 6 sounds total that formed a continuum (differing in either pitch or duration). One of the two sounds in the pair was always one of the endpoint sounds from the continuum (see **Figure 1**). The other sound was either identical or differed by an acoustic distance of 1-5 steps on the continuum. There were two trials at each "distance" (i.e., identical or 1-5 steps apart). In each trial, children heard 2 sounds, played

1 second apart. They were instructed to listen to both sounds and then say whether they were the same or different. Six different trial orders were used, and children were assigned to a different trial order for pitch vs. duration.

In each of the 12 trials in each sound-discrimination task, children responded "same" or "different" depending on whether they noticed a difference between the two sounds. We then converted these responses to D' scores (see *Statistical Design*, below) as an index of children's sensitivity to differences between sounds. We were interested in whether the likelihood of responding "different" differed across groups of children (TLD vs. DLD) or across cues (pitch vs. duration). We were also interested in whether children's sensitivity to differences increased significantly as the acoustic distance between the two sounds increased (from 1 to 5 steps on the continuum).

Explicit Sound-Meaning-Mapping Task

Immediately after one of the sound-discrimination tasks (either pitch or duration), children participated in an explicit sound-meaning-mapping task, in which they learned to map the same 6 sounds from the discrimination task to objects. Half of children in each group (TLD vs. DLD) mapped pitch-differentiated sounds to objects. Half of children in each group instead mapped duration-differentiated sounds to objects. The order of completion of the discrimination tasks was counterbalanced across children, so that half of children completed the pitch-discrimination task in the first experimental session (half of whom then immediately completed the explicit pitch-mapping task, and the other half of whom completed explicit duration-mapping in the following experimental session), while the other half completed the duration-discrimination task first (half

of whom then immediately completed the explicit duration-mapping task, and the other half of whom completed explicit pitch-mapping in the following experimental session).

The explicit sound-meaning-mapping task was designed to encourage children to attend to sounds and reason explicitly about their links to objects. Several task parameters were manipulated to tap explicit learning. First, sounds were emphasized in the instructions. In the training, a monster, "Shelly," appeared in the center of the screen with toys on either side (see **Figure 2**). Children were told that Shelly "talks in a funny way" to ask for the toys she likes to play with, and she wants them to learn the sounds for her toys. Second, children made an explicit choice to categorize each sound, by pressing the left arrow key (labeled with a picture of the left object) or the right arrow key (labeled with a picture of the right object). Finally, children received explicit feedback: a smiley face if they selected the correct toy and a frowny face if they did not.

Figure 2: Visual stimuli used in the experiments. A monster, Shelly, asked for one of two toys in each trial.

Each toy was depicted using the same visual image throughout the experiment. In the training, repeating each toy twice, the experimenter said, "If Shelly asks for this toy [pointing to the screen], which button do you push?" Children were required to understand that the buttons matched the toys on the screen in order to proceed (though one child responded via pointing in both tasks).

In the main experiment, across two blocks of trials, the six sounds from the continuum were played four times each, for 24 trials in each block and 48 total trials. Children listened to each sound and then pressed the left or right arrow key to categorize it as referring to one toy or the other. Children's responses were compared to the correct answer for each sound, as the 3 low-pitched or short-duration sounds corresponded to one object, and the 3 high-pitched or long-duration sounds corresponded to the other object. Each child's proportion of correct categorization responses was calculated for the 1st and 2nd trial blocks. These accuracy scores were interpreted as evidence of children's success in mapping each category of sounds to the correct object.

At the end of the experiment, participants also completed a production task. They were asked, "What sound did Shelly make for her red/blue toy?" Sample sizes for children's productions were not large enough for conducting inferential statistical comparisons between groups and cues, but a description of the task is provided in **Supplemental Materials** and descriptive statistics are provided in **Table S1**.

Statistical Design

For discrimination data, "same" and "different" responses were first converted to D' scores. The D' sensitivity index was calculated as $z(H)-z(F)$, where H (hits) was the proportion of "different" responses in trials in which sounds differed (by 1-5 steps), and F (false alarms) was the proportion of "different" responses in trials in which sounds were identical. To conduct inferential statistics on both discrimination and sound-meaning-mapping data, we employed a multivariate approach rather than univariate analyses of variance (ANOVA). This is because in preliminary ANOVAs on discrimination data, we found that Mauchley's test of sphericity was violated (e.g., for the main effect of Distance on the Continuum, Mauchley's $W = 0.30$, $p < .001$). Thus, we

conducted factorial multivariate analyses of variance (MANOVAs) for repeated measures (and multivariate analyses of covariance, or MANCOVAs, when covariates were included). Wilks' lambda (Λ) was used as the test statistic due to the prevalence of its use in randomized trials (Bodner, 2018).

Inferential statistics were conducted on sound-discrimination scores by conducting a MANOVA comparing sensitivity across Groups (TLD vs. DLD), Cues (pitch vs. duration), and Distance on the Continuum (i.e., how distinct the two sounds were, ranging from 1-5, with 1 being less distinct). We also included the variable First Cue (pitch-first vs. duration-first), to check for effects of the fact that half of children completed the pitch-discrimination task first and half the duration-discrimination task.

