Cross-situational Statistical learning in children with Developmental Language Disorder

Nadia Ahufinger[1*], Ernesto Guerra[2], Laura Ferinu[1], Llorenç Andreu[1]
& Mònica Sanz-Torrent[3]

[1]Universitat Oberta de Catalunya
[2]Universidad de Chile
[3]Universitat de Barcelona

[1]Universitat Oberta de Catalunya:

Rambla del Poblenou, 156, 08018 Barcelona (Spain)

- Nadia Ahufinger: nadiahufinger@uoc.edu **(Corresponding author)***
- Laura Ferinu: lferinu@uoc.edu
- Llorenç Andreu: landreub@uoc.edu

[2]Universitadad de Chile:

Av. Libertador Bernardo O'Higgins 1058, Santiago de Chile (Chile)

- Ernesto Guerra: ernesto.guerra@ciae.uchile.cl

[3]Universitat de Barcelona

Passeig de la Vall d'Hebron, 171, 08035 Barcelona (Spain)

- Mònica Sanz-Torrent: monicasanz@ub.edu

Funding: This work was supported by the Ministerio de Economía y Competitividad [2016EDU2016-75368-P y BES-2014-070511], the Agència de Gestió d’Ajuts Universitaris i de Recerca (AGAUR) de la Generalitat de Catalunya [2017SGR387] and support from ANID/PIA/Basal Funds for Centers of Excellence FB0003.

Abstract

Children with Developmental Language Disorder (DLD) need more exposures to learn new words in an unambiguous context compared to children with typical development (TD). However, it remains unclear whether they would be able to learn new words by extracting frequencies over multiple word-object encounters in ambiguous situations. The present study examines this question through a cross-situational statistical-learning task (CSSL). Thirty-eight school-aged children with DLD and thirty-eight age/sex-matched TD children completed a CSSL eye-tracking experiment. Participants' responses show that children with DLD had significantly poorer accuracy compared to TD children. However, both groups performed above chance. While the eye-tracking record evidenced no distinctive pattern between groups as children learnt new words, we observed a larger target visual preference in TD children when they were asked to find the referent for those new words. We discuss these findings in light of existing accounts for memory and language deficits in DLD.

Keywords: Word learning, Cross-situational statistical learning (CSSL); Developmental Language Disorder (DLD); Eye-tracking.

Introduction

Lexical acquisition is a complex cognitive process that requires different steps for words to be acquired. The first step required for vocabulary learning is the discovery of the sounds and phonological structure of words, that is, words have to be segmented from the oral fluent speech that is heard. The next step is to link each word (i.e., phonological label) to a visual referent (e.g., object). Research has shown that children have the ability to establish an initial link between a new word and its referent, even after a single exposure, or a few exposures. However, some clinical populations, such as children with developmental language disorder (DLD), exhibit a later and poorer acquisition of new words compared to the typically developing (TD) population.

DLD, previously known as Specific Language Impairment or SLI, is a neurodevelopmental language disorder that affects about 7.5% of the general population (Norbury et al., 2016). Children with DLD have a severe and persistent neurodevelopmental disorder in the acquisition and development of oral language, not associated with a medical condition (Bishop et al., 2016, 2017). The disorder may involve one or more components of language to different degrees in both language production (e.g., Bishop, 1979; Sanz-Torrent, Serrat et al., 2008; Schuele & Nicholls, 2000; Owen & Leonard, 2006; van der Lely, 1996) and comprehension (e.g., Montgomery & Evans, 2009; Robertson & Joanisse, 2010; van der Lely & Harris, 1990). Although there is agreement that the most affected language component in children with DLD is morphosyntax (Moscati et al., 2020; van der Lely, 1998; van der Lely et al., 2004), word-learning deficits in this population are also well-established, and may affect the acquisition, storage, and retrieval of new words. For example, children with DLD have lower vocabulary test scores (Gray et al., 1999) and show reduced receptive word learning in naturalistic contexts (Rice et al., 1994). In comparison to TD peers, children with DLD use high frequency verbs and nouns more often, and they are slower to acquire functional words (Eyer & Leonard, 1995; Leonard, 1995). They also present weaker semantic representations of words, as evidenced by their shallower and sparser vocabulary (Andreu et al., 2012; Kail et al., 1984; McGregor et al., 2002; 2011), and have smaller vocabularies than expected for their age (Rice et al., 1990). In addition, it has been shown that they have reduced sensitivity to phonological and semantic features of words (Alt & Plante, 2006).

Different studies suggest that children with DLD have difficulties in mapping labels to new objects (Dollaghan, 1987; Ellis Weismer & Hesketh, 1993, 1996; Gray, 2003, 2004, 2006, 2011; Kiernan & Gray 1998; Rice et al., 1992), that they typically require more exposures and find it more difficult to learn new lexical labels in comparison with TD children (Alt & Plante, 2006; Gray, 2004; Rice et al., 1994). Moreover, they do not fast map non-verbal semantic features associated with lexical labels as TD children (Alt et al., 2004; Rice et al., 1994). Although different methods have been used to study word-learning in the DLD (see Jackson et al., 2019), most studies have assessed fast-mapping abilities through a task in which an exposure phase consists in the presentations of novel words and visual referents pairs, followed by a testing phase that assesses their knowledge about the correct visual referent of words (Alt & Plante, 2006; Dollaghan, 1987; Haebig et al., 2017; McGregor, 2020; Rice et al.; 1990; 1992; 1994). Difficulties in children with DLD have been observed in both recognition test tasks, where children must choose, among many, the correct visual referent for a phonological label previously heard (i.e., Alt et al., 2004, Alt & Plante, 2006), and production test tasks (e.g., Jackson et al., 2016) where children are asked to name different newly learned objects.

Different theories have attempted to explain the deficits in word-learning in DLD. Some argue for a generalized processing deficit (Nation, 2014), while others suggest more specific ones (e.g., phonological processing deficits; Ellis Weismer & Hesketh 1998). It has also been proposed that difficulties to learn new words in this population could be due to poor storage of information in phonological working memory (Alt, 2011; Jackson et al., 2016; Montgomery, 1995). In fact, it has been shown that word learning difficulties can result from a deficient working memory encoding as children with DLD need more exposures to the word-referents pair to improve learning in comparison to children with TD (Gray 2003, 2004, 2005, 2006; Kiernan & Gray 1998; Rice et al. 1994). However, the optimal number of word-referent pairs presentations that children with DLD need to compensate for working memory problems has not yet been established.

Word learning in ambiguous contexts: the cross situational statistical learning task

Although new phonological labels usually appear together with their referent in naturalistic contexts, it is unlikely that they appear in isolation and the absence of ambiguity. Instead, names of new objects typically appear situated in a context with several other objects as well. In this regard, a growing body of research has explored whether a more complex mapping task can approach a real-world scenario in which children constantly learn new words through an unconscious mechanism. This mechanism is based on the tracking frequencies over multiple encounters with different words and possible referents in ambiguous situations. With this in mind, Yu and Smith (2007) designed the cross-situational statistical learning task (CSSL), which assesses the ability to learn arbitrary mappings between sounds and referents. Participants are exposed to a set of trials, each containing multiple spoken words and multiple pictures of individual objects, where the same number of words and pictures are presented (e.g., two words and two pictures, 2x2) on every trial. There is no information about the correspondence between the word-picture pairs within a trial. Yet, the authors suggested participants learn the word-picture mappings through cross-trial statistical relations. In this sense, the CSSL task was designed to assess one kind of statistical word learning ability.

In Yu and Smith (2007), a group of adults learnt new words and their referents in a CSSL task, as shown by an alternative forced choice test (AFC), where participants performed above chance and correctly identified up to 88% of the referents in the easiest (2x2) experiment and 53% in the hardest experiment (4x4). Different studies have since shown successful CSSL in adults (Fitneva & Christiansen, 2011; Yu et al., 2012; Yurovsky et al., 2012) and infants (Smith & Yu, 2008). A previous study (Suanda et al., 2014), tested typically developing 5- to 7-year-old children in a CSSL task. On each trial, the experiment presented two depicted objects in a visual context (simultaneously) and two auditory stimuli (sequentially). Participants were told that they would learn the names of new toys, yet no explicit cue with regard to the correspondence between sound and object was given within a trial. The results of a post-exposure test (i.e., 4-AFC) showed that typically developing children did learn the new labels of the objects by observing the co-occurrence regularities across these ambiguous naming events. To prove that, children were divided into three different conditions according to the contextual diversity of the learning environment (high, moderate, or low). That means that each contextual diversity condition was created by considering that the accompanying word-picture pair for any given word-picture pairing was always different (high condition) or in some cases repeated (moderate and low conditions). Using this design, the authors demonstrated that children were using the computation of the embedded frequencies (i.e., statistical learning) to learn that the highest frequencies were correct word-picture matches and to reject the less frequent co-occurrent pairs. While this is true for TD school-age children, no studies have yet addressed this issue in a population known to have word learning difficulties, such as DLD.

Statistical learning has been broadly defined as the incidental ability and sensitivity to detect regularities from the environment (Arciuli & Conway, 2018). Originally, Saffran et al. (1996), examined word segmentation through statistical learning (described as *statistical word learning*) and suggested that learners, including infants, may detect word boundaries, in part, by tracking the statistical properties of the sound combinations that they hear. Since then, several subsequent studies have addressed statistical learning in tasks that involve learning of sequential structure such as artificial grammar learning (e.g., Reber, 1967; Onnis, et al., 2003), serial reaction time task (SRT; Lum et al., 2019; Robertson, 2007), or visual shape sequences (Fiser & Aslin, 2002). However, after years of research on statistical learning, it has been demonstrated that not all statistical learning abilities are related to sequential stimuli only. Instead, statistical learning can also involve the extraction of distributional probabilities related to the frequency and variability of exemplars in the input, as it is the case CSSL task.

Statistical learning is closely linked to other constructs such as *procedural memory*, *implicit memory* and *implicit learning*. Sawi and Rueckl (2019) show in their revision the similarities and differences of these three constructs, which are part of different frameworks, used to dichotomize memory systems. They described the critical concepts as:

- **Declarative/procedural memory.** “Characterized by dependence on specific anatomical regions associated of conscious [declarative] and unconscious [procedural] access” (p.7, see also Squire, 2004; Ullman, 2004)
- **Explicit/implicit memory.** “Associated to memory retrieval processes that can occur intentionally [explicit] or incidentally [implicit]” (p.7, see also Schacter, 1987)
- **Explicit/implicit learning.** “Related to the processes involved in the initial encoding and storage of information that can be deliberated and occurs on the basis of a single event [explicit] or incidentally after extended practice [implicit]” (p.7, see also Reber et al., 2003).

In their review, Sawi and Rueckl (2019) recognized that even though these descriptions are conceptually different, there is a substantial overlap between these different frameworks. Thus, they opted for using a nomenclature that unifies the procedural memory, implicit memory, and implicit learning under the label “implicit/procedural memory” (IPM) and the declarative memory, explicit memory, and explicit learning under the label “explicit/declarative memory” (EDM). We shall adopt the same nomenclature in the present paper.

Many authors assume that statistical learning is an IPM process (e.g., Conway & Christiansen, 2006; Perruchet & Pacton, 2006; Thiessen, 2017), because most of the statistical learning tasks do not give explicit instructions; participants are not consciously aware of the patterns in the input (e.g., Fiser & Aslin, 2001) or they are presented with a cover task to get distracted while a pattern is presented (e.g., Arciuli & Simpson, 2011; Evans et al., 2009). Other authors, instead, suggest that multiple memory systems support statistical learning, because participants can retain statistical patterns learned after a single exposure (Durrant et al., 2011; Kim et al., 2009), and because explicit instructions can facilitate performance due to the encouragement of participants to attend to regularities in the input (e.g., Arciuli et al., 2014; Batterink et al., 2015; Gómez, 2017; Hamrick & Rebuschat, 2012). Thus, some authors have questioned the assumption that participants rely only in the IPM to learn new words in the CSSL task (e.g., Berens, et al., 2018).

Currently there are two opposite existing accounts that attempt to explain which kind of processes of learning are implicated in the CSSL task. On one hand, there is the *gradual associative account*, based on multiple hypotheses done through implicit statistical computations (thus, the IPM is assumed to be involved) where participants aggregate information over time (Chen et al., 2018; Fitneva & Christiansen, 2011; Monaghan et al., 2015). This is the original conceptualization of the task, which assumes that the learning of pairing involves the accumulation of word-object-occurrence statistics across the training trials. On the other hand, the *hypothesis testing account* based on a type of learning that is more in accord with intentional and strategic learning processes (thus, presumably involving the EDM). This account assumes that learning is based on a single hypothesis that participants entertain on each exposure to words and referents (Bloom, 2000; Trueswell et al., 2013). Therefore, subsequent exposures could trigger the rejection or confirmation of previous hypotheses based on the additional evidence.

