Distributional learning of novel visual object categories in children with and without developmental language disorder

Iris Broedelet
Paul Boersma
Judith Rispens
Amsterdam Center for Language and Communication,
University of Amsterdam, the Netherlands

Abstract: It has been proposed that a deficit in statistical learning contributes to problematic language acquisition in children with developmental language disorder (DLD), but at the same time the nature and extent of this relationship is not clear. This paper focuses on the role of statistical learning in lexical-semantic development by investigating visual distributional learning of novel object categories in children with and without DLD and its relation to vocabulary knowledge. Distributional learning is a form of statistical learning and entails the learning of categories based on the frequency distribution of variants in the environment. Fifty children (25 DLD, 25 TD) were tested on a visual distributional learning task. Results indicate that children can learn novel object categories on the basis of distributional information. We did not find evidence for a deficit in visual distributional learning in children with DLD. To investigate whether visual distributional learning ability is related to vocabulary knowledge, the children with DLD were tested on different measures of vocabulary. Phonological processing ability and non-verbal intelligence were taken into account as control variables. Multiple linear regression analyses did not reveal evidence for a relationship between distributional learning and vocabulary in DLD.

Keywords: developmental language disorder; statistical learning; distributional learning; lexical-semantic knowledge.

Corresponding author(s): Iris Broedelet, Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam. E-mail: I.R.L.Broedelet@uva.nl.

Citation: Broedelet, I., Boersma, P., & Rispens, J. (2023). Distributional learning of novel visual object categories in children with and without developmental language disorder. *Language Development Research*, *2*(1), 1-43, https://doi.org/10.34842/2023.0528.

Introduction

Most children acquire their native language(s) without many major difficulties, but this is different for children with developmental language disorder (henceforth: DLD). These children do not present major neurological deficits, hearing disabilities or low overall intelligence, nor is a lack of language input the underlying problem. DLD occurs in approximately 7% of school-aged children (Bishop, 2006), and the problems often last into adulthood. Social–emotional difficulties occur in this group as well: individuals with DLD have greater risk of depression disorders (Westby, 2019) and even have a lower quality of life compared to typically developing peers (Eadie et al., 2018).

Morphosyntactic impairments are viewed as a hallmark of DLD, while lexical abilities are often seen as a relative strength (e.g. Ullman & Pierpont, 2005). However, there is ample clinical evidence for a disadvantage in lexical skills as well (for reviews: Brackenbury & Pye, 2005; Nation, 2014). Recently, researchers have proposed that an impairment in statistical learning, a learning ability that is important for the discovery of patterns and sequences in sensory input (Siegelman et al., 2017), contributes to the language difficulties in children with DLD (Arciuli & Conway, 2018; Hsu & Bishop, 2010; Saffran, 2018). Experimental results suggest that a deficit in statistical learning (partly) explains lexical deficits (Evans et al., 2009; Mainela-Arnold & Evans, 2014), but the relationship between statistical learning and the development of lexical knowledge, especially lexical-semantic knowledge, requires more investigation.

Distributional learning, which plays a role in the categorization of sensory stimuli such as speech sounds (Maye et al., 2008; Maye et al., 2002) and novel visual objects (Junge et al., 2018) has never been investigated in children with DLD. Categorizing novel visual stimuli might be an important skill that is required when mapping new words to new objects. In our study we aim to investigate if this type of visual distributional learning is affected in children with DLD, and whether this ability relates to different types of lexical(-semantic) knowledge.

Background

Statistical learning deficit hypothesis

Although the main aspect of DLD is problematic language acquisition, children with DLD experience difficulties outside the linguistic domain as well. For example, there is evidence for deficits in motor skills (Sanjeevan & Mainela-Arnold, 2019), working memory (Montgomery et al., 2010), attention (Ebert & Kohnert, 2011) and processing visual information (Collisson et al., 2015). These findings have led to the idea that a deficit in a more general learning mechanism might be at the core of the disorder, as opposed to an impairment specific to linguistic representations (Arciuli & Conway, 2018; Hsu & Bishop, 2010).

Statistical learning is such a learning mechanism (Siegelman, 2020). Statistical learning underlies the extraction of regularities and patterns from sensory input and has been shown to correlate with or predict language ability in children and adults (Conway et al., 2010; Ellis et al., 2014; Hamrick et al., 2018; Kaufman et al., 2010; Kidd & Arciuli, 2016; Kidd, 2012; Misyak et al., 2010; Newman et al., 2006; Shafto et al., 2012; Spencer et al., 2015).

Results from several studies point towards a disadvantage in different types of statistical learning in individuals with DLD (Evans et al., 2009; Haebig et al., 2017; Hsu & Bishop, 2010; Hsu et al., 2014; Lammertink et al., 2019; Lian, 2017; Lukács & Kemény, 2014; Lum et al., 2014; Mainela-Arnold & Evans, 2014; Obeid et al., 2016; Plante et al., 2002; Tomblin et al., 2007); for a review see Saffran (2018). Please note that null results (Aguilar & Plante, 2014; Noonan, 2018) and even evidence of intact statistical learning in children with DLD (Lammertink et al., 2020) have also been reported. Importantly, several meta-analyses point to a statistical learning deficit in children with DLD (Lammertink et al., 2017; Lum & Conti-Ramsden, 2013; Obeid et al., 2016). Moreover, studies have suggested that statistical learning ability is related to different types of language skills in children with DLD: for example grammatical ability (Hedenius et al., 2011; Misyak et al., 2010; Tomblin et al., 2007) and lexical skills (Evans et al., 2009; Mainela-Arnold & Evans, 2014). Thus, accumulated evidence indicates that children with DLD are compromised in their statistical learning ability, which might (partly) explain their problematic language acquisition.

Lexical difficulties in children with DLD

Children with DLD may have difficulty with several aspects of language acquisition, such as vocabulary, morphology, syntax and phonology, and there is a large amount of heterogeneity within this population (Bishop, 2006; Leonard, 2014). Many studies have focused on morphosyntactic difficulties, for example a child saying she walk instead of she walks. However, these children also show evident difficulties in the development of lexical knowledge (Brackenbury & Pye, 2005; Nation, 2014). Research indicates that lexical difficulties impact social and academic development (Aguilar et al., 2017).

Studies suggest that children with DLD have a smaller vocabulary size and more shallow knowledge of words relative to TD children (McGregor et al., 2013). For example, they make semantic substitutions (confusing towel and blanket) and use more “allpurpose verbs" like go instead of more specific verbs like run, skip, sail, swim, etc. When naming objects, they are slower and make more phonological and semantic errors (Dockrell et al., 2001; Lahey & Edwards, 1999; Leonard et al., 1983; McGregor et al., 2002; McGregor, 1997). These errors reflect impoverished semantic representations. Dockrell et al. (2003) tested semantic knowledge of children with word-finding difficulties, and found that they provide less accurate definitions of objects and actions: their definitions often contained less information about the semantic category of an object, and more perceptual and redundant information compared to TD children. Moreover, compared to controls, children with DLD provide poor, incomplete definitions of common words (Mainela-Arnold et al., 2010; Marinellie & Johnson, 2002), and provide fewer semantic details in drawings (McGregor & Appel, 2002; McGregor et al., 2002).

On word association tasks, which are viewed as a measure of lexical-semantic organization, children with DLD produce fewer semantically related words than TD peers (Drljan & Vuković, 2019; McGregor et al., 2012; Sandgren et al., 2020; Sheng & McGregor, 2010). A less efficient lexical organization could have a negative effect on subsequent vocabulary development (Beckage et al., 2010). Finally, children with DLD also show difficulties on word learning tasks, both with learning phonological and semantic properties of words (Alt & Plante, 2006; Kan & Windsor, 2010; Nash & Donaldson, 2005) and fast mapping (Haebig et al., 2017; Kapa & Erikson, 2020).

Thus, children with DLD have lexical difficulties that go beyond word access, word retrieval and the phonological representations of words, pointing to suboptimal semantic representations. Little is known about the underlying cause of lexical-semantic deficits in children with DLD. Often put forward as a possible cause is poor phonological short-term memory, which is considered an important prerequisite for vocabulary acquisition (Melby-Lervåg et al., 2012). There is extensive evidence of deficits in phonological short-term memory and verbal working memory in children with DLD (for a review, see Montgomery et al., 2010). Phonological short-term memory is often measured using a non-word repetition (NWR) task. Studies show that performance on NWR tasks correlates with word-learning skills in TD children (Gathercole et al., 1997) and in children with DLD (Alt & Plante, 2006).

The causal direction of the relationship between phonological short-term memory and word learning is not clear. Difficulties with phonological processing might lead to poor phonological representations of words, which in turn may have a negative influence on the building of strong semantic representations. Indeed, NWR ability predicts vocabulary in young children between 4 and 5 years, but this relationship gets weaker in older children between 6 and 8 years (Gathercole et al., 1992; Gathercole, 2006). Furthermore, it has been found that vocabulary size is an important predictor of NWR ability, which could be explained as follows: as vocabulary size grows, phonological representations strengthen, which would improve non-word repetition ability (Metsala, 1999). Other studies fail to find evidence for a causal relationship between NWR ability and vocabulary. For example, Melby-Lervåg et al. (2012) carried out a large longitudinal study and did not find evidence for a causal relationship between NWR skills and vocabulary development in 4 to 7-year-old children. The authors also re-analyzed data from a similar longitudinal study (Gathercole et al., 1992), and failed to find the causal relationship that the authors of the original study had claimed. Finally, intervention studies have failed to find an effect of phonological memory-training on vocabulary knowledge (Melby-Lervåg et al., 2012: Dahlin et al., 2008, Schmiedek et al., 2010). Thus, although the difficulties in phonological processing in DLD are well-established, the role they play in vocabulary development remains unclear.

Statistical learning and the development of the lexicon

To summarize, a large body of studies points towards an important role for statistical learning in the acquisition of language. In children with DLD, the ability of extracting regularities from input seems to be affected, which could explain their language deficits. In this section we discuss the relationship between statistical learning and the development of the lexicon. Specifically, we look at the link between statistical learning and lexical-semantic knowledge.

Children with better statistical learning skills often have a larger vocabulary (Spencer et al., 2015), and Shafto et al. (2012) and Ellis et al. (2014) report a predictive relationship between TD infants' performance on a visual statistical learning task and their vocabulary size at a later point in time. In another longitudinal infant study, Singh et al. (2012) found that statistical learning ability in a word segmentation task at 7 months predicts productive vocabulary at 24 months.

Evidence also suggests a relationship between statistical learning and vocabulary in children with DLD. Evans et al. (2009) reported a correlation between statistical learning ability and vocabulary knowledge and claimed that lexical impairments might be explained by statistical learning difficulties. In another study, Mainela-Arnold and

Evans (2014) report a significant correlation between statistical learning ability on a word segmentation task and performance on a lexical-phonological access task. During this forward gating task, children heard increasingly longer parts of a word and had to guess which word they heard. On the other hand, no evidence was found for a relationship between statistical learning and performance on a word definition task. The authors suggest (from a comparison of their two p-values) that statistical learning underlies the acquisition of sequential lexical-phonological knowledge, but that lexical-semantic abilities might depend on other learning/memory systems.

