


ORIGINAL ARTICLE

Grammatical performance in children with dyslexia: the contributions of individual differences in phonological memory and statistical learning

Merel van Witteloostuijn^{1*} , Paul Boersma¹, Frank Wijnen² and Judith Rispens¹

¹Department of Linguistics, Amsterdam Center for Language and Communication, University of Amsterdam, Amsterdam, the Netherlands and ²Department of Linguistics, Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, the Netherlands

*Corresponding author. Email: m.t.g.vanWitteloostuijn@gmail.com

(Received 1 November 2019; revised 18 February 2021; accepted 9 March 2021; first published online 08 April 2021)

Abstract

Several studies have signaled grammatical difficulties in individuals with developmental dyslexia. These difficulties may stem from a phonological deficit, but may alternatively be explained through a domain-general deficit in statistical learning. This study investigates grammar in children with and without dyslexia, and whether phonological memory and/or statistical learning ability contribute to individual differences in grammatical performance. We administered the CELF “word structure” and “recalling sentences” subtests and measures of phonological memory (digit span, nonword repetition) and statistical learning (serial reaction time, nonadjacent dependency learning) among 8- to 11-year-old children with and without dyslexia ($N = 50$ per group). Consistent with previous findings, our results show subtle difficulties in grammar, as children with dyslexia achieved lower scores on the CELF (word structure: $p = .0027$, recalling sentences: $p = .053$). While the two phonological memory measures were found to contribute to individual differences in grammatical performance, no evidence for a relationship with statistical learning was found. An error analysis revealed errors in irregular morphology (e.g., plural and past tense), suggesting problems with lexical retrieval. These findings are discussed in light of theoretical accounts of the underlying deficit in dyslexia.

Keywords: morphology; syntax; phonological memory; statistical learning; dyslexia

Introduction

Developmental dyslexia (henceforth “dyslexia”) is a learning disability that is characterized by impaired reading and spelling in spite of normal intelligence and educational opportunities and in absence of sensory impairments (American Psychiatric Association, 2013; Snowling, 2001). Individuals with (a familial risk

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

of dyslexia are known to experience difficulties in the area of phonological skills (Vellutino *et al.*, 2004; see Melby-Lervåg *et al.*, 2012 for a meta-analysis), which has led to the predominant view that the literacy impairments in dyslexia stem from an underlying phonological deficit. When learning to read and spell, children must acquire the correspondences between letters and sounds (i.e., graphemes and phonemes). If, however, the processing, storage, and/or representation of phonological information is impaired, children experience difficulties in the acquisition of grapheme–phoneme mappings that in turn result in problems with literacy acquisition (e.g., Ramus & Szenkovits, 2008).

Over time, researchers have uncovered impairments in cognitive abilities other than literacy and phonological skills. These may involve general cognitive skills such as the processing of visual and auditory input (e.g., Stein & Walsh, 1997; Tallal, 2004), attention (e.g., Facoetti *et al.*, 2000), and motor functioning (e.g., Ramus, 2003), but may also extend to language domains other than literacy and phonology, such as inflectional morphology and syntax (e.g., Rispens & Been, 2007; Robertson & Joanisse, 2010; Scarborough, 1990). Together, these observations have led to suggestions of a more general learning deficit in dyslexia (Nicolson & Fawcett, 2007, 2011), that is, a deficit of the domain-general ability to detect statistical patterns in sensory input, including spoken and written language (henceforth “statistical learning”, but note that Nicolson and Fawcett [2007, 2011] use the term “procedural learning deficit”; Gabay *et al.*, 2015; Lum *et al.*, 2013, but see, e.g., Schmalz *et al.*, 2017, and Van Witteloostuijn *et al.*, 2017).

The aim of the present study is twofold: (1) to investigate the performance of children with and without dyslexia on measures assessing inflectional morphology and syntax and (2) to examine whether children’s performance in these domains can be explained by individual differences in phonological processing and memory and/or statistical learning ability. In doing so, we contribute to the existing literature on grammatical ability in children with dyslexia and enhance our understanding of their difficulties in this area. Most importantly, we hope to provide novel insights into the underlying cause of the linguistic difficulties observed in dyslexia by investigating two opposing theories (i.e., phonological or statistical learning deficit).