However, findings in the literature regarding the gradual associative account or hypothesis testing account for CSSL are not definitive. There is growing evidence that CSSL needs both IPM and EDM mechanisms to be solved. Evidence suggests that these two processes are not mutually exclusive (Kachergis et al., 2014; Smith et al., 2011; Yurovsky & Frank, 2015, Warren et al., 2020) and that in one form or another, the learning of pairing in a CSSL involves the accumulation of word-object co-occurrence statistics across the training trials in addition to the intentional selection of particular pairs for storage and testing (Kachergis et al., 2014).

Statistical learning and DLD

Since Ullman and Pierpont (2005) proposed the Procedural Deficit Hypothesis (PDH) to attempt to explain language difficulties in children with DLD, statistical learning has been measured in different ways in this population. The PDH draws from the dual declarative/procedural memory neurological system (Ullman, 2001), initially proposing that vocabulary acquisition and semantic knowledge is supported by the declarative memory system (EDM) that stores word-specific knowledge while grammar, syntax and phonology are supported by the procedural memory system (IPM).

According to the PDH, the procedural memory system is involved in implicit learning, control, and memorization and execution of motor and cognitive skills. It is "particularly important for acquiring and performing skills involving sequences—whether the sequences are serial or abstract, or sensory-motor or cognitive" (Ullman & Pierpont, 2005, p. 401). The PDH proposes that IPM is affected in children with DLD and it may explain most of their common grammar difficulties. By contrast, according to this hypothesis the EDM remains largely intact in children with DLD, since vocabulary is less affected compared to morphosyntax development in this population (e.g., Clarke & Leonard, 1996; Ellis Weismer & Hesketh, 1996; Rice, et al., 2008). However, the PDH also recognizes that some aspects of lexical learning and knowledge difficulties in DLD appear to rely on IPM deficits. In this regard, in the updated version of the PDH, Ullman et al. (2020) argued that lexical deficits are predicted by the level of lexical-phonological sequential information that can impact in spoken word recognition, word learning, lexical retrieval, and sensitivity to word frequency. In this view, IPM deficits might be related to the capacity to hold in memory the sequential order of phonemes and syllables extracted from the speech stream to implicitly track and compute the probabilities of adjacent sounds for a successful word segmentation. Consequently, in this hypothesis, statistical word learning tasks are assumed to depend on the IPM system.

Another key aspect of the PDH is the assumption that EDM can function as a *compensatory mechanism* in children with DLD. Specifically, it is argued that the degree to which the language deficits improve as children with DLD mature, is in part due to the EDM compensation when learning language (Ullman & Pullman, 2015). The vast majority of studies that have examined IPM deficits in children with DLD have used the visuospatial-sequential SRT task (see Hedenius et al., 2011; Hsu & Bishop, 2010, Lum & Bleses, 2012; Lum, et al., 2010; Lum et al., 2014; Tomblin et al., 2007) and most of these studies have shown lower levels of learning of sequential patterns in relation to typical population. Yet, studies assessing word learning through statistical learning (Saffran et al., 1996) in children with DLD are scarce. The few that exist have evaluated the ability of children with and without DLD to extract linguistic and non-linguistic transitional probabilities from a continuous stream of auditory stimuli (see Ahufinger et al., under review; Evans et al., 2009; Haebig et al., 2017; Mayor-Dubois et al., 2014).

These studies have shown that children with DLD performed less accurately than TD children evidencing difficulties to successfully discriminate words from nonwords in a 2-AFC test. Moreover, previous statistical word learning research in DLD has relied on response accuracy only (thus, observing the end-product of a learning process) and has not yet examined the learning process as it occurs.

Only a few experiments have used online measures, such as eye-tracking, to assess the learning processes during CSSL tasks. These studies report contrasting findings, yet they are difficult to compare, in part because these studies differed in the age groups they evaluate, the dependent variables they use, and the research question they were aiming to answer. For instance, Fitneva and Christiansen (2011) wanted to explore whether accuracy of initial word-referent associations is critical for word learning in a CSSL task. They reported that a group of adults who had *worse* initial mapping (as reflected in their gaze pattern) performed *better* in a 2-AFC subsequent test. By contrast, Yu et al., (2012) found that there were no differences in the initial mapping between the groups of adults that performed a 4x4 CSSL task. That is, participants that acquired fewer words in the test-phase task did not show differences in their looking behaviors at the beginning of training compared to those people that had better performance in the test phase.

In other study, Yu and Smith (2011) wanted to explore through different eye-tracking measures (e.g., looking duration and shift rate trial by trial) whether there were differences in the underlying mechanisms displayed to solve a CSSL between infants (14-month-old) who learn more words (i.e. strong learners) and those who learn fewer words (i.e. weak learners). They found that infants who exhibited a *more uneven* distribution of looking duration, containing longer fixations and several shorter attention switches, were *less* accurate in identifying the correct referent in a post-exposure test (a measure that was derived from the percentage of looks to the target). The authors interpreted their results as reflecting more stable patterns of visual exploration in strong learners. More recently, Venker (2019) investigated word learning in a CSSL task in children (2 to 7 years old) with autism spectrum disorder (ASD) and TD children, matched on vocabulary knowledge. In two experiments, they compared the performance between groups in a fast-mapping task (i.e., only a single label-object pairing at a time in the exposure phase, Experiment 1) and in a 2x2 word learning CSSL task (Experiment 2). The online eye-tracking measure was applied to a 2AFC test phase where the author examined whether children looked more often at the object that was named on each test trial, as opposed to the object that was not named. Results showed that both groups looked to the named image above chance level and no differences between proportion of look to the named objects were found between groups (TD= .60 / ASD = .61). To the best of our knowledge, there are no studies that have explored online behavior in a CSSL in TD school age children (i.e., from 6 to 12) nor children with DLD.

Current study

While sequential statistical learning is key to the child's ability to discover the lexical-phonological form from a stream of speech, word learning requires more than just segmenting words from a fluent speech, e.g., mapping phonological forms to visual objects. In the present study we aim to extend the research on word learning difficulties in children with DLD by examining their ability to map words in an ambiguous context, which does not contain embedded sequential patterns but rather a variability of frequency on label-object pairings. Specifically, we aim to investigate how children with DLD perform in a 2x2 CSSL compared to TD children. Indeed, there is an open debate about the mechanisms implicated in solving the CSSL task, and we do not intent to disentangle the implication of the IPM and the EDM system in this process. Our question is, instead, whether children with DLD show significantly lower accuracy in the task compared to TD children. If so, we expect that the group of TD children, but not the group of children with DLD, will identify that the labels and objects that appear more frequently together correspond to the correct word-picture pairs and reject the labels and objects that appeared together less frequently as incorrect word-picture pairs. To this end, we manipulated the contextual diversity (a feature of the learning context that is cross-situational in nature, see Suanda et al., 2014), since this manipulation could only affect learning when children extract the frequencies across trials.

In addition to the coarse behavioral measure, we also recorded participants' eye movements, which allow the continuous monitoring of participants' gaze behavior as they learn the mapping between words and pictures, as well as their gaze behavior while they answer the test phase (4-AFC). Concretely, in the exposure phase we evaluated whether participants would synchronize their gaze to target objects in the visual context with the auditory stimuli (cf. Yu & Smith, 2011). We expect a stronger synchronization (i.e., higher proportion of looks to the target) in the TD group compared to the DLD group. During the testing, and as far as gaze behavior is in line with children's explicit response, we should also see a larger preference for the target object in the TD group compared with the DLD group. Relating the gaze behavior with the accuracy of the groups should reflect how confident children were of their choice when answering correctly. In other words, gaze preference could be interpreted as an index of certainty in children's response. Our analysis of the eye movements during the testing phase will allow us to distinguish the magnitude of the preference for the target object upon hearing the critical spoken word and shortly after, as well as the shape of this preference over time. We hypothesize that the more linear the increase in preference for the target, the more confident the group is in their response. In turn, the less linear the eye movements trajectory over time, the less confident the participants are.

3. Methods

3.1. Participants

A total of 140 children (79 children with language difficulties and 61 children with typical language development and a standard academic level for their age) were initially screened with a battery of tests by two trained researchers. These children were contacted and recruited through a number of local institutions and schools from different areas of Barcelona and the surrounding area (metropolitan area of Barcelona). All the participants were native simultaneous bilingual speakers of Spanish and Catalan (i.e., all of them were exposed to both languages from birth). According to the parental survey, all families of the participants in this study reported that their children speak both languages. In the school system, Catalan is the primary language of instruction. According to Alarcón and Garzón (2011), children in Barcelona are equally proficient in both Spanish and Catalan, although the use of Spanish is more popular. For further information about Catalan and Spanish bilingualism and DLD, see Sanz-Torrent, Badia et al. (2008).

After our inclusion/exclusion criteria was applied (see below), our final sample for the experimental task included 38 children with DLD (12 girls, mean age=8.7 years; SD=1.10 years, range=5.6-12.11 years), and 38 age- and sex-matched TD children (12 girls, mean age=8.9 years; SD=1.10 years, range=5.7-12.9 years).

Inclusion/exclusion criteria: The inclusion criteria for children with DLD were defined following the DLD diagnostic criteria recommended by a Spanish expert committee that reached a consensus in 2015 (Aguado et al., 2015): (a) a non-verbal intellectual quotient (NVIQ) > 75 (Kaufmann Brief Intelligence Test Matrices section; K-BIT Mat; Kaufman & Kaufman, 2004); (b) a score of 1.25 SD below the mean on one of the three scales of the Clinical Evaluation of Language Fundamentals - Fourth Edition, Spanish (CELF-4, Semel et al., 2006): core language, expressive language, receptive language scales (because there are no normative language tests for the Catalan language, the standard clinical practice in Catalonia is to present all the stimuli of CELF-4 in Spanish, but if children answer correctly in Catalan, they are given credit for their answer. Children in this study were assessed following this protocol) (c) normal hearing at 500, 1000, 2000, and 4000 Hz at 20 dB based on the on the American Speech-Language Hearing Association (ASHA) 1997 guidelines for hearing screening; (d) normal or corrected-to-normal vision; (e) normal oral and speech motor abilities; and (f) absence of other medical or neurological conditions. The information about vision and the presence or absence of other medical or neurological conditions was provided by families through a background information questionnaire. With respect to oral structure and motor function, speech and language therapists examined the children to assess the shape, size, and motor function of the speech organs, both active (tongue, lips, and jaw) and passive (buccal cavity, palate, and teeth), as well as respiratory dynamics, exhalation, and rhythm. Motor function was assessed according to a protocol that used different practical exercises to verify that mobility was normal.

The inclusion/exclusion criteria for the group of TD children were: (a) a NVIQ >75 (K-BIT Mat, Kaufman & Kaufman, 2004); (b) scores around the mean on three scales of the Clinical Evaluation of Language Fundamentals - Fourth Edition, Spanish (CELF-4, Semel et al., 2006): core language, expressive language, receptive language scales and (c) absence of prior history of speech or psychological therapy. The descriptive data of the two experimental groups are presented in Table 1. Each child with DLD was matched to a TD child of the same sex and age (+/- 3 months) at the time of the experimental tasks. The children in the control (TD) group were administered the same tests as those in the experimental group (DLD).

(Table 1 about here)

3.2. Material and Methods

3.2.1. Apparatus

The stimuli were presented in a 800 x 600 pixels format and appeared on the integrated 17" TFT monitor of the Tobii T120 Eye Tracker at an horizontal distance of approximately 22" from the eyes of the participant. Both the presentation of the stimuli and the collection of the eye movement data were carried out using Tobii Studio software. At the beginning of the experiment a calibration of 20 sec was carried out in order to validate the tracking and registration of the eye movement.

3.2.2. Stimuli

The stimuli for this study were adapted following the design used by Suanda et al. (2014) to render a task suitable for young children. Eight recorded bi-syllabic CV-CV non-words paired with eight pictures of robot-like-cartoons (see Figure 1A) resulted in eight to-be-learned word-object pairs during the exposure phase. The novel words were recorded by a native Catalan-Spanish speaker (i.e., “pimo”, “lasi”, “zepi”, “rile”, “teco”, “mepo”, “buna” and “datu”). Four additional novel word-object pairings were used for the practice phase (“bose”, “sime”, “coti”, “fela”; see Figure 1B).

(*Figure 1 about here)*

3.3. Experimental design

In the exposure phase, thirty-two instances of word-object pairings were presented. This phase consisted of 16 total exposure trials. On each learning trail, two spoken novel words and two potential referents were presented. Every object was always presented together with its assigned label, yet no information about the link between the word and referent was available on a single trial. Instead, this information could be extracted across trials. The study uses a moderate level of contextual diversity, which follows the design of Suanda et al. (2014) and assesses both the learning of the associations as well as the strength of the representations. Contextual diversity refers to the number of different pairs of objects presented across the trials. Across the 16 exposure trials, each word-object pair appeared together four times. Therefore, over the 16 trials, each word-object pair co-occurred with one word-object pair in two trials, and two different word-picture pairings on the other two trials (see Table 2).