Accuracy in the explicit mapping task was evaluated via a MANOVA that compared sound-meaning-mapping accuracy across Blocks (1st vs. 2nd half of the experiment), Groups, and Cues. We also included the variable First Task (explicit-first vs. implicit-first), to check for effects of the fact that roughly half of children had completed a test of implicit sound-meaning-mapping prior to the test of explicit sound-meaning mapping reported here. Finally, a MANCOVA evaluated whether discrimination scores were associated with later sound-meaning mapping for sounds differentiated on the same dimension. Group, Cue, Blocks, and First Task were again included as factors.

Using the MANOVA family enabled us to minimize Type I error inflation by minimizing the number of tests (over running several separate ANOVAs). For follow-up t-tests, we took several steps to minimize Type I error inflation. First, for sound discrimination, in follow-up analyses of effects of distance on the continuum, we minimized the number of comparisons by comparing only adjacent distances. Second, in general, we employed Bonferroni corrections for

multiple comparisons in follow-up tests of statistically significant multivariate effects with more than one degree of freedom in the numerator (e.g., if there were 3 or more levels of a factor in a main effect).

Results

Sound Discrimination   **Figures 3a** and **b** depict D' scores for sounds differing in pitch vs. duration, respectively, by a distance of 1-5 steps on each continuum. Means are reported in **Table 2**. A MANOVA evaluated the impacts on D' scores of the within-subjects predictors Cue (pitch vs. duration) and Distance on the Continuum (1-5) and the between-subjects predictors Group (TLD vs. DLD) and First Cue Tested (pitch-1st or duration-1st). The MANOVA revealed a significant main effect of Group ($F(1,48) = 4.27, p = .044$), indicating overall higher D' scores for children with TLD than children with DLD; a significant main effect of Cue (Wilks' $\Lambda = .816$, $F(1,48) = 10.82, p = .002$), indicating overall higher difference scores for pitch than duration, and a significant main effect of Distance (Wilks' $\Lambda = .442$, $F(4,45) = 14.20, p < .001$). As D' scores would be predicted to increase incrementally as sounds become more distinct (as Distance increases), we investigated the main effect of Distance by conducting planned comparisons (paired t-tests) between adjacent distances. T-tests were Bonferroni corrected to minimize Type I error inflation for multiple comparisons. These revealed that D' scores were significantly higher for Distance 2 than for Distance 1 ($t(51) = 5.38, p < .001$) and significantly higher for Distance 3 than for Distance 2 ($t(51) = 2.83, p = .007$), but Distance 4 did not differ from Distance 3 or Distance 5, reflecting the fact that discrimination scores had largely asymptoted by Distance 3.
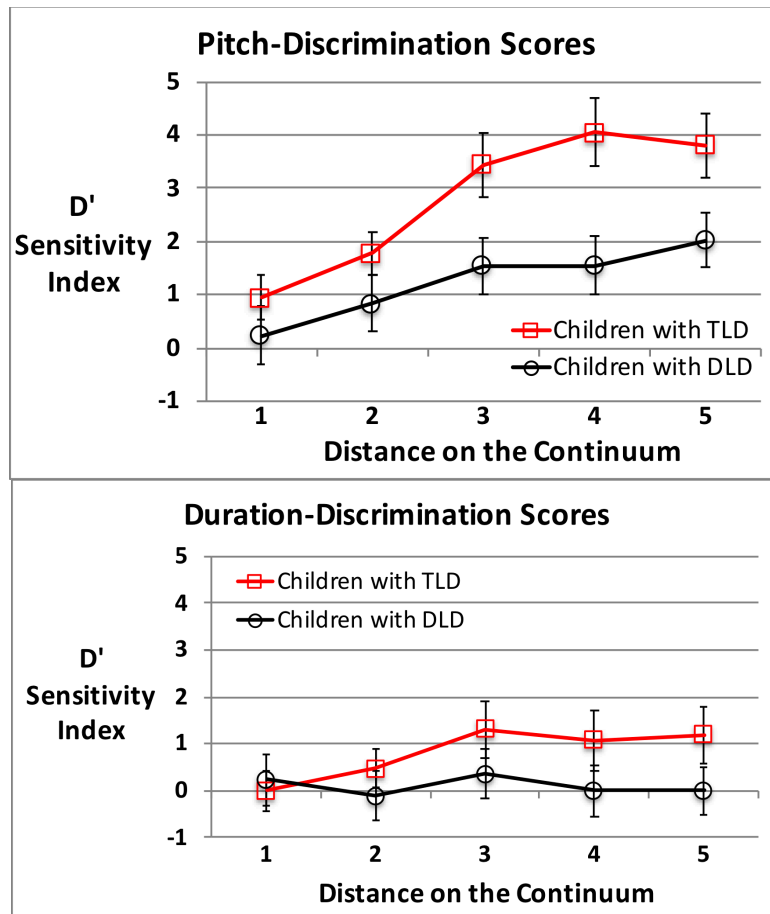
Figure 3: D' sensitivity index for discrimination of sounds differing by a distance of 1 to 5 steps on the pitch (top) and duration continua (bottom), for children with TLD (red squares) vs. DLD (black circles).