Grammatical performance in children with dyslexia

Besides pronounced deficiencies in the areas of literacy skills and phonology, individuals with dyslexia or with an elevated, familial risk of dyslexia (FRdys) have been shown to experience delays in oral language development in early childhood (see Snowling & Melby-Lervåg, 2016 for a meta-analysis). Studies of spoken language skills in young FRdys children have shown that they produce shorter sentences of lower syntactic complexity and achieve lower vocabulary scores than low-risk children (e.g., van Alphen *et al.*, 2004; Chen *et al.*, 2017; Koster *et al.*, 2005; Lyytinen *et al.*, 2001; Ramus *et al.*, 2013; Scarborough, 1990; Snowling & Melby-Lervåg, 2016). Furthermore, when school-aged children with dyslexia or FRdys children are compared to their TD peers, they are found to achieve lower scores on standardized tests of grammar (e.g., the Clinical Evaluation of Language Fundamentals [CELF]; total language score: McArthur *et al.*, 2000; word structure subtest: Joanisse *et al.*, 2000; recalling sentences subtest: Ramus *et al.*, 2013; Finnish inflectional morphology test: Aro *et al.*, 2012; Lyytinen *et al.*, 2001; cf., e.g., Carroll &

Myers, 2010). Similarly, there are indications that (pre)school-aged children with dyslexia and FRdys children perform more poorly on experimental tasks that assess inflectional morphology, including pluralization (Joanisse et al., 2000), subject–verb agreement (Jiménez et al., 2011; Nash et al., 2013; Rispens et al., 2004; Rispens & Been, 2007), and past tense formation (Joanisse et al., 2000; Nash et al., 2013). The same holds for the comprehension of sentences (Robertson & Joanisse, 2010) and the correct interpretation (and production) of complex syntactical structures such as passive sentences (Reggiani, 2010; Shankweiler et al., 1995), relative clauses (Bar-Shalom et al., 1993; Mann et al., 1984; Shankweiler et al., 1995), and referential pronouns (Waltzman & Cairns, 2000). Even though overt morphological and grammatical difficulties (in oral language) appear to attenuate with age, some studies suggest that such difficulties may be persistent (Cantiani et al., 2013; Leikin and Zur Hagit 2006; Rispens et al., 2006; Vender et al., 2017).

Note, however, that group comparisons of children with dyslexia with age-matched TD children have yielded null findings as well, both on standardized tests of grammar (Carroll & Myers, 2010; Ramus et al., 2013) and on experimental tasks examining inflectional morphology and syntax (e.g., Rispens et al., 2014; Ramus et al., 2013). It is currently unclear how this should be interpreted. Conceivably, the reported difficulties in the area of inflectional morphology and syntax in individuals with dyslexia may be restricted to specific grammatical processes (or may only affect subgroups of children with dyslexia; see Rispens et al., 2004). On the other hand, it is also possible that tasks employed in these studies were insufficiently sensitive. In any case, it is relevant and necessary to contribute new data.

The abovementioned oral language difficulties in children with dyslexia are reminiscent of developmental language disorder (DLD; previously known as specific language impairment; Bishop et al., 2017), a disorder that is defined by oral language problems and pronounced difficulties in the areas of morphology and syntax. Dyslexia and DLD are distinct diagnoses that can co-occur within a single child. The behavioral overlap between the two disorders is known to be high (e.g., McArthur et al., 2000; Catts et al., 2005), which has raised the question of whether the two disorders should be viewed as distinct or as two points on a single continuum (e.g., Bishop & Snowling, 2004). Although we are aware that the debate on this matter continues, we leave it aside here. We focus on grammatical performance in children who have been diagnosed with dyslexia and do not meet the diagnostic criteria for DLD.

Theories of dyslexia: Phonological deficit and statistical learning deficit

Theories of the underlying cause of dyslexia should not only account for the impairments in the area of reading and spelling, but should also be able to explain the abovementioned difficulties with inflectional morphology and syntax. In line with the dominant view that dyslexia originates from a deficit in phonological skills, grammatical problems in dyslexia have been theorized to be “further symptoms of an underlying phonological weakness” (Shankweiler et al., 1995, p. 149). This idea is supported by evidence that children with dyslexia are especially impaired in morpho-phonology – morphological processes that interact with phonology