(Table 2 about here)

Each visual object had a different role depending on the order of the presentation of spoken words. Thus, if the first word referred to the object on the left side, then that object assumed the role of target during the first time window (first spoken word). In turn, the object on the right assumed the competitor role in that first time window. On that same trial, the object on the left side assumed the role of competitor during the second time window (second spoken word) and the object on the left assumed the target role. In a similar way, the role of the objects switched from target to competitor (and vice versa) during the second time window, when the first word referred to the object on the left side during the first spoken word (see Figure 3A).

(Figure 2 about here)

In the testing phase, we evaluated whether participants learned the word-object pairs during the exposure phase. This phase consisted in 16 trials (considering that each word was assessed twice), in which participants were presented with four objects in the display and a single spoken word that referred to one of the objects (see Figure 2B). Participants were instructed to select a visual referent from four alternatives that corresponded to the spoken word by clicking on the picture with the computer mouse. On each testing trial, three kind of objects accompanied the target object. One object acted as strong competitor of the target since this object co-occurred with the spoken word twice during the exposure phase. Another object acted as weak competitor since it co-occurred only once together with the spoken word. Finally, the last object was a distractor since it never co-occurred with the spoken word during the exposure phase.

3.4. Procedure

All the families who agreed to participate in the study were asked to sign an informed consent form and fill in a background information questionnaire, which was approved by the Ethics Committee of the Universitat Oberta de Catalunya. A final report containing the results of all the tests administered to the children was given to the family as compensation for their participation to the study, and children received a toy as a gift.

3.4.1. Practice phase

Before the experiment began, the experimenter showed each participant a picture of a cartoon dog on the screen and said: ‘This is Bobby and he is really happy because he has new robot toys! We are going to learn the names of his robots! Look! These are Bobby’s new funny robot toys!’. A picture of the twelve cartoon robots appeared on the screen. Then, the experimenter said: ‘Ok, now you are going to hear all the names of the new toys’. An audio recording of the twelve novel words ensued, played in a random order after a black screen replaced the picture of the objects. Consequently, the experimenter said: ‘Ok, now we are going to learn which name goes with each robot. First you will see a green dot. Click on the green dot to see a picture of the toy and hear its name’.

Two exposure practice trials with a single picture of a robot each was presented. One second after picture onset, the corresponding spoken word was played. After those two practice trials, the experimenter said: ‘Let’s see if you learned the names of these new toys! You will see a picture with four of the new toys. When you see the green dot, I want you to click on it, then I want you to click on the picture of the toy that was named. I want you to click on the picture of that toy as quickly as you can. Are you ready to try some?’

Two test practice trials then commenced. Both trials consisted in the presentation of four objects on the screen followed by a single spoken word. Participants had to click on the object they thought corresponded to the word they heard. During these two practice test trials, the experimenter was allowed to reinforce the child's performance through comments and verbal feedback. After the practice phase finished, participants began with the experiment.

3.4.2. Exposure phase

The first phase started with the presentation of a picture of Bobby and the following instructions: 'Okay! You did it very well! Now we are going to learn all the names of Bobby's new robots! Just like before, you are going to click on the green dot to hear the names of the toys. Are you ready?'

Each exposure trial began with a green fixation point in the middle of the screen that had to be clicked on. After that, the two images were presented side by side on the screen simultaneously (see Figure 3). One second later, the two words were played consecutively with a 750 ms silence interval between them. After the 16 exposure trials, the experimenter introduced the testing phase in the same way as in the practice trials.

3.4.3. Testing phase

The testing phase consisted of the same procedure as in the practice test trials but included the 16 test trials. During the test trials, the experimenter was not allowed to provide any feedback or reinforcement while the participant was making her choices. The experimenter wrote down the child's performance on a sheet.

3.5. Data analysis

All data and analysis scripts are available online in https://osf.io/pht7u/.

3.5.1. Behavioral data

To assess the potential differences between the two groups of children, a Generalized Linear Mixed Model (GLMM) was applied with the participants' response to the 4-AFC during the testing phase, as the dependent variable (Binomial distribution, Logit link), and group (TD, DLD) and choice (Target, 50% Competitor, 25% Competitor, 0% Competitor) as fixed factors. For the group factor we used a treatment contrast to set the DLD group as intercept. For the choice factor, we rotated the intercept to each level also using a treatment contrast. Random intercept for subjects and its corresponding random slope for choice were included in the model. We further evaluated whether the mean accuracy for each group was above, below or at chance level (25%). To do this, we aggregated by participant, computing the proportion of test trials answered correctly, and then we conducted a one-sample *t*-test for each group against 25% accuracy.

3.5.2. Eye tracking data

For the exposure and the testing experimental trials, we defined two and four areas of interest, respectively. These areas corresponded to the location and size of the displayed pictures. We obtained the participants' gaze location at the horizontal and vertical axes at a sample rate of 120 Hz (approximately every 8 ms). Consequently, it was possible to determine, for each gaze sample, whether it was located inside any of the areas of interest at any given time. Using the R Project software (R Core Team, 2018), steps of one ms were inspected per participant, trial, and object along the time windows of interest. A value of 1 was given to the area of interest that the participant was fixating on at each time step. We then calculated both the fixation counts and the proportion of fixations per participant on a trial basis for the areas of interest in 50 ms time bins preserving both participants' and items indexes. Subsequently, we computed the mean proportion of fixation and the corresponding 95% confidence intervals by participant, adjusted for within-subject designs and multiple comparisons, for each time bin of 50 ms. This provides a detailed description of the time course of the visual preferences (see Cumming, 2014; Huettig & Janse, 2016; Huettig & Guerra, 2019; Huettig et al., 2020) that children exhibit during the exposure phase and the test phase. Inferential analysis was implemented in R using the lme4 package (Bates et al., 2015) and was based on linear mixed-effects regressions (LMER) on log-ratios between objects (Arai et al., 2007) in the exposure phase, and a quasi-logistic on empirical logit transformation of the fixation proportion (Barr, 2008) in the testing phase. These analyses are presented in detail in the following sections.

3.5.2.1. Exposure phase

To assess whether children synchronized their visual preference with stimuli presentation, we created two time windows of 800 ms each, which corresponded to the two auditory stimuli appearing on every trial. It takes approximately 200 ms to plan and launch an eye movement in response to an auditory cue (Viviani, 1990). Thus, the first time window starts 200 ms after the onset of the first spoken word and ends with the onset of the second spoken word in the same trial (1000 ms after word onset). The second window started 200 ms after the onset of the second spoken word in a trial to 1000 ms after word onset. To obtain the log-ratio between targets and competitors, we divided the proportion of looks towards the target plus a constant value (i.e., 1) by the proportion of looks towards the competitor plus that constant.

We compared the log-ratios for the experimental group (DLD) against the group of TD children with the time window as a factor (first vs. second). The analysis used a deviation contrast for the time window factor and a treatment contrast for the comparison between groups. This coding scheme compares the mean of both time windows for the reference group (i.e., DLD group) against the other group (i.e., the TD group). This means that in the analysis, the intercept of the model represented the mean log-ratio between objects across time windows for the group of children with DLD. The coding was accomplished by assigning 0 to the reference level and 1 to the other group. Time window factor levels were coded as -1 and 1.

Our regression model used a “maximal” random structure (Barr et al., 2013) including both subjects and items random intercepts, as well as random slopes for every within-subject fixed effect included in the regression and the interaction between those fixed effects. Between-subject factor (such as group), were included as a random slope for items only. Finally, we removed random correlations to ease model converges (see Barr et al., 2013). The output of the LMER produces estimates, standard error of the mean, *t*-values, and *p*-values (the latter calculated with the *lmerTest* package, Kuznetsova et al., 2017).

3.5.2.2. Testing phase

For the testing phase eye tracking data, we first calculated the proportion of looks to each object in the visual context (in 50 ms time bins), and then divided these data based on participants' response (i.e., Target, 50% competitor, 25% competitor, 0% competitor, see Figure 5). We also calculated the corresponding within-subject 95% confidence intervals, obtaining, in this way, a detailed description of the gaze pattern over time as a function of participants' response and group.

For correct trials (i.e., Target choice), we implemented a quasi-logistic growth curve analysis (GCA; Mirman et al., 2008; Mirman, 2014) on empirical logit transformation of the proportion of looks to the target (Barr, 2008) with group as predictor. A reviewer suggested we should include participants' age in the model. Thus, participants' age (in months, range = 66-155 months) was first centered around zero and then included in the GCA model as a continuous variable. The GCA approach explicitly integrates time as a continuous variable into a single analysis avoiding multiple comparisons and reducing power loss. This approach uses orthogonal higher-order polynomials as predictors of the time course, accommodating the non-linear changes of proportion of looks over time that characterize visual attention when language is involved. We determined the polynomial predictors by model comparison (see Mirman, 2014:p.46). The analyzed time window created for the GCA analysis started at 200 ms after the onset of the word and ended 1700 ms after the onset of the word, when looks to the target peaked.

We compared four models that differed only in their number of polynomials, from a single linear term to a quartic term in ascending order. Before analysis, an empirical logit transformation was applied to the proportion of looks for each time window, scaling binary data to a continuous variable (Barr, 2008; Mirman, 2014). All four models had the empirical logit for the target object as the dependent variable, group (DLD vs. TD), age (as a continuous variable), and the polynomial as fixed effects, as well as the interaction between group, age, and each polynomial. The random structure of the models included cross-random intercepts for participants and items, and random slopes for each polynomial predictor. To facilitate convergence, the models did not include random correlations between random factors (see Barr et al., 2013).

4. Results

4.1. Behavioral data

With correct responses as the intercept (i.e., Target choice), the results of the GLMM analysis indicated a significant difference between the groups (β = 0.50, *se* = 0.23, *z* = 2.10, *p* = 0.036), indicating that the group of children with DLD was significantly less accurate (*M* = .37; *SD* = .48) than the TD group (*M* = .48; *S D*= .5). When we changed the intercept to the 50% or the 25% competitor, models showed no differences between groups (*z*-values < |2|), but with the 0% competitor as intercept we observed a significant difference between groups (β = -0.52, *se* = 0.21, *z* = -2.49, *p* = 0.013), reflecting a higher preference for this competitor by the DLD group. Furthermore, children with DLD exhibited a significant difference between the target and the 50% competitor (β = -0.66, *se* = 0.28, *z* = -2.41, *p* = 0.016), the 25% competitor (β = -0.66, *se* = 0.23, *z* = -2.85, *p* = 0.004) and the 0% competitor (β = -1.12, *se* = 0.27, *z* = -4.19, *p* < 0.001), in favor of the correct responses. Moreover, their preference for the 50% competitor was not reliably different than their preference for the 25% competitor (*z*-values < |2|), however, their preference for the 0% competitor was significantly smaller relative to both the 50% (β = 0.46, *se* = 0.16, *z* = 2.88, *p* = 0.004) and the 25% (β = 0.46, *se* = 0.15, *z* = 2.97, *p* = 0.003) competitor. Two interaction effects between group and choice (25% and 0% competitors), revealed that the difference between the target and each competitor was larger for the TD group. Finally, the one-sample *t*-tests (one-tailed) calculated for each group individually indicated that both groups performed significantly better than would be expected by chance: DLD, *t*(37) = 4.23, *p*<.001, TD group, *t*(37) = 5.82, *p*<.001. Participants' average response are presented in Figure 3 as a function of children group and choice.

(*Figure 3 about here)*

4.2. Eye tracking results

4.2.1. Results exposure phase

The LMER results for the analysis of visual preference for target vs. competitor during the two time-windows in the exposure phase are presented in Table 3. Overall, the group of children with DLD did not prefer the target object or the competitor across time windows. Moreover, the results also revealed that there were no differences in the visual preference for target or competitor between the group of children with DLD and TD children across time window. Similarly, no effect of time window was observed and finally, no significant interaction effect between groups and time windows was found (all t-values < |2|). This pattern of results is shown in Figure 4, which presents the fixation proportion towards the target and the competitor over time for both groups and time windows. In summary, there was no preference for any of the targets or competitors for either of the groups and no group differences were found.