Table 3: Mean d' scores in sound-discrimination task (standard deviations in parentheses).

STEP	1	2	3	4	5
Overall	-1.34 (1.77)	0.58 (1.59)	1.33 (1.82)	1.14 (2.22)	1.45 (2.11)
Children With TLD	-1.93 (1.64)	0.84 (1.51)	2.06 (1.46)	1.74 (2.14)	2.11 (1.93)
Children With DLD	-0.76 (1.73)	0.33 (1.65)	0.60 (1.87)	0.53 (2.16)	0.78 (2.10)
Pitch	-1.44 (2.34)	0.96 (2.52)	2.11 (2.79)	2.05 (2.97)	2.24 (3.16)
Duration	-1.25 (2.22)	0.21 (1.93)	0.55 (2.69)	0.22 (3.14)	0.60 (3.02)

There was also a significant interaction of Distance and Group (Wilks' $\Lambda = .779$, $F_{(4,41)} = 3.19$, $p = .022$). To investigate the interaction, for each Group separately (TLD vs. DLD), we conducted planned comparisons (paired t-tests) between adjacent Distances (as for the

investigation of the main effect of Distance, above). Bonferroni corrections were conducted separately within each Group for the four Distance comparisons. Paired t-tests revealed that children with TLD showed patterns across Distances that matched the patterns for children overall: D' scores (see **Table 2** for means) were significantly higher for Distance 2 than for Distance 1 (t(25) = 5.38, p < .001) and significantly higher for Distance 3 than for Distance 2 (t(25) = 2.83, p < .001), but Distance 4 did not differ from Distance 3 or Distance 5. Children with DLD showed numerically similar differences between distances (in particular, a numerically large difference between Distance 1 and Distance 2 (t(25) = 2.22, p = .035), but none reached the threshold for statistical significance after Bonferroni correction (p < .0125). Thus, children with DLD's sensitivity to sound distinctions numerically increased as differences between sounds increased, but this pattern did not reach statistical significance.

Finally, there was a significant interaction of Cue and First Cue (Wilks' Λ = .904, F(1,48) = 5.09, p = .029), reflecting the fact that children overall showed stronger pitch discrimination when the pitch-discrimination task was at the start of the first experimental session (M= 1.39, SD= 1.76) than when it followed a previous test date consisting of the duration-discrimination task and a duration-to-meaning mapping task (M= 0.98, SD= 1.75). Likewise, they showed stronger duration discrimination when the duration-discrimination task was at the start of the first experimental session (M= 0.59, SD= 1.44) than when it followed the pitch tasks (M= -0.46, SD= 1.77).

To rule out alternative explanations for the results, such as the significant difference between groups in maternal education, we conducted additional analyses including the variables Maternal Education, Age, and Gender. The main effect of Group, the Distance by Group interaction, and the Cue by First Cue interaction were not meaningfully affected by the inclusion

of these other variables (F-values were always above 2.5 and p-values ranged from .009-.064). However, the main effect of Cue and the main effect of Distance became non-significant in the models that included Age or Maternal Education (Fs < 1, *p's* > .5). In the model that included Gender, there was a significant Group by Gender interaction (F(1,44) = 4.32, *p* = .044), indicating that in the TLD group, girls (*M*= 1.45, *SD*= 1.37) had numerically stronger discrimination scores than boys (*M*= 0.79, *SD*= 0.88; t(24) = 1.79, *p* = .086), whereas in the DLD group, boys (*M*= 0.54, *SD*= 1.12) had numerically stronger scores than girls (*M*= -0.36, *SD*= 1.18; t(24) = 1.46, *p* = .156).[5]

We also addressed whether individual differences in speech-production accuracy, as measured by the GFTA-2, might predict sound-discrimination scores for children with DLD. Among children with DLD in our sample, 17 of 26 children had GFTA-2 scores calculated, which ranged from standard scores of 50 to 105. In Pearson's correlation tests, GFTA-2 scores did not correlate significantly with pitch D' scores or with duration D' scores. However, there was a non-significant tendency for higher pitch D' scores for children with GFTA-2 scores above the 16th percentile (*M* = 0.64, *SD* = 1.15) than for children with GFTA-2 scores at or below the 16th percentile (*M* = 0.56, *SD* = 1.96; see Zuk, Iuzzini-Seigel, Cabbage, Green, & Hogan, 2018, for use of a similar threshold). Thus, it is possible that pitch discrimination might show relationships with

