Nonverbal cognition as a predictor of linguistic abilities in children with and without developmental language disorder

Ashley Blake
University of Birmingham, UK

Richenda Wright
Ewa Dąbrowska
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Nick Riches
Newcastle University, UK

Elodie Winckel
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Abstract: This study examined the relationship between cognitive abilities and linguistic abilities in children with and without developmental language disorder (DLD). The cognitive abilities we investigated were scene analogies (a form of nonverbal analogical reasoning), grammatical analogies (using a grammatical sensitivity subtest from a foreign language aptitude test battery), implicit statistical learning, and inhibition. 39 language typical children and 34 children with DLD (M = 10;8 years) took part in a moderated online study. Our results revealed large between-group differences in all linguistic abilities. Regarding cognitive measures, there were large between-group differences in performance for both analogical reasoning tasks, but similar performance between groups in inhibition and implicit statistical learning. The most striking finding was that analogical reasoning is the cognitive ability that predicted most of the variance in all linguistic measures, suggesting that analogical reasoning could play a key role in linguistic abilities in children with and without DLD.

Keywords: developmental language disorder; analogical reasoning; implicit statistical learning, inhibition

Corresponding author(s): Ashley Blake, Department of Linguistics and Communication, University of Birmingham, Edgbaston, B15 2TT, UK. E-mail: arb988@alumni.bham.ac.uk
Richenda Wight, Department of English and American Studies, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany, Bismarckstr.6, 91054, Erlangen. E-Mail: richenda.wright@fau.de
NB. Ashley Blake and Richenda Wright are joint first authors.

ORCID ID(s): Ashley Blake: https://orcid.org/0000-0001-5963-9632; Richenda Wright: https://orcid.org/0000-0002-0091-4959; Ewa Dąbrowska: https://orcid.org/0000-0002-3926-9555; Nick Riches: https://orcid.org/0000-0002-1867-3014; Elodie Winckel: https://orcid.org/0000-0001-5721-3160

Citation: Blake, A., Wright, R., Dąbrowska, E., Riches, N., & Winckel, E. (2026). Nonverbal cognition as a predictor of linguistic abilities in children with and without developmental language disorder. *Language Development Research, 6*(1), 259–309. http://doi.org/10.34842/ldr2026-928

Introduction

Developmental language disorder (DLD) is a persistent condition which affects children's understanding and use of language (Bishop et al., 2017; Leonard, 2014) and occurs in approximately 7.5% of the population (Norbury et al., 2016; Tomblin et al., 1997). Although language is the primary area of difficulty, many children with DLD also have problems with domain-general cognitive processes, including working memory (for reviews, see: Archibald, 2017; Henry & Botting, 2017; Montgomery et al., 2010), executive function (Pauls & Archibald, 2016), speed of processing (Zapparrata et al., 2023), temporal processing (Tallal & Piercy, 1973a, 1973b, 1974, 1975), and implicit statistical learning[1] (see meta-analyses by Lammertink et al., 2017; Lum et al., 2014; Obeid et al., 2016). Furthermore, a meta-analysis by Gallinat and Spaulding (2014) shows that children with DLD often have lower scores in nonverbal intelligence tests than their language typical (LT) peers (see Lancaster et al., 2024 for possible reasons for this). Lower scores could be indicative of underlying deficits in nonverbal cognitive abilities, or alternatively, may reflect a bidirectional relationship between language and cognition, where weak nonverbal intelligence influences language, and vice versa (Griffiths et al., 2022).

Due to the heterogeneity observed in DLD, it has been suggested that this condition should be conceptualised as a spectrum disorder, with different but overlapping strengths and weaknesses (Lancaster & Camarata, 2019). DLD is a complex, polygenic disorder (Mountford et al., 2022), but despite extensive research to date, we lack a comprehensive understanding of the associations between cognitive and linguistic factors in individuals with DLD. Therefore, the present study aims to explore this relationship further. We investigate several cognitive abilities, some of which are novel in the context of DLD, for example, verbal and nonverbal analogical reasoning. We discuss language acquisition from a usage-based perspective and specific cognitive abilities that potentially contribute to linguistic abilities in these populations.

Analogical Reasoning

According to usage-based theories of language acquisition, children's grammar develops through linguistic experience and domain-general cognitive capacities (Behrens, 2009; Bybee, 2010; Goldberg, 2006; Tomasello, 2003). These abilities include categorization, chunking, memory, analogy, and the ability to create cross-modal

[1] In this paper we use the term "implicit statistical learning" as a broader term which includes procedural learning.

associations between form and meaning (Bybee, 2010). Language is intrinsically pattern based, and children observe abstract patterns of similarities between utterances. Through analogy, children identify similarities between the utterances that they encounter and can generalize to build their own mental representations of more schematic patterns (Goldberg, 1995; Tomasello, 2003). For example, consider a simple low-level schema such as *Can I get down?* (Can I VP). These chunks are generalized through slot-and-frame patterns to create new abstract utterances, e.g., a construction with two slots: *Can Mummy do it?* (Can NP VP), and to a fully general schema such as *Will Mummy do that?* (Aux NP VP) (see examples in Dąbrowska, 2010b, p.698). Ultimately, the process of language development occurs in piecemeal fashion through repeated exposure until linguistic knowledge is entrenched. The kinds of generalizations about which lexical items can fit into specific slots in a construction rely on analogy. It has been suggested that there is a reciprocal relationship between analogical reasoning ability and language abilities (Christie & Gentner, 2014; Gentner & Christie, 2010). Since analogical reasoning relies on the ability to establish correspondences across forms, problems with forming analogies may lead to difficulties with language; and indeed, some studies have found that children with DLD have analogical reasoning deficits in comparison to their LT peers (Krzemien et al., 2017, 2019, 2020; Leroy et al., 2012, 2014; Nippold et al., 1988). Leroy et al. (2012, 2014) suggest that children with DLD may over rely on perceptual similarity when forming analogies, and when such cues are not available (i.e., when the analogy is purely relational), their difficulties may be more evident.

Grammatical Analogies

Research on second language acquisition has consistently found robust relationships between performance on language aptitude tests such as the Modern Language Aptitude Test (MLAT: Carroll & Sapon, 1959) and language learning success (Carroll, 1981; Carroll & Sapon, 2002). Language aptitude tests assess, among other things, sensitivity to grammatical structure and the ability to reason about it, without relying on the use of grammatical terminology such as 'verb' or 'direct object'. For example, in the 'Words and Sentences' subtest of the MLAT, participants are presented with two sentences and are tasked with identifying a word in the second sentence that fulfils the same grammatical function as a specified word in the first (see Figure 2 in Materials for an example). Thus, this task essentially involves analogical reasoning about sentences. In fact, Carroll (1990, p.19) describes the grammatical sensitivity task as a 'grammatical analogies test'.

Language aptitude is typically not considered relevant to L1 acquisition. This assumption stems from the belief among many researchers that L1 grammar acquisition

relies almost exclusively on implicit learning mechanisms (DeKeyser et al., 2010; Ellis, 1996; Ullman, 2001). However, if, as suggested above, child language learners extract rules by recognizing analogical relationships between utterances, then the explicit reasoning abilities assessed by traditional language aptitude tests should also be relevant for first language acquisition (see Dąbrowska 2010 for further discussion). Indeed, several recent studies have revealed robust relationships between 'foreign' language aptitude and grammar in adults and older children (Dąbrowska, 2018; Llompart & Dąbrowska, 2023; Prela et al., 2022; Winckel & Dąbrowska, 2024; Wright et al., 2024), as well as in adults with DLD (Blake, Dąbrowska & Llompart, 2025). To the best of our knowledge, language aptitude, specifically verbal analogical reasoning as it pertains to grammatical sensitivity, has not been studied in children with DLD.

Inhibitory Control

Inhibitory control is a core executive function which involves controlling one's attention while ignoring irrelevant stimuli (Diamond, 2013). Children with DLD are known to have difficulties with executive functions, including inhibitory control (see meta-analysis by Pauls & Archibald, 2016), which may contribute to their challenges with processing language (Boerma et al., 2017). For example, inhibition is one of the key factors in pragmatic development, as it allows individuals to suppress points of view that are egocentric, allowing individuals to demonstrate better communicative perspective taking (Matthews et al., 2018).

Similar to tasks measuring inhibition, analogical reasoning requires recognizing structural similarities while filtering out irrelevant information. This fact has led to the suggestion that performance in analogical reasoning tasks may be linked to maturation of inhibitory control (Richland et al., 2006). If children with DLD have weaknesses in inhibition, as previous research suggests (Pauls & Archibald, 2016), this could be a contributing factor to poor performance on analogical reasoning tasks. This hypothesis was originally put forward by Leroy et al. (2012; 2014), and later investigated by Krzemien et al. (2020), suggesting a potential link between poor inhibition and analogical reasoning in children with DLD. Weakness in inhibitory control (the ability to suppress irrelevant or competing information) may also contribute to the difficulties children with DLD experience when generalising novel constructions to new contexts (see Krzemien et al., 2021). Furthermore, evidence from Ibbotson and Kearvell-White (2015), suggests that differences in grammatical ability, particularly the ability to produce correct irregular past tense forms, can be predicted by variation in inhibitory control. In their study, children who performed better on a Stroop task made fewer overgeneralisation errors (e.g., *flyed* for *flew*).

Implicit Statistical Learning

Children are able to learn the complex patterns underlying the language spoken around them relatively early, at a time when their explicit learning and reasoning abilities are rather rudimentary. Implicit statistical learning (ISL), on the other hand, develops early: for example, infants as young as 8 months can detect statistical regularities such as word boundaries from a continuous pseudoword speech stream (Saffran et al., 1996). This fact has led many researchers to propose that child language acquisition relies predominantly on implicit processes; and indeed, a number of studies appear to support this proposal (DeKeyser et al., 2010; Ellis, 1996; Ullman, 2001).

Implicit learning or statistical learning has been implicated in the acquisition of several kinds of linguistic abilities (Isbilen & Christiansen, 2022), such as learning written orthographic and morphological regularities, speech segmentation, the formation of phonetic and syntactic categories, and learning artificial phonotactic patterns (Misyak & Christiansen, 2012). The role of ISL in a usage-based account of language acquisition is also clear: it is through repeated exposure that individuals are able to track the co-occurrence statistics of linguistic forms, which causes frequent forms to be entrenched and accessed easier and faster. More frequent words are acquired earlier (Ambridge et al., 2015) and even larger units, such as phrases, are processed faster if they occur together more frequently (Arnon & Snider, 2010). Thus, ISL may provide processing advantages for the recognition or judgement of linguistic stimuli. However, some language tasks require learners to make inferences about linguistic forms or functions and studies with adults show that general problem-solving abilities as well as aspects of foreign language aptitude that depend on explicit reasoning are relevant for native grammar acquisition (Dąbrowska, 2018). The interplay between implicit and explicit learning (in the language aptitude sense) and language abilities is widely studied in second language acquisition (see Ellis, 2007), leaving a research gap in the roles that these aspects play in the native language acquisition process.

If typical language learning depends crucially on implicit statistical learning, then the difficulties that children with DLD experience might be plausibly explained by deficits in ISL—and several studies provide evidence of ISL deficits in children with DLD (see meta-analyses: Lammertink et al., 2017; Lum et al., 2014; Obeid et al., 2016). However, the evidence is not entirely clear-cut. Another meta-analysis by West et al. (2021) found between-group differences in ISL (on the serial reaction time task), in clinical groups (DLD and dyslexia), compared to typically developing children (g = -.30). However, there was no relationship between ISL and language when the groups were combined (i.e., unselected for language ability) (r = .03).

A subsequent meta-analysis conducted by Oliveira et al. (2023) found a significant relationship between the two variables, but the effect size was negligible (r = .06). The variability in ISL research findings has been attributed to task reliability, particularly in studies involving children (Arnon, 2020). Given the inconsistent findings to date, further research is needed to clarify the role of ISL in language acquisition.

The above literature review establishes the context for several cognitive abilities that are important to language learning in the context of pattern finding. In the next section, we will consider specific linguistic abilities of interest in the present study, namely vocabulary, grammar, and collocational knowledge.