(Table 3 about here)

(Figure 4 about here)

4.2.2. Results testing phase

The results based on the confidence interval approach are presented first, followed by those from the GCA analysis. Figure 5 shows the mean proportion of fixations to each object as function of participants group and choice. Gray-shaded areas represent the upper and lower boundaries of the corresponding 95% confidence intervals, adjusted for within-subject designs (see Cousineau & O'Brien, 2014) in time steps of 50 ms. Plots are time-locked to the onset of the spoken word and divided into panels per group (DLD, TD) and choice. Figure 5 reveals that, overall, children preferred to look at the object they chose compared to the discarded objects in both groups (with the exception of 0% competitor choice in the TD group). This trend was most pronounced for correct trials (target choice), yet the lower panels of Figure 5 shows that the difference between the target and distractors is larger in the group of TD children compared to the group of children with DLD.

(*Figure 5 about here)*

We turned to the GCA approach using a quasi-logistic regression model to examine whether gaze preference for the target on correct trials differed between groups and the effect of age as a continuous value (in months, range = 66-155 months). We opted for a model comparison approach in which we increased the number of polynomial terms incrementally. As described by Mirman (2014), it is important to introduce polynomial predictors that are justified by the data. Thus, the model comparison was centered around the polynomial predictors, with other fixed (group, age) and random effects (by participants and items) kept constant. Four models were created to conduct the comparison: First, we ran a linear term model, second the linear and quadratic term model, then the linear, quadratic and cubic term model and finally the linear, quadratic, cubic and quartic term model. We compared the four models that increasingly included higher-order polynomials as predictors through a likelihood-ratio test using the *anova* function (R Core Team, 2020). The results of these model comparisons showed that the inclusion of each polynomial term increased the fit of the model (all χ^2-values>52.92, df = 6, all p-values<.001). Consequently, we reported the results of the linear mixed model fit based on maximum likelihood with fourth-order orthogonal polynomials as time course predictors. Table 5 shows fixed effects parameter estimates, standard errors, t-values, and p-values for main and interaction terms in the GCA model.

(Table 4 about here)

The results show that the linear component had a significant main effect (β = -2.56, se= 0.63, t= 4.02, p <.001). This means that linear polynomial is the best predictor to account for the time course, regardless of the group and age. The model also shows a reliable difference between groups (β = -0.69, se = 0.28., t = -2.42, p=.018), and a significant effect of age (β = 0.41, se = 0.20., t = -2.04, p=.045). This effect reflects that the older the children were, the larger the preference for the target in correct trials (see Figure 7, middle panel). Finally, GCA analysis also showed a three two-way interaction effects that pertains to the polynomial predictors: The linear term interacted with group, while the cubic and the quartic term interacted with age. The first interaction reflected more linear target fixation proportion trajectory over time for the TD compared to the DLD group (see Figure 6). The other two interactions suggest that the older the children, the more cubic that proportion of fixation to the target was over time, while the younger the children, the more quartic that proportion of fixation to the target was over time.

(*Figure 6 about here)*

(*Figure 7 about here)*

5. Discussion

5.1. Behavioral data

The present study is the first to examine whether children with DLD can learn novel labels for unknown objects in a 2x2 CSSL task. Children with and without DLD were exposed to pairs of novel words and objects, which in a single trial provided no information about the correspondence between these word-object pairs. Participants' accuracy showed that both groups of children performed above chance (25%) in the 4-AFC test that followed the CSSL task. However, the accuracy for the group of children with DLD (*M*=37%) was significantly poorer than that for the group of TD children (*M*=48%). These findings extend previously documented word learning difficulties on this population to difficulties in a task that demands participants' ability to match words and pictures based on cross-trial relations. We interpret the current results exploring why DLD may be associated with learning difficulties in a word learning CSSL task by contrasting the observed behavioral results with existing accounts that have attempted to explain the mechanism involved in this process.

Initially, the CSSL task was assumed to be governed by a gradual associative mechanism that relies purely on statistical learning (i.e., IPM) in which learners track the across-trial statistics that lead to map the correct word–referent pairs (Yu & Smith, 2007). However, there is an ongoing discussion about the actual learning processes underlying participants' accuracy in these tasks. A first account, the *gradual associative account*, assumes that a learner can store multiple associations between words and referents but each of these associations has a relative strength learned across trials (Scott & Fisher, 2012; Smith & Yu, 2008; Suanda et al., 2014). By contrast, other authors have questioned this assumption, arguing that successful learning in the CSSL is the product of a one-trial procedure in which a single hypothesized word-referent pairing is retained across learning instances and abandoned only if the subsequent instance fails to confirm the pairing (Aravind et al., 2018; Trueswell et al., 2013; Woodard et al., 2016). This second account, so-called *hypothesis testing account*, is more related to EDM than IPM. Finally, other authors have argued that these two hypothesized learning processes need not to be mutually exclusive and could function in parallel (Yurovsky & Frank, 2015, Warren et al., 2020) depending on different factors such as the degree of the contextual diversity, the number of pairs to be retained on each trial or the presence or absence of social cues (MacDonald et al., 2017; Yurovsky & Frank, 2015).

Our experimental task was an adaptation from Suanda et al. (2014) that was designed to assess the degree of participants' sensibility to detect target and competitor frequencies embedded in the task (i.e., contextual diversity) in a post-exposure test. In that sense, our design is somewhat ill-equipped to assess whether participants relied on a single-trial hypothesis on each exposure (see Trueswell et al., 2013). Furthermore, in following Suanda et al. (2014), we used the minimal pairings that can be presented in one trial (i.e., 2x2). According to Yurovsky and Frank (2015), participants in a CSSL task are more inclined to maintain multiple possible associations when the number of possibilities is more limited. Thus, we first assume that at least some of the children in our study stored multiple associations between words and referents as a strategy to learn (see also Venker, 2019).

Consequently, to accommodate the accuracy difference between groups, we hypothesize that in the present task participants relied at least in part on statistical learning. The contextual diversity results in the 4-AFC might be taken as evidence for the presence of statistical learning deficits in DLD, well documented in previous studies (Evans et al., 2009; Lammertink, et al., 2017; 2020; Ullman & Pierpont, 2005; Ullman et al., 2020). Although the 0% competitor was the least frequently selected option for both groups, the DLD group selected it significantly more often than the group of TD children. In this sense, TD children showed a clearer pattern of correct extraction and computation frequencies. While the percentage of choice for the 50% and the 25% competitors was not significantly different between groups, children with TD exhibited a numerically higher percentage of selection (relative to the DLD group) for those two competitors. In turn, children with DLD showed a (significantly) larger percentage of 0% competitor selection contrasted with the TD group. Moreover, our behavioral results from the TD group are similar to those previously reported in English-speaking TD school-age children without DLD (see Suanda et al., 2014) and are also in line with CSSL studies conducted on typical populations (Fitneva & Christiansen, 2011; Yurovsky et al., 2012; Yu et al., 2012; Smith & Yu, 2008). Thus, we can infer that DLD has an important impact on the capacity of children to retain word-object pairs. The results of the present study suggest that difficulties to extract probabilities are not exclusive to sequential patterns but can also impact learning based on co-occurrence of label-object pairings in children with DLD.

Despite the difference between groups, children with DLD had an overall performance above chance. One possible hypothesis is that the EDM system might be also implicated in the present CSSL task as several previous studies have demonstrated a relatively intact EDM in children with DLD (Bishop & Hsu, 2015; Conti-Ramsden et al., 2015; Lum et al., 2015). In our study, participants were presented with an explicit instruction that demanded active participation (i.e., they were asked to learn the names of new objects). This contrasts with previous research, such as statistical word learning (Evans et al., 2009), where participants are presented with a cover task and are exposed passively to a background stimulus. In this sense, these explicit instructions may have triggered a compensatory mechanism (Ullman & Pullman, 2015) to help children with DLD to perform above chance. This hypothesis, however, should be further investigated by assessing the accuracy in this population in a CSSL task with no explicit instructions and no explicit response. Alternatively, it still is plausible that some children relied on a more intentional and strategic learning strategy to solve the CSSL (i.e., *hypothesis testing account*; Gleitman et al., 2005; Trueswell et al., 2013) despite (or in addition) to the effect of the explicit instructions. We can speculate that this strategy might have benefited children with DLD with better EDM abilities, an idea that requires further empirical testing. Finally, we cannot discard the possibility that in some of the children in our DLD sample, the IPM is not fully affected.

The compensatory role of the EDM as a benefit for children with DLD in a CSSL task could be a bit controversial, since some authors have proposed that children with DLD might have EDM deficits (Baird et al., 2010; Duinmeijer et al., 2012; Lum et al., 2010). Others, however, have questioned this proposition, arguing that preserved working memory abilities are necessary for the initial holding and learning process sustained by the EDM system. For instance, Lum and Bleses (2012) tested children with DLD and TD children using a variety of memory tests involving the EDM, IPM, and working memory systems. A subsequent analysis of covariates demonstrated non-significant effects for group (DLD and TD) on verbal EDM performance after controlling for verbal working memory, suggesting that EDM depends on working memory to work properly. Subsequent studies by Bishop and Hsu (2015), Lum et al. (2012, 2013, 2015) and Jackson et al. (2020) reported similar results showing that, once verbal working memory is controlled for, the verbal domain of EDM is unaffected in children with DLD.

Indeed, there is a large consensus about deficits in the working memory system in children with DLD (Archibald, 2018; Dollaghan & Campbell, 1998; Gathercole et al., 1994; Marton, 2008; Vugs et al., 2013) to the extent that some authors have even suggested that lexical difficulties could be related to encoding the new words in the working memory (Alt, 2011; Jackson et al., 2016). In our study, the length of the exposure phase is brief (5 min) and the number of repetitions of the stimulus is small (i.e., four repetition per pair). Learning new words only with four exposures might be challenging for working memory. Furthermore, the behavioral 4-AFC that tests learning after the whole exposure phase requires holding the phonological-visual paired representations in working memory, which could be also highly demanding. These aspects of the task could be related to the observed differences between groups. Future studies could be carried out to test school-age children with and without DLD performing a CSSL task with more exposure trials which would enable researchers to determine if a specific number of presentations of word-referents allows children with DLD to increase accuracy.

Finally, we should ponder the potential role of attention. As we argued before, the instructions of the CSSL task prompted children to actively attend to learn new words, a process that needs attention control to activate optimal vigilance during the task (Fortenbaugh et al., 2017; Langner and Eickhoff, 2013; see also Tomas & Vissers, 2019). Several studies have reported attentional difficulties in DLD. For example, the study by Finneran et al. (2009) finds sustained selective attention deficits in visual modality. Other studies show sustained selective attention deficits in visual and verbal modalities (Ebert

& Kohnert, 2011; Spaulding et al., 2008). Although it is still unclear how the different types of attention interact to satisfy the cognitive demands presented by the CSSL task, potential attentional deficits could be contributing to the group differences.

It is worth mentioning that all our participants were simultaneous bilinguals. Research in CSSL in the bilingual population is sparse, and it is unclear whether bilingualism, DLD and CSSL could interact. However, as we previously noted, bilingual TD children in our study performed at similar levels as monolingual TD children of similar age (Suanda et al., 2014). In addition, evidence from adults showed that English monolinguals and two bilingual groups (i.e.,Chinese-English and English-Spanish) did not differ in their learning rates in a CSSL task (see Poepsel & Weiss, 2012). Future research in children with and without DLD should address whether differences in proficiency in different languages might have an effect on word learning in a CSSL task.

5.2. Eye-tracking data

In the present study we also aimed to examine the learning process underlying the performance of both groups as it occurs by recording participants' eye movements. In the exposure phase, we examined whether children in both groups synchronized their preference for the objects in the visual context with the onset of the auditory stimuli, evidenced as an overall preference for the target object. We predicted a higher proportion of looks to the target, (i.e., more looks to the visual object that represents the spoken word in a specific time window), in the TD group relative to the DLD group, since we also predicted that TD children should exhibit a higher response accuracy. We expected to find distinctive online visual patterns between the two groups, if their eye movements reflect their learning process during the exposure phase. The eye-tracking data analysis for the exposure phase showed that there was no preference for any of the targets or competitors for either of the groups and no group differences were found. Eye movement data during the exposure phase reveal that as children learned associations between novel words and novel objects, they did not exhibit an increasing preference for the correct object. In this sense, it is difficult to draw a main conclusion about the online cognitive process used by the two groups tested when learning new words in this CSSL task. Future research might reveal underlying differences by modifying the design (e.g., including more items and exposures) for the exposure phase.

The eye-tracking analyses for the testing phase examined whether gaze behavior was in line with children's explicit responses. To do so, we analyzed the looking proportion to the target and competitors as children performed the 4-AFC task. Overall, both groups preferred to look at the object they chose compared to the other objects. This reflects that the visual preference was associated with the explicit responses (i.e., mouse click). TD group however, showed no preference to the 0% competitor when it was chosen (see Figure 5), and a significantly larger preference for the target (when chosen) relative to the DLD group. Eye-movements in the testing phase can be taken as an index of how confident children were when answering. In this sense, the results showed that when TD children chose the correct visual object (i.e., target), they showed rapid and strong (linear) increase in preferential looks for the target over looks to the competitors. Instead, the eye movements trajectory over time in the DLD group appeared to reflect less confidence in their choice (i.e., cubic shape) and more hesitation relative to the TD children. In sum, it appears that our online measure in the testing phase faithfully reflects children's knowledge about these novel words-object pair, thus, their learning process.