---

[5]In the model including Gender, the Cue by First Cue interaction entered into a 3-way interaction with Distance (Wilks' Λ = .753, F(4,41) = 3.36, *p* = .018). We investigated the interaction by first splitting the data into 4 subsets by Cue and First Cue (1: pitch discrimination when pitch was the 1st cue; 2: pitch discrimination as 2nd cue; 3: duration discrimination as 1st cue; 4: duration discrimination as 2nd cue). For each of these 4 subsets in turn, we then conducted pair-wise comparisons of the 5 distances, as for the main effect of Distance. For pitch as the 1st cue, patterns matched the patterns across children and cues overall: D' scores were significantly higher for Distance 2 than for Distance 1 (t(25) = 2.98, *p* = .006) and significantly higher for Distance 3 than for Distance 2 (t(25) = 3.21, *p* = .004), but Distance 4 did not differ from Distance 3 or Distance 5. For pitch as the 2nd cue, Distance 2 was significantly higher than Distance 1 (t(25) = 3.94, *p* = .001), but no other distances differed. The patterns for duration were analogous to those for pitch. For duration as the 1st cue, patterns matched the overall patterns, with Distance 2 exceeding Distance 1 (t(25) = 2.45, *p* = .022) and Distance 3 exceeding Distance 2 (t(25) = 2.5, *p* = .019). For duration as the 2nd cue, Distance 2 was significantly higher than Distance 1 (t(25) = 2.10, *p* = .046), but no other distances differed.

the presence or absence of co-occurring speech delay in a larger sample.

It was also important to address the possibility that non-verbal cognitive abilities could account for some of the variance in sound-discrimination and sound-meaning-mapping scores, given the significant group difference in KABC-II scores reported in **Table 1**. In Pearson's correlation tests, we found that across all 52 children, KABC-II standard scores were significantly correlated with pitch D' average (r = .397, p = .004), but not with duration D' average or overall D' average. The significant correlation with pitch discrimination warranted conducting an additional MANCOVA including the same predictors as before, but with KABC-II scores included as a covariate. KABC-II scores were not allowed to interact with other factors, as we were interested specifically in accounting for direct effects of KABC-II scores on sound discrimination. There was no significant main effect of KABC-II scores and results did not meaningfully change when KABC-II scores were included in the model.

Another possible predictor is the type and intensity of intervention children had received prior to participation. We did not collect detailed intervention histories, but 4 children with DLD were recruited at an intensive language-intervention preschool. Exploratory examination of mean scores suggests that these 4 children had more robust pitch D' scores than the other 22 children (Intervention-preschool group M = 1.57; all other children's M = 0.70).

Explicit Sound-Meaning Mapping   We compared the sound-meaning-mapping accuracy of children with TLD vs. DLD. A MANOVA included the within-subject factor Trial Block (1st half vs. 2nd half of the experiment) and the between-subjects predictors Group (TLD vs. DLD), Cue (pitch vs. duration), and First Task (explicit-1st vs. implicit-1st—i.e., the implicit sound-meaning-mapping task not reported here). The MANCOVA revealed no significant main effects, but a significant interaction of Group and Cue (F(1,44) = 4.33, p = .043), reflecting that children with

TLD learned pitch categories (*M* = 65.9%, *SD* = 20.8%) significantly better than children with

DLD (*M* = 47.8%, *SD* = 17.1%; two-tailed t(24) = 2.43, *p* = .023), but for duration categories,

children with TLD's accuracy (*M* = 50.6%, *SD* = 19.7%) was no different from children with DLD

(*M* = 53.8%, *SD* = 10.1%; t < 1, *p* > .5). There were no significant effects or interactions with Trial

Block or First Task. Thus, typically developing children outperformed children with DLD only for

pitch.

These results held when we controlled for additional variables including age and maternal

education.[6] We also considered effects of speech-production accuracy, non-verbal cognitive

abilities, and intervention history. For children with DLD, individual differences in speech-

production accuracy, as measured by the GFTA-2, did not correlate in Pearson's correlation tests

with sound-meaning-mapping accuracy overall or for pitch or duration separately. For children

overall, KABC-II scores were significantly correlated with pitch-meaning mapping (r = .401, *p* =

.042) but not with duration-meaning mapping or overall sound-meaning mapping. In a

MANCOVA with KABC-II scores included as a covariate, there was no significant main effect of

KABC-II scores and the effects of the other predictors did not meaningfully change.

Finally, regarding effects of prior intervention history, an exploratory comparison

suggested that the 4 children recruited at the language-intervention preschool had more robust

overall sound-meaning-mapping scores than the other 22 children (intervention-preschool group

M = 61%; all other children's *M* = 50%).

Linking Sound Discrimination and Mapping of Sounds to Objects   A   MANCOVA   evaluated

whether discrimination scores were associated with later sound-meaning mapping for sounds

---

[6] Additional analyses were conducted including the additional variables Age (in days), Maternal Education (in years), and Gender (M vs. F). There were no significant effects of (or interactions with) any of these additional variables, and the interaction of Group and Cue was not meaningfully affected by the inclusion of these other variables (F-values for the interaction were all above 4, and *p*-values ranged from .043-.05).

differentiated on the same dimension. Group, Cue, Trial Block (1st half of mapping task, 2nd half

of mapping task), and First Task (explicit-1st vs. implicit-1st—i.e., the implicit mapping task not

reported here) were again included as categorical predictors, with D' Score included as a

continuous predictor. To include discrimination scores as a predictor, we computed an average D'

Score across all 5 distances.[7] The model revealed a significant main effect of D' Score ($F(1,43) =$

16.97, $p < .001$), indicating a positive association between discrimination scores and mapping

accuracy. It also revealed a significant interaction of D' Score with Trial Block ($F(1,43) = 4.15$, p