Language Abilities in Children with DLD

Children with DLD acquire their first words later than language typical peers (Leonard, 2014). There is evidence that children with DLD have problems with encoding words, referring to the initial memory trace linking a word form to its referent, prior to integration into long term memory (see Jackson et al., 2021). While some research finds that children with DLD have similar word learning retention to their LT peers (Haebig et al., 2017; Jackson et al., 2021; Leonard et al., 2019), other studies report poor retention (Rice et al., 1994; Riches et al., 2005). Overall, children with DLD often have vocabularies that are limited in breadth (how many words are in the child's vocabulary), and depth (how well the child knows the words), when compared to their LT peers (McGregor et al., 2013). This limitation negatively affects grammar, as vocabulary and grammar are closely interconnected during early development. Children require a substantial amount of vocabulary knowledge to recognize grammatical patterns (Bates et al., 1988; Marchman & Bates, 1994). It is, therefore, not surprising that children with DLD have syntactic deficits (Leonard, 2014; Rice et al., 1994, 1995, 1998; Rice & Wexler, 1996; van der Lely, 2005).

While expressive and receptive vocabulary difficulties are well documented in DLD research, an area that, to our knowledge, has yet to be studied is collocational knowledge. Collocations refers to two or more words that occur together or near each other more frequently that can be expected by chance (Riches et al., 2022). For example, *strong wind* and *severe storm* are established collocations; we tend not to say, *severe wind* or *strong storm* (Herbst, 2011). In adults, collocational knowledge is related to performance in receptive vocabulary and grammar (Dąbrowska, 2014a; 2019; Llompart & Dąbrowska, 2020), suggesting that grammar, vocabulary, and collocational knowledge may rely on similar learning mechanisms in accordance with usage-based theories of language acquisition (see Llompart & Dąbrowska, 2020). Collocational knowledge is correlated with other language measures, including vocabulary

size, however, it is only through exposure that individuals can learn that certain words frequently occur together (Dąbrowska, 2014b). Thus, knowledge about collocations could provide a relatively pure measure of sensitivity to frequency in language. This is especially the case when measures of collocational knowledge contain basic vocabulary, acquired early in development, so that only exposure to the word order found in a particular combination determines if the collocation has been acquired (e.g. *cats and dogs* and not *dogs and cats*).

To the best of our knowledge, only two studies have explored collocational knowledge in children (Riches et al., 2022; Smith & Murphy, 2015). Smith and Murphy (2015) assessed performance on a multi-word phrase test in children aged 7-10 years and found that collocational knowledge was significantly associated with both receptive and expressive vocabulary scores. More recently, Riches et al. (2022) explored collocational knowledge in monolingual English children and children learning English as an additional language (EAL), to assess whether knowledge of collocation predicted performance in other linguistic abilities. Performance on the collocation task was associated with vocabulary and grammar in both groups, although correlations were slightly weaker among EAL learners, possibly reflecting more limited exposure to English. It is possible that children with DLD may be less sensitive to frequently occurring patterns (see Evans et al. 2009), which may influence their collocational knowledge.

Aims and Research Questions

The aim of the study is to examine the differences between LT and DLD children in specific cognitive abilities. These are:

1. Verbal analogical reasoning (grammatical analogies task)
2. Nonverbal analogical reasoning (scene analogies task)
3. Inhibition (Go/No-Go task)
4. Implicit statistical learning (visual search paradigm)

Given that children with DLD have challenges in both linguistic and non-linguistic tasks, we anticipate that the DLD group will perform below the LT children on these measures.

We investigate the extent of these cognitive differences and how they predict language abilities in both groups. The linguistic measures of interest are vocabulary, receptive grammar, recalling sentences, and collocational knowledge. While vocabulary and grammar are known to be problematic for children with DLD, we are not

aware of any research exploring collocational knowledge in DLD populations.

This study was pre-registered:
(https://osf.io/uc2qa/?view_only=e0e2327d46de49abb14da339688717ae).

Methods

Participants

Participants were recruited through advertising with organizations such as: 'Engage with Developmental Language Disorder', (E-DLD: https://www.engage-dld.com), and The DLD Project in Australia (https://thedldproject.com), and via social media platforms (e.g., Facebook groups 'Raising awareness for DLD', 'Developmental Language Disorder (formerly SLI) Support', and 'RADLD: Raising Awareness for DLD'). We recruited children with and without DLD, specifying the following inclusion criteria: 1) ages 9 to 12 years, and 2) monolingual English speakers. For the DLD group, a confirmed diagnosis was required. Additionally, all participants (both language typical and DLD) were required to have no hearing impairments or co-occurring biomedical conditions, such as autism spectrum disorder or Down syndrome. Our study was open to English-speaking children from the United Kingdom, Australia, New Zealand, Canada, and the United States of America. Participants received a voucher as compensation for their participation. Ethical approval for this study was obtained from the University of Birmingham (ERN-2022-0365).

Overall, we recruited 86 children to the study (43 LT children and 40 DLD children). We excluded three participants (two LT and one DLD) who completed less than half the tasks in the study. Parents reported whether their child had a diagnosis of DLD, and we conducted an additional classification measure to verify accuracy of our groupings. This involved measuring children's performance on the Recalling Sentences subtest of the Clinical Evaluation of Language Fundamentals: CELF-5 (Wiig et al., 2013). Sentence repetition tasks are considered reliable clinical markers of DLD (Conti-Ramsden et al., 2001), and the difficulties with sentence repetition persist into adulthood (Poll et al., 2016), making it a stable diagnostic criterion. Participants with DLD were only included in the DLD group if they scored more than 1.25 SDs below the mean on this measure, which excluded six participants. Four additional participants originally in the LT group were excluded because they also scored more than 1.25 SDs below the mean and, thus, partially meet the criteria for the diagnosis of a language disorder. Therefore, our final sample, post-exclusion, consisted of 73 children (37 female and 36 male) comprising 39 LT children and 34 DLD children, with a mean age of 10;8 (range: 8;11 – 12;11). The mean age in the LT group was 10.79 years

and 10.64 in the DLD group. The difference between the groups was not significant (p = .59).

Our participants were age-matched as some of the tasks (grammatical analogies, receptive grammar) required participants to be independent readers. Matching according to ability may have required us to recruit LT participants who may be matched on some of the other tasks but not be able to read well enough for the tasks that require a certain level of literacy skills.

We examined parental education levels for each parent/caregiver, coding education on a scale from 1 (primary school) to 5 (postgraduate degree). For the first parent/caregiver, there was a small difference between LT and DLD groups, with the LT group showing slightly higher educational attainment (Fisher's exact test, p = .08; d = 0.43, 95% CI [–0.06, 0.90]). For the second parent/caregiver, no group difference was observed (Fisher's exact test, p = .70; d = 0.16, 95 CI [–0.32, 0.63]). While these results suggest that there are slight differences in parental education (particularly for the first parent/caregiver), the differences are not statistically significant.

Materials

Language Tasks

Receptive Grammar. The Pictures and Sentences task (Dąbrowska, 2018; 2019) is a forced-choice receptive grammar task similar to the Test for Reception of Grammar (TROG; Bishop & Garsell, 2003) but more appropriate for our participants (LT children in the age group studied here perform close to ceiling on the TROG). Each test item comprises two pictures and a sentence read aloud by the experimenter. The child's task is to select the picture that matches the sentence. There are 88 items in total representing 11 different constructions (see Table 1); items are presented in a semi-random order with the constraint that no two items representing the same construction appear consecutively. An example test item is shown in Figure 1.

Table 1. ***Constructions in the Receptive Grammar task***

Construction	Example
Active	The boy scratched the dancer.
Passive	The dancer was scratched by the boy.
Subject cleft	It was the boy that scratched the dancer.
Object cleft	It was the dancer that the boy scratched.
Subject relative	The boy was the one that scratched the dancer.

Object relative	The dancer was the one that the boy scratched.
Quantifier + is	Every lamp is on a table.
Quantifier + has	Every table has a lamp on it.
Postmodifying prepositional phrase	The lamp on the table is white.
Complex postmodifying preposition phrase	The window in the room with the chair is broken.
X-Is-Difficult-to-Answer	The doll is hard to see

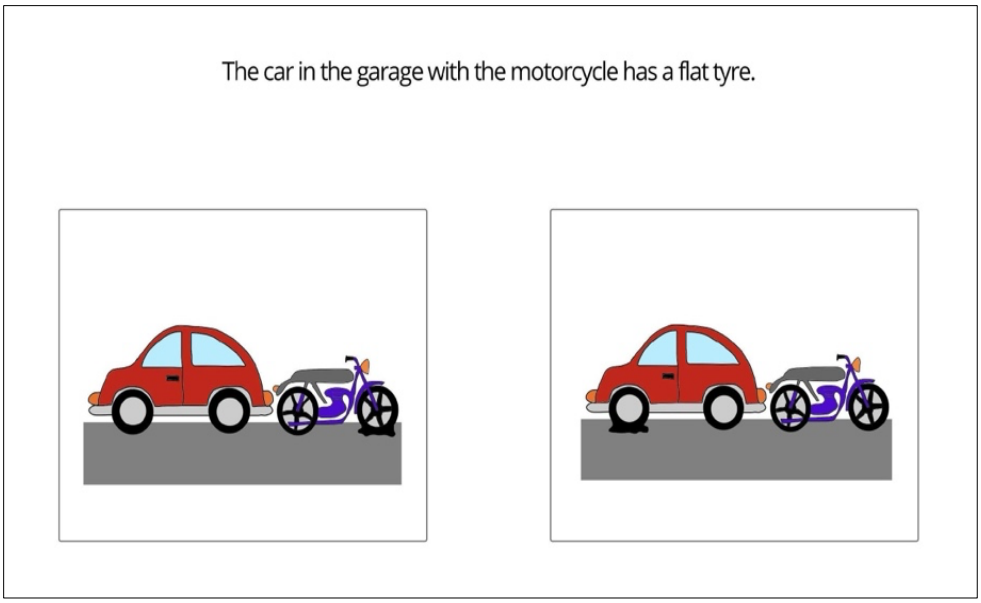

Figure 1. ***Test item from the Receptive Grammar task.***

Expressive Grammar. The Recalling sentences subtest of Clinical Evaluation of Language Fundamentals (CELF-5; Wiig et al., 2013) was used as a measure of expressive grammar. Sentence repetition tasks are considered a reliable clinical marker of DLD (Conti-Ramsden et al., 2001). In this task, the child listens to sentences, and their task is to repeat the sentence back to the experimenter. The sentences increase in length and complexity during the task. The test is discontinued after five consecutive sentences containing four or more errors in each. The maximum possible raw score is 78.

Vocabulary. The British Picture Vocabulary Scale (BPVS3) (Dunn et al., 2009) is a standardized receptive vocabulary task. Each test item comprises four pictures, and the participant is asked to select the one that best matches a word read aloud by the

experimenter. There are 14 sets of 12 items, increasing in difficulty as the test progresses. The test is discontinued after eight or more errors in a set. The maximum possible raw score is 168.

Collocations. The Language Detective game (Riches et al., 2022) is used to measure collocational knowledge. The task is presented as a game and the child is instructed that they will listen to some phrases, some of which are 'good', and some that are 'not so good'. Their task is to choose which phrase sounds 'better', by repeating the phrase, or by pointing to the picture associated with the 'better' phrase. In each trial, the experimenter reads two expressions aloud (one target and one distractor, e.g. 'do homework' or 'do a picture'), each with an accompanying visual representation.

There are three practice trials in which experimenter reads two sentences out loud, such as 'eat cake' and 'cake eat' (see Figure 2 for an example), and the child is asked 'which one sounds better?'.

Figure 2. ***Sample practice item from the collocation task.***

The task contains 89 items in total and consisted of 47 binomials (e.g. *left and right, fish and chips*), 9 verb + noun collocations (e.g. *take a break*), 2 verb + adj collocations (e.g. *go crazy*), 7 Adj + noun collocations (e.g. *strong winds*), and 24 similes (e.g. *as busy as a bee*). The maximum score is 89.

Cognitive Tasks

Grammatical analogies. The Matching Words subtest of the Modern Language

Aptitude Test: Elementary (MLAT-E; Carroll & Sapon, 1967) assesses grammatical sensitivity without making use of traditional parts of speech terminology. Participants see 30 pairs of sentences. The first sentence in each pair contains a key word (printed in capitals). The participants' task is to choose the word from the second sentence that 'does the same job' as a key word in the first (see Figure 3 for an example of an item[2]). The task included eight sample sentences to assess children's understanding before beginning the test items. The maximum possible score is 30.