In addition, we observed a significant effect of age in the looks to the target. While we did not find a reliable interaction effect between children's group and age, Figure 7 shows that this effect is mainly brought about by the TD group. We tentatively interpret these findings as a clear developmental effect for children with TD, absent in children with DLD (see Figure 7) within the range of ages in our sample (age range=5- to 12-year-old). This is coherent with a vast amount of literature that describes DLD as a persistent but heterogeneous disorder (Bishop et al., 2016, 2017; Blom & Boerma, 2020; Laasonen et al., 2018), in which we can see children of the same age, performing differently. In other words, some children with DLD might increase performance with time, while others do not. Overall, these results are coherent with the behavioral data and reflect that the difference in learning between the children groups can be observed both in online implicit behavioral measures (i.e., eye movements) as well as overt responses (i.e., accuracy).

6. Conclusion

The present results shed light on the difficulties that children with DLD have in the context of a statistical word learning from non-sequential ambiguous situations. To the best of our knowledge, this is the first time that a group of children with DLD have been tested in a CSSL task using both behavioral and online measures. The mechanisms involved in this task remain an open issue. However, we demonstrated that children with DLD can learn in this context, but not as successfully as their TD peers. Future research should attempt to understand which are the impaired mechanisms that do not allow children with DLD to learn new words from an ambiguous context at a similar level as TD children, while at the same time learning above chance.

7. References

Aguado, G., Coloma, C. J., Martínez, A. B., Mendoza, E., & Montes, A. (2015). Documento de consenso elaborado por el comité de expertos en TEL sobre el diagnóstico. *Revista de logopedia, foniatría y audiología*, *35*(4), 147-149. https://doi.org/0.1016/j.rlfa.2015.06.004

Ahufinger, N., Ferinu, L., Sanz-Torrent., M., Andreu., L., & Evans, J.L. (under review). Statistical Word Learning in Catalan-Spanish and English-speaking children with and without Developmental Language Disorder. *International Journal of Language & Communication Disorders*

Alarcón, A., & Garzón, L. G. (Eds.). (2011). *Language, migration and social mobility in Catalonia.* Brill

Alt, M. (2011). Phonological working memory impairments in children with specific language impairment: Where does the problem lie?. *Journal of communication disorders*, *44*(2), 173-185. https://doi.org/10.1016/j.jcomdis.2010.09.003

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)

Alt, M., Plante, E., & Creusere, M. (2004). Semantic features in fast-mapping performance of preescholers with specific Language impairment versus preschoolers with normal Language. *Journal of Speech Language and Hearing Research, 47*(2), 407-420. https://doi.org/10.1044/1092-4388(2004/033)

American National Standard Institute (1997). *American National Standard: Methods for Calculation of the Speech Intelligibility Index*. Acoustical Society of America.

Andreu, L., Sanz-Torrent, M., Legaz, L. B., & MacWhinney, B. (2012). Effect of verb argument structure on picture naming in children with and without specific language impairment (SLI). *International journal of language & communication disorders, 47*(6), 637-653. https://doi.org/10.1111/j.1460-6984.2012.00170.x

Arai, M., Van Gompel, R. P., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive psychology*, *54*(3), 218-250. https://doi.org/10.1016/j.cogpsych.2006.07.001.

Aravind, A., de Villiers, J., Pace, A., Valentine, H., Golinkoff, R., Hirsh-Pasek, K., Iglesias, A., & Wilson, M. S. (2018). Fast mapping word meanings across trials: Young children forget all but their first guess. *Cognition*, *177*, 177-188. https://doi.org/10.1016/j.cognition.2018.04.008

Archibald, L. M. (2018). The reciprocal influences of working memory and linguistic knowledge on language performance: Considerations for the assessment of children with developmental language disorder. *Language, Speech, and Hearing Services in Schools*, *49*(3), 424-433. https://doi.org/10.1044/2018_LSHSS-17-0094

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), 464-473. https://doi.org/10.1111/j.1467-7687.2009.00937.x

Arciuli, J., Torkildsen, J. V. K., Stevens, D. J., & Simpson, I. C. (2014). Statistical learning under incidental versus intentional conditions. *Frontiers in Psychology*, *5*, 747. https://doi.org/10.3389/fpsyg.2014.00747

Baird, G., Dworzynski, K., Slonims, V., & Simonoff, E. (2010). Memory impairment in children with language impairment. *Developmental Medicine & Child Neurology*, *52*(6), 535-540. https://doi.org/10.1111/j.1469-8749.2009.03494.x

Barr, D. J. (2008). Analyzing 'visual world'eyetracking data using multilevel logistic regression. *Journal of memory and language*, *59*(4), 457-474. https://doi.org/10.1016/j.jml.2007.09.002.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, 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.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of memory and language*, *83*, 62-78. https://doi.org/10.1016/j.jml.2015.04.004

Berens, S. C., Horst, J. S., & Bird, C. M. (2018). Cross-situational learning is supported by propose-but-verify hypothesis testing. *Current Biology*, *28*(7), 1132-1136. https://doi.org/10.1016/j.cub.2018.02.042

Bishop, D. V. M. (1979). Comprehension in developmental language disorders. *Developmental Medicine and Child Neurology, 21*, 225–238. https://doi.org/10.1111/j.1469-8749.1979.tb01605.x

Bishop, D. V., & Hsu, H. J. (2015). The declarative system in children with specific language impairment: a comparison of meaningful and meaningless auditory-visual paired associate learning. *BMC psychology*, *3*(1), 1-12. https://doi.org/10.1186/s40359-015-0062-7

Bishop, D.V.M, Snowling, M.J, Thompson, P.A, Greenhalgh, T., CATALISE consortium (2016). CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children. *PLoS ONE*, 11(7): e0158753. https://doi.org/10.1371/journal.pone.0158753

Bishop, D.V.M, Snowling, M.J, Thompson, P.A, Greenhalgh, T., CATALISE consortium (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, 58(10), 1068-1080. https://doi.org/10.1111/jcpp.12721

Blom, E. & Boerma, T. (2020). Do Children With Developmental Language Disorder (DLD) Have Difficulties With Interference Control, Visuospatial Working Memory, and Selective Attention? Developmental Patterns and the Role of Severity and Persistence of DLD. *Journal of Speech, Language, and Hearing Research*, *63*(9), 3036-3050. https://doi.org/10.1044/2020_JSLHR-20-00012

Bloom, L. (2000). The intentionality of word learning: How to learn a word, any word. In R. Micknick Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a Word Learner: A Debate on Lexical Acquisition* (pp. 19-50). Oxford University Press. https://doi.org/10.7916/D8H70G6K

Chen, C. H., Zhang, Y., & Yu, C. (2018). Learning object names at different hierarchical levels using cross-situational statistics. *Cognitive science*, *42* (Suppl 2), 591-605. https://doi.org/10.1111/cogs.12516

Clarke, M. G., & Leonard, L. B. (1996). Lexical comprehension and grammatical deficits in children with specific language impairment. *Journal of communication disorders*, *29*(2), 95-105. https://doi.org/10.1016/0021-9924(94)00036-0

Conti-Ramsden, G., Ullman, M. T., & Lum, J. A. (2015). The relation between receptive grammar and procedural, declarative, and working memory in specific language impairment. *Frontiers in psychology*, *6*, 1090. https://doi.org/10.3389/fpsyg.2015.01090

Conway, C.M., & Christiansen, M.H. (2006). Statistical learning within and between modalities. *Psychological Science*, *17*(10), 905–912. doi:10.1111/j.1467-9280.2006.01801.x

Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: a comment on Baguley

(2012). *Behavior Research Methods*, *46*(4), 1149-1151. https://doi.org/10.3758/s13428-013-0441-z

Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science (0956-7976)*, *25*(1), 7–29. https://doi.org/10.1111/jcpp.12721/10.1177/0956797613504966.

Dollaghan, C. A. (1987). Fast mapping in normal and language-impaired children. *Journal of speech and hearing disorders*, *52*(3), 218-222. https://doi.org/218-222. 10.1044/jshd.5203.218

Dollaghan, C. A., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136–1146. https://doi.org/10.1044/jslhr.4105.1136

Duinmeijer, I., de Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language & Communication Disorders*, *47*(5), 542-555. https://doi.org/10.1111/j.1460-6984.2012.00164.x

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*(5), 1322-1331. https://doi.org/10.1016/j.neuropsychologia.2011.02.015

Ebert, K. D., & Kohnert, K. (2011). Sustained attention in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 54*(5), 1372–1384. https://doi.org/10.1044/1092-4388(2011/10-0231)

Ellis Weismer, S., & Hesketh, L. J. (1993). The influence of prosodic and gestural cues on novel word acquisition by children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 36*(5), 1013-1025. https://doi.org/10.1044/jshr.3605.1013

Ellis Weismer, S., & Hesketh, L. J. (1996). Lexical learning by children with specific language impairment: Effects of linguistic input presented at varying speaking rates. *Journal of Speech, Language, and Hearing Research*, *39*(1), 177-190. https://doi.org/10.1044/jshr.3901.177

Ellis Weismer, S., & Hesketh, L. J. (1998). The impact of emphatic stress on novel word learning by children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(6), 1444-1458. https://doi.org/10.1044/jslhr.4106.1444

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).

Eyer, J. A., & Leonard, L. B. (1995). Functional categories and specific language impairment: A case study. *Language Acquisition*, *4*(3), 177-203. https://doi.org/10.1207/s15327817la0403_1

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *52*(4), 915-929. https://doi.org/10.1044/1092-4388(2009/07-0053)

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological science*, *12*(6), 499-504. https://doi.org/0.1111/1467-9280.00392

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 458-467. https://doi.org/10.1037//0278-7393.28.3.458

Fitneva, S. A., & Christiansen, M. H. (2011). Looking in the wrong direction correlates with more accurate word learning. *Cognitive Science*, *35*(2), 367-380. https://doi.org/10.1111/j.1551-6709.2010.01156.x.

Fortenbaugh, F. C., DeGutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Annals of the New York Academy of Sciences*, *1396*(1), 70. https://doi.org/10.1111/nyas.13318

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), https://doi.org/103-127. 10.1080/09658219408258940

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language learning and development*, *1*(1), 23-64. https://doi.org/10.1207/s15473341lld0101_4

Gómez, R. L. (2017). Do infants retain the statistics of a statistical learning experience? Insights from a developmental cognitive neuroscience perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160054. https://doi.org/10.1098/rstb.2016.0054

Gray S. (2003) Word-learning by preschoolers with specific language impairment: what predicts success? *Journal of Speech, Language, and Hearing Research, 46*(1), 56-67. https://doi.org/10.1044/1092-4388(2003/005).

Gray, S. (2004). Word learning by preschoolers with Specific Language Impairment: Predictors and poor learners. *Journal of Speech, Language & Hearing Research*, *47*(5), 1117-1132. https://doi.org/10.1044/1092-4388(2004/083)

Gray, S. (2005). Word learning by preschoolers with specific language impairment: Effect of phonological or semantic cues. *Journal of Speech, Language & Hearing Research*, *48*(6), 1452- https://doi.org/1467. 10.1044/1092-4388(2005/101)

Gray, S. (2006). The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with specific language impairment. *Journal of Speech, Language & Hearing Research*, *49*(6), 955-969. https://doi.org/10.1044/1092-4388(2006/069)

Gray, S., & Brinkley, S. (2011). Fast mapping and word learning by preschoolers with SLI in a supported learning context: Effect of encoding cues, phonotactic probability and object familiarity. *Journal of speech, language, and hearing research, 54*(3), 870. https://doi.org/10.1044/1092-4388(2010/09-0285)

Gray, S., Plante, E., Vance, R., & Henrichsen, M. (1999). The diagnostic accuracy of four vocabulary tests administered to preschool-age children. *Language, Speech, and Hearing Services in Schools*, *30*(2), 196-206. https://doi.org/10.1044/0161-1461.3002.196

Haebig, E., Saffran, J. R., i Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. Journal of Child Psychology and Psychiatry, 58(11), 1251-1263. https://doi.org/10.1111/jcpp.12734

Hamrick, P., Rebuschat, (2012). How implicit is statistical learning. In P. Rebuschat, P., & J. Williams. *Statistical learning and language acquisition* (pp. 365-382). De Gruyter Mouton https://doi.org/10.1515/9781934078242.365

Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Veríssimo, J., Dye, C. D., ... & Ullman, M. T. (2011). Grammar predicts procedural learning and consolidation deficits in children with specific language impairment. *Research in developmental disabilities*, *32*(6), 2362-2375. https://doi.org/10.1016/j.ridd.2011.07.026.