$= .048$). To investigate the interaction, we examined the Parameter Estimates table (which reports

the underlying regression output from the MANCOVA). For Block 1, D' Scores were significantly

associated with mapping accuracy ($\beta = 0.037$, $t(43) = 2.75$, $p = .009$). For Block 2, D' Scores were

also significantly associated with mapping accuracy ($\beta = 0.067$, $t(43) = 4.32$, $p < .001$). The two

associations differed significantly from one another in that the association was stronger in Block

2. There were no main effects or interactions with Group or with Cue, indicating the patterns held

for children overall and for both pitch and duration. **Figure 4** plots accuracy in Block 1 (left) and

accuracy in Block 2 (right) against D' scores, with best-fit lines.

---

[7] We also conducted the same analyses with D' Score at Distance 5 included as the continuous predictor (instead of the average of all 5 distances). Results were highly convergent with those reported in the main text. The MANCOVA revealed a significant main effect of D' Score ($F(1,43) = 31.08$, $p < .001$). It also revealed a significant interaction of D' Score with Trial Block ($F(1,43) = 13.47$, $p = .001$). The Parameters Estimate table indicated that For Block 1, D' Scores were significantly associated with mapping accuracy ($\beta = 0.037$, $t(43) = 2.75$, $p = .009$). For Block 2, D' Scores were also significantly associated with mapping accuracy ($\beta = 0.067$, $t(43) = 4.32$, $p < .001$). The two associations differed significantly from one another in that the association was stronger in Block 2. Unlike in the test reported in the main text, there was also a significant interaction between Trial Block and First Task, indicating that children whose first task was the Explicit task showed a non-significant tendency for accuracy to decrease from Block 1 (*M*= 60.0%, *SD*= 12.4%) to Block 2 (*M*= 55.4%, *SD*= 20.8%; t(19) = -1.43, *p* = .170). By contrast, children whose first task was the Implicit task showed a non-significant tendency for accuracy to *increase* from Block 1 (*M*= 50.5%, *SD*= 14.6%) to Block 2 (*M*= 54.0%, *SD*= 16.8%; t(31) = 1.38, *p* = .179).There were no main effects or interactions with Group or with Cue, indicating the patterns held for children overall and for both pitch and duration.
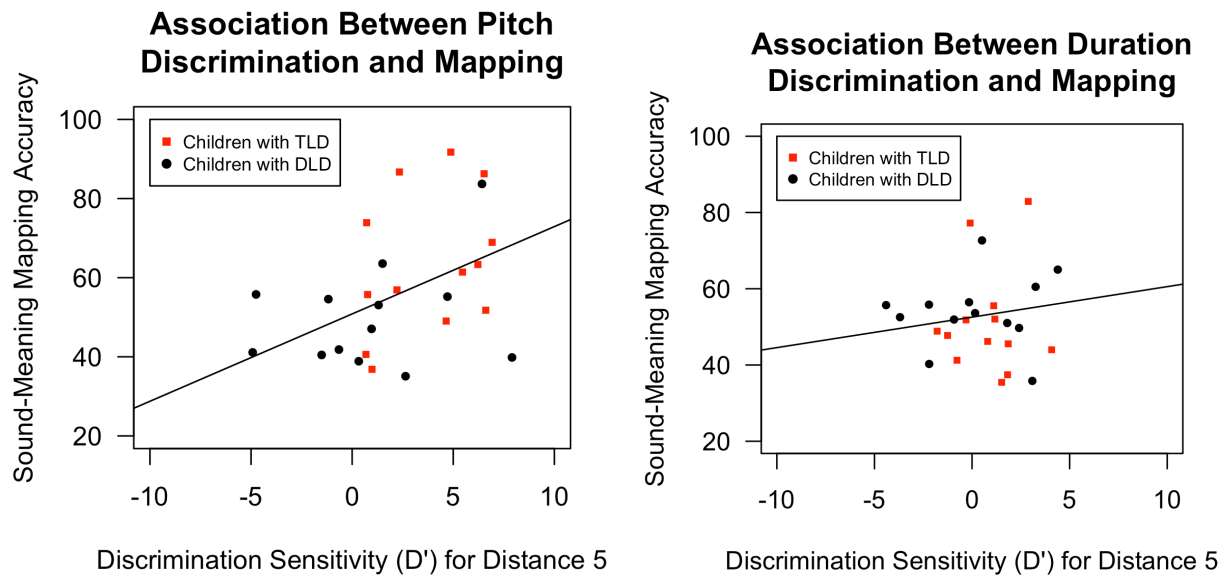
Figure 4: Sound discrimination was significantly associated with sound-meaning mapping in both Block 1 (left) and Block 2 (right). The association was significantly stronger in Block 2. The association did not significantly interact with group membership or cue. It held overall across children with both TLD (red squares) and DLD (black circles) and across both pitch (filled shapes) and duration (unfilled shapes).