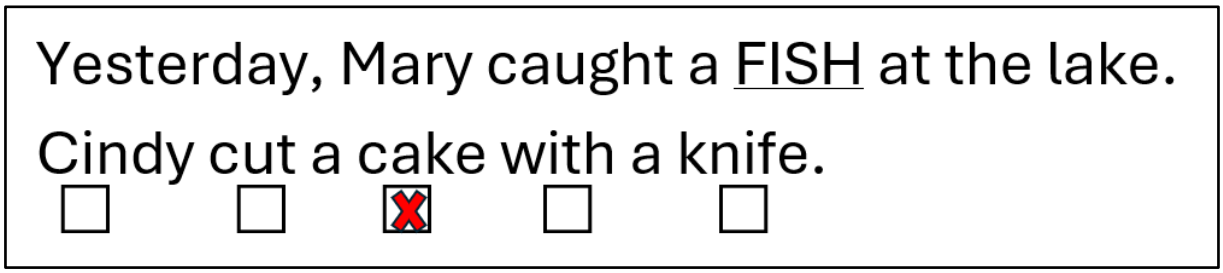

Figure 3. ***Sample item from the grammatical analogies task. The correct answer is 'cake'.***

Scene Analogies. The Scene Analogy task (Richland et al., 2006) consists of 20 scene analogy problems, adapted for online use with permission from the authors. The task involves pairs of pictures and the child's task is to observe the role of an object in the first picture (the top picture) and use this information to select an object in the second picture (the bottom picture) that plays the same role. The picture sets vary in the number of relational mappings (one relation versus two relation problems), and the presence of featural distractors (distractor versus no distractor problems). Figure 4 is an example of an item with two relations (an adult reading to a child who reads to their doll) with no distractor (an unrelated action).

The task was presented as a PowerPoint using two functions on Zoom: screen sharing and mouse control. The experimenter read instructions to the child, as per Richland et al. (2006: p.263-264). First, the experimenter explains the one-relation example picture set by saying: 'There is a certain thing that happens in the top picture, and the same thing happens in the bottom picture, but it looks different […]. I am going to put a blue star on one thing in the top picture, and your job is to tell me *what is the same thing happening* in the bottom picture.'

The two-relation problems are explained as follows: 'Now sometimes what is

[2] The example given is available on the internet as a sample item of the 'Matching Words' subtest and is not an item from the test (https://lltf.net/mlat-sample-items/mlat-e-part-2/)

happening will have two parts, like the one you just saw [as per first example] and sometimes there will be three parts to what is happening. Let me show you what I mean.' The experimenter then explains the actions in both pictures, and says: 'If I put the blue star on [...], what is the same thing happening in the bottom picture?' The task begins with two practice trials which helped the experimenter to check that the child understood the instructions, before beginning the test trials. For every set of pictures, the experimenter says: 'If I put my blue star on [...] in the top picture, where will you put your yellow star in the bottom picture? *What is the same thing happening*?' Children then moved the yellow star to their chosen location in the bottom picture, using mouse control in Zoom. The example in Figure 4 shows the target item in the top picture (indicated by a blue star), with the correct item in the bottom picture (indicated by a yellow star). The maximum score in the scene analogies task was 20, which was converted to the percentage of correct responses.

Figure 4. ***Example of a sample item from the Scene Analogy task: two relations/no distractor (with permission from Richland et al., 2006).***

Inhibition. We designed a Go/No-Go (GNG) game modelled on an original task designed by Geurten et al. (2016), with permission from the authors. Children are introduced to Nellar, a friendly cartoon alien (see Figure 5). The storyline is presented as two aliens who are playing hide-and-seek. When the child sees Nellar (a green

alien), they should help him to hide by pressing the spacebar. But when they see Nellar's cousin, Neera (a red alien), they should not press anything. The task begins with a practice block of eight trials with feedback provided for correct and incorrect trials (for example, if the child does not press the spacebar when they see Nellar, they hear a noise and see the words: 'Uh oh, remember when you see Nellar, press the space bar'). Likewise, in the practice block, if the child presses the space bar when they see Neera, they hear a noise and are reminded: 'Uh oh, remember, when you see Neera, do not press anything'. The child is advised that they will gather gold coins as they play, and they will see their score at the end. There are 40 test trials, and no feedback is given in the test phase. As per Geurten et al. (2016), there are 14 Go trials and 26 No-Go trials. The stimuli are presented for 350ms, with varied interstimulus intervals ranging from 1900 to 4100ms so that the child could not anticipate the presentation of the stimulus. Responding to the distractor stimulus is a commission error whereas not responding to a target stimulus is an omission error. The inhibition task is scored by accuracy of response where each correct response (i.e. responding or not responding as appropriate for the presented stimuli) is awarded one point. The maximum score is 40 and was converted to the percentage correct responses.

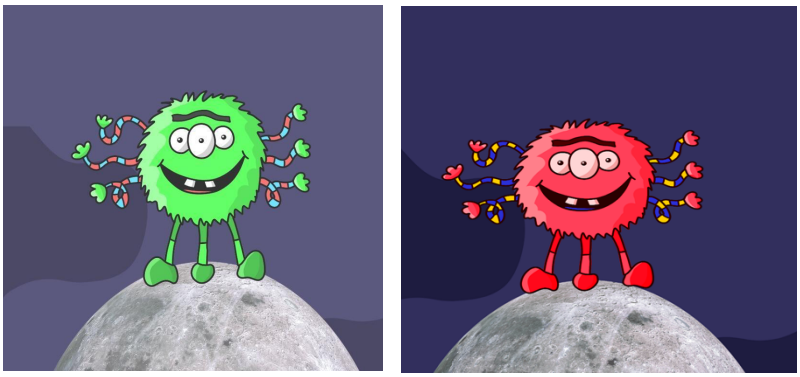

Figure 5. ***Example trials from the Go No-Go task. A 'Go' trial is depicted on the left with the green alien, and a 'No-Go' trial on the right with a red alien.***

Implicit statistical learning. "Where's Nellar?" (Wright, Prela, Riches, Blake & Dąbrowska, in preparation) is an implicit statistical learning task based on a visual search paradigm. The task is a variation of an alternating serial reaction time (ASRT) task where stimuli that alternate between predictable and random positions are presented within a search paradigm. Individuals become sensitive to the stimuli that follow the predictable pattern and, over time, respond faster to these than the stimuli in the random positions. Learning is quantified as the difference in reaction time between the random and predictable stimuli. ASRTs are widely used to measure implicit

learning, which is an aspect of procedural learning (Hedenius et al., 2011). However, the inclusion of a distractors in a search paradigm typically results in longer reaction times between random and predictable trials, which should make the task more sensitive.

Children meet Nellar, the friendly alien, once again, and they also see Nellar's friends (distractor aliens). The children are told that they will play a game of hide-and-seek, and they should try and catch Nellar by clicking on him as fast as they can. The child is then presented with a 3 x 3 grid of nine aliens, including Nellar, who moves positions in each trial (see Figure 6). During the practice trials, if the child clicks on an incorrect alien, they hear a noise and see the words 'Uh oh'. If they click correctly, they hear a sound and see the words 'Well done'. After three practice trials, they are informed that they will try to catch Nellar on every planet.

The task is divided into six planets, representing artificial levels between blocks of trials to keep participants motivated to complete the task. Each planet consists of four blocks, with a total of 24 blocks in the task. Each block contains four triplets (a fixed sequence of three locations where Nellar appears e.g. locations 1-7-3 in the 3x3 grid) and the presentation of each triplet is preceded by the appearance of Nellar in a random filler position, separating the triplets. Thus, each sequence of a filler-triplet-combination is a sequence of two random and two predictable locations given that the filler position is random, as is the move to the first position in the triplet. The following two locations are predictable from the initial triplet position. There were no consecutive appearances in the same location. In the first block, participants start off with an additional three fillers to ensure that mistakes on the initial trials do not influence the results. Participants start on level 1 and, after every four blocks, see a level-up screen to mark progress and motivate them. A timeout of 2000ms per trial was imposed, after which participants heard negative audio feedback before Nellar moved to the next position. This was in order to motivate participants to respond quickly. We obtained two reaction time-based scores from this measure: a motor learning score and a sequence learning score (see the Data Processing section for details). Two measures were derived to disentangle the unique contributions they make to overall task performance, allowing for a more precise assessment of implicit learning mechanisms.

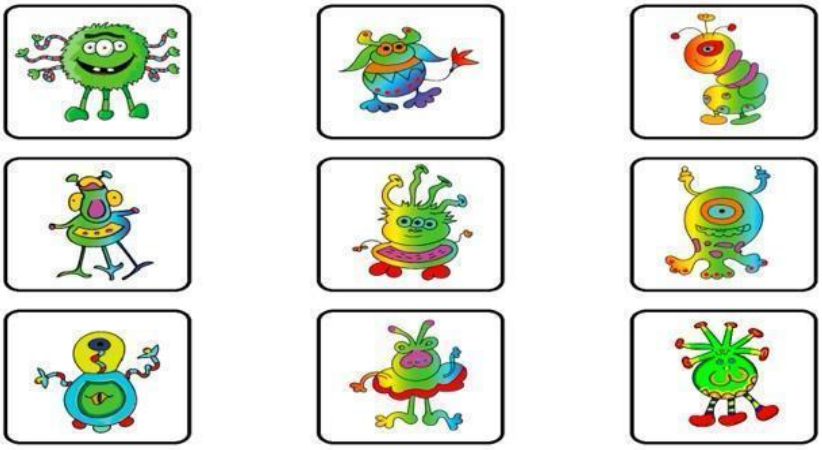

Figure 6. ***Example trial from 'Where's Nellar.' The target alien is in the top left corner, location 1.***

Procedure

Before the testing sessions, parents and caregivers completed a consent form and a background questionnaire, which gathered information on the child's language ability and parental education and employment. Participants were instructed to complete the tasks on a desktop computer or laptop. They were not accessible via mobile devices. Parents and caregivers were then invited to schedule their child's online testing sessions, which was conducted via Zoom.

Each child participated in two moderated sessions, hosted by one of two experimenters. Moderation allowed the experimenters to ensure task comprehension, provide clarification, and offer breaks as needed. Two tasks —Go/No-Go (GNG) and "Where's Nellar?"—were conducted asynchronously using the Gorilla Game Builder Platform (Anwyl-Irvine et al., 2020). These tasks included written instructions supplemented by automatically playing audio recordings to support comprehension.

In session one we administered the vocabulary test, the scene analogy task, and the recalling sentences task. After completing this session, participants received the login information to complete the two online tasks for inhibition (Go/No-Go) and implicit statistical learning (Where's Nellar?). The final session consisted of the receptive grammar, collocations, and grammatical analogies tasks.

Data Processing

The tests for expressive grammar (recalling sentences), vocabulary, and grammatical analogies were administered and scored according to the testing manuals.

For the scene analogy task, the receptive grammar task (Pictures and Sentences) and the task for collocations (The Language Detective Game) we computed the percentage of correct responses.

For the measure of inhibition, we extracted a response accuracy score based on the number of correct responses (i.e. responding or not responding as required by the type of trial). Six trials were excluded from data processing for being above 4500ms (longer than the timeout built-in to the experiment) or below 5ms (impossibly short), representing 0.20% of the data. Reaction times were collected to monitor the administration of the task but were not used for analysis as only Go trials have an active response, the No-Go trials are successful when no response is made.

In the implicit statistical learning task, the data from two participants were excluded from analysis (0.8%): one participant from the LT group who only completed 72 trials (less than 20% of the task), and one participant from the DLD group who had ten or more consecutive timeouts. 138 trials with reaction times over 3000ms were excluded (0.47%). These trials were more than 1000ms over the programmed timeout for the online task. The reason is that for tasks conducted online, connection or server problems can occasionally result in reaction times that are longer than the programmed timeout. An additional 238 trials under 100ms were excluded (0.82%) as these are likely errors due to inattention (Farkas et al., 2024). We excluded 1274 responses due to multiple answers on one trial (for example, when participants first selected the wrong location, then the correct one), which represented 4.37% of the data. Furthermore 1100 trials that were more or less than three times the mean absolute deviation (MAD) by participant-specific mean reaction time were excluded (3.77%). This means that outliers were excluded by comparing each participant's reaction times to only their own mean reaction time, ensuring that individuals with DLD were not unfairly disadvantaged for potentially having slower processing. After these exclusions, a total of 603 timeouts remained (308 occurred in random trials and 295 occurred in predictable trials) which were also excluded from analysis as there was no response to analyse. The excluded timeouts constitute 2.30% of the trials, of which 436 came from the DLD group and 167 from the LT group.