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.

Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain research*, *1706*, https://doi.org/196-208. 10.1016/j.brainres.2018.11.013.

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80-93. https://doi.org/10.1080/23273798.2015.1047459.

Huettig, F., Guerra, E., & Helo, A. (2020). Towards understanding the task dependency of embodied language processing: The influence of colour during language-vision interactions. *Journal of Cognition*, *3*(1), 1-19. http://doi.org/10.5334/joc.135

Jackson, E., Leitao, S., & Claessen, M. (2016). The relationship between phonological short-term memory, receptive vocabulary, and fast mapping in children with specific language impairment. *International Journal of Language & Communication Disorders*, *51*(1), 61-73. https://doi.org/10.1111/1460-6984.12185

Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2019). The evaluation of word-learning abilities in people with developmental language disorder: a scoping review. *International journal of language & communication disorders*, *54*(5), 742-755. https://doi.org/10.1111/1460-6984.12490

Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2020). Working, Declarative, and Procedural Memory in Children With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, *63*(12), 4162-4178. https://doi.org/10.1044/2020_JSLHR-20-00135

Kachergis, G., Yu, C., & Shiffrin, R. M. (2014). Cross-situational word learning is both implicit and strategic. *Frontiers in psychology*, *5*, 588. https://doi.org/10.3389/fpsyg.2014.00588

Kail, R., Hale, C. A., Leonard, L. B., & Nippold, M. A. (1984). Lexical storage and retrieval in language-impaired children. *Applied Psycholinguistics*, *5*(1), 37-49. https://doi.org/10.1017/S0142716400004823

Kaufman, A.S. and Kaufman N.L. (2004) *Test Breve de Inteligencia Kaufman. K-BIT*. Spanish adapation: Cordero- Pando, A. & Calonge Romano, I. Pearson.

Kiernan, B., & Gray, S. (1998). Word learning in a supported-learning context by preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(1), 161-171. https://doi.org/10.1044/jslhr.4101.161

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit?. *Neuroscience letters*, *461*(2), 145-149. https://doi.org/10.1016/j.neulet.2009.06.030

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Laasonen, M., Smolander, S., Lahti-Nuuttila, P., Leminen, M., Lajunen, H. R., Heinonen, K., Pesonen, A. K., Bailey, T. M., Pothos, E. M., Kujala, T., Leppänen, P., Bartlett, C. W., Geneid, A., Lauronen, L., Service, E., Kunnari, S., & Arkkila, E. (2018). Understanding developmental language disorder - the Helsinki longitudinal SLI study (HelSLI): a study protocol. *BMC psychology*, *6*(1), 24. https://doi.org/10.1186/s40359-018-0222-7

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*, *60*(12), 3474-3486. doi.org/10.1044/2017_JSLHR-L-16-0439

Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2020). Children With Developmental Language Disorder Have an Auditory Verbal Statistical Learning Deficit: Evidence From an Online Measure. *Language Learning*, *70*(1), 137-178. https://doi.org/10.1111/lang.12373

Langner, R., and Eickhoff, S. B. (2013). Sustaining attention to simple tasks: a metaanalytic review of the neural mechanisms of vigilant attention. *Psychological Bulletin*, *139*(4), 870–900.

https://doi.org/10.1037/a0030694

Leonard, L. B. (1995). Functional categories in the grammars of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *38*(6), 1270-1283. https://doi.org/10.1044/jshr.3806.1270

Lum, J. A., & Bleses, D. (2012). Declarative and procedural memory in Danish speaking children with specific language impairment. *Journal of Communication Disorders*, *45*(1), 46-58. https://doi.org/10.1016/j.jcomdis.2011.09.001.

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. 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*, *51*, 1-10. https://doi.org/10.1016/j.cortex.2013.10.011.

Lum, J. A., Gelgic, C., & Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders*, *45*(1), 96-107. https://doi.org/10.3109/13682820902752285.

Lum, J. A., Lammertink, I., Clark, G. M., Fuelscher, I., Hyde, C., Enticott, P. G., & Ullman, M. T. (2019). Visuospatial sequence learning on the serial reaction time task modulates the P1 event-related potential. Psychophysiology, 56(2), e13292. https://doi.org/10.1111/psyp.13292

Lum, J. A., Ullman, M. T., & Conti-Ramsden, G. (2015). Verbal declarative memory impairments in specific language impairment are related to working memory deficits. *Brain and Language*, *142*, 76-85. https://doi.org/10.1016/j.bandl.2015.01.008

MacDonald, K., Yurovsky, D., & Frank, M. C. (2017). Social cues modulate the representations underlying cross-situational learning. *Cognitive psychology*, *94*, 67-84. https://doi.org/10.1016/j.cogpsych.2017.02.003

Marton, K. (2008). Visuo-spatial processing and executive functions in children with specific language impairment. *International Journal of Language & Communication Disorders*, *43*(2), 181-200. https://doi.org/10.1080/16066350701340719

Mayor-Dubois, C., Zesiger, P., Van der Linden, M., & Roulet-Perez, E. (2012). Nondeclarative learning in children with specific language impairment: predicting regularities in the visuomotor, phonological, and cognitive domains. *Child Neuropsychology*, *20*(1), 14-22. https://doi.org/10.1080/09297049.2012.734293.

McGregor, K. K., Eden, N., Arbisi-Kelm, T., & Oleson, J. (2020). The Fast-Mapping Abilities of Adults With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, *63*(9), 3117-3129. https://doi.org/10.1044/2020_JSLHR-19-00418

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*, 45(5), 998-1014. 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

Mirman, D. (2014). *Growth curve analysis and visualization using R* (pp. 109-112). Boca Raton, FL: CRC Press.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of memory and language*, *59*(4), 475-494. https://doi.org/10.1016/j.jml.2007.11.006.

Monaghan, P., Mattock, K., Davies, R. A., & Smith, A. C. (2015). Gavagai is as gavagai does: Learning nouns and verbs from cross-situational statistics. *Cognitive science*, *39*(5), 1099-1112. https://doi.org/10.1111/cogs.12186

Montgomery, J. W. (1995). Examination of phonological working memory in specifically language-impaired children. *Applied Psycholinguistics*, *16*(4), 355-378. https://doi.org/10.1017/S0142716400065991

Montgomery, J. W., & Evans, J. L. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. https://doi.org /10.1044/1092-4388(2008/07-0116).

Moscati, V., Rizzi, L., Vottari, I., Chilosi, A., Salvadorini, R., & Guasti, M. (2020). Morphosyntactic weaknesses in Developmental Language Disorder: The role of structure and agreement configurations. *Journal of Child Language, 47*(5), 909-944. https://doi.org /1010.1017/S0305000919000709

Nation, K. (2014). Lexical learning and lexical processing in children with developmental language impairments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120387. https://doi.org /10.1098/rstb.2012.0387

Norbury, C.F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. J*ournal of Child Psychology and Psychiatry,* 57 (11), 1247-1257. https://doi.org/10.1111/jcpp.12573

Onnis, L., Christiansen, M. H., Chater, N., & Gómez, R. (2003). Reduction of uncertainty in human sequential learning: Evidence from artificial grammar learning. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 25, No. 25). https://escholarship.org/content/qt7r51f2n0/qt7r51f2n0.pdf

Owen, A. J., & Leonard, L. B. (2006). The production of finite and nonfinite complement clauses by children with specific language impairment and their typically developing peers. *Journal of Speech, Language, and Hearing Research*, *49*(3), 548-571. https://doi.org/10.1044/10902-4388(2006/040).

Perruchet P, & Pacton S (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10*(5), 233–238. https://doi.org/10.1016/j.tics.2006.03.006

Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, *152*, 9-19. https://doi.org/10.1016/j.cognition.2016.03.001

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reber, P.J, Gitelman, D.R., Parrish, T.B. & Mesulam, M.M. (2003). Dissociating explicit and implicit category knowledge with fMRI. *Journal of Cognitive Neuroscience, 15*(4), 574–583. https://doi.org/10.1162/089892903321662958

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*(6), 855–863. https://doi.org/10.1016/S0022-5371(67)80149-X

Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech and Hearing Disorders*, *55*(1), 33-42. https://doi.org/10.1044/jshd.5501.33

Rice, M. L., Buhr, J., & Oetting, J. B. (1992). Specific-language-impaired children's quick incidental learning of words: The effect of a pause. *Journal of Speech, Language, and Hearing Research*, *35*(5), 1040-1048. https://doi.org/10.1044/jshr.3505.1040

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *37*(1), 106-122. https://doi.org/10.1044/jshr.3701.106

Rice, M. L., Taylor, C. L., & Zubrick, S. R. (2008). Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2008/029)

Robertson, E. M. (2007). The serial reaction time task: implicit motor skill learning?. *Journal of Neuroscience*, *27*(38), 10073-10075. https://doi.org/10.1523/JNEUROSCI.2747-07.2007

Robertson, E. K., & Joanisse, M. F. (2010). Spoken sentence comprehension in children with dyslexia and language impairment: The roles of syntax and working memory. *Applied Psycholinguistics*, *31*(1), 141-165. https://doi.org/10.1017/S0142716409990208.

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.

Sanz-Torrent, M., Badia, I., & Serra, M. (2008). Contributions from Bilingual Specific Language Impairment in Catalan and Spanish to the Understanding of Typical and Pathological Language Acquisition. In C. Pérez-Vidal, M. Juan-Garau and A. Bel (Eds.), *A portrait of the young in the new multilingual Spain* (Vol. 6, pp. 135-158). Multilingual Matters. https://doi.org/10.21832/9781847690241-010

Sanz-Torrent, M., Serrat, E., Andreu, L., & Serra, M. (2008). Verb morphology in Catalan and Spanish in children with specific language impairment: a developmental study. *Clinical Linguistics & Phonetics*, *22*(6), 459-474. https://doi.org/10.1080/02699200801892959

Sawi, O. M., & Rueckl, J. (2019). Reading and the neurocognitive bases of statistical learning. *Scientific Studies of Reading*, *23*(1), 8-23. https://doi.org/10.1080/10888438.2018.1457681

Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(3), 501–518. https://doi.org/10.1037//0278-7393.13.3.501

Schuele, C. M., & Nicholls, L. M. (2000). Relative clauses: Evidence of continued linguistic vulnerability in children with specific language impairment. *Clinical Linguistics & Phonetics*, *14*(8), 563-585. https://doi.org/10.1080/026992000750048116.

Scott, R. M., & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, *122*(2), 163-180. https://doi.org/10.1016/j.cognition.2011.10.010

Semel, E.M., Wiig, E.H., & Secord, A.W. (2006). *Clinical evaluation of language fundamentals 4: Spanish edition.* CELF-4-Spanish. Pearson.

Smith, K., Smith, A. D., & Blythe, R. A. (2011). Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science*, *35*(3), 480-498. https://doi.org/10.1111/j.1551-6709.2010.01158.x

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.

Spaulding, T. J., Plante, E., & Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: Evidence for separate attentional capacities. *Journal of Speech, Language, and Hearing Research*, *51*(1), 16-34. https://doi.org/10.1044/1092-4388(2008/002)

Squire L.R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82*(3), 171–177. https://doi.org/10.1016/j.nlm.2004.06.005

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. https://doi.org/10.1016/j.jecp.2014.06.003.

Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160056. https://doi.org/10.1098/rstb.2016.0056

Tomas, E., & Vissers, C. (2019). Behind the scenes of developmental language disorder: Time to call neuropsychology back on stage. *Frontiers in human neuroscience*, *12*, 517. https://doi.org/10.3389/fnhum.2018.00517

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

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. Cognitive psychology, 66(1), 126-156. https://doi.org/10.1016/j.cogpsych.2012.10.001

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of psycholinguistic research*, *30*(1), 37-69. https://doi.org/10.1023/a:1005204207369

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1-2), 231-270. https://doi.org/10.1016/j.cognition.2003.10.008

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.

Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews*, *51*, 205-222. https://doi.org/10.1016/j.neubiorev.2015.01.008

Ullman, M.T., Earle, F.S., Walenski, M., Janacsek, K. (2020). The neurocognition of developmental disorders of language. *Annual Review of Psychology*. *71*, 389–417. https://doi.org/10.1146/annurev-psych-122216-011555

van der Lely, H. (1996). Specifically language impaired and normally developing children: Verbal passive vs. adjectival passive sentence interpretation. *Lingua*, *98*, 243–272. https://doi.org/10.1016/0024-3841(95)00044-5.

van der Lely, H. K.J., (1998). SLI in Children: Movement, Economy, and Deficits in the Computational-Syntactic System. *Language Acquisition*, *7*(2-4), 161-192, https://doi.org/10.1207/s15327817la0702-4_4

van der Lely, H., & Harris, M. (1990). Comprehension of reversible sentences in specifically language-impaired children. *Journal of speech and hearing disorders*, *55*(1), 101-117.

van der Lely, H. K.J., Rosen, S. & Adlard, A. (2004) Grammatical language impairment and the specificity of cognitive domains: relations between auditory and language abilities. *Cognition*, *94* (2), 167-183. https://doi.org/10.1016/j.cognition.2004.01.003

Venker, C. E. (2019). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. *Cognition*, *183*, 181-191. https://doi.org/10.1016/j.cognition.2018.10.025

Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual, and motor control aspects. *Eye movements and their role in visual and cognitive processes*, 353-383.