The link between statistical learning and lexical-semantic knowledge requires further investigation. In the study of Mainela-Arnold and Evans (2014), the status of a potential relation cannot be concluded from comparing a null result with a statistically significant result. Moreover, they used a word definition task to measure lexical-semantic knowledge, which requires very explicit semantic knowledge. It could be the case that statistical learning is related to more implicit forms of semantic knowledge. Furthermore, statistical learning in this and many other studies was measured using a word segmentation task. It is not unexpected that this type of sequential statistical learning contributes to lexical-phonological knowledge due to the nature of the task. However, as Mainela-Arnold and Evans (2014) also state, it is possible that other types of (non-sequential) statistical learning that were not taken into account play a role in the building of a semantically rich lexicon.

Statistical learning mechanisms indeed seem to be sensitive to semantic information (see Paciorek & Williams (2015) for a review). For example, the mapping of newly learned words to their corresponding referents is suggested to be a gradual statistical learning process named cross-situational learning, which entails the (implicit) tracking of co-occurrences between words and their visual referents (Kachergis et al., 2014; Smith & Yu, 2008; Suanda et al., 2014; Yu & Smith, 2011). In another strand of research, Goujon (2011) showed that adults implicitly learn that the semantic categories of real-world scenes predict the position of the following target in a visual search task, indicating that semantic information is processed automatically and can be facilitated to make unrelated decisions. Similarly, Rogers et al. (2020) report that higher-order categories influence the learning of visual statistical regularities: people learn implicit mappings between visual stimuli better when the stimuli belonged to the same category rather than two different categories.

An important phenomenon in the development of the lexicon is shape bias. This entails the tendency for children to extend the use of newly learned object names to objects that share the same shape with the original object rather than the same color or size. The emergence of this shape bias might depend on statistical learning mechanisms: if children pick up the regularity that early learned object categories often share the same shape, they learn to consider shape as an important cue when learning new object labels. Results from a novel object name learning experiment of Collisson et al. (2015) indicate that 3-to-4-year-old children with DLD do not show shape bias to a similar extent as TD children. Moreover, children with DLD perform more poorly on a task that measures visual paired-associate learning, and this performance predicts the strength of their shape bias. This finding suggests that an impairment in visual statistical learning might underlie the lagging development of shape bias in these children, which in turn may hinder their lexical development.

Another process in the development of the lexicon that could be supported by statistical learning mechanisms is learning to categorize and name the enormous number of different objects in the visual world. For example, a child needs to learn which round fruits are called apples and which ones are called peaches. Studies point out that infants automatically track the co-occurrence of visual features of objects in visual statistical learning tasks (Wu et al., 2011, 2010). This ability of learning which object features co-occur and which do not, plays an important role in learning about visual categories (Palmeri & Gauthier, 2004). Similarly, Younger (1985) and Plunkett et al. (2008) showed that statistical learning may underlie semantic category learning, as infants learn object categories based on the co-occurrence of features.

Distributional learning

A specific type of non-sequential statistical learning, distributional learning, plays a role in the formation of new categories as well. Maye et al. (2002) showed that infants can pick up speech sound categories based on the frequency distribution of speech sound exemplars. Their infants were exposed to variants from the /ta/-/da/ continuum. The distribution of the variants was either bimodal or unimodal: in the bimodal condition there were two distributional peaks, reflecting two distinct sound categories /t/ and /d/, while in the unimodal condition there was only one peak reflecting one broad category. After familiarization it was tested whether the infants could discriminate the endpoint tokens of the continuum. Maye et al. found that only their participants in the bimodal condition had statistically significantly formed two distinct categories, as they were able to discriminate the two endpoint tokens, while infants in the unimodal condition did not reach significance. This result indicated to Maye et al. that infants can learn phonetic categories based on distributional information. Although Maye et al.'s claim was based on a p-value comparison (a direct comparison between the two groups gave a non-significant p-value of 0.063), together with later findings of distributional learning of sound categories (Escudero et al., 2011; Hayes-Harb, 2007; Maye et al., 2008; Vandermosten et al., 2019; Wanrooij et al., 2014) the results point towards a distributional learning mechanism underlying bottom-up categorization of speech sounds.

More recent studies have shown that distributional learning mechanisms also play a role in the visual domain, for example in categorizing new faces. In the study of Altvater-Mackensen et al. (2017), infants were subjected to a familiarization phase in either a unimodal or a bimodal condition. They saw tokens from a continuum that was created from two female faces. After familiarization, results from a discrimination task indicated that infants in the bimodal condition form two distinct categories of faces, while infants in a unimodal condition form one broad category. The same result has been shown in a novel visual object category learning experiment (Junge et al., 2018): infants in the bimodal condition showed better discrimination of two endpoint tokens than infants in the unimodal condition. Distributional learning thus seems to be important for the categorization of different types of sensory stimuli: speech sounds, faces and novel objects. To our knowledge, children with DLD have never been tested on such distributional learning tasks. In the current study we aim to investigate whether these children have a deficit in visual distributional learning and whether this ability correlates with their lexical-semantic knowledge, as a lessened sensitivity to regularities in object categories could contribute to their problems in building strong semantic representations.

The current study

Children with DLD have previously displayed difficulties with verbal and visual statistical learning which could hinder their ability to pick up language efficiently. Indeed, statistical learning ability correlates with or even predicts different types of linguistic skills, such as lexical skills. However, the relationship between statistical learning and the development of vocabulary skills in children with and without DLD is not well understood. In the current study we want to explore this relationship further by investigating visual distributional learning and its relation to vocabulary in children with and without DLD.

Our first research question was: are children with DLD less sensitive to distributional cues compared to TD children when learning novel visual object categories in an experiment? Distributional learning has never been investigated in individuals with DLD, but one study shows that distributional learning of speech sounds is impaired in children with dyslexia (Vandermosten et al., 2019). Developmental dyslexia and DLD are distinct but overlapping disorders (Snowling et al., 2020) and together with previous evidence showing that both verbal and visual statistical learning is impaired in children with DLD, we expected that they show less proficiency in visual distributional learning as well.

Our second research question was: Does the ability of visual distributional learning contribute to lexical knowledge in children with DLD? The underlying cause of the lexical-semantic difficulties in this group is not clear. There is extensive evidence for problems with phonological short-term memory, but this does not seem to be an adequate explanation. We expected that visual distributional learning contributes to these lexical-semantic difficulties, as it could be important for learning semantic information about (the use of) words, object categories and how to map words to objects. Difficulties with processing visual patterns in the environment might result in problems with building a semantically rich lexicon.

To answer our research questions we constructed a visual distributional learning task based on Junge et al. (2018) to test novel object categorization in children with and without DLD. Moreover, we measured lexical knowledge comprehensively in the children with DLD: besides productive and receptive vocabulary size, we tapped the organization of the lexicon and the knowledge of relationships between concepts/words. Finally, we control for variation in phonological processing, as children with DLD are known to have difficulties with this ability and because it is probably related to lexical knowledge. We also controlled for variation in non-verbal intelligence.

Wanrooij et al. (2015) discuss potential pitfalls in the typical design employed when comparing a unimodal with a bimodal familiarization phase in distributional learning tasks. We therefore adapted a different design. In the usual design there might be a confounding factor at play: besides the number of distributional peaks in the input, the spreading of variants (or *dispersion*) also differs between conditions. This difference might result in easier discrimination of endpoint tokens for individuals who had been familiarized with the bimodal condition, as spreading of the variants is higher in that condition. Chládkova et al. (2020) designed a (auditory) distributional learning task that tackled this problem: they constructed two bimodal learning conditions which differed in the position of the distributional peaks, ensuring that spreading of the variants was not different in the two conditions. We applied this design to the visual distributional learning task of Junge et al. (2018).

Method

Participants

27 children diagnosed with DLD participated in our research. One child did not finish the statistical learning task and another child was removed because of bilingualism, resulting in a final sample of 25 children with DLD (17 male, 8 female) between the ages of 7;2 and 9;3 (years;months). For the control group we used previously collected data from a study in which TD children were tested on the same task (Broedelet et al., 2021).[1] We selected 25 children (15 male, 10 female) from a larger sample that

[1] We had planned to test a new group of TD children matched to the DLD group. Unfortunately, we were unable to administer the tests as all primary schools in the Netherlands were closed from March

matched the DLD group best regarding age and gender. Their ages varied between 7;6 and 8;9. Age did not differ significantly between groups (TD age in months $M = 97.64$, $SD = 4.99$, DLD age in months $M = 96.56$, $SD = 6.49$), as tested with a two-sample t-test: $t = 1.864$, $p = 0.063$.

The children with DLD were recruited via different institutions in the Netherlands: Pento, Royal Dutch Auris Group and VierTaal. All children had been officially diagnosed with DLD by a professional clinician and were included if they met the standard DLD inclusion and exclusion criteria used within the institution. All children met the following criteria: they scored at least 1.5 standard deviations below the age norm on at least two of the four language domains (speech, auditory processing, grammar, lexical-semantic development), tested with standardized tests like the CELF; their language disorder was not secondary to a physiological or neurological disorder such as ASD, ADHD or hearing difficulties; they did not have a severe form of dyspraxia and at least one of their caretakers had acquired Dutch as a native language. Data from one child was removed because he was growing up bilingually and answered multiple questions on a vocabulary task in English.

The TD children were recruited via two primary schools in the Netherlands and met the following criteria: they had not been diagnosed with hearing difficulties, language disorders, dyslexia, ADHD or ASD and had at least one caretaker that was a native speaker of Dutch. Our study was approved by the Ethical Committee of the Faculty of Humanities of the University of Amsterdam. The parents/caretakers of all children filled in an informed consent form prior to their participation.

To get a general estimate of the language ability in our DLD subgroup, we administered the Sentence Recalling subtask from the CELF (Clinical Evaluation of Language Fundamentals: Core Language Scales, Dutch version; Semel et al., 2010). In this task, children are asked to repeat sentences of increasing complexity, measuring their morphosyntactic abilities. The Raven Progressive Matrices task was administered to measure non-verbal intelligence (Raven, et al., 2003). One of the children could not finish the Sentence Recalling task due to time constraints. The children's scores (raw, percentile and if available norm and age-equivalent scores) on these two tasks are

---

to June 2020 due to the outbreak of COVID-19. After the reopening of the schools many restrictions still applied, making it impossible to enter schools for testing participants. We therefore decided to use a subset of an already collected dataset as control data. This dataset was previously used for an article about visual distributional learning in TD children (Broedelet et al., 2021). The decision to use previously collected data was taken only because of this circumstance, and *not* because we found a significant effect in this group and deemed it sufficient to use this data. As a result of this reuse, the control group, unlike the DLD group, was not tested on the background tasks measuring vocabulary, morphosyntactic skills, phonological processing and non-verbal intelligence. This means the control group could unfortunately not be matched on vocabulary skills to the DLD group.

shown in Table 1. The children with DLD had low scores on the Sentence Recalling task and performed on average 50 months below their age level, confirming that our sample indeed had difficulty with language acquisition, while they scored within the average range on non-verbal intelligence. This discrepancy between language skills and non-verbal cognitive skills is typical for children with DLD.

Table 1 – Scores of the children with DLD on the sentence recalling and non-verbal intelligence task.