(Shankweiler *et al.*, 1995; Rispens & Been, 2007). In such processes, the selection between allomorphs depends on the phonological characteristics of the stem. For example, the selection of the /t/, /d/, or /ɪd/ (or /əd/, depending on the region) allomorph in English past tense verb inflection, as in *bake – baked*, *try – tried*, and *bait – baited*, depends on the final phoneme of the verb (e.g., Joanisse & Seidenberg, 1998; Joanisse *et al.*, 2000). More generally, problems with the processing of phonological information may affect the acquisition of morphological paradigms (i.e., verb inflection; Joanisse & Seidenberg, 1999), which in its turn may give rise to difficulties in detecting and grasping syntactic agreement. In addition, difficulties with syntactic structures have been linked to limitations in phonological short-term memory in dyslexia (see Melby-Lervåg *et al.*, 2012; Snowling & Melby-Lervåg, 2016 for meta-analyses): if the processing and storage of phonological information is impaired or limited, this is likely to affect syntactical processing of speech. In support of this idea, Robertson and Joanisse (2010) showed that when memory demands are high, children show poorer syntactical processing of spoken sentences, and this effect is more pronounced in children with dyslexia than in TD children.

Alternatively, the grammatical difficulties may be explained through an underlying deficit in implicit learning. Nicolson and Fawcett (2007, 2011) and Ullman and Pierpont (2005) argue that procedural learning difficulty is associated with grammatical impairment. Statistical learning, which may be seen as a specific instantiation of procedural learning (e.g., Qi *et al.*, 2019; Steacy *et al.*, 2019; Ullman *et al.*, 2019), has been found to be deficient in individuals with dyslexia (e.g., Gabay *et al.*, 2015; Sigurdardottir *et al.*, 2017; Singh *et al.*, 2018), although the evidence is not unequivocal (Schmalz *et al.*, 2017; Van Witteloostuijn *et al.*, 2019). Statistical learning, that is, detecting distributional and sequential patterns in (linguistic) input, has been argued to support the acquisition of syntactic categories (Mintz, 2002, 2003; Wijnen, 2013) as well as the rules of morphology and syntax (e.g., Bannard *et al.*, 2009; Evans *et al.*, 2009; Kidd & Kirjavainen, 2011). For example, the acquisition of nonadjacent patterns in language, such as the relationship between auxiliaries and inflections on the verb (e.g., the boy *is* running, where the intervening verb may vary), may be supported by a mechanism that tracks co-occurrence statistics (e.g., Gómez, 2002). The hypothesized relationship between statistical learning and grammatical acquisition is supported by research that has shown that performance on statistical learning tasks is related to grammatical abilities in TD children. Studies have established such relationships between statistical learning and syntactic priming (Kidd, 2012), grammatical processing (Clark & Lum, 2017), and the comprehension of complex syntactical structures such as passives and relative clauses (Kidd & Arciuli, 2016). Likewise, individual differences in the statistical learning ability of adults have been shown to correlate with the comprehension of relative clauses (Misyak *et al.*, 2010) and the comprehension of written sentences in general (Misyak & Christiansen, 2012). Moreover, studies have demonstrated impaired statistical learning in children with DLD who are known to experience grammatical difficulties (see Lammertink *et al.*, 2017, and Lum *et al.*, 2014 for meta-analyses). No studies to date have explored the relationship between grammatical performance and statistical learning ability in individuals with and without dyslexia.

The current study

In the present study, we tested the grammatical abilities of 100 school-aged Dutch-speaking children with and without dyslexia. This was done using two standardized tests of grammar, which target different levels of grammatical knowledge: inflectional morphology and syntax (the “word structure” and “recalling sentences” subtests of the Dutch version of the CELF; Kort et al., 2008). Furthermore, we aimed to highlight specific areas of difficulty through an exploratory analysis of error patterns. Most importantly, we tested two accounts of dyslexia that make predictions about the relationship between grammar on the one hand and underlying problems in either phonological memory or statistical learning ability on the other hand. Thus, we aimed to answer the following three research questions (note that whereas research questions 1 and 3 are confirmatory, research question 2 is of an exploratory nature):

- (1) Do children with dyslexia perform worse than their TD peers on grammar as measured through standardized tests (CELF word structure and CELF recalling sentences)?
- (2) Do children with dyslexia make different errors than their TD peers on the CELF word structure and/or CELF recalling sentences?
- (3) Do phonological memory and/or statistical learning ability contribute to individual differences in the CELF word structure and/or CELF recalling sentences? And, if so,
 - (a) is this contribution different for dyslexia versus TD?
 - (b) is this contribution different for the CELF word structure versus the CELF recalling sentences?