We used the remaining trials to compute two RT-based scores for each child: the sequence learning score and the motor learning score. The sequence learning score

captures an individual's ability to detect and exploit predictable patterns in the task. The score was derived from the difference between the mean reaction time on random and predictable trials across all blocks. Thus, a positive score means that the individual was faster on predictable trials and, hence, presumably learned the sequence implicitly. The motor learning score, on the other hand, accounts for improvements in general motor execution over time. Following Divjak et al. (2022) and Llompart and Dąbrowska (2020), we calculated individual slopes for the first 15 blocks. The first 15 blocks were selected as the DLD group showed a plateau and then an increase in RT after this point, indicating that motor learning was no longer taking place. The reaction time data was transformed by inverting and scaling to facilitate interpretation. Trial numbers were standardized by converting the values to z-scores. This mean-centres the trial numbers and improves model conversion, removes potential collinearity issues, and makes it easier to interpret regression results. A linear mixed-effects model was fitted to predict the transformed reaction time with random slopes and intercepts for participants. Coefficients from the model were extracted to obtain individual participant intercepts and slopes, of which the latter was used as the measure of ISL motor learning.

By analysing these two complementary learning processes separately, we gain a clearer understanding of whether performance differences arise from difficulties in extracting sequential patterns or from more general limitations in motor adaptation.

The data were imported to R (ver. 4.2.1, R Core Team, 2022) for processing. Data were processed using the tidyr package (Wickham et al., 2023). Graphs were generated using ggplot2 (Wickham, 2016). For scored measures, internal reliability was calculated using the splithalf package (Parsons, 2021) using 5000 random splits, and we report the Spearman-Brown corrected reliability estimate. Regression models were run using the stats package in base R. The relative contribution (lmg) of predictors to the variance in the outcome variable in the regression models was calculated with the relaimpo package (Grömping, 2006).

Results

This section is divided into three parts.

In Part 1, we report descriptive statistics and t-tests addressing our primary research question concerning potential between-group differences. We also report the reliability of the assessment measures. The results of selected individual tasks (namely scene analogies, receptive grammar, and statistical learning) are also discussed in more detail, to aid interpretation of our results.

Part 2 presents the results of the correlational analyses. These analyses serve two purposes: (a) to explore patterns of association among the variables, and (b) to inform subsequent regression modelling.

In Part 3, we report the results of regression analyses examining between-group differences in cognitive abilities and the extent to which these abilities predict language abilities.

Raw scores were used for all analyses. As the LT and DLD groups were closely matched in age, raw scores allow for direct comparison of performance without the need for standardization.

Descriptive Statistics and Reliability

Descriptive statistics and reliability for the assessment measures are presented in Table 2. Reliability was not calculated for the standardised measures as these are available in the testing manuals. Although the reliability is good for the receptive grammar, collocations, implicit statistical learning, and grammatical analogies tasks, we have poor reliability on the inhibition and scene analogies tasks. The low reliability for scene analogies is due to ceiling effects in the LT group, which result in very low variability in scores and low signal-to-noise ratios. The reliability on the inhibition task is also lower than expected. This may be due to the age of our participants, who are likely still developing executive function skills that include inhibition and working memory. Additionally, the low number of trials in the task (15 Go trials and 25 No-Go trials) may have negatively influenced the reliability.

The between group *t*-tests and effect sizes for the different measures are reported in Table 2. There are large between-group differences on all language measures as well as on both analogical reasoning tasks. These differences are reflected in the distribution plots in Figure 7 and 8. There are no significant differences in performance on the motor learning, sequence learning, and inhibition measures.

Table 2. ***Mean, median, standard deviation (SD), reliability, t-tests***

Measure	LT Mean (SD)	DLD Mean (SD)	*t*-test	*p*-value	Cohen's *d*	Reliability: LT [95% CI]	Reliability: DLD [95% CI]	Reliability: All participants
Expressive grammar (raw)	58.41 (7.01)	29.38 (8.14)	$t(65.65) = -16.20$	< .001	–3.84	-	-	-
Expressive grammar (scaled score)	11.70 (2.83)	5.29 (1.62)	$t(62) = -12.00$	< .001	–2.72	-	-	-
Vocabulary (raw)	139 (12.45)	115.20 (16.02)	$t(61.98) = -7.28$	< .001	–1.74	-	-	-
Vocabulary (standardized)	103.00 (1.79)	81.80 (9.18)	$t(70) = -8.60$	< .001	–1.98	-	-	-
Receptive grammar	89.48 (6.04)	73.06 (8.59)	$t(58.17) = -9.32$	< .001	–2.24	.72 [.56, .84]	.71 [.55, .83]	.87 [.82, .91]
Collocations	80.84 (7.29)	65.86 (7.79)	$t(68.12) = -8.45$	< .001	–1.99	.68 [.51, .81]	.59 [.38, .77]	.82 [.75, .87]
Grammatical analogies	62.56 (19.53)	33.63 (10.65)	$t(60.26) = -7.99$	< .001	–1.81	.83 [.73, .89]	.36 [.02, .63]	.85 [.80, .89]
Scene analogies	92.18 (6.16)	83.82 (11.01)	$t(50.19) = -3.92$	< .001	–0.95	.06 [–.31, .43]	.47 [.16, .71]	.47 [.27, .64]

Inhibition	85.45 (9.71)	86.54 (9.23)	$t(70.45) = 0.49$	.62	0.12	.69 [.52, .82]	.7 [.51, .83]	.69 [.57, .79]
ISL: Motor learning	-0.02 (0.04)	-0.02 (0.03)	$t(66.46) = 0.73$	.47	0.17	-	-	-
ISL: Sequence learning	14.19 (25.98)	10.30 (33.37)	$t(60.14) = –0.54$	.59	–0.13	1[.99, 1]	.99 [.99, 1]	.99 [.99, 1]

Scene analogies: percentage correct responses. Sequence learning: mean random RT – mean predictable RT. Motor learning: individual slope. Inhibition: percentage correct responses. Grammatical analogies: percentage correct responses. Collocations: percentage correct responses. Receptive grammar: percentage correct responses. Vocabulary: raw score. Standardized scores are based on a mean of 100 and SD of 15; scaled scores are based on a mean of 10 and SD of 3.

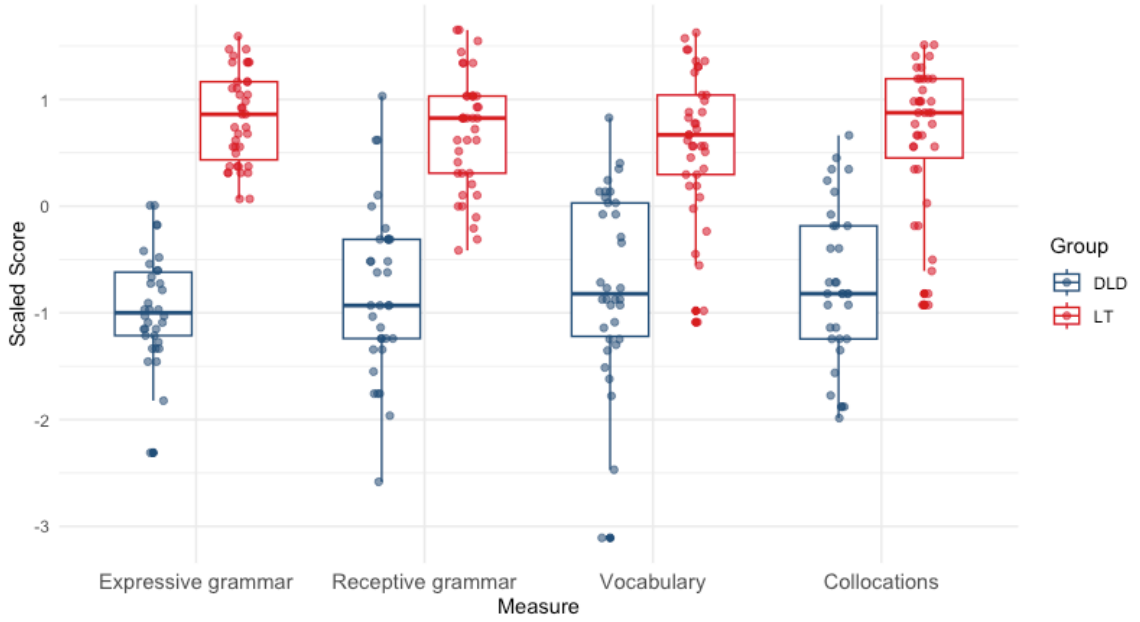

Figure 7. ***Distribution of scores by group on the language tasks.***

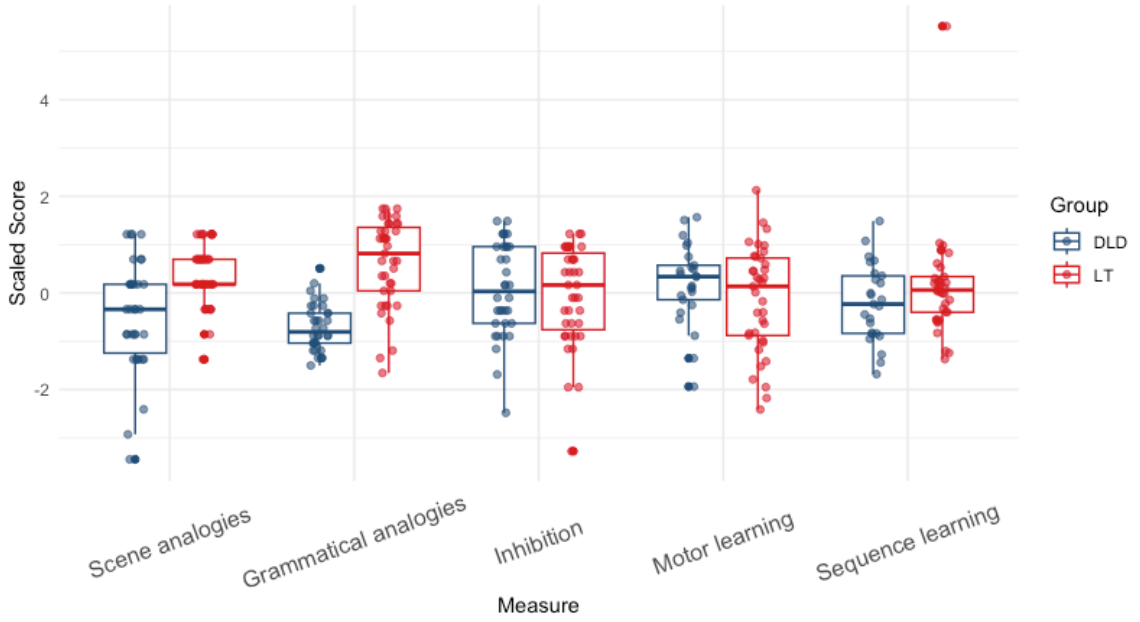

Figure 8. ***Distribution of scores by group on the predictor measures.***

Scene Analogies

In analysing the results from the scene analogy task, we followed the method from Richland et al. (2006) to look at accuracy across different trial types. Trials vary by number of relations and the presence or absence of a distractor. The LT group had higher accuracy across all trial types and gaps between performances for the two groups were largest when a distractor was present. The majority of errors made by both groups were relational errors (75.44% of all errors), followed by distractor errors (15.20%) and 'other' errors (9.36%). In both groups errors were made more often in scenes that are more complicated to process in terms of the number of relations and the presence of a distractor.

Receptive Grammar

Figure 9 shows the mean score of each group by construction. The maximum score for each construction is 8 and chance is indicated by the dashed line. Both groups are at or near ceiling on actives, the control condition. The group differences vary across constructions, with groups performing similarly on subject relatives and subject clefts while larger differences can be seen for complex NP, object relatives, and object clefts. The LT group performed at chance only on the Q-has construction, while the DLD group performed below chance on both the Q-has and object relative constructions. The comparisons to chance are numerical rather than statistical. Chance is 50% as there are two choices in each item presented.