Task	Raw scores	Norm scores	Percentile scores	AES	Diff.
Sentence Recalling (N=24)	4 .. 42 *M* = 18.46 *SD* = 9.27	1 .. 8 *M* = 3.58 *SD* = 2.02	0.1 .. 25 *M* = 4.07 *SD* = 6.37	36 .. 83 *M* = 45.79 *SD* = 13.05	-68 .. -21 *M* = -50.46 *SD* = 14.43
Raven's progressive Matrices	11 .. 38 *M* = 23.24 *SD* = 7.41		5 .. 95 *M* = 41.04 *SD* = 26.25		

Notes: AES = Age-equivalent score (months). Diff. = Difference AES and chronological age. The chronological age (months) is subtracted from the age equivalent score (months). A negative value means that the age-equivalent score was lower than the actual age (M = 96.56, SD = 6.61, range 86 - 111). Scale used for interpreting percentile scores: 0-3 Very low, 3-10 Low, 10-16 Below average, 16-84 Average, 84-90 Above average, 90-98 High, 98-100 Very high. The Sentence Recalling percentile score is in the low range; the Raven's percentile score is in the average range.

Stimuli and design distributional learning task

The design of this experiment follows Junge et al. (2018) and Chládkova et al. (2020), and was previously reported in Broedelet et al. (2021). The aim of our experiment was to measure whether the frequency distribution of tokens along a continuum influenced categorization of those tokens. To this end we constructed an 11-step continuum by morphing two pictures in equal steps using the Sqirlz 2.1 software (Xiberpic.com). We obtained permission to use the pictures of two cuddly toys from Giant Microbes (www.giantmicrobes.com) that were also used in the study of Junge et al. (2018). See Figure 1.

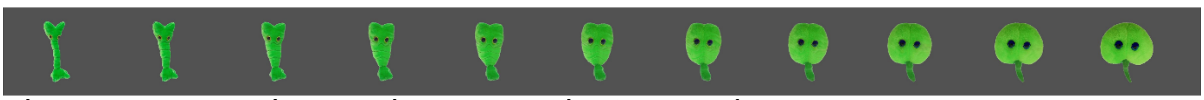

Figure 1 - Novel object continuum used in the experiment.

In the familiarization phase of the experiment, stimuli from the continuum were presented to the children. Two different between-participant familiarization conditions were constructed (see Figure 2). Both conditions contained a bimodal distribution, but the conditions differed concerning the position of the peaks in the continuum. Three of the 11 tokens, which were all equally frequent in both conditions, were used to measure categorization in the test phase: 6, 4 and 8, hereafter referred to as S (standard), D1 (deviant 1) and D2 (deviant 2).

In Condition 1 (Figure 2, blue line), token S and token D2 belonged to the same peak, while token 5 was shown less frequently, creating the perception of a category boundary. In Condition 2 (Figure 2, orange line), token S and token D1 belonged to the same peak and token 7 was shown less frequently. Our hypothesis was that our participants would learn that tokens in one distributional peak belong to one category while tokens from different peaks belong to two different categories. Therefore we predicted that children in Condition 1 learn that tokens S and D2 belong to one category while children in Condition 2 learn that tokens S and D1 belong to one category.

Children were shown 12 blocks of 24 stimuli each (288 stimuli in total), as well as 2 filler stimuli per block (see Figure 4). In each block, the tokens of the continuum were presented one by one following the frequency distribution shown in Figure 2, in a randomized order. Each stimulus was shown for 800 ms and the interstimulus interval was 200 ms (based on the results of Turk-Browne et al. (2005) and Arciuli & Simpson (2011)). Stimuli were shown against a gray background (see Figure 3). A cover task was added to the task to make it more engaging: the filler stimuli jumped across the screen and children were instructed to click on them as fast as possible.

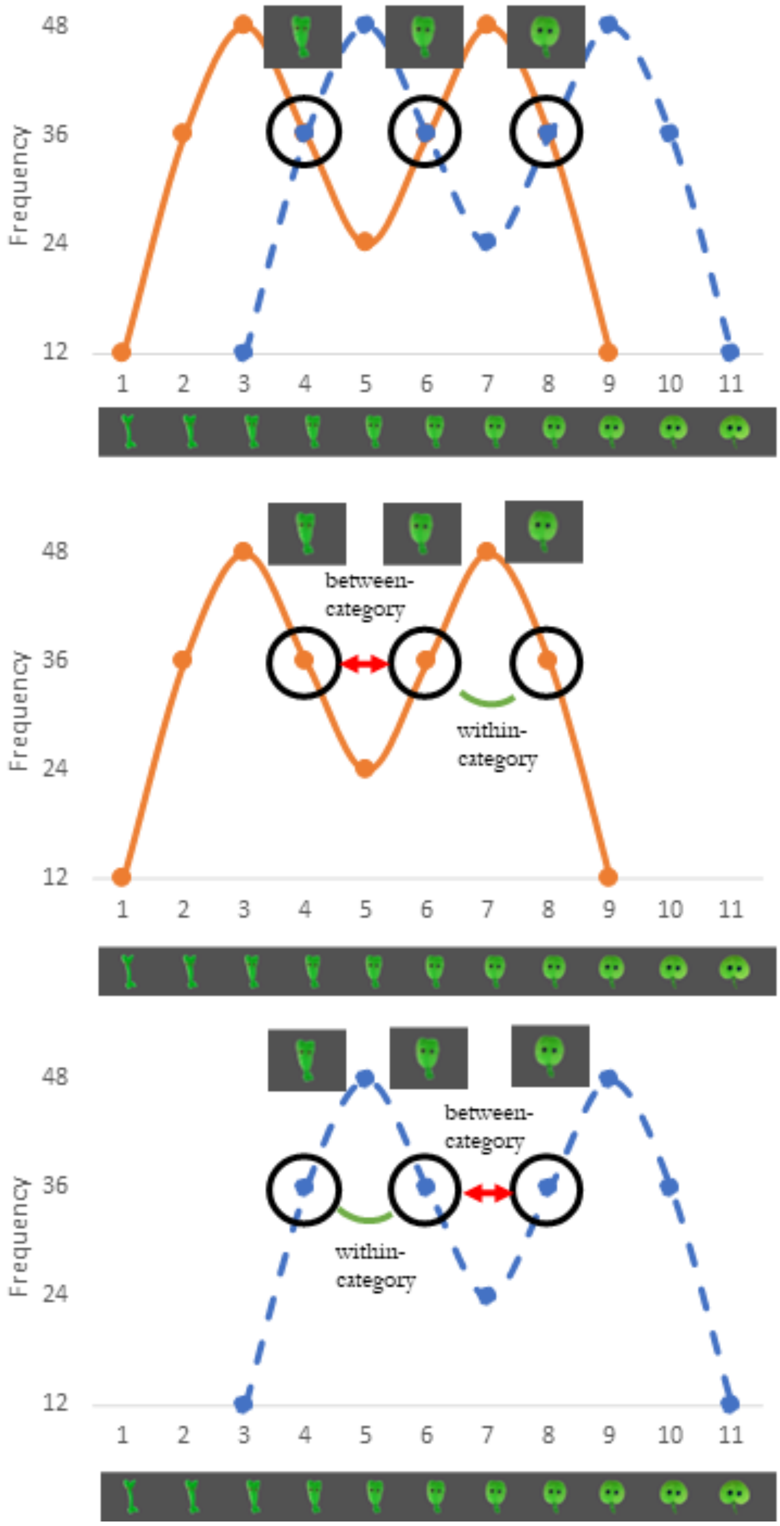
Frequency
48
36
24
12
1 2 3 4 5 6 7 8 9 10 11
between-category
within-category

Figure 2 - Familiarization conditions in the experiment. In Condition 1 (blue line), tokens S and D2 belong to one distributional peak while D1 lies in another peak. On the other hand, in Condition 2 (orange line), tokens S and D1 belong to one distributional peak while D2 lies in another peak. We hypothesize that participants in Condition 1 will learn that S and D2 belong to one category and thus will look more alike than S and D1, and the reversed for participants in Condition 2.

Figure 3 – A familiarization trial.

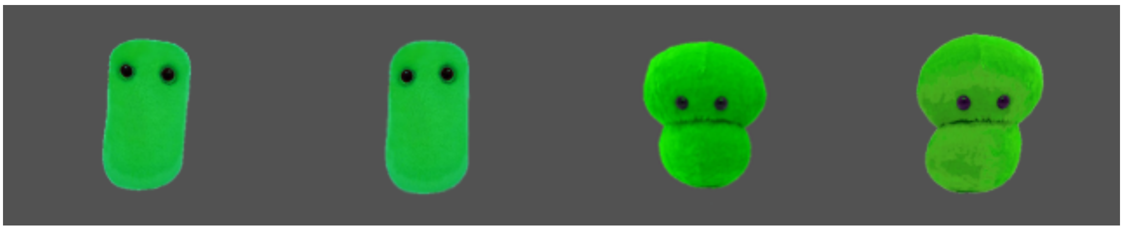

Figure 4 – Stimuli that were used as fillers/cover task.

Categorization was tested after familiarization using AXB-type questions. Children were asked to choose whether stimulus D1 or D2 looked more like stimulus S. In the eight questions, stimulus S was shown above a white stripe and stimuli D1 and D2 were shown below the stripe (see Figure 5). The position of D1 and D2 (left/right) was counterbalanced. Four filler questions were included to add some variation to the test phase, as well as a practice question. For these questions the stimuli that functioned as fillers in the familiarization phase were used and there was a clearly correct answer. The test phase was identical for every child, except that the order of the test questions was randomized. We hypothesized that children that underwent Condition 1 of the familiarization phase would choose stimulus D2 more often than children in Condition 2. This effect of Condition would be considered a learning effect.

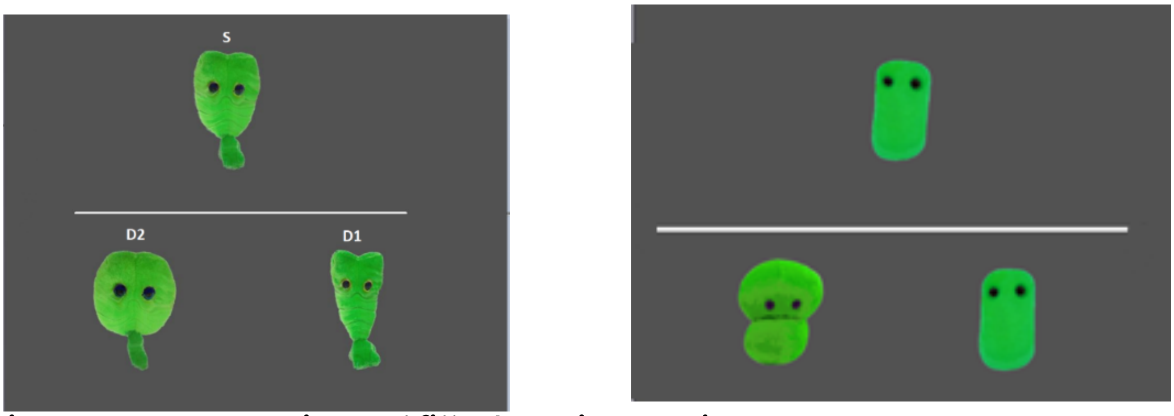

Figure 5 – A test question and filler/practice question.

Measures of vocabulary, phonological processing, non-verbal intelligence and socio-economic status

To investigate the relationship between visual distributional learning and lexical skills in children with DLD[2], we administered several subtests of the CELF (Active Vocabulary, Word Classes 1 or 2 (depending on the age of the child) and Word Associations, as well as the Peabody Picture Vocabulary Task (PPVT; Schlichting, 2005). The tasks were used as measures of receptive and productive vocabulary size (PPVT, Active Vocabulary), the ability to find and express semantic relations between words/concepts (Word Classes 1 and 2) as well as the ability to name words of a semantic category as an indicator of lexical-semantic organization (Word Associations). See Table 2 for more information about the vocabulary tasks.