In relation to research question 3, we focused on measures of phonological memory, since individuals with dyslexia are typically impaired in this area (Melby-Lervåg et al., 2012; Snowling & Melby-Lervåg, 2016) and phonological memory is theorized to contribute to grammatical abilities (e.g., Robertson & Joanisse, 2010). Digit span forward and nonword repetition tasks were used to assess phonological short-term memory (i.e., immediate recall), while the digit span backward was used as a measure of verbal working memory (i.e., the manipulation of verbal information prior to recall; e.g., Baddeley, 2012; Alloway et al., 2009). Naturally, these memory tasks also rely on the processing of phonological information and (already established) phonological representations, which is especially true for nonword repetition (Rispen & Baker, 2012). In addition to relations between the measures of phonological short-term memory (digit span forward), verbal working memory (digit span backward), and phonological processing (NWR), receptive vocabulary (measured in our study using the PPVT; Dunn et al., 2005) is also assumed to be related to the ability to repeat nonwords (Bowey, 2001; Edwards et al., 2004; Munson et al., 2005; Gathercole, 2006, Rispen et al., 2015). In young children, phonological processing capacity contributes to vocabulary acquisition (Bowey, 2001; Gathercole, 2006). During development, lexical knowledge in turn has a predictive effect on nonword repetition (Metsala et al., 2009). Lexical representations stored in long-term memory predict NWR: chunks from existing lexical representations can be used for nonword repetition. The more detailed and “robust” the lexical representations are, the

better phonological information can be flexibly used for novel phonological representations as is the case in nonword repetition (Edwards *et al.*, 2004, Munson *et al.*, 2005, Metsala, *et al.*, 2009). For this study, we wanted to test vocabulary, phonological processing, and (short-term/working) memory separately in order to investigate the relation to grammar, even though we expect some overlap in performances on these tasks.

Statistical learning was tested using two experimental tasks that targeted different aspects of the domain-general ability to detect statistical regularities: visuomotor sequence learning (SRT task) and auditory nonadjacent dependency learning (A-NADL task). Although we are aware that the SRT task is typically considered a sequence learning or procedural learning task, the type of learning that takes place in an SRT task can also be described as statistical: it involves tracking the co-occurrences of adjacent elements (see also, *e.g.*, Kidd, 2012). Both statistical learning measures have previously been related to grammatical performance in children and/or have demonstrated impaired learning ability in children with DLD (SRT: *e.g.*, Clark & Lum, 2017; Kidd, 2012; Lammertink *et al.*, 2020a; A-NADL: *e.g.*, Iao *et al.*, 2017; Lammertink *et al.*, 2020b). Besides phonological memory and statistical learning measures, our regression analysis includes other potential sources of variance in grammatical performance (children's age, gender, and SES, and their scores on measures of nonverbal reasoning, vocabulary, and sustained attention).

It is important to note here that any statistical analyses were done in order to answer research question 2 are exploratory, since the tasks used to measure grammatical performance were not designed for error analysis specifically. The results from these analyses may further our understanding of the grammatical problems associated with dyslexia and may thereby serve to highlight potentially interesting directions for future research. Moreover, it should be noted that group comparisons regarding statistical learning ability in the present sample have already been discussed in detail elsewhere (van Witteloostuijn *et al.*, 2019).

Methods

Participants

The ethics review board of the University of Amsterdam approved this study. One-hundred 8- to 11-year-old children were included: 50 children with a prior diagnosis of dyslexia (26 girls, 24 boys, mean age in years:months = 9:10) and 50 individually age-matched TD children (24 girls, 26 boys, mean age = 9:8). To confirm participation as (non-)dyslexic, word (*EMT*; Brus, & Voeten, 1972) and pseudowords (*Klepel*; van den Bos *et al.*, 1994) reading tests were administered. All included children with dyslexia had a maximum norm score of 6 (*i.e.*, 10th percentile) on word and pseudoword reading, while TD children had a minimum norm score of 8 (*i.e.*, 25th percentile). Ten additional children with dyslexia and four additional TD children did not meet these predetermined inclusion criteria regarding their reading scores and were therefore excluded from the final sample. Parental (in the case of dyslexia) and teacher (in the case of TD) reports confirmed that all 100 participants in the final sample were native speakers of Dutch and none had diagnoses of (other) developmental disorders such as DLD. All participants completed each of the tasks

included in the present study. Again, please note that the sample is identical to the one described by Van Witteloostuijn et al. (2019; 2021), which focuses on group comparisons on statistical learning measures. Similarly, the group of TD children partly overlaps with studies examining language and statistical learning in children with DLD (Lammertink et al., 2020a, 2020b). These previous reports thus have a different focus than this study and there is no overlap in the interpretation of the data.