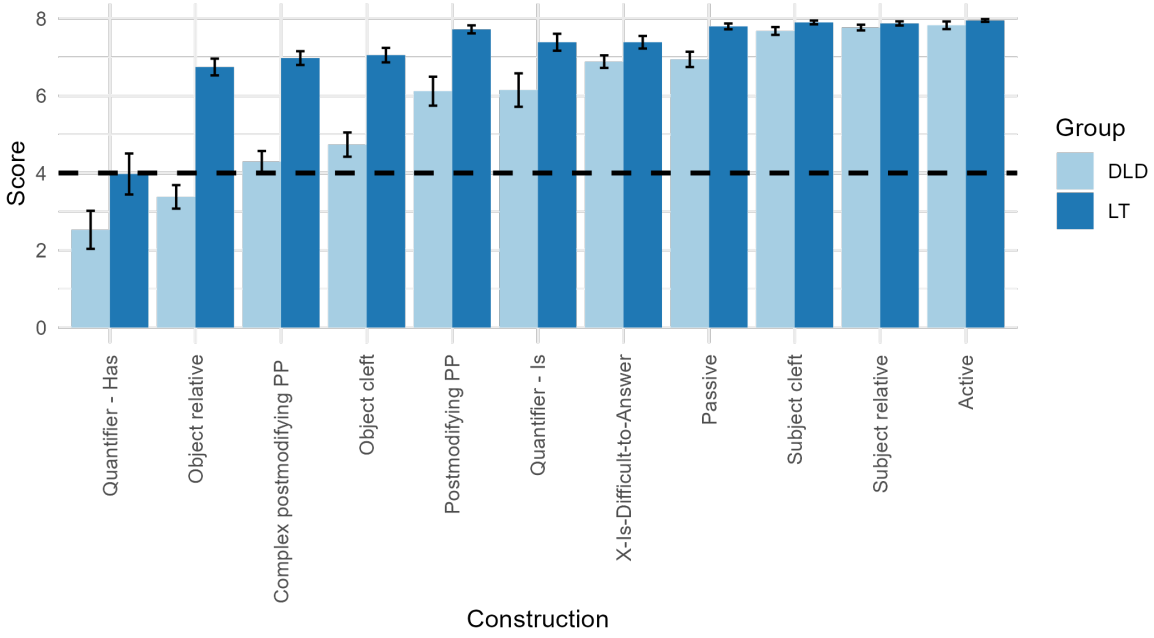

Figure 9. ***Accuracy rates by construction for the receptive grammar task.***

Implicit Statistical Learning RT

Figure 10 shows the mean reaction time by block number for all participants, separated by trial type. There is a visible downward slope in reaction time from Block 1 to Block 9 for both random and predictable trials, likely due to practice effects or motor learning taking place. Divergence of the slopes for random and predictable trials is visible in all figures, showing faster reaction times for predictable than for random trials. This shows a trend that participants have implicitly learned the sequences. Although reaction times increase towards the end of the task, presumably due to fatigue, the difference in reaction time between the random and predictable trials remains.

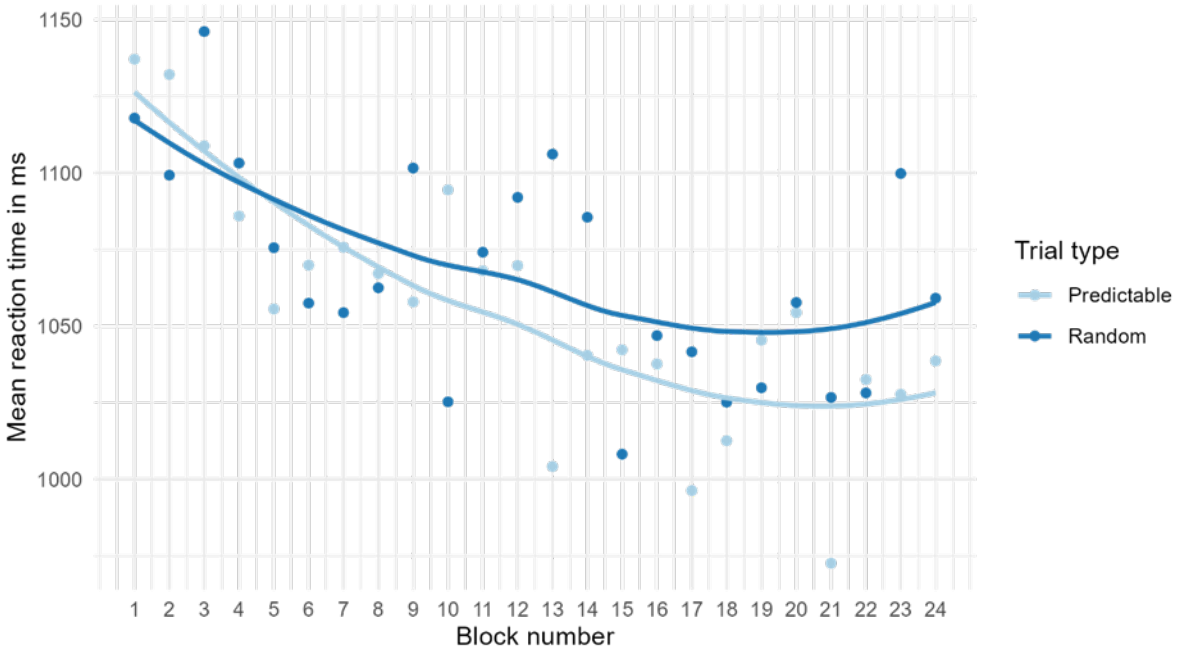

Figure 10. ***Reaction time on the implicit memory task by block for all participants.***

If we compare LT children with DLD children (Figure 11), we can see that the starting point where the lines begin to diverge between random and predictable trials differs. This takes place from Block 7 onwards for the participants with DLD, but sooner for LT participants. Although DLD participants have longer RTs overall, the difference in the amount of learning (difference in reaction time between random and predictable trials) between groups is not significant (see t-test in Table 2). Thus, while the LT group is faster to respond to both random and predictable stimuli than the DLD group, the RT gap between the two types of trials, which is what implicit statistical learning scores are based on, is not significantly different between the two groups. This indicates that they have similar amounts of learning by the end of the task.

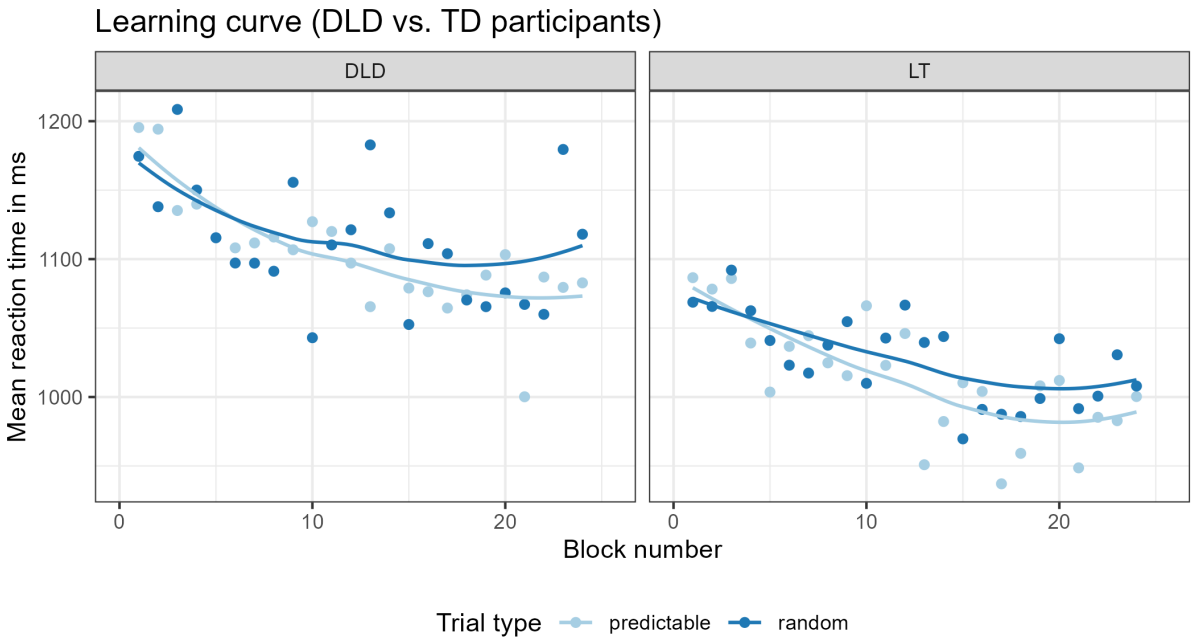

Figure 11. ***Reaction time on the implicit memory task by block for LT participants and participants with DLD.***

Correlational Analysis

We ran correlations to explore patterns in our data. Correlations for the DLD group are reported in Appendix 1, the LT group in Appendix 2, and both groups combined in Appendix 3. In the correlations for the LT participants, grammatical analogies correlated with all the language measures as well as inhibition. In the DLD group, grammatical analogies only correlated with scene analogies. Scene analogies did not correlate with any language or cognition measures in the LT group, whereas in the DLD group, scene analogies correlated with receptive grammar.

Inferential Analysis

We conducted a series of regression models to predict the language abilities (vocabulary, receptive grammar, collocations, and expressive grammar) from grammatical analogies, scene analogies, inhibition, and implicit statistical learning and ran models with and without Group as a variable. Numerical predictor variables were centre-scaled.

Table 3 provides a summary of the R-squared of the models with and without Group

as a variable. ANOVA comparisons between models with and without Group indicated that adding Group significantly improved model fit for all dependent variables. These results suggest that Group explains a significant portion of variance in all four measures beyond the significant contribution of the cognitive predictors, which is also elaborated on in the discussion. As the main purpose of the study was to investigate the role of the cognitive factors in linguistic abilities, the results and discussion focus on the models without Group. Full results for the models with Group and the ANOVA comparisons are available in the appendix (Appendices 4-8).

Table 3. ***R-squared for inferential models with and without Group***

	With Group	Without Group
Vocabulary	.54	.44
Receptive grammar	.69	.55
Collocations	.65	.55
Expressive grammar	.82	.57

As shown in Table 4, the overall model for vocabulary accounts for 44% of the variance. Grammatical analogies is the only significant predictor, contributing 31% of the total variance.

Table 4. ***Model parameters for vocabulary***

	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	Relative Importance
Multiple R-squared: 0.44					
(Intercept)	128.34	1.76	73.00	< .001***	
Grammatical analogies	10.92	2.10	5.19	< .001***	.31
Scene analogies	3.05	2.04	1.50	.14	.11
Sequence learning	–2.77	1.87	–1.48	.14	.01
Motor learning	–0.81	1.79	–0.45	.65	< .01
Inhibition	–1.09	1.85	–0.59	.56	< .01

Table 5 presents the model output for receptive grammar. The model accounts for 55% of the overall variance with the largest contribution made by grammatical analogies (32%) and scene analogies (22%).

Table 5. ***Model parameters for receptive grammar***

	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	Relative Importance
Multiple R-squared: 0.55					
(Intercept)	81.91	0.91	89.61	< .001***	
Grammatical analogies	5.62	1.09	5.14	< .001***	.32
Scene analogies	3.79	1.06	3.57	< .001***	.22
Sequence learning	–0.09	0.97	–0.10	.92	< .01
Motor learning	–0.15	0.93	–0.16	.87	< .01
Inhibition	0.20	0.96	0.21	.83	< .01

Table 6 presents the model for collocations, explaining 55% of the overall variance. The significant predictors in this model are grammatical analogies and inhibition, contributing 43% and 4% of the variability respectively.

Table 6. ***Model parameters for collocations***

	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	Relative Importance
Multiple R-squared: 0.55					
(Intercept)	73.67	0.88	83.39	< .001***	
Grammatical analogies	7.87	1.06	7.44	< .001**	.43
Scene analogies	0.30	1.03	0.29	.77	.08
Sequence learning	–1.03	0.94	–1.10	.27	.01
Motor learning	–0.44	0.90	–0.49	.63	< .01
Inhibition	–2.49	0.93	–2.68	.01**	.04

The model for expressive grammar (Table 7) explains 57% of the variance. Significant contributions are made by grammatical analogies (43%) and inhibition (2%).

Table 7. ***Model parameters for expressive grammar***

	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	Relative Importance
Multiple R-squared: 0.57					
(Intercept)	44.82	1.34	33.36	< .001***	
Grammatical analogies	11.75	1.61	7.31	< .001***	.43
Scene analogies	1.24	1.56	0.79	.43	.09
Sequence learning	0.32	1.43	0.23	.82	.01
Motor learning	–1.76	1.37	–1.29	.20	.01
Inhibition	–3.39	1.41	–2.40	.02*	.02

Discussion

The aim of this study was to explore potential differences between children with and without DLD in specific cognitive and metalinguistic abilities: analogical reasoning (nonverbal analogy assessed using a scene analogy task and verbal analogy by a grammatical analogies task), implicit statistical learning, and inhibition. If language is dependent upon these abilities (as discussed in the introductory section), we would expect children with DLD to have weaknesses in these areas. We also examine the extent to which these abilities predict language abilities in children with and without DLD.

Group Comparisons on Language Abilities

Although comparing the language abilities of the two groups was not our primary research aim (given that children with DLD are known to have language difficulties), a comparison of the group results in the context of the current literature helps to interpret the findings more effectively. We find large differences between the groups for recalling sentences, receptive grammar, vocabulary and collocational knowledge. The largest effect is observed for expressive grammar (d = -3.84), while the smallest effect is for vocabulary (d = -1.74). The effect size for collocational knowledge falls between these two measures (d = -1.99). The differences in performance on linguistic assessments are consistent with the diverse linguistic profiles seen in children with DLD.