As control tasks, the children were tested on phonological short-term memory using the digit span task Number Repetition 1 from the CELF, on verbal working memory

[2] Our original plan was to investigate this relationship in both groups of children. Unfortunately, as is mentioned in our first footnote, we were not able to test the TD children on these tasks.

using the Number Repetition 2 task (digit span backwards) from the CELF and on verbal short-term memory using the non-word repetition task (Rispens & Baker, 2012). Moreover, performance on the Raven Progressive Matrices task was used as a control variable for non-verbal intelligence. See Table 3 for more information about the control tasks. Finally, as socio-economic status (SES) may play a role in vocabulary development (e.g. Hoff, 2003), we took the SES of the children into account using a database from Sociaal en Cultureel Planbureau (2018). In this database, socio-economic scores are computed on the basis of the average education level and income in a particular zip code. The SES scores are based on the home addresses of the children.

Table 2 – Vocabulary measures administered to the children with DLD.

Construct	Task	Description	Scoring	Score range
Vocabulary size	Receptive vocabulary (PPVT)	Children heard a word and had to point to one of the four pictures.	Correct: 1 point Incorrect: 0 points	0 .. 204
Vocabulary size	Productive vocabulary (CELF)	Children saw a picture and had to name it.	2 points for a correct answer, for some items there were 1-point answer possibilities	0 .. 56
Semantic knowledge	Word Classes 1 (7 y.o. children) (CELF)	Children had to choose which two out of three/four pictures were related and why.	1 point for choosing the correct picture, 1 point for expressing the relationship correctly	0 .. 38
Semantic knowledge	Word Classes 2 (8+) (CELF)	Children had to choose which two words out of four were related and why.	1 point for choosing the correct word, 1 point for expressing the relationship correctly	0 .. 40
Lexical-semantic organization	Word Associations (CELF)	Children had to name as many words as they could in a semantic category: food, clothes and professions.	1 point for every related word	0 .. ∞

Table 3 – Control measures administered to the children with DLD.

Construct	Task	Description	Scoring	Score range
Verbal short-term memory	Digit span forwards	Children had to repeat strings of number increasing in length.	Correct: 1 point Incorrect: 0 points	0 .. 16
Verbal working memory	Digit span backwards	Children had to repeat strings of number backwards increasing in length.	Correct: 1 point Incorrect: 0 points	0 .. 14
Phonological short-term memory	Non-word repetition	Children had to repeat non-words.	Correct: 1 point Incorrect: 0 points	0 .. 22
Non-verbal intelligence	Raven Progressive Matrices	Children had to complete a visual pattern.	Correct: 1 point Incorrect: 0 points	0 .. 60

Procedure

Testing took place in a quiet room in the school or in the home of the child. The distributional learning experiment was run on a laptop computer using E-Prime 3.0 (Psychology Software Tools, Pittsburgh, PA). Children wore headphones. We had recorded the instructions in advance, in a child-directed manner. Before the experiment started, the children were instructed to look at the images on the screen and click on moving images as fast as they could if they saw one. They were told to watch carefully as there would be questions about the images later on, but the type of questions was not specified. The experiment started when the child confirmed that s/he understood the task. Familiarization condition was counterbalanced between participants. There was a short break halfway the familiarization phase and the child could indicate when s/he wanted to continue. The test phase started immediately after the familiarization phase with a practice question. Children were instructed to carefully look at the image above the white stripe, and to indicate which of two images below the stripe they thought looked more like the upper image. The experimenter repeated the question while pointing out the images. The experiment had a total duration of approximately 10 minutes.

Besides the distributional learning task, the children with DLD did two other statistical learning tasks (results are not discussed in this paper) as well as the aforementioned background tasks. For those children, testing was divided over two separate test sessions on different days; the second session usually took place within a few days or one week. The order of the tasks within the sessions as well as the order of the sessions was counterbalanced across participants. Each test session took approximately 50 to 60 minutes.

Results

Split-half reliability distributional learning task

Split-half reliability was computed as a measure of reliability of the distributional learning task. Two separate generalized mixed effect models were run with only the odd or even test items included. Then, the correlation between the answers to even and odd test items was computed, using the random slopes of the intercept for the even/odd test items. After the application of the Spearman-Brown correction, the split-half reliability of the task turned out to be $r = 0.73$ (95% CI 0.52 .. 0.85), approaching the value of $r = 0.80$ which is considered the standard that reliable tests should meet (Nunnally & Bernstein, 1994).

Group comparison distributional learning task[3]

See Table 4 and Figure 6 for the descriptive data. As a first step in our analysis, we removed all practice and filler items from the data. A generalized mixed effect model was run with the package *lme4* (Bates et al., 2015) in R (R Core Team, 2020) to test whether familiarization condition and participant group influenced categorization. The choice for stimulus D2 (which could either be 1 or 0) was the dependent variable. Between-participant predictors were Condition (Condition 1 or 2), Group (TD or DLD) and Age (in months). PositionD2 was a within-participant predictor reflecting the position of token D2 (left or right) that varied between test items. We chose the maximal model that is still correctly computable and that keeps all its included predictors and interactions reportable (by including random slopes for all within-participant predictors and interactions). The model includes main effects for Condition, Group, Age and PositionD2, all two- and three-way interactions between Condition, Group and Age as well as the simple interaction between Condition and PositionD2. Moreover, we included random intercepts by participant as well as by-subject random slopes for PositionD2. Sum-to-zero orthogonal coding (Kraemer & Blasey, 2004) was applied to the predictors Condition (-½ for Condition 2 and +½ for Condition 1), Group (-½ for DLD and +½ for TD) and Position D2 (-½ for right and +½ for left). The predictor Age was centered by subtracting its average.

We predicted that if children are sensitive to the distributional cues in the familiarization phase, our children in Condition 1 would prefer the combination S + D2, while

[3] The TD children of whom results are reported here are a subgroup of the sample reported in Broedelet et al. (2021).

our children in Condition 2 would prefer the combination S + D1 - in other words, a stronger preference for D2 in Condition 1 than Condition 2. This could manifest as a significant effect of Condition on the dependent variable. Moreover, we expected that our children with DLD would be less sensitive to the distributional cues in the familiarization phase than our TD children, which could manifest as a significant interaction between the effects of Condition and Group on the dependent variable, indicating that the Condition effect is not equally strong in the two subpopulations.

Confirmatory results

In our sample, as determined by our model, Condition influenced the choice for stimulus D2: children in Condition 1 were 4.04 times more likely to choose stimulus D2 than children in Condition 2, and this effect was significantly above 1: $z = 2.758$, $p = 0.006$, 95% CI 1.497 .. 10.9. This is in line with our prediction and indicates that school-aged children can learn novel visual object categories based on distributional properties. Our second prediction is not confirmed: although the effect of Condition was 1.007 times stronger in the TD group compared to the DLD group, this interaction between Condition and Group was not significantly above 1: $z = 0.007$, $p = 0.994$, 95% CI 0.15 .. 6.8. We thus cannot conclude anything about a difference in distributional learning in children with DLD compared to TD children: the confidence interval tells us that children with DLD could be up to 6.7 times better or 6.8 times weaker on the visual distributional learning task than TD children. We therefore cannot conclude whether children with DLD do or do not have a distributional learning deficit.

Exploratory results

To explore whether children with DLD show a distributional learning effect, we ran a separate model which only included the children with DLD. This model included the main effects for Condition, Age and PositionD2 as well as all three-way interactions between those predictors. According to the model, our children with DLD in Condition 1 were 3.75 times more likely to choose D2 than our children in Condition 2, but the effect was not significantly above 1: $z = 1.788$, $p = 0.074$, 95% CI 0.86 .. 19.4[4]. On the basis of this result we cannot conclude whether children with DLD are able to learn novel visual object categories based on distributional information[5].

[4] When we ran a model which included random slopes per participant for PositionD2 (as we did in our first model with all participants), the effect of Condition was 4.11 (95% CI 1.01 .. 16.7): $z = 1.977$, $p = 0.048$. However, as this model had a singular fit, we chose to report the results of a simplified model without random slopes for PositionD2 (this makes the effect of PositionD2 unreportable, but as we are not directly interested in this effect, this is not problematic). Note that neither the p-value of 0.074 neither the p-value of 0.048 can be called statistically significant, because this exploratory test came on top of the earlier confirmatory test, for which we already used a preset p-value criterion of 0.05.

[5] We also ran an analysis that only included the TD children, which yielded a significant effect of Condition ($z = 2.047$, $p = 0.04$). However, please note that this finding cannot be interpreted as a difference

Table 4 - Descriptive data for the choice of stimulus D1 or D2.

	TD children		Children with DLD	
	D1	D2	D1	D2
Condition 1	55	49 target	61	35 target
Condition 2	71 target	25	84 target	20

in distributional learning between children with DLD and TD children, as the effect of Group was not significant in our first model.

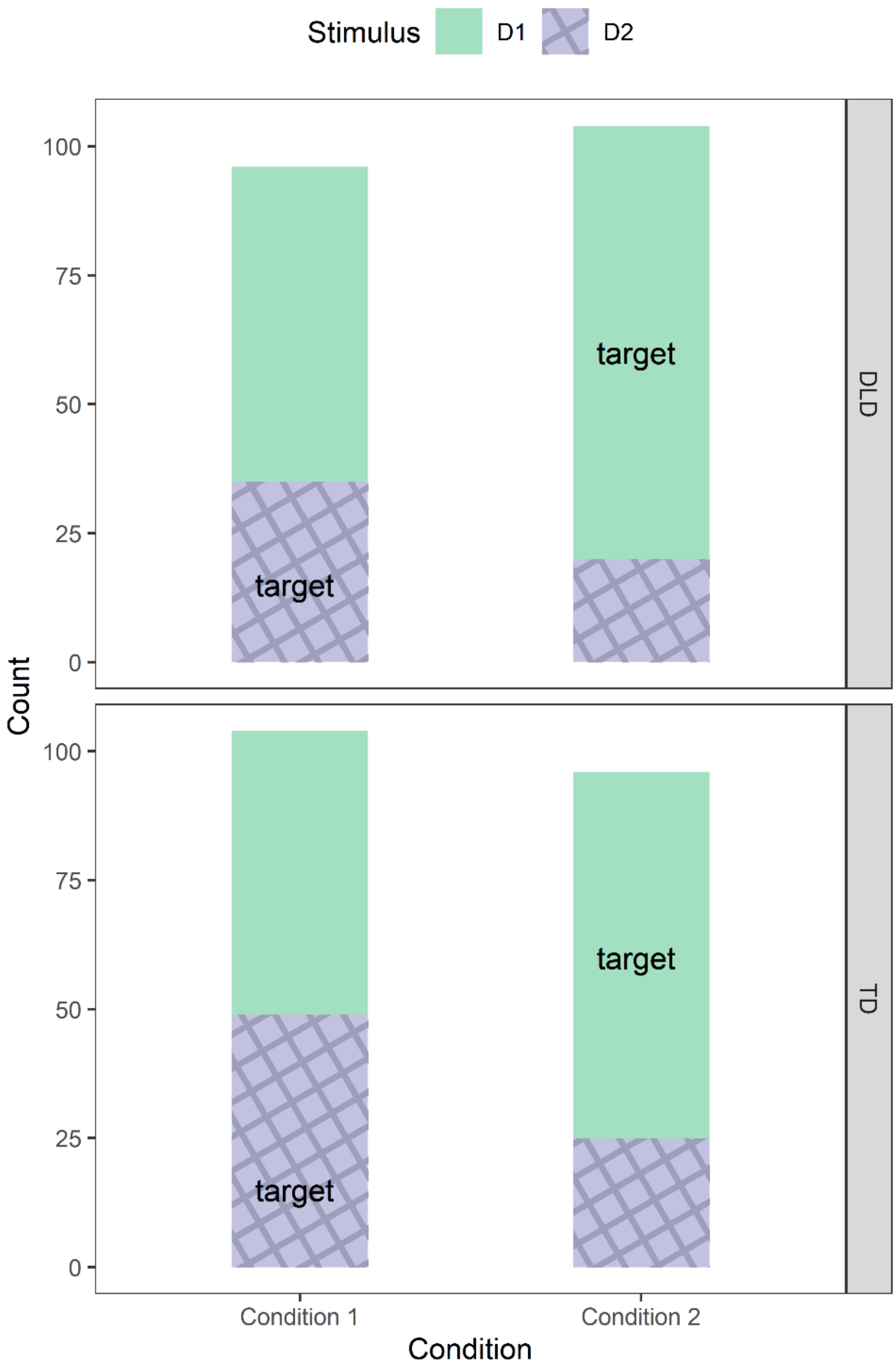

Figure 6 - Choice for stimulus D1 or D2 depending on condition and group.