Besides reading and spelling, a range of background measures were collected (Table 1). These included standardized measures of receptive vocabulary (*Peabody Picture Vocabulary Test* (PPVT-III-NL; Dunn et al., 2005), nonverbal reasoning (*Raven's Standard Progressive Matrices*; Raven, 2003), and sustained attention (the *Score!* subtest of the *Dutch Test of Everyday Attention for Children* (Schittekatte et al., 2007). Also, an indication of children's socioeconomic status (SES) was determined on the basis of their home or school postal codes through open data that was published by the *Netherlands Institute for Social Research* (NISIR, 2017). These SES scores reflect the status of a given postal code in comparison to other Dutch postal codes. Open source data can be accessed through the following (Dutch) link: https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores.

In line with their diagnosis, children with dyslexia were found to perform significantly worse than the TD children on reading words, reading pseudowords, and spelling. No evidence for a difference between the two child groups was found regarding their age, SES, vocabulary, or nonverbal reasoning. Children with dyslexia scored lower than TD children on our measure of sustained attention, although this effect did not reach significance. Individual differences in age, SES, vocabulary, and nonverbal reasoning are included as control predictors in our regression model that investigates the contribution of phonological memory and statistical learning ability to grammatical performance (research question 3).

Materials

Measures of grammatical performance

Children's grammatical abilities were assessed through two subtests of the Dutch version of the standardized CELF language assessment battery (CELF-4-NL; Kort et al., 2008): the word structure and recalling sentences subtests. The word structure task measures children's ability to apply word formation rules (i.e., inflectional morphology), while the recalling sentences task tests children's ability to listen to and repeat sentences, thereby considering grammatical performance at different levels (i.e., semantics, morphology, and syntax). It is important to note here that a sentence recall task is not a purely syntactic task; although CELF recalling sentences performance is influenced by syntactic complexity and long-term memory representations of language (i.e., lexical knowledge; Klem et al., 2015), it is also affected by a child's verbal short-term memory (Alloway & Gathercole, 2005). Nevertheless, it is typically considered a measure of syntactic competence (e.g., Blom & Boerma, 2019; Frizelle & Fletcher, 2014).

In the CELF word structure task, children were shown pictures and were instructed to finish sentences read out by the experimenter. The task consists of

Table 1. Children with and without dyslexia's mean (and *SD*) age and SES, and results from reading, spelling, nonverbal reasoning, and sustained attention: raw scores, standardized scores, and group comparisons

	Dyslexia (<i>N</i> = 50)		TD (<i>N</i> = 50)		Group comparisons (<i>t</i> (98), <i>p</i>)
	Raw	Standardized	Raw	Standardized	
Age	9:10 (0:9)	N/A	9:8 (0:10)	N/A	.84, .40
SES	0.2 (1.2)	N/A	0.2 (1.1)	N/A	.17, .86
Reading words ^a	34.1 (11.7)	3.3 (2.1)	66.3 (11.6)	10.5 (2.2)	13.83, 9.1×10^{-25}
Reading pseudowords ^a	22.0 (8.0)	4.4 (1.6)	61.0 (14.4)	11.1 (2.2)	16.75, 1.6×10^{-30}
Spelling ^b	8.4 (4.6)	11.8 (13.7)	18.6 (4.7)	49.9 (24.7)	11.41, 1.1×10^{-19}
Nonverbal reasoning ^b	37.2 (6.6)	55.7 (25.0)	37.3 (8.1)	60.1 (28.1)	.04, .97
Vocabulary ^b	117.3 (9.7)	54,8 (21,3)	118.0 (9.4)	58,7 (20,6)	.37, .71
Sustained attention ^a	7.0 (2.5)	7.4 (3.3)	7.8 (1.8)	9.1 (3.0)	1.94, .055