Discussion

This study investigated two primary questions. First, it asked whether children with DLD would have more difficulty linking speech-sound categories to meanings in an explicit-learning task than TLD peers. Second, it asked whether links would emerge between sound discrimination and sound-meaning mapping. The results provided affirmative answers to both questions. Regarding the first question, children with TLD explicitly mapped pitch categories to meanings significantly more accurately than children with DLD. Children with TLD's advantage mapping sounds to objects emerged quickly in the task and was robust even in the first half of the experiment. This result is compatible with auditory-deficit accounts of DLD (Wright et al., 1997, Tallal et al., 1996). It does not seem to be compatible with the Procedural Deficit Hypothesis (PDH; Ullman & Pierpoint, 2005), which would predict impairments in implicit-learning tasks but

not in explicit-learning tasks. However, recent evaluations of the PDH also report declarative-memory impairments in language-impaired learners (e.g., Lum & Conti-Ramsden, 2013). In addition, a procedural-learning deficit could interact with auditory-processing or other impairments depending on the demands of the particular task.[8] By examining explicit mapping of sounds to meanings, this study represents the first step in a line of research that will contrast explicit vs. implicit learning of both auditory and visual categories, in order to examine interactions between general learning deficits and auditory-processing deficits in DLD.

Regarding the second question, across both groups and both cues, discrimination sensitivity was significantly associated with sound-meaning-mapping accuracy. Children with TLD showed stronger overall sound discrimination than children with DLD. Taken together, these two results suggest that the difficulty children with DLD had learning pitch categories explicitly was driven by their significantly weaker sound-discrimination scores relative to their typically developing peers. Pitch-discrimination scores were also numerically higher for children with DLD who had stronger speech-production skills than for those who had weaker speech-production skills, providing tentative evidence that sound discrimination may be particularly impacted in children with DLD who have a co-occurring speech-sound disorder.

The reason we predicted that sound-discrimination might predict sound-meaning mapping—despite differences in the processes underlying the two tasks—was that achieving high accuracy in the sound-meaning-mapping task requires that children discriminate small acoustic differences between sounds that straddle the category boundary. Still, it is noteworthy

---

[8] One aspect of the explicit-learning task used here that may have been challenging is the need to integrate visual feedback (smiley/frowny faces) to self-correct responses. Integrating feedback requires processing it visually, storing it in working memory, and using it to self-correct hypotheses about sound-object mappings. The memory component of integrating feedback could pose difficulty for children with DLD, given evidence of working-memory impairments (e.g., Leonard et al., 2007; note, however, that most findings of impaired working memory are in the verbal domain, not the visual domain; see Mainela-Arnold, Evans, & Coady, 2010, for discussion).

that sound-discrimination skills for children overall were associated with sound-meaning-mapping accuracy, considering that these two tasks rely on somewhat distinct sets of cognitive and linguistic processes. While the same inventory of sounds was used in both tasks, the nature of the tasks was different by design. Our goal was to look for links between two tasks that were not identical: a task akin to speech-sound processing (sound discrimination) and a task akin to word learning (mapping sounds to meanings). Both tasks involve processing sounds and attending to their similarities and differences. Both tasks also rely on auditory memory, as the listener must hold each sound in working memory in order to make a judgment about it. However, working-memory demands are likely greater in the sound-discrimination task, where children have to remember the first sound as they are hearing the second sound, in order to compare the two. Another difference is that in our discrimination task, any detectable distinction between sounds might lead a listener to judge them as different. The sound-meaning-mapping task, by contrast, requires categorization. The presence of only two categories means that learners must respond the same way to some sounds even if they can distinguish them.

The link we found between discrimination and sound-meaning mapping makes a novel contribution to the literature, considering that multiple prior studies have not found associations between speech-sound processing and language-learning outcomes in DLD (Tuomainen, Stuart, & van der Lely, 2015; Evans, Viele, Kass, & Tang, 2002). The absence of demonstrated links between speech-sound processing and expressive vocabulary and grammar has led to arguments that auditory difficulties are not a primary contributor to expressive language impairments in DLD (e.g., Leonard, Eyer, Bedore, & Grela, 1997), despite abundant evidence of auditory-processing impairments in DLD (e.g., Wright et al., 1997; Dailey, Plante, & Vance, 2013; Tallal et al., 1996; Basu, Krishnan, & Weber-Fox, 2010; Elliott & Hammer, 1988; Schwartz, Scheffler, & Lopez,

2013). Another limitation of prior work is the existence of many studies of DLD that only minimally consider speech skills and, conversely, many studies of speech-sound disorders (SSD) that only minimally consider language skills. Our results call for future work to better differentiate impacts of DLD, SSD, and DLD with co-occurring SSD on both speech and language processing.

The present study cannot speak directly to links between sound discrimination and learning of words or grammar, as children in our tasks were mapping synthesized, isolated vowel sounds to objects. However, it does provide preliminary evidence that sound-discrimination impairments in DLD may cascade to impact later learning. The scope of this cascade—in particular, whether it extends to learning of morphological and syntactic categories—should be investigated in future work.