Regression analyses

Descriptive data

To investigate the relationship between distributional learning and vocabulary, we administered tasks measuring several types of lexical knowledge to the children with DLD, as well as several control tasks (see section 4.3). In Table 5 we present the scores of the children with DLD on the vocabulary tasks and in Table 6 their scores on the control tasks: the raw scores, the norm and percentile scores (if available), and the age-equivalent scores. The raw scores are used in our statistical analysis. The norm, percentile and age-equivalent scores are presented to illustrate the abilities of the children with DLD.

Table 5 – Children with DLD's scores on the vocabulary task.

Task	Subtask	Raw scores	Norm scores	Percentiles	AES	Diff.
Productive vocabulary		8 .. 41 *M* = 28.16 *SD* = 8.94	2 .. 12 *M* = 6.84 *SD* = 2.46	0.4 .. 75 *M* = 20.46 *SD* = 21.33	36 .. 98 *M* = 73.24 *SD* = 16.69	-62 .. 7 *M* = -23.32 *SD* = 15.83
Receptive vocabulary		70 .. 119 *M* = 90.48 *SD* = 13		0 .. 91 *M* = 27.36 *SD* = 26.82		
Word associations		10 .. 42 *M* = 23.92 *SD* = 6.37	2 .. 15 *M* = 7.48 *SD* = 2.45	0.4 .. 95 *M* = 24.22 *SD* = 19.77	42 .. 133 *M* = 77.2 *SD* = 18.42	-56 .. 42 *M* = -19.36 *SD* = 19.19
Word classes	Receptive	2 .. 19 *M* = 11.2 *SD* = 6.95	3 .. 12 *M* = 7.24 *SD* = 2.63	1 .. 75 *M* = 24.92 *SD* = 22.6	36 .. 109 *M* = 70.36 *SD* = 22.85	-68 .. 18 *M* = -26.2 *SD* = 25.59
	Expressive	0 .. 18 *M* = 8.8 *SD* = 5.95	1 .. 13 *M* = 6.88 *SD* = 2.71	0.1 .. 84 *M* = 21.68 *SD* = 22.84	36 .. 116 *M* = 71.92 *SD* = 19.19	-68 .. 25 *M* = -24.64 *SD* = 21.06
	Total	2 .. 37 *M* = 20 *SD* = 12.79	2 .. 13 *M* = 6.88 *SD* = 2.60	0 .. 84 *M* = 21.3 *SD* = 22.57	36 .. 116 *M* = 71.6 *SD* = 19.4	-68 .. 25 *M* = -24.96 *SD* = 21.72

Notes: AES = Age-equivalent score (months). Diff. = Difference AES and chronological age. The chronological age (months) is subtracted from the age equivalent score (months). A negative value means that the age-equivalent score was lower than the actual age (M = 96.56, SD = 6.61, range 86 - 111). Scale used for interpreting percentile scores: 0-3 Very low, 3-10 Low, 10-16 Below average, 16-84 Average, 84-90 Above average, 90-98 High, 98-100 Very high. The scores for the vocabulary tasks fall within the average range.

Table 6 – Children with DLD's scores on the control tasks

Task	Subtask	Raw scores	Norm scores	Percentile scores	AES	Diff.
Raven's progressive Matrices		11 .. 38 *M* = 23.24 *SD* = 7.41		5 .. 95 *M* = 41.04 *SD* = 26.25		
Digit Span	Forwards	3 .. 9 *M* = 5.36 *SD* = 1.58	1 .. 12 *M* = 6 *SD* = 2.8	0.1 .. 75 *M* = 16.6 *SD* = 21.27	50 .. 103 *M* = 68.76 *SD* = 16.87	-52 .. 12 *M* = -27.8 *SD* = 17.33
	Backwards	0 .. 4 *M* = 2.72 *SD* = 1.02	2 .. 11 *M* = 7.52 *SD* = 2.35	0.4 .. 63 *M* = 26.06 *SD* = 19.34	57 .. 101 *M* = 79.52 *SD* = 13.97	-43 .. 14 *M* = -17.04 *SD* = 13.73
	Total	4 .. 12 *M* = 8.08 *SD* = 1.91	1 .. 10 *M* = 5.68 *SD* =2.39	0.1 .. 50 *M* = 12.9 *SD* = 14.67	48 .. 102 *M* = 71.8 *SD* = 11.81	-56 .. -2 *M* = -24.76 *SD* = 12.31
Non-word repetition		0 .. 9 *M* = 3.36 *SD* = 2.36		Low		

Notes: AES = Age-equivalent score (months). Diff. = Difference AES and chronological age. The chronological age (months) is subtracted from the age equivalent score (months). A negative value means that the age-equivalent score was lower than the actual age (M = 96.56, SD = 6.61, range 86 - 111). Scale used for interpreting percentile scores: 0-3 Very low, 3-10 Low, 10-16 Below average, 16-84 Average, 84-90 Above average, 90-98 High, 98-100 Very high. The scores for the Raven, and digit span backwards fall within the average range, the scores for digit span forwards and total digit span score fall in the below average range.

In contrast to their scores on the sentence recall task (see Table 1), the children with DLD scored within the average range (low end of the continuum) on the measures of vocabulary. The age-equivalent scores on these subtasks were between 19.36 and 26.2 months below their chronological age. Their non-verbal intelligence scores are also within the average range (see Table 5). However, the children showed below-average scores on the digit span forward task, which presumably reflect limitations in phonological short-term memory, which are reported often in DLD (Montgomery et al., 2010). Norm scores are available for the non-word repetition task for TD children of 7 (*N* = 96) years old, 8 years old (*N* = 82) and 9 years old (*N* = 208)[6]. The mean raw scores

[6] https://progracy.com/normscores/

for these age groups are 8.03, 8.83 and 9.07 out of 22 words correct respectively. Compared to that, the average score of 3.36 out of 22 in our group of children with DLD (see Table 6) can be considered as low. The children's age in months was on average 96.56 (SD = 6.61, range 86 .. 111), and their SES score on average -0.37 (SD = 1.04, range = -1.96 .. 1.52).

Principle component analysis

Prior to the regression analysis, all variables were centered around zero and scaled to a standard deviation of 1. To reduce the number of predictor variables, we ran a principal component analysis (PCA) in R using the raw scores on the digit span forward, digit span backward, non-word repetition and non-verbal intelligence tasks. The PCA analysis yielded four components, which explained 44%, 36%, 15% and 5% of the variance respectively. On the basis of this outcome, we decided to use three components, as they together explained 95% of the variance in the data. After varimax rotation, the three components explained 46%, 27% and 26% of the variance respectively. These components were saved and used for further analysis. See Table 7 for the component loadings. The first component represented phonological processing (mainly digit span forward and non-word repetition scores, the scores of which strongly correlated (r = 0.77, p = 0.0001)), the second component non-verbal intelligence (mainly Raven scores), and the third component verbal working memory (mainly digit span backward scores).

Table 7 – Standardized loadings of varimax-rotated PCA.

	Component 1 (phonological processing)	Component 2 (non-verbal intelligence)	Component 3 (verbal working memory)
Digit span forwards	<u>0.93</u>	-0.22	0.05
Digits span backwards	0.05	0.20	<u>0.98</u>
Non-word repetition	<u>0.95</u>	0.13	0.03
Non-verbal intelligence	-0.05	<u>0.97</u>	0.21

Predictor variables

The predictor variables were accuracy on the distributional learning task, age, SES, and the three component scores representing phonological processing, non-verbal intelligence and verbal working memory respectively. There were no significant correlations between the predictor variables (see Table 8). Accuracy on the distributional learning task was used as the measure for distributional learning ability, and was computed by comparing the answer to every test question to the target answer. For Condition 1, the target answer was D2 while it was D1 for Condition 2. This variable thus reflects sensitivity to the distributional properties in the familiarization phase. See Figure 7 for the distribution of the accuracy scores.

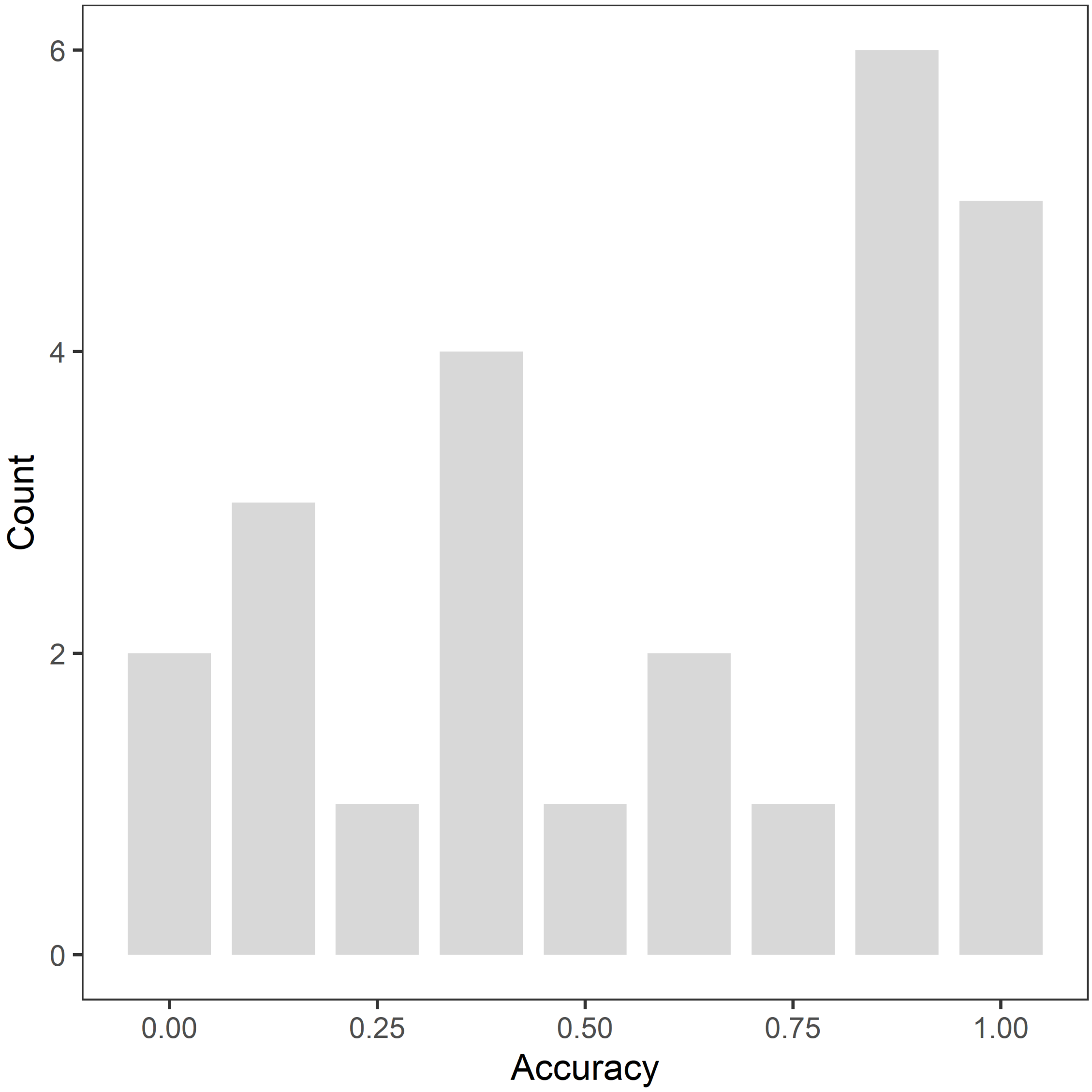

Figure 7 – Distribution of accuracy scores on the distributional learning task.