Vugs, B., Cuperus, J., Hendriks, M., & Verhoeven, L. (2013). Visuospatial working memory in specific language impairment: A meta-analysis. *Research in Developmental Disabilities*, *34*(9), 2586-2597. https://doi.org/10.1016/j.ridd.2013.05.014

Warren, D. E., Roembke, T. C., Covington, N. V., McMurray, B., & Duff, M. C. (2020). Cross-situational statistical learning of new words despite bilateral hippocampal damage and severe amnesia. *Frontiers in Human Neuroscience*, *13*. https://doi.org/10.3389/fnhum.2019.00448

Woodard, K., Gleitman, L. R., & Trueswell, J. C. (2016). Two-and three-year-olds track a single meaning during word learning: Evidence for Propose-but-verify. *Language Learning and Development*, *12*(3), 252-261. https://doi.org/10.1080/15475441.2016.1140581

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological science*, *18*(5), 414-420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

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

Yu, C., Zhong, Y., & Fricker, D. (2012). Selective attention in cross-situational statistical learning: evidence from eye tracking. *Frontiers in psychology*, *3*, 148. https://doi.org/10.3389/fpsyg.2012.00148.

Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, *145*, 53-62. https://doi.org/10.1016/j.cognition.2015.07.013

Yurovsky, D., Yu, C., & Smith, L. B. (2012). Statistical speech segmentation and word learning in parallel: scaffolding from child-directed speech. *Frontiers in psychology*, *3*, 374. https://doi.org/10.3389/fpsyg.2012.00374.

Acknowledgments: We thank Dr. Julia L. Evans for their assistance with this research. We especially want to thank the children and their families who generously contributed with their time. We would also like to thank the teachers, speech-language pathologists, and other professionals that contributed to the study. We thank Associació del Trastorn Específic del Llenguatge de Catalunya (ATELCA) and all the schools that participated in the study. Especially, we thank the implication of Escola Santa Anna and the Centres de Recursos per a Deficients Auditius de Catalunya (CREDA) and the Equips d'Assessorament Psicopedagògic de Catalunya (EAP). Finally to Coral Pérez and Joan Tarrida who contributed to the data collection.

Table 1. Age and standardized scores for language and cognitive assessment measures for children with specific language impairment (DLD) and typically developing (TD) children.

	DLD (*n*=38)			TD (*n*=38)			Comparison	
Variable	***Mean***	***SD***	***Range***	***Mean***	***SD***	***Range***	**t(74)**	***p***
Age in months	103.15	21.82	66-155	105.47	21.95	67-153	-.46	*p*=.64
K-BIT mat (NVIQ)[a]	99.28	11.52	82-119	103.36	9.41	88-125	-1.69	*p*=.09
CELF- CLS[b]	73.31	10.84	45-89	108.68	5.9	95-125	17.65	***p*<.01**
CELF- ELS[c]	73.60	8.60	52-87	108.42	7.83	89-128	-34.81	***p*<.01**
CELF -RLS[d]	78.73	10.83	59-100	105.57	6.04	94-118	13.33	***p*<.01**

Note. For each variable, age-scaled scores have a mean of 100 and an SD of 15 (except age in months).
[a] K-BIT mat=Kaufman Brief Intelligence, Spanish version: Non-verbal intelligence score (Kaufman & Kaufman, 2004)
NVIQ=non-verbal intellectual quotient
[b] CELF-4 CLS=Spanish Clinical Evaluation of Language Fundamentals,Fourth Edition: Core Language score (Semel et al., 2006).
[c] CELF-4 ELS=Spanish Clinical Evaluation of Language Fundamentals,Fourth Edition: Expressive Language score (Semel, et al.,2006).
[d] CELF-4 RLS=Spanish Clinical Evaluation of Language Fundamentals,Fourth Edition: Receptive Language score (Semel et al., 2006).

Table 2. Contextual diversity in the CSSL task: total frequencies of the co-occurrences between the spoken words (columns) and the visual objects (rows).

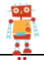
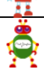
	(a) PIMO	**(b) LASI**	**(c) ZEPI**	**(d) BUNA**	**(e) DATU**	**(f) TECO**	**(g) MEPO**	**(h) RILE**
A 	4	2	1	1				
B 	2	4	1	1				
C 	1	1	4	2				
D 	1	1	2	4				
E 					4	2	1	1
F 					2	4	1	1
G 					1	1	4	2
H 					1	1	2	4

Table 3. Main and interaction effects in the linear mixed-effects regression on log-transformed proportion of looks ratios between target and competitor throughout the exposure phase.

Effect	***β***	***se***	***t*-value**	***p*-value**
Intercept (children with DLD)	-0.006	0.012	74.000	-0.53
Group	0.011	0.017	61.289	0.66
Time window	0.018	0.011	73.998	1.59
Group*time window	0.002	0.016	73.998	0.11

$^*p<.05$; $^{**}p<.01$; $^{***}p<.001$

Table 4. Main and interaction effect in the quasi-logistic GCA mixed model analysis.

	β	***se***	***t*-value**	***p*-value**	
(Intercept)	0.792	0.416	1.90	0.068	.
Linear	2.563	0.637	4.03	0.000	***
Quadratic	-0.663	0.483	-1.37	0.174	
Cubic	-0.077	0.300	-0.26	0.799	
Quartic	0.027	0.304	0.09	0.930	
Group	-0.689	0.284	-2.43	0.018	*
Age	0.408	0.200	2.04	0.045	*
Group * Age	-0.375	0.284	-1.32	0.190	
Linear * Group	-1.611	0.779	-2.07	0.042	*
Linear * Age	0.763	0.549	1.39	0.168	
Quadratic * Group	0.215	0.583	0.37	0.714	
Quadratic * Age	-0.383	0.411	-0.93	0.355	
Cubic * Group	-0.483	0.425	-1.14	0.260	
Cubic * Age	-0.774	0.300	-2.58	0.012	*
Quartic * Group	0.512	0.381	1.34	0.183	
Quartic * Age	0.607	0.269	2.26	0.027	*
Linear * Group * Age	-0.563	0.779	-0.72	0.472	
Quadratic * Group * Age	0.157	0.583	0.27	0.789	
Cubic * Group * Age	0.498	0.425	1.17	0.245	
Quartic * Group * Age	-0.433	0.382	-1.14	0.260	

*p<.05; **p<.01; ***p<.001

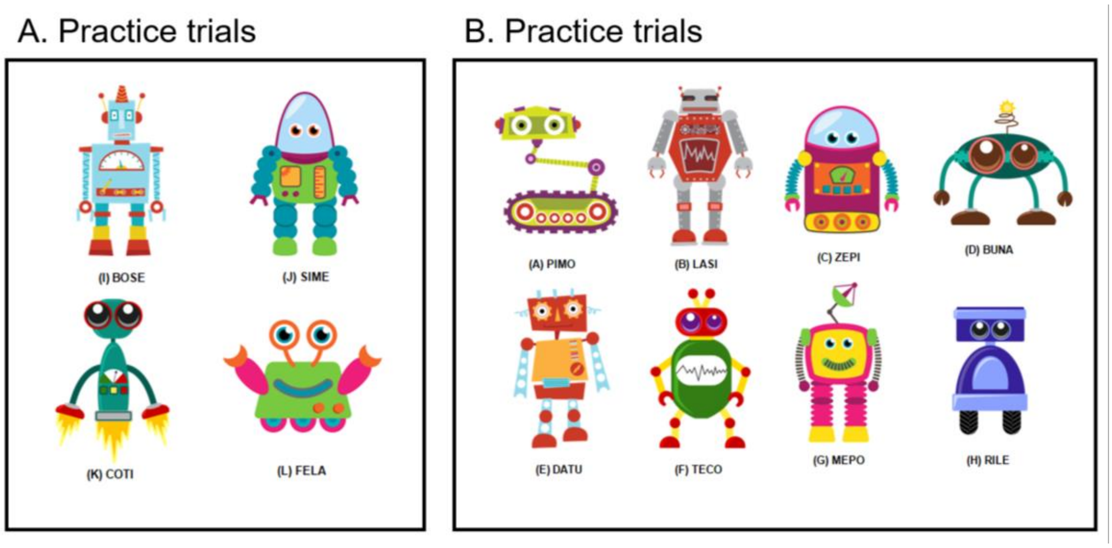


Figure 1. The eight to-be-learned word-object pairs used in the exposure phase (panel A) and the four word-object pairs used in the practice phase (panel B).

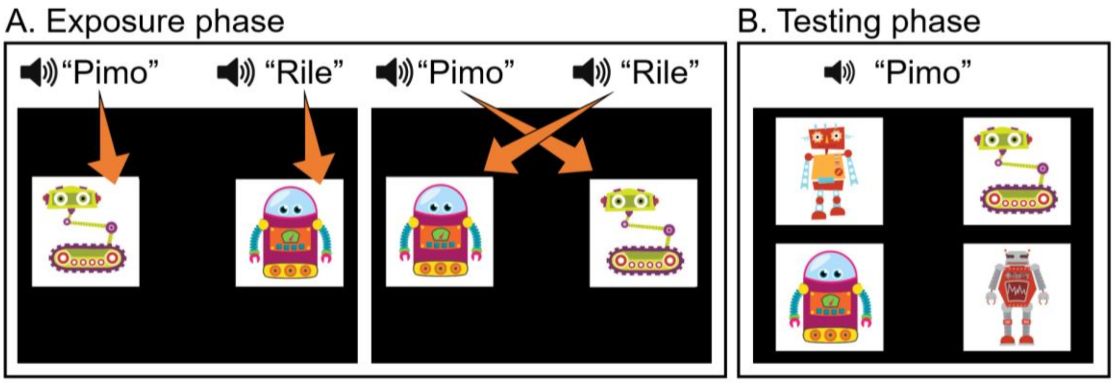


Figure 2. Visual context example for exposure trials (A) and testing trials in the four-alternative forced-choice task (B).

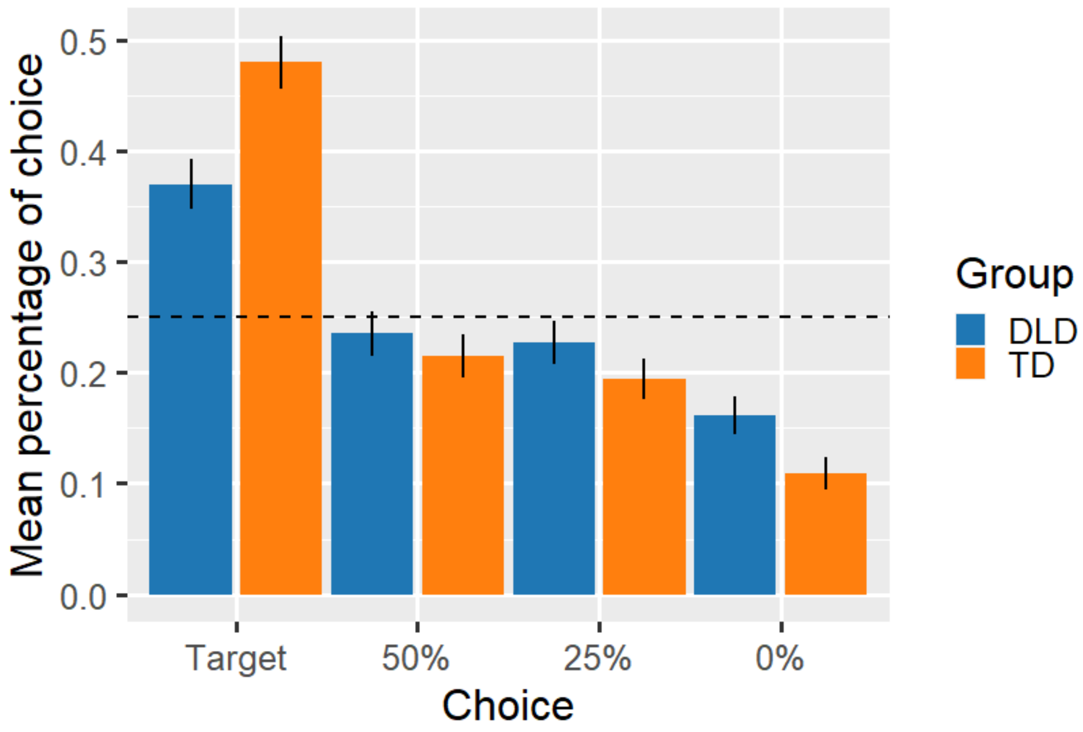


Figure 3. Accuracy data per group (color) and response (x-axis). Bar plots show the mean percentage of choice with error bars representing within-subject adjusted standard error of the mean. The dashed horizontal line represents the 25% chance level.