In the receptive grammar task, disparities between the two groups become larger in less frequent structures (e.g. object relatives and object clefts). Given that some children with DLD have difficulties with written language (Tucci & Choi, 2023), it is not surprising that performance is poorer on structures, such as relative clauses, that are found more frequently in written than spoken language (Cilibrasi et al., 2019), highlighting the importance of exposure in the acquisition of linguistic structures.

One possible explanation for the weaker performance of children with DLD on the collocations task is the relationship between collocational knowledge and vocabulary. While children with DLD may be just as good at inferring the meanings of new words from context as their LT peers (see Dąbrowska, 2009), they may struggle with tracking statistical probabilities in language (see Evans et al., 2009).

In the scene analogy task, the DLD group performed below the LT children, with a between-group effect size of d = -0.95. Even larger differences were observed in the grammatical analogies task (d = -1.81). These differences are consistent with prior studies that report difficulties in analogical reasoning in children with DLD

(Krzemien et al., 2017, 2019, 2020; Leroy et al., 2012, 2014; Nippold et al., 1988). However, we did not find significant group differences in implicit statistical learning or inhibition.

The Effects of Cognitive Abilities on Linguistic Abilities

The regression models show that grammatical analogies accounts for most of the variance in the linguistic measures. We observe larger effect sizes for grammar (especially expressive grammar), which is in line with previous research suggesting that sentence repetition is a reliable clinical marker for DLD (Conti-Ramsden et al., 2001; Poll et al., 2016). For grammar, there is quite a lot of variance that our model accounts for, and although we did not assess working memory limitations in this study, it is possible that this variance is due to weaknesses in phonological short-term memory, typically seen in children with DLD (Gathercole and Baddeley, 1990; Bishop & Hsu, 2015; Gordon et al., 2021; Haebig et al., 2019, Jackson et al., 2021; Leonard et al., 2019). Inhibition is a significant contributor to the collocations and expressive grammar models whereas motor learning and sequence learning explain little variance. Although the reliability of the sequence learning task is high, it still does not contribute significantly to the language abilities we measured in our study. Similarly, the inhibition task demonstrates fair reliability (α = .69, 95% CI [.57, .79]), suggesting that the lack of a significant effect is likely genuine.

Analogical Reasoning

The most striking finding from our results is that analogical reasoning predicts language abilities, but the two different tasks (grammatical analogies and scene analogies) do so in different ways. Scene analogies is only a significant predictor in the model predicting receptive grammar, while grammatical analogies is a significant predictor for all linguistic measures (vocabulary, collocations, receptive grammar, and expressive grammar).

We note that performance on the two analogical reasoning tasks were correlated in the DLD group but not in the LT group. This is likely due to ceiling effects in the LT group on the scene analogies task whereas in the DLD group, scores in the grammatical analogies task are just above chance (mean score of 33%). Although performance in the grammatical analogies task is much higher in the LT group (mean score of 62%), the range of scores (13-86%) suggests that the task is also difficult for LT children. This performance, however, is unlikely to be attributed solely to the linguistic complexity of the task. The vast majority of the stimulus sentences are simple or compound sentences. Approximately 17% of the sentences contain some form of

subordination (e.g. because) and 8% of the sentences are in the form of a simple question. Based on their performance in the receptive grammar task, the participants have little difficulty processing simple sentences such as actives, as well as those containing subject clefts, passives, or complex NPs. Rather, the grammatical analogies task demands metalinguistic processing that is still developing in children of this age. The weaker performance on the grammatical analogies task in the children with DLD compared to the LT children is likely reflective of their general difficulty with analogical reasoning.

To the best of our knowledge, this is the first study to explore the relationship between grammatical analogies (a form of metalinguistic ability that is typically measured with foreign language aptitude tests) and language in DLD children. Although the grammatical analogies task makes use of language, it is not a linguistic task as it does not involve comprehension or production, only reasoning about linguistic stimuli, making it a metalinguistic task. In fact, using the examples from the task description, it would be possible to complete the task even if the content words were replaced with nonsense words (Blank, blank blanked a BLANK at the blank. Blank blanked a blank with a blank). As discussed in the introduction, grammatical analogies is not considered relevant to L1 language acquisition because it is widely assumed that grammatical development is implicit (DeKeyser et al., 2010; Ellis, 1996; Ullman, 2001). However, recent studies provide evidence that grammatical analogies is related to grammar attainment in the L1 in adults (Blake, Dąbrowska & Llompart, 2025; Llompart & Dąbrowska, 2023; Winckel & Dąbrowska, 2024) and language typical children (Wright et al., 2024). Our research adds to the growing body of evidence that these explicit reasoning skills may also be relevant for native language acquisition of children, including individuals with DLD.

While language aptitude and intelligence are viewed as distinct constructs, there may be an overlap between the two, given that both IQ and aptitude tests commonly assess similar components such as working memory and vocabulary (Winckel & Dąbrowska, 2024). It is interesting to note that a meta-analysis by Li (2016) found a significant correlation between performance on language aptitude tests and IQ (r = .64), and similarly, Dąbrowska (2018) found a correlation of r = .50 between language aptitude (measured by the Language Analysis subtest from the Pimsleur Language Aptitude Battery: Pimsleur, 1966) and IQ. The correlations indicate that the constructs could be closely related.

Implicit Statistical Learning

Turning to performance in implicit statistical learning (ISL), we expected to find a

difference between the two groups in our ISL task, based on evidence from prior research that suggests an ISL deficit in children with DLD. Our results show similar amounts of sequence learning in both groups on the task overall. Response times between the groups diverge early in the task (from approximately block eleven onwards), and from thereon, children with DLD are consistently slower throughout the task. The fact that children with DLD appear to have overall slower processing speed in this task is consistent with studies that show slower RTs in ISL tasks in children with DLD (Gabriel et al., 2013; Lum et al., 2010; Tomblin et al., 2007).

We anticipated that ISL would predict linguistic abilities - however, this hypothesis was not confirmed. Some studies find minimal or no evidence of an ISL deficit in DLD (Lammertink et al., 2020; Oliveira et al., 2023; West et al., 2021), and many researchers express concerns regarding poor reliability of ISL tasks, especially in children (Arnon, 2020). We designed a new ISL task for this study and report overall split-half reliability of .99 (95% CI .99, 1).

Inhibition

We expected that children with DLD would show effects of interference in the inhibition task. This hypothesis was not confirmed, and our data revealed similar performance in both groups. These results are contrary to prior research which suggests that children with DLD have weaknesses in inhibition (Pauls & Archibald, 2016), or that inhibition could be related to analogical reasoning performance in children with DLD (see Krzemien et al., 2020). However, reliability for the inhibition task was very low in the DLD group. Low reliability is not uncommon in executive function tasks and one of the reasons given is that these tasks often tap into other cognitive abilities (see Pauls & Archibald, 2016).

Inhibition is a significant predictor for collocational knowledge (accounting for 4% of the overall variance) and expressive grammar, where it explained 2% of the overall variance. The effect of inhibition in the collocations task may be driven by the nature of the task itself. Unlike tasks with clear right or wrong answers, the collocation task requires participants to select the more natural-sounding option (e.g., *left and right* vs. *right and left).* Inhibition may play a role in suppressing the less conventional choice, thereby influencing performance on this task.

Hypotheses Summary

We predicted that we would find between-group differences in all cognitive predictors. This hypothesis was partially met, with large between-group differences in the

grammatical analogies task (d = -1.81), and the scene analogy task (d = -.95). Our remaining hypotheses were not met, as we did not find interference effects in the DLD group in the inhibition task, and there were no significant differences between groups on implicit statistical learning. Both groups showed similar amounts of learning in ISL, although the DLD group were slower overall. The key finding to emerge from the present study is that analogical reasoning consistently predicts language abilities and this warrants further research in DLD populations.

Overall, analogical reasoning (grammatical and scene analogies) is the cognitive predictor that explains the majority of the variance in linguistic performance. However, the contribution of each may be different in the two populations. Based on the correlations, children with DLD appear to rely more on scene analogies while LT children tap into grammatical analogies. This may also be due to the ceiling effects in the LT group on the scene analogies task. The differences between our models and the additional regression models with group in the appendix indicate that there is an additional source of variance that contributes to the language difficulties that children with DLD have, which was not measured in our study. The most likely factor is phonological short-term memory, which may contribute to language difficulties observed in children with DLD (Bishop & Hsu, 2015; Gathercole & Baddeley, 1990; Gordon et al., 2021; Haebig et al., 2019; Jackson et al., 2021; Leonard et al., 2019).

Limitations

There are limitations that should be considered in generalizing the results of this study. We did not assess nonverbal IQ and working memory, and it is highly possible that these may be mediating factors for some of the tasks (as per Blake, Dąbrowska, & Riches, 2025).

The language difficulties of children with DLD were considered in the assessment process and both visual and auditory instructions were given for all tasks. The vast majority of the tests did not rely on reading as a skill and where reading was required in the grammatical analogies task, the researcher was present to ensure that the sentence was read correctly. Although we did not examine literacy-based skills, including a measure of reading ability would have allowed for a more fine-grained analysis of the language abilities of the participants. Reading ability is closely tied to print exposure, with the direction of causation likely running from ability to exposure (Van Bergen et al., 2018). Print exposure is also an important source of exposure to lexically rich and syntactically complex language. This means that the extent to which reading difficulties are present may have a significant impact on the quantity of complex language that children are exposed to. This effect is hinted at in the receptive grammar

task where children with DLD show weaker performance on the structures that are more common in written language.

The demographics of our sample should be taken into consideration when interpreting the results. For example, there were higher education levels in the first parent/carer in the LT group in comparison to the DLD group. While this difference was not statistically significant, it may have influenced the findings. We did not request information on ethnicity of our participants. However, given that cultural differences may influence analogical reasoning, this should be considered in future research.

Lastly, with 73 participants in this study (39 LT children and 34 children with DLD), we recognise the limitation of statistical power, particularly given the number of predictors examined. Furthermore, whilst children with DLD were matched to language typical children based on age, the inclusion of an additional language matched group would have enhanced comparability and strengthen the findings reported in this study.

Conclusion

Our findings confirm that children with DLD show delayed language development when compared to their LT peers across several linguistic tasks. However, no significant differences were observed in inhibitory skills or ISL motor learning. Overall, the results of this study contribute to our understanding of the cognitive processes underlying language acquisition in children with and without DLD. According to usage-based theories of language acquisition, analogy is a key mechanism that enables children to generate novel utterances of their own. Our results suggest that challenges with analogical reasoning may help explain the language difficulties observed in DLD. Both general scene analogies and language-specific grammatical analogies play important roles in the acquisition of linguistic knowledge, and in fact, differences in performance on these two tasks account for most, though not all, of the difference in linguistic performance between DLD and language-typical children.

References

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language, 42*(2), 239–273. https://doi.org/10.1017/S030500091400049X

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Archibald, L. M. (2017). Working memory and language learning: A review. *Child Language Teaching and Therapy, 33*(1), 5–17. https://doi.org/10.1177/0265659016654206

Arnon, I. (2020). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods, 52*(1), 68–81. https://doi.org/10.3758/s13428-019-01205-5

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*(1), 67–82. https://doi.org/10.1016/j.jml.2009.09.005

Bates, E., Bretherton, I., & Snyder, L. S. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge University Press.

Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition. *Linguistics, 47*(2), 383–411. https://doi.org/10.1515/LING.2009.014

Bishop, D. V., & Garsell, M. (2003). *Test for reception of grammar: Version 2: TROG-2 manual*.

Bishop, D. V. M., & 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), Article 3. https://doi.org/10.1186/s40359-015-0062-7

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

Blake, A., Dąbrowska, E., & Llompart, M. (2025). Toward an understanding of developmental language disorder in adults: Investigating the relationship between cognitive abilities and linguistic outcomes. *Journal of Speech, Language, and Hearing Research, 68*(8), 3895–3918. https://doi.org/10.1044/2025_JSLHR-24-00557

Blake, A., Dąbrowska, E., & Riches, N. (2025). Exploring cognitive predictors of language in children with developmental language disorder: The role of nonverbal IQ,

working memory, implicit statistical learning, and speed of automatization. *Journal of Communication Disorders, 116*, Article 106541. https://doi.org/10.1016/j.jcomdis.2025.106541

Boerma, T., Leseman, P., Wijnen, F., & Blom, E. (2017). Language proficiency and sustained attention in monolingual and bilingual children with and without language impairment. Frontiers in Psychology, 8, Article 1241. https://doi.org/10.3389/fpsyg.2017.01241

Bybee, J. (2010). Language, usage and cognition. Cambridge University-Press. https://doi.org/10.1017/CBO9780511750526

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Newbury House.