Table 8 - Correlations between the predictor variables.

	Comp 1 (phonological processing)	Comp 2 (non-verbal intelligence)	Comp 3 (verbal working memory)	Age	SES
Distributional learning	r = -0.17 p = 0.426	r = -0.05 p = 0.819	r = -0.24 p = 0.256	r = 0.09 p = 0.677	r = 0.03 p = 0.881
Component 1 (phonological processing)		r = 0 p = 1	r = 0 p = 1	r = 0.09 p = 0.675	r = 0.02 p = 0.917
Component 2 (non-verbal intelligence)			r = 0 p = 1	r = -0.21 p = 0.323	r = 0.02 p = 0.923
Component 3 (verbal working memory)				r = 0.26 p = 0.217	r = -0.22 p = 0.299
Age					r = 0.04 p = 0.866

Dependent variables

We ran four separate multiple linear regression analyses in R to test the relationship between distributional learning and different measures of vocabulary. The dependent measures were raw scores on the tasks measuring receptive vocabulary size, productive vocabulary size, and word associations. For the scores on the word classes tasks (part 1 and part 2) we decided to use the norm total scores (receptive + expressive) instead of raw scores (see Table 5)[7].

Regression analyses

The first model was run with receptive vocabulary size as the dependent variable and the five predictors as predictor variables. The model did not explain variation in receptive vocabulary size better than the null model (F = 0.59, p = 0.734) and none of the predictors were significant (see Table 9). The second model with productive vocabulary size as the dependent variable also was not significant (F = 1.693, p = 0.18) and contained no significant predictors (see Table 10). The third model with word classes total score as the dependent variable was not significant (F = 1.604, p = 0.2033), but component 2 (non-verbal intelligence) significantly predicted word classes score (t = 2.156, p = 0.045), indicating that the ability of completing non-verbal patterns might

[7] We felt it was not possible to use the raw scores, as the two parts of the task (part 1 for children up until 7 years old and part 2 for 8+ children) yielded different ranges of scores, while they were meant to measure the same underlying skill. Using the norm scores enabled us to use word category score as one variable for the whole group of children. The norm scores were computed from tables provided in the manual of the test, based on a sample of 1336 Dutch children (5-16 years old). The norm score provides information about a child's performance compared to the age norm.

explain unique variance in semantic knowledge about words, but please note that this result is exploratory. None of the other predictors were significant (see Table 11). The last model with word association score as the dependent variable was not significant ($F = 0.827$, $p = 0.564$), and none of the variables significantly predicted the dependent variable (see Table 12). In none of the models distributional learning significantly predicted vocabulary scores. Based on this null result, we cannot conclude anything about the relationship between visual distributional learning and vocabulary knowledge.

Table 9 – Results from the first linear model predicting receptive vocabulary size.

Predictor	Estimate (log odds) [95% CI]	Std. error (log odds)	*t*	*p*
Age	0.18 [-0.31 .. 0.67]	0.234	0.782	0.444
SES	0.09 [-0.37 – 0.56]	0.222	0.420	0.680
Component 1 (phonological processing)	0.27 [-0.20 .. 0.73]	0.220	1.204	0.244
Component 2 (non-verbal intelligence)	0.17 [-0.30 .. 0.63]	0.221	0.747	0.465
Component 3 (verbal working memory)	0.12 [-0.38 .. 0.62]	0.239	0.501	0.622
Distributional learning	0.15 [-0.33 .. 0.63]	0.229	0.674	0.509
Comparison with null model: $F = 0.59$, $p = 0.734$				

Table 10 – Results from the second linear model predicting productive vocabulary size.

Predictor	Estimate (log odds) [95% CI]	Std. error (log odds)	*t*	*p*
Age	0.32 [-0.11 .. 0.75]	0.205	1.584	0.131
SES	0.36 [-0.05 .. 0.77]	0.194	1.869	0.078
Component 1 (phonological processing)	0.29 [-0.11 .. 0.70]	0.193	1.520	0.146
Component 2 (non-verbal intelligence)	0.14 [-0.27 .. 0.55]	0.193	0.729	0.476
Component 3 (verbal working memory)	-0.03 [-0.47 .. 0.41]	0.209	-0.126	0.901

Distributional learning	0.08 [-0.34 .. 0.50]	0.200	0.401	0.693
Comparison with null model: $F = 1.693$, $p = 0.18$				

Table 11 – Results from the third linear model predicting word classes total score.

Predictor	Estimate (log odds) [95% CI]	Std. error (log odds)	*t*	*p*
Age	-0.19 [-0.63 .. 0.24]	0.207	-0.930	0.365
SES	0.04 [-0.38 .. 0.45]	0.196	0.180	0.8595
Component 1 (phonological processing)	-0.25 [-0.66 .. 0.16]	0.195	-1.301	0.2098
Component 2 (non-verbal intelligence)	0.42 [0.01 .. 0.83]	0.195	2.156	0.045*
Component 3 (verbal working memory)	0.03 [-0.42 .. 0.47]	0.211	0.124	0.903
Distributional learning	-0.17 [-0.60 .. 0.25]	0.202	-0.859	0.402
Comparison with null model: F = 1.604, p = 0.2033				

Table 12 – Results from the fourth linear model predicting word association score.

Predictor	Estimate (log odds) [95% CI]	Std. error (log odds)	*t*	*p*
Age	0.14 [-0.33 .. 0.62]	0.227	0.630	0.536
SES	0.23 [-0.23 .. 0.68]	0.215	1.049	0.308
Component 1 (phonological processing)	0.10 [-0.34 .. 0.55]	0.214	0.486	0.633
Component 2 (non-verbal intelligence)	0.13 [-0.32 .. 0.57]	0.214	0.584	0.567
Component 3 (verbal working memory)	-0.30 [-0.79 .. 0.18]	0.232	-1.307	0.208
Distributional learning	0.88 [-0.38 .. 0.55]	0.222	0.398	0.695
Comparison with null model: $F = 0.827$, $p = 0.564$				

Discussion

In the current study we aimed to shed more light on the relationship between statistical learning ability and lexical-semantic skills in children with and without DLD. Specifically, we investigated whether children with DLD are sensitive to distributional information in a visual distributional learning task, and whether this ability is related to different types of lexical knowledge. Our results show that, overall, school-aged children learn novel visual object categories based on distributional information. We cannot answer our first research question as we did not find evidence for or against a visual distributional learning deficit in children with DLD. The confidence interval of our group comparison shows that children with DLD could be between 6.8 times weaker and 6.7 times better on the visual distributional learning task than TD children. The finding of a non-significant group difference could be due to chance. It is possible that the true effect is zero, but we can only speculate about possible underlying reasons.

It could be the case that children with DLD have no disadvantage in visual distributional learning compared to TD children. Previous evidence has suggested that visuomotor statistical learning is impaired in children with DLD (Lum et al., 2014; Obeid et al., 2016; Tomblin et al., 2007). However, null results have also been found (Aguilar & Plante, 2014; Noonan, 2018) and Lammertink et al. (2020) report evidence for visual statistical learning in children with DLD. Intact visual statistical learning cannot be concluded from our null result, but accumulated evidence could point towards a specifically verbal statistical learning deficit in children with DLD, as opposed to a domain-general deficit. Statistical learning is often characterized as a domain-general ability, but research suggests the existence of different domain-specific components of statistical learning (Siegelman, 2020). It is also possible that sequential statistical learning as is tested with for example word segmentation tasks is problematic for children with DLD, while specifically distributional learning is not. More research is necessary to disentangle these possibilities. For example, it would be interesting to investigate whether *verbal* distributional learning is problematic for children with DLD.
The absence of a significant DLD–TD difference could also be due to a lack of statistical power. We tested 25 children in both participant groups, but the between-participants design of our experiment results in relatively limited number of participants per subgroup. Future studies should test larger participant groups and/or change the design such that multiple between-participant comparisons are avoided. Another option would be to test categorization in a way that would provide more data, for example by using an online behavioral measure or an neurological measure like EEG (Altvater-Mackensen et al., 2017), which could make the task more sensitive to potential DLD–TD differences.

To answer our second research question, we investigated whether distributional learning ability predicted vocabulary knowledge in children with DLD, while controlling for variation in phonological processing, verbal working memory, non-verbal intelligence, SES and age. We did not find any evidence for or against this relationship in our sample of children with DLD. Apart from chance, several factors could underlie this null-result. It could be the case that, as statistical learning tasks are designed to measure group-level performance, they are not suitable for measuring individual differences reliably and thus should not be used to predict differences in language outcome.(Arnon, 2019; Siegelman et al., 2017; Siegelman et al., 2017). For example, Arnon (2019) showed that three different statistical learning tasks had a low test-retest reliability and internal consistency in children, illustrating that they did not capture individual statistical learning ability reliably. This is a serious problem in the field of statistical learning research, as correlations between statistical learning ability and language proficiency might have been both overestimated and underestimated in previously reported studies (Siegelman, 2020). The split-half reliability of our visual distributional learning task was $r = 0.73$, approaching the standard of $r = 0.80$. This suggests that the test is a fairly reliable test of categorization. However, test-retest reliability should still be investigated to find out whether this task is able to capture individual differences reliably.

Another phenomenon that could occur when investigating individual differences in statistical learning is a large portion of the participants performing around chance level. Variation around chance level is not meaningful variation, which could result in the absence of significant correlations. However, this does not seem to be the case for our sample (see Figure 7). Another problem with this type of tasks might be that implicit knowledge that is built during familiarization does not transfer to the more explicit test questions in the test phase. Introducing more implicit and/or online measures of statistical learning could address this problem.