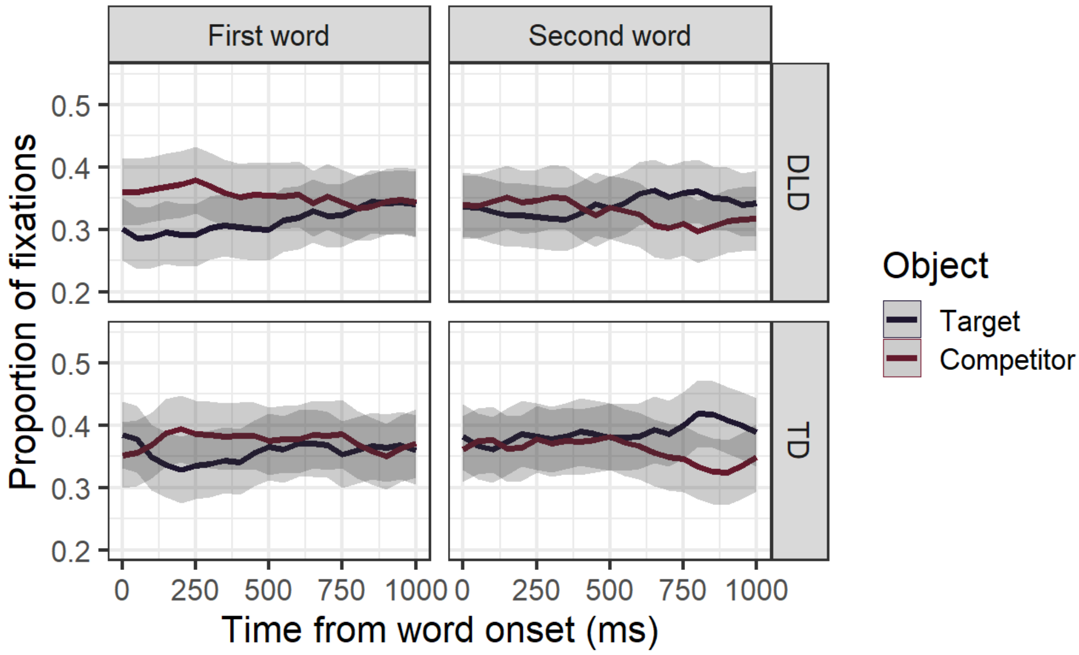


Figure 4. Mean proportion of fixation over time as a function of objects' role, word time window and children's groups. Different objects are represented by different color lines, while shaded gray areas around the lines represent 95% CI adjusted for within-subject design.

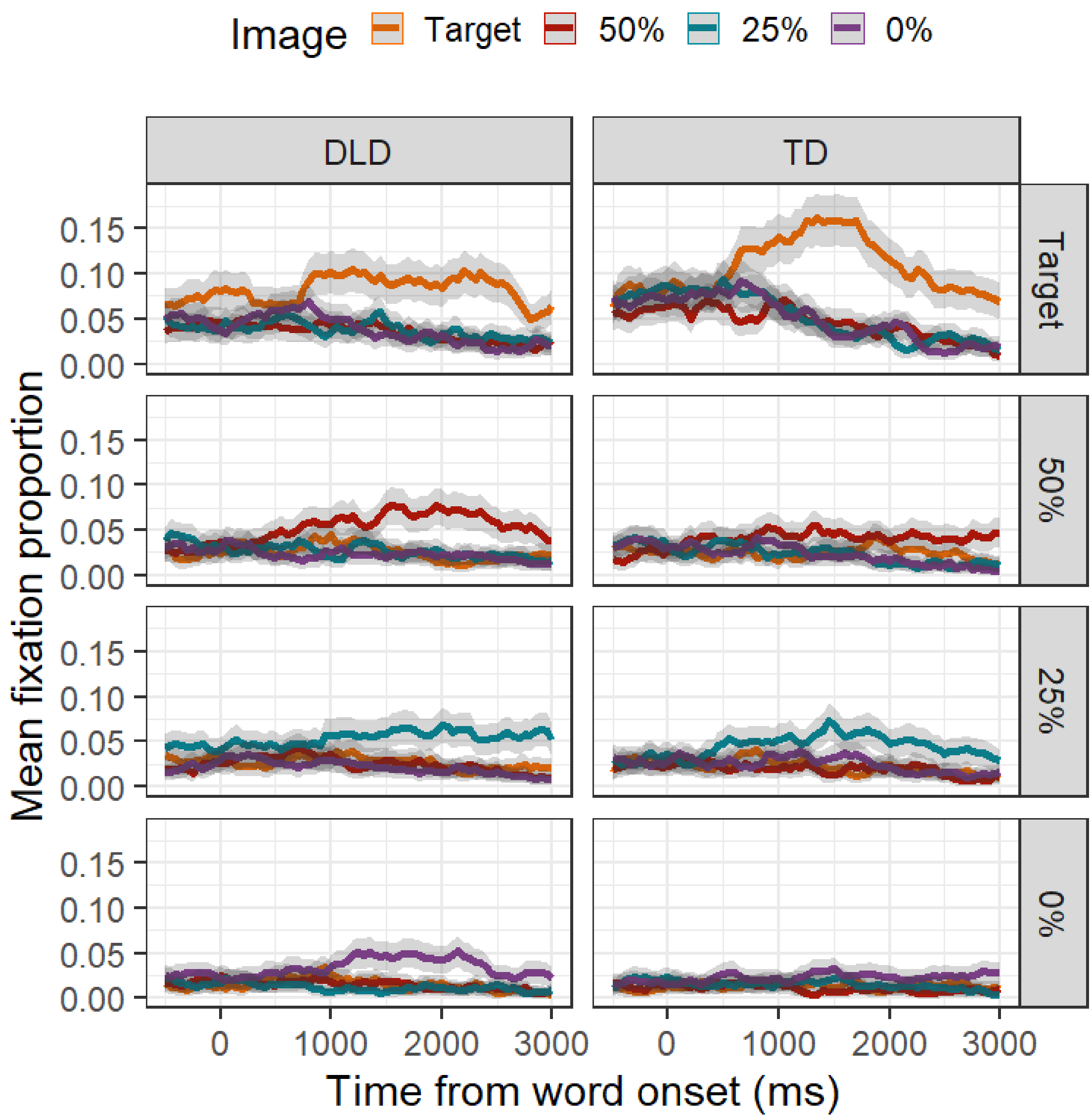


Figure 5. Mean proportion of fixation (aggregated by participants) to the target and distractors and as function of groups (TD, DLD) and choice (Target, 50%, 25%, 0%), time-locked to the onset of the critical word. Shaded areas around the lines represent 95% confidence intervals adjusted for within-subject designs and multiple time windows.

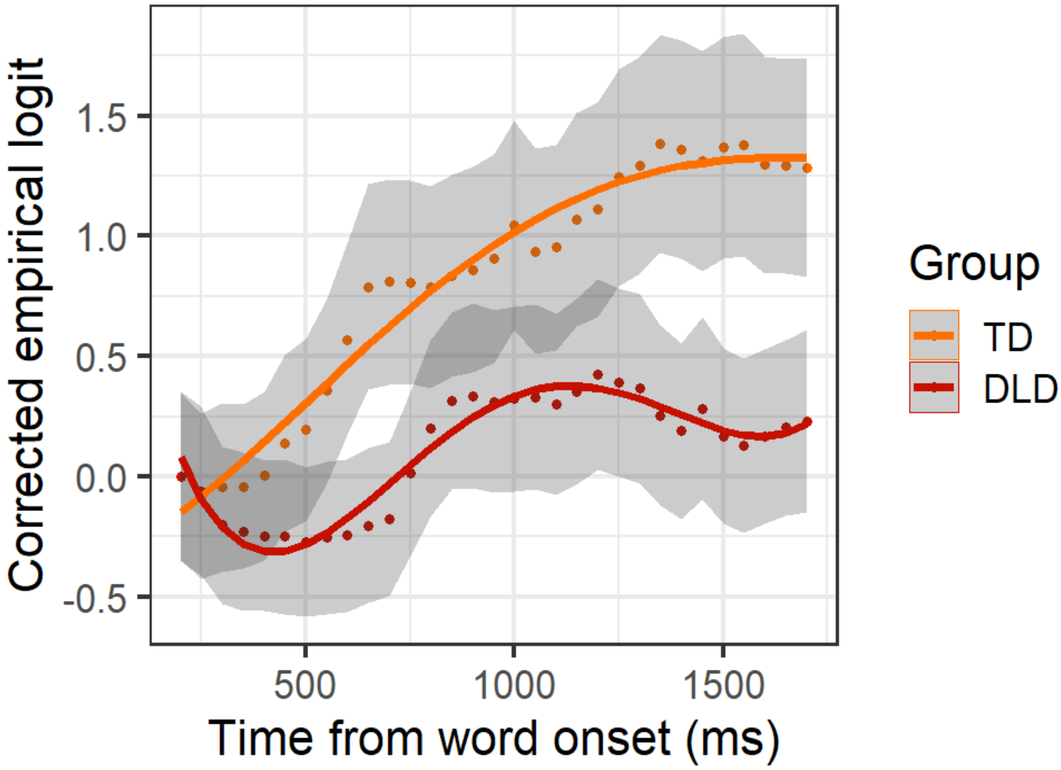


Figure 6. GCA model fit (lines) of baseline-corrected empirical logit (points) to the target for both groups. Shaded grey area around the points represent 95%CI adjusted for within-subject designs.

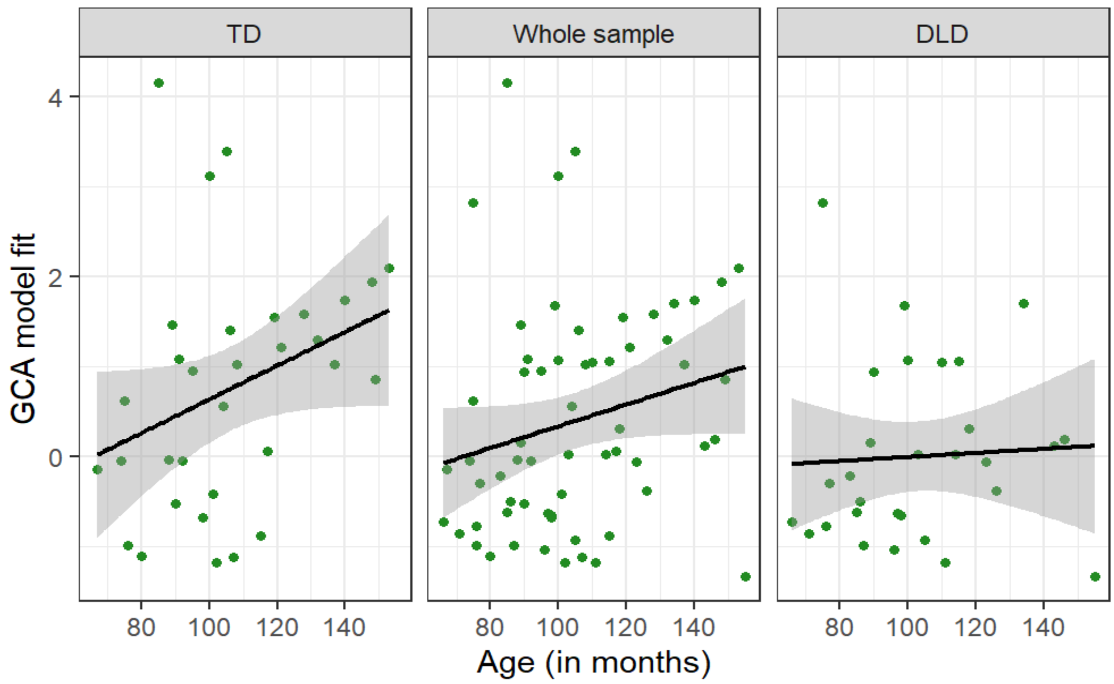


Figure 7. GCA model fit (points) of the baseline-corrected empirical logit, aggregated by participants as a function of participants age (in moths). The scatter plots show a linear function (with shaded grey area representing the standard error of the mean) between empirical logit and participants' age. The panels shows (from left to right) the TD children data, the data from the DLD group, and the data from both groups.