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. S. Parry & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11–29). Prentice Hall Regents.

Carroll, J. B., & Sapon, S. M. (1959). *Modern Language Aptitude Test*. Second Language Testing.

Carroll, J. B., & Sapon, S. M. (1967). *Modern Language Aptitude Test-Elementary Manual.* San Antonio, Texas: The Psychological Corporation.

Carroll, J. B., & Sapon, S. M. (2002). *Modern Language Aptitude Test, MLAT: Manual.* Second Language Testing.

Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science, 38*(2), 383–397. https://doi.org/10.1111/cogs.12099

Cilibrasi, L., Adani, F., & Tsimpli, I. (2019). Reading as a predictor of complex syntax. The case of relative clauses. *Frontiers in Psychology, 10*, Article 1450. https://doi.org/10.3389/fpsyg.2019.01450

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42*(6), 741–748. https://doi.org/10.1111/1469-7610.00770

Dąbrowska, E. (2009). Words as constructions. In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics* (Vol. 24, pp. 201–223). John Benjamins. https://doi.org/10.1075/hcp.24.16dab

Dąbrowska, E. (2010a). Formulas in the acquisition of English interrogatives: A case study. In D. Stanulewicz, T. Z. Wolanski, & J. Redzimska (Eds.), *LinguaTerra cognita II: A festschrift for Roman Kalisz* (pp. 675–702). Wydawnictwo Uniwersytetu Gdańskiego.

Dąbrowska, E. (2010b). Productivity, proceduralization and specific language impairment. *Human Development, 53*(5), 278–286. https://doi.org/10.1159/000321290

Dąbrowska, E. (2014a). Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics, 25*(4), 617–653. https://doi.org/10.1515/cog-2014-0057

Dąbrowska, E. (2014b). Words that go together. *The Mental Lexicon, 9*(3), 401–418. https://doi.org/10.1075/ml.9.3.02dab

Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition, 178*, 222–235. https://doi.org/10.1016/j.cognition.2018.05.018

Dąbrowska, E. (2019). Experience, aptitude, and individual differences in linguistic attainment: A comparison of native and nonnative speakers. *Language Learning, 69*(S1), 72–100. https://doi.org/10.1111/lang.12323

Dąbrowska, E., & Blake, A. (2022). Speed of automatization predicts performance on "decorative" grammar in second language learning. In T. Piske & A. K. Steinlen (Eds.), *Cognition and second language acquisition: Studies on pre-school, primary school and secondary school children* (pp. 285-309). Narr Francke Attempto.

DeKeyser, R., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics, 31*(3), 413–438. https://doi.org/10.1017/S0142716410000056

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Divjak, D., Milin, P., Medimorec, S., & Borowski, M. (2022). Behavioral signatures of memory resources for language: Looking beyond the lexicon/grammar divide. *Cognitive Science, 46*(11), Article e13206. https://doi.org/10.1111/cogs.13206

Dunn, L. M. (2009). *The British Picture Vocabulary Scale* (3rd ed.). GL Assessment.

Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). *The British Picture Vocabulary Scale* (3rd ed.). GL Assessment Ltd.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18*(1), 91–126. https://doi.org/10.1017/S0272263100014698

Ellis, N. C. (2007). Implicit and explicit knowledge about language. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed., Vol. 6, Knowledge about language, pp. 119–132). Springer.

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)

Farkas, B. C., Krajcsi, A., Janacsek, K., & Nemeth, D. (2024). The complexity of measuring reliability in learning tasks: An illustration using the Alternating Serial Reaction Time Task. *Behavior Research Methods,* 57(7), 185 https://doi.org/10.3758/s13428-022-02038-5

Gabriel, A., Maillart, C., Stefaniak, N., Lejeune, C., Desmottes, L., & Meulemans, T. (2013). Procedural learning in specific language impairment: Effects of sequence complexity. *Journal of the International Neuropsychological Society, 19*(3), 264–271. https://doi.org/10.1017/S1355617712001270

Gallinat, E., & Spaulding, T. J. (2014). Differences in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 57*(4), 1363–1382. https://doi.org/10.1044/2014_JSLHR-L-12-0363

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*(3), 336–360. https://doi.org/10.1016/0749-596X(90)90004-J

Gentner, D., & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition, 2*(2), 261–283. https://doi.org/10.1515/langcog.2010.011

Geurten, M., Catale, C., & Meulemans, T. (2016). Involvement of executive functions in children's metamemory. *Applied Cognitive Psychology, 30*(1), 70–80. https://doi.org/10.1002/acp.3168

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Gordon, K. R., Storkel, H. L., Lowry, S. L., & Ohlmann, N. B. (2021). Word learning by preschool-age children with developmental language disorder: Impaired encoding and robust consolidation during slow mapping. *Journal of Speech, Language, and Hearing Research, 64*(11), 4250–4270. https://doi.org/10.1044/2021_JSLHR-21-00046

Griffiths, S., Kievit, R. A., & Norbury, C. (2022). Mutualistic coupling of vocabulary and non-verbal reasoning in children with and without language disorder. *Developmental Science, 25*(3), Article e13208. https://doi.org/10.1111/desc.13208

Grömping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software, 17*(1). https://doi.org/10.18637/jss.v017.i01

Haebig, E., Leonard, L. B., Deevy, P., Karpicke, J., Christ, S. L., Usler, E., Kueser, J. B., Souto, S., Krok, W., & Weber, C. (2019). Retrieval-based word learning in young typically developing children and children with development language disorder II: A comparison of retrieval schedules. *Journal of Speech, Language, and Hearing Research, 62*(4), 944–964. https://doi.org/10.1044/2018_JSLHR-L-18-0071

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

Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Veríssimo, J., Dye, C. D., Alm, P., Jennische, M., Tomblin, B. J., & 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

Henry, L. A., & Botting, N. (2017). Working memory and developmental language impairments. *Child Language Teaching and Therapy, 33*(1), 19–32. https://doi.org/10.1177/0265659016655378

Herbst, T. (2011). Choosing sandy beaches – Collocations, probabemes and the idiom principle. In T. Herbst, S. Faulhaber, & P. Uhrig (Eds.), *Phraseological view of language: A tribute to John Sinclair* (pp. 27–57). De Gruyter Mouton.

Ibbotson, P., & Kearvell-White, J. (2015). Inhibitory control predicts grammatical ability. PLoS ONE, 10(12), Article e0145030. https://doi.org/10.1371/journal.pone.0145030

Isbilen, E. S., & Christiansen, M. H. (2022). Statistical learning of language: A meta-analysis into 25 years of research. *Cognitive Science, 46*(9), e13198. https://doi.org/10.1111/cogs.13198

Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2021). Word learning and verbal working memory in children with developmental language disorder. *Autism & Developmental Language Impairments, 6*, Article 23969415211004109. https://doi.org/10.1177/23969415211004109

Krzemien, M., Jemel, B., & Maillart, C. (2017). Analogical reasoning in children with specific language impairment: Evidence from a scene analogy task. *Clinical Linguistics & Phonetics, 31*(7–9), 573–588. https://doi.org/10.1080/02699206.2017.1302509

Krzemien, M., Maillart, C., Parisse, C., & Leroy, S. (2019). Impact of processing load on analogical mapping with visual sequences in children with developmental language disorders (DLD). *International Journal of Language & Communication Disorders, 54*(3), 417–429. https://doi.org/10.1111/1460-6984.12446

Krzemien, M., Thibaut, J.-P., & Maillart, C. (2020). How language and inhibition influence analogical reasoning in children with or without developmental language disorder? *Journal of Clinical and Experimental Neuropsychology, 42*(1), 76–89. https://doi.org/10.1080/13803395.2019.1676881

Krzemien, M., Seret, E., & Maillart, C. (2021). The generalisation of linguistic constructions in children with or without developmental language disorders. *Journal of Child Language, 48*(2), 413–427. https://doi.org/10.1017/S0305000920000288

Lammertink, I., Boersma, P., Rispens, J., & Wijnen, F. (2020). Visual statistical learning in children with and without DLD and its relation to literacy in children with DLD. *Reading and Writing, 33*(6), 1557–1589. https://doi.org/10.1007/s11145-020-10018-4

Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2017). Statistical learning in specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 60*(12), 3474–3486. https://doi.org/10.1044/2017_JSLHR-L-16-0439

Lancaster, H. S., & Camarata, S. (2019). Reconceptualizing developmental language disorder as a spectrum disorder: Issues and evidence. *International Journal of Language & Communication Disorders, 54*(1), 79–94. https://doi.org/10.1111/1460-6984.12433

Lancaster, H. S., Milne, A., Gordon, K., Emerson, S. N., & Selin, C. (2024). *Analyzing the impact of explicit and implicit cognitive constructs on NVIQ test performance* [Preprint]. OSF. https://doi.org/10.31219/osf.io/fusr6

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

Leonard, L. B., Karpicke, J., Deevy, P., Weber, C., Christ, S., Haebig, E., Souto, S., Kueser, J. B., & Krok, W. (2019). Retrieval-based word learning in young typically developing children and children with developmental language disorder I: The benefits of repeated retrieval. *Journal of Speech, Language, and Hearing Research, 62*(4), 932–943. https://doi.org/10.1044/2018_JSLHR-L-18-0070

Leroy, S., Maillart, C., & Parisse, C. (2014). Analogical mapping across modalities in children with specific language impairment (SLI). *Research in Developmental Disabilities, 35*(9), 2158–2171. https://doi.org/10.1016/j.ridd.2014.05.005

Leroy, S., Parisse, C., & Maillart, C. (2012). Analogical reasoning in children with specific language impairment. *Clinical Linguistics & Phonetics, 26*(4), 380–396. https://doi.org/10.3109/02699206.2011.641059

Li, S. (2016). The construct validity of language aptitude: A meta-analysis. *Studies in second language acquisition, 38*(4), 801–842. https://doi.org/10.1017/S027226311500042X

Llompart, M., & Dąbrowska, E. (2020). Explicit but not implicit memory predicts ultimate attainment in the native language. *Frontiers in Psychology, 11*, Article 569586. https://doi.org/10.3389/fpsyg.2020.569586

Llompart, M., & Dąbrowska, E. (2023). "Foreign" language aptitude predicts individual differences in native grammatical proficiency. *Linguistics,* 61*(5), 1165-1193.* https://doi.org/10.1515/ling-2022-0009

Lum, J. A. G., 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. G., 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

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*(2), 339–366. https://doi.org/10.1017/S0305000900009302

Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: A review of associations with formal language, social cognition, and executive functions. *Language Learning and Development, 14*(3), 186–223. https://doi.org/10.1080/15475441.2018.1455584

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

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning, 62*(1), 302–331. https://doi.org/10.1111/j.1467-9922.2010.00626.x

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology, 19*(1), 78–94. https://doi.org/10.1044/1058-0360(2009/09-0028)

Mountford, H. S., Braden, R., Newbury, D. F., & Morgan, A. T. (2022). The genetic and molecular basis of developmental language disorder: A review. *Children, 9*(5), Article 586. https://doi.org/10.3390/children9050586

Nippold, M. A., Erskine, B. J., & Freed, D. B. (1988). Proportional and functional analogical reasoning in normal and language-impaired children. *Journal of Speech and Hearing Disorders, 53*(4), 440–448. https://doi.org/10.1044/jshd.5304.440

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. *Journal of Child Psychology and Psychiatry, 57*(11), 1247–1257. https://doi.org/10.1111/jcpp.12573

Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. G. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Frontiers in Psychology, 7*, Article 1245. https://doi.org/10.3389/fpsyg.2016.01245

Oliveira, C. M., Henderson, L. M., & Hayiou-Thomas, M. E. (2023). Limited evidence of an association between language, literacy, and procedural learning in typical and atypical development: A meta-analysis. *Cognitive Science, 47*(7), Article e13310. https://doi.org/10.1111/cogs.13310

Parsons, S. (2021). splithalf: Robust estimates of split half reliability. *Journal of Open Source Software, 6*(60), Article 3041. https://doi.org/10.21105/joss.03041

Pauls, L. J., & Archibald, L. M. D. (2016). Executive functions in children with specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 59*(5), 1074–1086. https://doi.org/10.1044/2016_JSLHR-L-15-0174

Pimsleur, P. (1966). *Pimsleur language aptitude battery (form S)*. Harcourt, Brace and World.