Importantly, although we did not compare the children with DLD to TD children on measures of vocabulary directly, it is striking that the percentile scores of the children with DLD in our sample are within the average range. Still, it is important to note that the ranges are wide and the children do fall behind same-aged peers if we consider the age-equivalent scores. The scores on the task measuring syntax and morphology do fall in the low range. This could mean that grammatical difficulties are more pronounced than vocabulary problems in our sample. Future studies could consider picking specific subgroups of children with DLD who have pronounced vocabulary problems to investigate the relationship between statistical learning and vocabulary development.

Although we cannot conclude this on basis of our results, there is also the possibility that there is no (strong) relationship between statistical learning and lexical-semantic knowledge. Perhaps statistical learning does contribute to more structural linguistic knowledge such as rules and regularities, but deeper (semantic) knowledge is subject to other types of learning mechanisms, although research did point out that statistical learning mechanisms are sensitive to semantic information (Goujon, 2011; Paciorek & Williams, 2015). Possibly, deficits in other cognitive mechanisms such as attention, inhibition or verbal short-term memory play a role in the lexical-semantic difficulties that are observed in children with DLD (Alt & Plante, 2006; Mainela-Arnold & Evans, 2014). More research into these difficulties and their underlying mechanisms is necessary.

We included measures of phonological processing, verbal working memory and non-verbal intelligence in our regression models as control variables. Somewhat unexpectedly, we did not find evidence for a contribution of phonological processing or verbal working memory ability to different types of vocabulary knowledge in our sample of children with DLD. Similarly, Rispens and Baker (2012) found no evidence for a relationship between non-word repetition and vocabulary size in TD children and children with DLD, and the longitudinal study of Melby-Lervåg et al. (2012) yielded no evidence of a causal relationship between non-word repetition and vocabulary acquisition in 4-7 year old TD children. A meta-analysis could shed light on the relationship between phonological processing and vocabulary development in children with and without DLD. Moreover, we found an indication that non-verbal intelligence contributes to word category knowledge in children with DLD. This might be explained by similarities between the tasks: in the Word Category task, children had to choose which two out of three pictures/words were related (and why), while in the Raven progressive matrices task children had to complete visual patterns (see Table 2). Still, it is an interesting finding that non-verbal intelligence could explain variation in the verbal (semantic) domain, although we want to emphasize that this is an exploratory finding.

A shortcoming of the visual distributional learning task we have used is the finding that children overall prefer the combination S + D1, which is a result we have also reported in Broedelet et al. (2021). In that study, we tested 32 adults in an online experiment to explore *a priori* preferences for either S+D1 or S+D2. We wanted to investigate how participants who had not been exposed to a familiarization phase would answer questions similar to the test phase of our experiment. Results showed that participants chose D1 to look more like S 75% of the time, which was significantly higher than chance level. This result implies that D1 looks more like S for most participants, which is not an ideal starting point for testing the influence of distributional learning on categorization. This a priori preference might have diminished the distributional learning effect as well as a potential group difference in learning. However, our results show that despite this preference for the combination of S+D1, exposure to a familiarization phase in which S and D2 belonged to one distributional peak still caused participants to categorize S and D2 more often. Future studies might choose to use different stimuli when testing visual distributional learning and test beforehand whether participants show any unexpected preferences.

Conclusion and future directions

Our study shows that school-aged children can learn novel visual object categories based on distributional information. We did not find evidence for or against a visual distributional learning deficit in children with DLD. Future research could use our results for meta-analyses. Moreover, it would be interesting to investigate whether children with DLD have a domain-general deficit in statistical learning or solely a verbal statistical learning deficit, for example by a comparison between visual and verbal distributional learning. The relationship between statistical learning and lexical-semantic knowledge should be examined further. It could be fruitful to investigate children who show difficulties with lexical-semantic skills. Finally, measuring statistical learning online could be beneficial for both group comparisons as well as studying individual differences.

Data, code and materials availability statement

R scripts, data and materials are available on FigShare: 10.21942/uva.c.5174660 (reserved DOI).

Ethics statement

Ethics approval was obtained from the ethics committee of the University of Amsterdam. The caretakers of all participants gave informed written consent before the participant took part in the study.

Authorship and contributor ship

Iris Broedelet: conceptualization, methodology, formal analysis, investigation, data curation, writing – original draft, visualization, project administration, funding acquisition; Paul Boersma: conceptualization, methodology, formal analysis, writing – review and editing, supervision; Judith Rispens: conceptualization, methodology, writing – review and editing, supervision.

Acknowledgements

We offer our sincere gratitude to all children who participated and their caretakers, all speech therapists and teachers from VierTaal, The Royal Auris Group and Pento for their extensive help on facilitating children's participation, as well as the staff members from the primary schools that participated in our research. We would also like to thank our test-assistant Maayke Sterk and our lab technician Dirk Jan Vet.

References

Aguilar, J. M., & Plante, E. (2014). Learning of grammar-like visual sequences by adults with and without language-learning disabilities. *Journal of Speech, Language, and Hearing Research, 57*(4), 1394–1404. https://doi.org/10.1044/2014_JSLHR-L-13-0124

Aguilar, J. M., Plante, E., & Sandoval, M. (2017). Exemplar Variability Facilitates Retention of Word Learning by Children With Specific Language Impairment. *Language, Speech, and Hearing Services in Schools,* 1–13. https://doi.org/10.1044/2017_LSHSS-17-0031

Alt, M., & Plante, E. (2006). Factors that influence lexical and semantic fast mapping of young children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 49*(5), 941–954. https://doi.org/10.1044/1092-4388(2006/068)

Altvater-Mackensen, N., Jessen, S., & Grossmann, T. (2017). Brain responses reveal that infants' face discrimination is guided by statistical learning from distributional information. *Developmental Science, 20*(2). https://doi.org/10.1111/desc.12393

Arciuli, J., & Conway, C. M. (2018). The promise—and challenge—of statistical learning for elucidating atypical language development. *Current Directions in Psychological Science, 27*(6), 492–500. https://doi.org/10.1177/0963721418779977

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: the role of age and speed of stimulus presentation. *Developmental Science, 14*(3), 464–73. https://doi.org/10.1111/j.1467-7687.2009.00937.x

Arnon, I. (2019). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods,* 1–14. https://doi.org/10.3758/s13428-019-01205-5

Bates D., Mächler M., Bolker B., & Walker S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software,* 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beckage, N., Smith, L., & Hills, T. (2010). Semantic network connectivity is related to vocabulary growth rate in children. In *Proceedings of the Cognitive Science Society* (Vol. 32, pp. 2769–2774).

Bishop, D. V. (2006). What causes specific language impairment in children? *Current Directions in Psychological Science, 15*(5), 217–221. https://doi.org/10.1111/j.1467-8721.2006.00439.x

Brackenbury, T., & Pye, C. (2005). Semantic deficits in children with language impairments. *Language, Speech, and Hearing Services in Schools, 36*, 5–16. https://doi.org/10.1044/0161-1461(2005/002)

Broedelet, I., Boersma, P., & Rispens, J. (2022). School-aged children learn novel categories on the basis of distributional information. *Frontiers in Psychology*, 12. https://doi.org/10.3389/fpsyg.2021.799241

Chládková, K., Boersma, P., & Escudero, P. (2022). Unattended distributional training can shift phoneme boundaries. *Bilingualism: Language and Cognition*, 1–14. https://doi.org/10.1017/S1366728922000086

Collisson, B. A., Grela, B., Spaulding, T., Rueckl, J. G., & Magnuson, J. S. (2015). Individual differences in the shape bias in preschool children with specific language impairment and typical language development: Theoretical and clinical implications. *Developmental Science, 18*(3), 373–388. https://doi.org/10.1111/desc.12219

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: word predictability is the key. *Cognition, 114*(3), 356–371. https://doi.org/10.1016/j.cognition.2009.10.009

Dockrell, J. E., Messer, D., & George, R. (2001). Patterns of naming objects and actions in children with word finding difficulties. *Language and Cognitive Processes, 16*(2-3), 261–286. https://doi.org/10.1080/01690960042000030

Dockrell, J. E., Messer, D., George, R., & Ralli, A. (2003). Beyond naming patterns in children with WFDs—Definitions for nouns and verbs. *Journal of Neurolinguistics, 16*(2), 191–211. https://doi.org/10.1016/S0911-6044(02)00012-X

Drljan, B., & Vukovic, M. (2019). Comparison of lexical-semantic processing in children with developmental language disorder and typically developing peers. *Govor, 36*(2), 119–138. https://doi.org/10.22210/govor.2019.36.07

Eadie, P., Conway, L., Hallenstein, B., Mensah, F., McKean, C., & Reilly, S. (2018). Quality of life in children with developmental language disorder. *International Journal of Language & Communication Disorders*, 1–12. https://doi.org/10.1111/1460-6984.12385

Ebert, K. D., & Kohnert, K. (2011). Sustained attention in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2011/10-0231)

Ellis, E. M., Gonzalez, M. R., & Deák, G. O. (2014). Visual prediction in infancy: what is the association with later vocabulary? *Language Learning and Development, 10*(1), 36–50. https://doi.org/10.1080/15475441.2013.799988

Escudero, P., Benders, T., & Wanrooij, K. (2011). Enhanced bimodal distributions facilitate the learning of second language vowels. *The Journal of the Acoustical Society of America, 130*(4), EL206–EL212. https://doi.org/10.1121/1.3629144

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52*(2), 321–335. https://doi.org/10.1044/1092-4388(2009/07-0189)

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*(4), 513. https://doi.org/10.1017.S0142716406060383

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33*(6), 966–79. https://doi.org/10.1037/0012-1649.33.6.966

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological Memory and Vocabulary Development During the Early School Years: A Longitudinal Study. *Developmental Psychology, 28*(5), 887–898. https://doi.org/10.1037/0012-1649.28.5.887

Goujon, A. (2011). Categorical implicit learning in real-world scenes: evidence from contextual cueing. *Quarterly Journal of Experimental Psychology (2006), 64*(5), 920–41. https://doi.org/10.1080/17470218.2010.526231

Haebig, E., Saffran, J. R., & Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry*, 1251–1263. https://doi.org/10.1111/jcpp.12734

Hamrick, P., Lum, J. A., & Ullman, M. T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences*, 1–6. https:///doi.org/10.1073/pnas.1713975115

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research, 23*(1), 65–94. https://doi.org/10.1177/0267658307071601

Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Veríssimo, J., Dye, C. D., ... Bruce Tomblin, J. (2011). Grammar predicts procedural learning and consolidation deficits in children with Specific Language Impairment. *Research in Developmental Disabilities*, 2362–2375. https://doi.org/10.1016/j.ridd.2011.07.026

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child development, 74*(5), 1368-1378. https://doi.org/10.1111/1467-8624.00612

Hsu, H. J., & Bishop, D. V. (2010). Grammatical difficulties in children with specific language impairment: Is learning deficient? *Human Development, 53*(5), 264–277. https://doi.org/10.1159/000321289

Hsu, H. J., Tomblin, J. B., & Christiansen, M. H. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Frontiers in Psychology, 5*, 1–10. https://doi.org/10.3389/fpsyg.2014.00175

Junge, C., van Rooijen, R., & Raijmakers, M. (2018). Distributional Information Shapes Infants' Categorization of Objects. *Infancy, 23*(6), 917–926. https://doi.org/10.1111/infa.12258

Kachergis, G., Yu, C., & Shiffrin, R. M. (2014). Cross-situational word learning is both implicit and strategic. *Frontiers in Psychology, 5*, 1–10. https://doi.org/10.3389/fpsyg.2014.00588