While our favored account of the results is that group differences in sound-meaning-mapping accuracy can be attributed to group differences in speech-sound-processing abilities, we also considered several other possible contributors to group and individual differences in performance. Examination of possible contributions of individual variation in speech-production accuracy (as measured on the GFTA-2) and non-verbal cognitive abilities (as measured on the KABC-II) did not reveal strong contributions of these individual-difference measures to performance in our tasks. However, effects of the GFTA-2 could have been tempered by limited statistical power, as children with higher GFTA-2 scores showed numerically higher pitch-discrimination performance than children with lower GFTA-2 scores.

Another potential contributor to group differences in our tasks is the significant difference between groups in maternal education. As reported in **Table 1**, mothers of children with TLD had on average 1.6 years more education than mothers of children with DLD. MANCOVAs including maternal education revealed no significant effects or interactions of the factor, and its inclusion

did not meaningfully impact group differences in sound discrimination or sound-meaning mapping. Nevertheless, the existence of group differences in maternal education is important to consider in the context of evidence about impacts of socioeconomic background on children's language outcomes (e.g., Hart & Risley, 2003; Rowe, 2012; Hoff, 2013; Romeo et al., 2017; but see Dudley-Marling & Lucas, 2013). The role of maternal education in DLD/SLI is controversial. Rice (2019) recently argued that low maternal education is not a "disconfirming diagnosis," and does not cause SLI, but rather is linked in some way to SLI.

We also considered the possibility that intervention history could predict the performance of children with DLD in our tasks. While we did not collect detailed intervention histories, an exploratory analysis indicated that the 4 children who were recruited at an intensive language-intervention preschool had higher mean pitch-discrimination and overall sound-meaning-mapping scores than the other 22 children with DLD. While inferential statistics were not possible given the small sample size recruited at the language-intervention preschool, these mean differences provide tentative evidence that intensive intervention may have boosted children's scores on our tasks.

Future Directions

A productive avenue for future research would be to explore factors that would boost performance for both groups of children. One means of increasing sound-meaning-mapping accuracy might be to provide additional training with categorization of the sounds. Another might be to embed sounds in word forms, such as CVC (consonant-vowel-consonant) syllables, and in sentence frames. Discrimination of synthetic nonsense syllables like the ones used here has proven particularly challenging for children with DLD (Evans, Viele, Kass, & Tang, 2002), perhaps because they impose a greater processing load than more naturalistic stimuli (Coady,

Evans, Mainela-Arnold, & Kluender, 2007).

Relatedly, another future direction would be to explore ways to improve discrimination and mapping of durations for both groups of children. Duration (and pitch to perhaps a lesser degree) is typically interpreted relative to its context. Listeners judge vowels and syllables as long or short with respect to speech rate and the lengths of nearby segments and syllables. Durations may be particularly difficult to judge without context, a potential explanation for why children overall discriminated pitch more robustly than duration.

It is possible that children ignored duration distinctions in the present study because they are not used contrastively in English (Dietrich, Swingley, & Werker, 2007). However, this is also true of pitch (Quam & Swingley, 2010). Duration is also critical to speech perception in English. It is a secondary cue to the voicing of coda consonants, and six-year-old (Krause, 1982) and adult listeners (Peterson & Lehiste, 1960) can exploit duration differences in consonant perception. The distance between the centers of the duration categories in our task was 450 ms. Prior work (Krause, 1982, Fig. 1) indicates that 3-year-olds can distinguish duration differences of roughly 200 ms., and the average difference between vowel lengths preceding voiced vs. voiceless consonants in English is only 100 ms. (Peterson & Lehiste, 1960). Thus, future work using more naturalistic vowel sounds embedded in words and carrier phrases may find improved duration discrimination and mapping performance.[9]

In summary, the present investigation of sound discrimination and explicit sound-meaning mapping revealed that children with DLD mapped pitch-differentiated sounds to meanings less successfully than their typically developing peers. Children with DLD also showed weaker sound discrimination. Sound discrimination was associated with sound-meaning-mapping accuracy for

---

[9] Future work using a similar discrimination task could also include contrasts in which both sounds are taken from the middle of the continuum, to promote generalizability.

children overall. Thus, impaired sound discrimination could at least partly account for weaker sound-meaning mapping in children with DLD. Thus, the results suggest potential contributions of auditory processing to language-learning difficulties in DLD.

Acknowledgements

Declaration of Interest Statement

All authors have no financial or non-financial conflicts of interest to report.

References

Basu, M., Krishnan, A., & Weber-Fox, C. (2010). Brainstem correlates of temporal processing in children with specific language impairment. *Developmental Science, 13,* 77-91.

Beyer, T., & Hudson Kam, C. L. (2009). Some cues are stronger than others: The (non)interpretation of 3rd person present -s as a tense marker by 6- and 7-year-olds. *First Language, 29,* 208-227.

Bishop, D.V.M., Snowling, M.J., Thompson, P.A., Greenhalgh, T., & the CATALISE-2 consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry, 58*, 1068–80.

Bodner, T.E. (2018). Estimating and testing for differential treatment effects on outcomes when the outcome variances differ. *Psychological Methods, 23,* 125–37.