Poll, G. H., Miller, C. A., & Van Hell, J. G. (2016). Sentence repetition accuracy in adults with developmental language impairment: Interactions of participant capacities and sentence structures. *Journal of Speech, Language, and Hearing Research, 59*(2), 302–316. https://doi.org/10.1044/2015_JSLHR-L-15-0020

Prela, L., Llompart, M., & Dąbrowska, E. (2022). Aptitude and experience as predictors of grammatical proficiency in adult Greek-English bilinguals. *Frontiers in Psychology, 13*, Article 1062821. https://doi.org/10.3389/fpsyg.2022.1062821

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., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research, 39*(6), 1239–1257. https://doi.org/10.1044/jshr.3906.1239

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language, and Hearing Research, 38*(4), 850–863. https://doi.org/10.1044/jshr.3804.850

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1412–1431. https://doi.org/10.1044/jslhr.4106.1412

Riches, N., Letts, C., Awad, H., Ramsey, R., & Dąbrowska, E. (2022). Collocational knowledge in children: A comparison of English-speaking monolingual children, and children acquiring English as an additional language. *Journal of Child Language, 49*(5), 1008–1023. https://doi.org/10.1017/S0305000921000490

Riches, N. G., Tomasello, M., & Conti-Ramsden, G. (2005). Verb learning in children with SLI: Frequency and spacing effects. *Journal of Speech, Language, and Hearing Research, 48*(6), 1397–1411. https://doi.org/10.1044/1092-4388(2005/097)

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*(3), 249–273. https://doi.org/10.1016/j.jecp.2006.02.002

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

Smith, S. A., & Murphy, V. A. (2015). Measuring productive elements of multi-word phrase vocabulary knowledge among children with English as an additional or only language. *Reading and Writing, 28*(3), 347–369. https://doi.org/10.1007/s11145-014-9527-y

Tallal, P., & Piercy, M. (1973a). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature, 241*(5390), 468–469. https://doi.org/10.1038/241468a0

Tallal, P., & Piercy, M. (1973b). Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia, 11*(4), 389–398. https://doi.org/10.1016/0028-3932(73)90025-0

Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia, 12*(1), 83–93. https://doi.org/10.1016/0028-3932(74)90030-X

Tallal, P., & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia, 13*(1), 69–74. https://doi.org/10.1016/0028-3932(75)90049-4

Tucci, A., & Choi, E. (2023). Developmental language disorder and writing: A scoping review from childhood to adulthood. *Journal of Speech, Language, and Hearing Research, 66*(8), 2900–2920. https://doi.org/10.1044/2023_JSLHR-22-00583

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

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

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*(6), 1245–1260. https://doi.org/10.1044/jslhr.4006.1245

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience, 2*(10), 717–726. https://doi.org/10.1038/35094573

Van Bergen, E., Snowling, M. J., De Zeeuw, E. L., Van Beijsterveldt, C. E. M., Dolan, C. V., & Boomsma, D. I. (2018). Why do children read more? The influence of reading ability on voluntary reading practices. *Journal of Child Psychology and Psychiatry, 59*(11), 1205–1214. https://doi.org/10.1111/jcpp.12910

van der Lely, H. K. J. (2005). Domain-specific cognitive systems: Insight from Grammatical-SLI. *Trends in Cognitive Sciences, 9*(2), 53–59. https://doi.org/10.1016/j.tics.2004.12.002

West, G., Melby-Lervåg, M., & Hulme, C. (2021). Is a procedural learning deficit a causal risk factor for developmental language disorder or dyslexia? A meta-analytic review. *Developmental Psychology, 57*(5), 749–770. https://doi.org/10.1037/dev0001172

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). Springer. https://ggplot2.tidyverse.org

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation* (Version 1.1.4) [Computer software]. https://CRAN.R-project.org/package=dplyr

Wiig, E. H., Semel, E. M., & Secord, W. (2013). *CELF-5: Clinical Evaluation of Language Fundamentals* (5th ed.). Pearson.

Winckel, E., & Dąbrowska, E. (2024). Language analytic ability, print exposure, memory and comprehension of complex syntax by adult native speakers. *Journal of Cognition, 7*(1), Article 7. https://doi.org/10.5334/joc.333

Wright, R., Geertsema, S., Le Roux, M., Winckel, E., & Dąbrowska, E. (2024). The acquisition of constructions: Does modality matter? *Reading & Writing, 15*(1), Article a489. https://doi.org/10.4102/rw.v15i1.489

Zapparrata, N. M., Brooks, P. J., & Ober, T. (2023). Developmental language disorder is associated with slower processing across domains: A meta-analysis of time-based tasks. *Journal of Speech, Language, and Hearing Research, 66*(1), 325–346. https://doi.org/10.1044/2022_JSLHR-22-00221

Data, Code and Materials Availability Statement

The preregistration information as well as the data and code used for the analysis are available in an online data repository on OSF (https://osf.io/uc2qa/?view_only=e0e2327d46de49abb14da339688717ae).

Ethics Statement

Ethics approval was obtained from the ethics committee of the University of Birmingham (ERN_0365).

Authorship and Contributorship Statement

Ashley Blake: conceptualization, pre-registration, design of online experimental tasks, data collection, writing: initial draft, reviewing and editing, contribution to data processing and analysis. **Richenda Wright**: conceptualization, pre-registration, design of online experimental tasks, data collection, writing: contribution to initial draft, reviewing and editing, data processing and analysis. **Ewa Dąbrowska**, **Nick Riches, Elodie Winckel**: conceptualization, writing: reviewing and editing, guidance on data analysis. The overall project was supervised by **Ewa Dąbrowska**.

Acknowledgements

The authors thank the children who participated in our study, and their families for their help in arranging and supporting the online sessions. The online games were presented on Gorilla Game Builder with a subscription and tokens funded by the 'Bishop' prize awarded to Ashley Blake. This study was funded the German Research Foundation (DFG) [Research Training Group "Dimensions of Constructional Space", project number 468527017], and by an Alexander von Humboldt Professorship (ID-1195918) awarded to Ewa Dąbrowska.

Appendices

Appendix 1. *Correlations for participants with DLD (N = 34).*

	Age	Vocab	RS	Coll	RGr	GA	SA	ISL	ML	Inhib
Age										
Vocabulary	.50**									
Recalling sentences	.31	.42*								
Collocations	.35*	.29	.51**							
Receptive grammar	–.20	.12	.22	.26						
Grammatical analogies	–.06	.07	.19	.20	.21					
Scene analogies	–.04	.26	.07	.14	.60**	.46**				
ISL - Sequence learning	–.14	–.15	0	–.14	–.04	.14	–.10			
Motor learning	.12	.24	.17	–.05	.32	.17	.24	.16		
Inhibition	–.11	.13	–.23	–.46**	.12	–.06	.13	–.30	.14	

RS: Recalling sentences; Coll: collocations; RGr: receptive grammar; GA: grammatical analogies (total score), SA: scene analogies, ISL: implicit sequence learning; ML: motor learning; Inhib: inhibition

Note: correlations of .44 and above are significant at $p < .01$** and correlations of .34 and above are significant at $p < .05$*.

Appendix 2. ***Correlations for the LT group (N = 39).***

	Age	Vocab	SR	Coll	RGr	GA	SA	ISL	ML	Inhib
Age										
Vocabulary	.44**									
Recalling sentences	.47**	.59**								
Collocations	.59**	.52**	.48**							
Receptive grammar	.32*	.56**	.47**	.41**						
Grammatical analogies	.48**	.47**	.44**	.48**	.54**					
Scene analogies	0	.19	.07	.11	.11	.24				
ISL - Sequence learning	.08	–.16	.17	-.08	.12	.23	–.10			
Motor learning	.05	–.05	-.17	.12	.06	.13	.10	–.16		
Inhibition	–.02	–.01	.10	.16	.39*	.49**	–.15	–.21	.06	

RS: Recalling sentences; Coll: collocations; RGr: receptive grammar; GA: grammatical analogies (total score), SA: scene analogies, ISL: implicit sequence learning; ML: motor learning; Inhib: inhibition

Note: correlations of .41 and above are significant at $p < .01$** and correlations of .32 and above are significant at $p < .05$*.

Appendix 3. ***Correlations for both DLD and LT groups together (N = 73).***

	Age	Vocab	SR	Coll	RGr	GA	SA	ISL	ML	Inhib
Age										
Vocabulary	.40**									
Recalling sentences	.23	.76**								
Collocations	.38**	.68**	.79**							
Receptive grammar	.06	.64**	.76**	.68**						
Grammatical analogies	.26*	.61**	.71**	.67**	.68**					
Scene analogies	.01	.44**	.41**	.39**	.59**	.48**				
ISL - Sequence learning	–.03	–.07	.09	–.02	.07	.18	–.06			
Motor learning	.07	0	–.09	–.02	.05	.05	.10	–.02		
Inhibition	–.07	.01	–.08	–.13	.11	.19	–.01	–.25	.09	

RS: Recalling sentences; Coll: collocations; RGr: receptive grammar; GA: grammatical analogies (total score), SA: scene analogies, ISL: implicit sequence learning; ML: motor learning; Inhib: inhibition

Note: correlations of .30 and above are significant at $p < .01$** and correlations of .24 and above are significant at $p < .05$*.

Appendix 4. ***Model parameters for vocabulary with Group as a variable.***

	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	Relative Importance
Multiple R-squared: 0.54					
(Intercept)	119.13	2.92	40.78	< .001***	
Grammatical analogies	5.27	2.44	2.16	.03*	.19
Scene analogies	1.95	1.89	1.04	.30	.08
Sequence learning	–2.74	1.70	–1.61	.11	.01
Motor learning	0.17	1.65	0.11	.92	< .01
Inhibition	0.33	1.73	0.19	.85	< .01
Group	17.37	4.61	3.77	< .001***	.26

Appendix 5. ***Model parameters for receptive grammar with Group as variable.***

	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	Relative Importance
Multiple R-squared: 0.69					
(Intercept)	75.59	1.39	54.46	< .001***	
Grammatical analogies	1.74	1.16	1.50	.14	.20
Scene analogies	3.04	0.90	3.39	.001**	.16
Sequence learning	–0.08	0.81	–0.09	.93	< .01
Motor learning	0.52	0.79	0.67	.51	< .01
Inhibition	1.18	0.82	1.44	.15	.01
Group	11.93	2.19	5.45	< .001***	.32

Appendix 6. ***Model parameters for collocations with Group as a variable.***

	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	Relative Importance
Multiple R-squared: 0.65					
(Intercept)	68.73	1.44	47.58	< .001***	
Grammatical analogies	4.83	1.20	4.01	< .001***	.26
Scene analogies	–0.29	0.93	–0.31	.76	.05
Sequence learning	–1.02	0.84	–1.21	.23	.01
Motor learning	0.09	0.82	0.11	.91	< .01
Inhibition	–1.72	0.85	–2.02	.05*	.02
Group	9.34	2.28	4.10	< .001***	.30

Appendix 7. ***Model parameters for recalling sentences with Group as a variable.***

	Parameter Estimate	Standard Error	*t* value	Pr(>|t|)	Relative Importance
Multiple R-squared: 0.82					
(Intercept)	32.14	1.58	20.33	< .001***	
Grammatical analogies	3.97	1.32	3.01	.004**	.24
Scene analogies	–0.28	1.02	–0.27	.79	.06
Sequence learning	0.36	0.92	0.39	.70	< .01
Motor learning	–0.41	0.89	–0.46	.65	< .01
Inhibition	–1.43	0.93	–1.53	.13	.01
Group	23.92	2.50	9.59	< .001***	.50

Appendix 8. ***ANOVA comparisons of models with and without Group as a variable.***

		Res.DF	RSS	DF	Sum of Squares	F	Pr(>|t|)
Vocabulary	Without Group	65	14252.0				
	With Group	64	11665.0	1	2587.8	14.20	< .001***
Receptive grammar	Without Group	65	3583.4				
	With Group	64	2633.0	1	1220.4	29.67	< .001***
Collocations	Without Group	65	3599.2				
	With Group	64	2851.8	1	747.4	16.77	< .001***
Recalling sentences	Without Group	65	8322.7				
	With Group	64	3417.2	1	4905.5	91.87	< .001***

License

Language Development Research (ISSN 2771-7976) is published by TalkBank and the Carnegie Mellon University Library Publishing Service. Copyright © 2026 The Author(s). This work is distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 International license (https://creativecommons.org/licenses/by-nc/4.0/), which permits any use, reproduction and distribution of the work for noncommercial purposes without further permission provided the original work is attributed as specified under the terms available via the above link to the Creative Commons website.