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: a meta-analysis. *Journal of Speech, Language, and Hearing Research : JSLHR, 53*(3), 739–756. https://doi.org/10.1044/1092-4388(2009/08-0248)

Kapa, L. L., & Erikson, J. A. (2020). The relationship between word learning and executive function in preschoolers with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research, 63*(7), 2293–2307. https://doi.org/10.1044/2020_JSLHR-19-00342

Kaufman, S. B., Deyoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*(3), 321–340. https://doi.org/10.1016/j.cognition.2010.05.011

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology, 48*(1), 171. https://doi.org/10.1037/a0025405

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development, 87*(1), 184–193. https://doi.org/10.1111/cdev.12461

Kraemer, H. C., & Blasey, C. M. (2004). Centring in regression analyses: a strategy to prevent errors in statistical inference. *International Journal of Methods in Psychiatric Research, 13*(3), 141–51. https://doi.org/10.1002/mpr.170

Lahey, M., & Edwards, J. (1999). Naming errors of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*(1), 195–205. https://doi.org/10.1044/jslhr.4201.195

Lammertink, I., Boersma, P., Rispens, J., & Wijnen, F. (2020). Visual statistical learning in children with and without DLD and its relation to literacy in children with DLD. *Reading and Writing*, 1–33. https://doi.org/10.1007/s11145-020-10018-4

Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2017). Statistical Learning in Specific Language Impairment: A Meta-Analysis. *Journal of Speech, Language, and Hearing Research*, 1–13. https://doi.org/10.1044/2017_JSLHR-L-16-0439

Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2019). Children with Developmental Language Disorder have an auditory verbal statistical learning deficit: evidence from an online measure. *Language Learning*. https://doi.org/10.1111/lang.12373

Leonard, L. B. (2014). *Children with specific language impairment*. MIT press.

Leonard, L. B., Nippold, M. A., Kail, R., & Hale, C. A. (1983). Picture naming in language-impaired children. *Journal of Speech and Hearing Research, 26*(4), 609–15. https://doi.org/10.1044/jshr.2604.609

Lian, A. (2017). Statistical learning and developmental language impairments. *Scandinavian Psychologist, 4*, 1–16. https://doi.org/10.1057/978-1-137-58746-6_8

Lukács, A., & Kemény, F. (2014). Domain-general sequence learning deficit in specific language impairment. *Neuropsychology, 28*(3), 472. https://doi.org/10.1037/neu0000052

Lum, J. A., & 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, 33*(4), 282–297. https://doi.org/10.1097/01.TLD.0000437939.01237.6a

Lum, J. A., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): a meta-analysis of serial reaction time task performance. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 51*, 1–10. https://doi.org/10.1016/j.cortex.2013.10.011

Mainela-Arnold, E., & Evans, J. L. (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without SLI? *Journal of Child Language, 41*(2), 327–351. https://doi.org/10.1017/S0305000912000736

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*(6), 1742–1756. https://doi.org/10.1044/1092-4388(2010/08-0198)

Marinellie, S. A., & Johnson, C. J. (2002). Definitional skill in school-age children with specific language impairment. *Journal of Communication Disorders, 35*(3), 241–259. https://doi.org/10.1016/S0021-9924(02)00056-4

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11*(1), 122–134. https://doi.org/10.1111/j.1467-7687.2007.00653.x
Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101–B111. https://doi.org/10.1016/s0010-0277(01)00157-3

McGregor, K. K. (1997). The nature of word-finding errors of preschoolers with and without word-finding deficits. *Journal of Speech, Language, and Hearing Research, 40*(6), 1232–1244. https://doi.org/10.1044/jslhr.4006.1232

McGregor, K. K., & Appel, A. (2002). On the relation between mental representation and naming in a child with specific language impairment. *Clinical Linguistics & Phonetics, 16*(1), 1–20. https://doi.org/10.1080/02699200110085034

McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A., J. & Lloyd, M. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *Journal of Autism and Developmental Disorders, 42*(1), 35–47. https://doi.org/10.1007/s10803-011-1210-4

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2002/081)

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language \& Communication Disorders, 48*(3), 307–319. https://doi.org/10.1111/1460-6984.12008

Melby-Lervåg, M., Lervåg, A., Lyster, S.-A. H., Klem, M., Hagtvet, B., & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychological Science, 23*(10), 1092–1098. https://doi.org/10.1177/0956797612443833

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*(1), 3. https://doi.org/10.1037/0022-0663.91.1.3

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology, 1*, 1–9. https://doi.org/10.3389/fpsyg.2010.00031

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/1058-0360(2009/09-0028)

Nash, M., & Donaldson, M. L. (2005). Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research, 48*(2), 439–458. https://doi.org/10.1044/1092-4388(2005/030)

Nation, K. (2014). Lexical learning and lexical processing in children with developmental language impairments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*, 1–10. https://doi.org/10.1098/rstb.2012.0387

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology, 42*(4), 643–655. https://doi.org/10.1037/0012-1649.42.4.643

Noonan, N. B. (2018). *Exploring the Process of Statistical Language Learning*.

Nunnally, J.C., & Bernstein, I.H. (1994). The Assessment of Reliability. *Psychometric Theory*, *3*, 248-292.

Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Frontiers in Psychology*, *7*, 1–18. https://doi.org/10.3389/fpsyg.2016.01245

Paciorek, A., & Williams, J. N. (2015). Semantic implicit learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 67–88). John Benjamins.

Palmeri, T. J., & Gauthier, I. (2004). Visual object understanding. *Nature Reviews Neuroscience*, *5*(4), 291–303. https://doi.org/10.1038/nrn1364

Plante, E., Gómez, R., & Gerken, L. (2002). Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders*, *35*(5), 453–462. https://doi.org/10.1016/s0021-9924(02)00094-1

Plunkett, K., Hu, J.-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, *106*(2), 665–81. https://doi.org/10.1016/j.cognition.2007.04.003

Psychology Software Tools, Inc. (2016). E-Prime 3.0. [Computer program].
Raven, J., Raven, J.C., & Court, J.H. (2003). Manual for Raven's progressive matrices and vocabulary scales. San Antonia: Harcourt Assessment.

R Core Team (2020). A language and environment for statistical computing. Vienna, Austria. https://www.r-project.org/. [Computer program].

Rispens, J., & Baker, A. (2012). Nonword repetition: The relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2011/10-0263)

Rogers, L., Park, S. H., & Vickery, T. (2020). *Visual Statistical Learning Is Modulated by Arbitrary and Natural Categories.* PsyArXiv. https://doi.org/10.31234/osf.io/9ca28

Saffran, J. R. (2018). Statistical learning as a window into developmental disabilities. *Journal of Neurodevelopmental Disorders*, *10*(1), 35. https://doi.org/10.1186/s11689-018-9252-y

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Sandgren, O., Salameh, E.-K., Nettelbladt, U., Dahlgren-Sandberg, A., & Andersson, K. (2020). Using a word association task to investigate semantic depth in swedish-speaking children with developmental language disorder. *Logopedics Phoniatrics Vocology*, 1–7. https://doi.org/10.1080/14015439.2020.1785001

Sanjeevan, T., & Mainela-Arnold, E. (2019). Characterizing the Motor Skills in Children with Specific Language Impairment. *Folia Phoniatrica et Logopaedica, 71*(1), 42–55. https://doi.org/10.1159/000493262

Schlichting, L. (2005). Peabody Picture Vocabulary Test-III-NL [Measurement instrument]. Amsterdam, The Netherlands: Harcourt.

Semel, E., Wiig, E., & Secord, W. (2010). Clinical evaluation of language fundamentals: Dutch Version (W. Kort, E. Compaan, M. Schittekatte, & P. Dekker, Trans.; Third Edition.) [Measurement instrument]. Amsterdam, The Netherlands: Pearson.

Shafto, C. L., Conway, C. M., Field, S. L., & Houston, D. M. (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy, 17*(3), 247–271. https://doi.org/10.1111/j.1532-7078.2011.00085.x

Sheng, L., & McGregor, K. K. (2010). Lexical-semantic organization in children with specific language impairment. *Journal of Speech, Language, and Hearing Research : JSLHR, 53*(1), 146–159. https://doi.org/10.1044/1092-4388(2009/08-0160)

Siegelman, N. (2020). Statistical learning abilities and their relation to language. *Language and Linguistics Compass, 14*(3), e12365. https://doi.org/10.1111/lnc3.12365

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B, 372*(1711), 1–10. http://doi.org/10.1098/rstb.2016.0059

Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods, 49*(2), 418–432. https://doi.org/10.3758/s13428-016-0719-z

Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2017). Redefining "learning" in statistical learning: what does an online measure reveal about the assimilation of visual regularities? *Cognitive Science*, 1–36. https://doi.org/10.1111/cogs.12556

Singh, L., Steven Reznick, J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science, 15*(4), 482–95. https://doi.org/10.1111/j.1467-7687.2012.01141.x

Snowling, M.J., Hayiou-Thomas, M.E., Nash, H.M., & Hulme, C. (2020). Dyslexia and Developmental Language Disorder: comorbid disorders with distinct effects on reading comprehension. *Journal of Child Psychology and Psychiatry, 61*(6), 672-680. https://doi.org/10.1111/jcpp.13140.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Sociaal en Cultureel Planbureau (2018). *Statusscores 2016* [report Social and Cultural Planning Bureau]. No longer available.

Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and Writing, 28*(4), 467–490. https://doi.org/10.1007/s11145-014-9533-0

Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology, 126*, 395–411. http://doi.org/10.1016/j.jecp.2014.06.0030022-0965/?

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development, 3*(4), 269–293. https://doi.org/10.1080/15475440701377477

Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology. General, 134*(4), 552–64. https://doi.org/10.1037/0096-3445.134.4.552

Ullman, M. T. (2016). The declarative/procedural model: a neurobiological model of language learning, knowledge and use. In G. Hickock & S. A. Small (Eds.), *The neurobiology of language* (pp. 953–968). Elsevier.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41*(3), 399–433. https://doi.org/10.1016/S0010-9452(08)70276-4

Vandermosten, M., Wouters, J., Ghesquière, P., & Golestani, N. (2019). Statistical learning of speech sounds in dyslexic and typical reading children. *Scientific Studies of Reading, 23*(1), 116–127. https://doi.org/10.1080/10888438.2018.1473404

Wanrooij, K., Boersma, P., & Benders, T. (2015). Observed effects of "distributional learning" may not relate to the number of peaks. A test of "dispersion" as a confounding factor. *Frontiers in Psychology, 6*, 1341. https://doi.org/10.3389/fpsyg.2015.01341

Wanrooij, K., Boersma, P., & van Zuijen, T. L. (2014). Fast phonetic learning occurs already in 2-to-3-month old infants: an ERP study. *Frontiers in Psychology, 5*, 77. https://doi.org/10.3389/fpsyg.2014.00077

Westby, C. (2019). Depression in Children With Developmental Language Disorders. *Word of Mouth, 30*(4), 1–4. https://doi.org/10.1177/1048395019833703

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology, 47*(5), 1220–9. https://doi.org/10.1037/a0024023

Wu, R., Gopnik, A., Richardson, D., & Kirkham, N. (2010). Social cues support learning about objects from statistics in infancy. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 32).

Younger, B. A. (1985). The segregation of items into categories by ten-month-old infants. *Child Development, 56*(6), 1574–83.

Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental Science, 14*(2), 165–180. https://doi.org/10.1111/j.1467-7687.2010.00958.x