Boersma, P., & Weenink, D. (2008). *Praat: doing phonetics by computer (Version 5.0. 30)[Computer program]. Retrieved from http://www.praat.org/.*

Coady, J. A., Evans, J. L., Mainela-Arnold, E., & Kluender, K. R. (2007). Children with specific language impairments perceive speech categorically when tokens are natural and meaningful. *Journal of Speech, Language, and Hearing Research, 50,* 41-57.

Creel, S.C., & Jimenez, S.R. (2012). Differences in talker recognition by preschoolers and adults. *Journal of Experimental Child Psychology, 113*, 487–509.

Dailey, N. S., Plante, E., & Vance, R. (2013). Talker discrimination in preschool children with and without specific language impairment. *Journal of Communication Disorders*, *46*, 330–337.

Dawson, J., Stout, C., Eyer, J., Tattersall, P., Fonkalsrud, J., & Croley, K. (2005). *Structured Photographic Expressive Language Test-Preschool 2*. DeKalb, Il.: Janelle Publications.

Dietrich, C., Swingley, D., & Werker, J. F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences of the USA, 104,* 16027–16031.

Dudley-Marling, C., & Lucas, K. (2009). Pathologizing the language and culture of poor children. *Language Arts, 86,* 362-370.

Dunn, D. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition*. Bloomington, MN: NCS Pearson, Inc.

Elliott, L. L., & Hammer, M. A. (1988). Longitudinal changes in auditory discrimination in normal children and children with language-learning problems. *Journal of Speech and Hearing Disorders, 53,* 467-474.

Evans, J. L., Viele, K., Kass, R. E., & Tang, F. (2002). Grammatical morphology and perception of synthetic and natural speech in children with specific language impairments. *Journal of Speech, Language, and Hearing Research, 45,* 494-504.

Goldman, D., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation-Second Edition*. Circle Pines, MN: American Guidance Services.

Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the Structured Photographic Expressive Language Test—Preschool: Second Edition. *Language, Speech, and Hearing Services in Schools*, *40*, 150–160.

Hart, B. & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Federation of Teachers*, *27,* 4–9.

Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Veríssimo, J., Dye, C. D., Alm, P., Jennische, M., Tomblin, J. B., & Ullman, M. T. (2011). Grammar predicts procedural learning and consolidation deficits in children with specific language impairment. *Research in Developmental Disabilities*, *32*, 2362–2375.

Hoff, E. (2013). Interpreting the early language trajectories of children from low SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology, 49,* 4-14.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children, Second Edition*. Circle Pines, MN: American Guidance Services.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America, 87,* 820-857.

Krause, S. E. (1982). Vowel duration as a perceptual cue to postvocalic consonant voicing in young children and adults. *The Journal of the Acoustical Society of America, 71,* 990.

Leonard, L. B., Eyer, J. A., Bedore, L. M., & Grela, B. G. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40,* 741-753.

Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50,* 408-428.

Lum, J. A. G., & Conti-Ramsden, G. (2013). Long term memory: A review and meta-analysis of studies of declarative and procedural memory in specific language impairment. *Topics in Language Disorders, 37,* 85-100.

Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2010). Explaining lexical-semantic deficits in specific language impairment: The role of phonological similarity, phonological working memory, and lexical competition. *Journal of Speech, Language, and Hearing Research, 53,* 1742-1756.

Nittrouer, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds. *Journal of Speech and Hearing Research, 39,* 278-297.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.

Peterson, G. E., and Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America, 32,* 673-703.

Quam, C., & Swingley, D. (2012). Development in children's interpretation of pitch cues to emotions. *Child Development, 83,* 236-250.

Quam, C., & Swingley, D. (2010). Phonological knowledge guides 2-year-olds' and adults' interpretation of salient pitch contours in word learning. *Journal of Memory and Language, 62,* 135-150.

Rice, M. (2019). What studies of twins tell us about Specific Language Impairment in children: Twinning effects & heritability at 2, 4, 6, and 16 years of age. *2019 Research Symposium at the ASHA Convention.*

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*, 700–710.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*, 1762–1774.

Schwartz, R. G., Scheffler, F. L. V., & Lopez, K. (2013). Speech perception and lexical effects in specific language impairment. *Clinical Linguistics & Phonetics*, *27*, 339–354.

Spaulding, T. J., Plante, E., & Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: Evidence for separate attentional capacities. *Journal of Speech, Language, and Hearing Research, 51,* 16-34.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W. M., & Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, *271*, 81–84.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*, 1245–1260.

Tuomainen, O., Stuart, N. J., & van der Lely, H. K. J. (2015). Phonetic categorization and cue weighting in adolescents with Specific Language Impairment (SLI). *Clinical Linguistics & Phonetics*, *29*, 557–572.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41,* 399-433.

Weenink, D. (2009). The KlattGrid speech synthesizer. *Interspeech*, *10*, 2059–2062.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature*, *387*, 176–178.

Zuk, J., Iuzzini-Seigel, J., Cabbage, K., Green, J. R., & Hogan, T. P. (2018). Poor speech perception is not a core deficit of childhood apraxia of speech: Preliminary findings. *Journal of Speech, Language, and Hearing Research, 61,* 583-592.

Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *The Journal of the Acoustical Society of America*, *33*, 248–248.