Note: Age in years:months. Data regarding SES by postal codes were obtained from the *Netherlands Institute for Social Research*. Raw scores on reading words (*Een minuut test*, Brus & Voeten, 1972) and pseudowords (*Klepel*, van den Bos et al., 1994) represent the number of words read within the time limit of 1 and 2 min, respectively. Raw scores on spelling represent the number of words spelled correctly out of 30 in a Dutch dictation test (Braams & de Vos, 2015). Nonverbal reasoning was measured using *Raven's Standard Progressive Matrices* (Raven, 2003); raw scores represent the number of items answered correctly out of 60. The *Peabody Picture Vocabulary Test* (PPVT-III-NL; Dunn et al., 2005) was used as a test of receptive vocabulary; raw scores represent the number of items answered correctly out of a maximum of 204 items. Finally, sustained attention was assessed by the *Score!* subtest of the *Dutch Test of Everyday Attention for Children* (Schittekatte et al., 2007); raw scores represent the number of items answered correctly out of 10. Standardized scores represent either ^a norm scores (norm = 10) or ^b percentile scores (norm = 50).

30 items that are divided into categories targeting different aspects of morphology including pronouns, nouns (i.e., diminutives and plurals), verbs (i.e., subject–verb agreement, tense, and compound verbs), and adjectives (i.e., comparatives and superlatives). Responses were coded as either correct or incorrect, with a maximum score of 30. Children’s scores were not converted to standardized (i.e., norm) scores, since norms are available up until the age of 8 and our sample consists of 8- to 11-year-old children.

The CELF recalling sentences task required children to repeat sentences of increasing length and complexity as dictated by the experimenter. In accordance with the CELF manual, 8-year-old children repeated a maximum of 31 sentences, while children aged 9 years or older were administered a maximum of 23 sentences (the first 8 sentences were not administered). Responses received a score of 3 (0 errors), 2 (1 error), 1 (2 or 3 errors), or 0 (4 or more errors) and testing was discontinued after five consecutive 0 scores. Children’s individual score was the total number of points awarded to the administered sentences.

Measures of phonological memory

Phonological processing and phonological short-term and working memory were assessed through two tasks: a digit span task (CELF-4-NL; Kort et al., 2008) and a shortened version of a nonword repetition task (NWR-S; le Clercq et al., 2017). Both the forward and backward digit span tasks were administered, in which children were required to repeat sequences of digits of increasing length either in the same order (forward digit span; 16 items) or in the reversed order (backward digit span; 14 items). In the NWR-S, 22 pre-recorded nonwords were played one at a time and children had to listen carefully and repeat each nonword as accurately as possible. Items in the digit span and NWR-S tasks were scored as either correct or incorrect.

Measures of statistical learning

In each trial of the SRT task, a single visual stimulus (smiley) appeared in one of the four marked locations on the screen of a tablet computer. There were 7 blocks of 60 trials each. Participants were required to press a button on a gamepad with a location corresponding to the location of stimulus on the screen as accurately and as quickly as they could upon the appearance of the visual stimulus. Without the participants’ knowledge, the successive locations of the visual stimulus (over trials) followed a predetermined sequence (4, 2, 3, 1, 2, 4, 3, 1, 4, 3; numbers indicate locations) that was repeated six times in blocks 2–5 and block 7 (“sequence blocks”). In contrast, in block 6 (“disruption block”), the 60 presentation locations of the stimulus were chosen randomly. The 7 blocks of the actual test were preceded by a practice block (28 trials). Learning in the SRT task is measured as the increase in reaction times (RTs) in disruption block 6 as compared to the surrounding sequence blocks.

In the A-NADL task, children listened to an artificial language consisting of a series of simple “sentences”, each built from three pseudowords: *a-X-b*. Unbeknownst to the participants, the first pseudoword (*a*) in each sentence was taken from a set of two monosyllables (*tep*, *sot*), each of which was associated with

a specific pseudoword (*b*) in the third position (*mip*, *lut*). Thus, the pseudowords *a* and *b* formed 2 dependency pairs, separated by another pseudoword *X*, which was taken from a set of 24 disyllabic items. The task was designed to assess if participants (implicitly) detect this recurrent nonadjacent *a–b* dependency. This dependency mirrors those found in natural languages, such as the morphosyntactic relationship between auxiliaries and inflections on the verb in English (e.g., “*is walking*”, where the *is–ing* relationship is nonadjacent and the intervening verb may vary).

The task was modelled on the SRT task and thus contained blocks in which the artificial language strings adhered to *a–X–b* nonadjacent dependency rules (i.e., rule blocks 1–3 and 5) and an intervening block in which strings were presented in which the dependency between the first and third pseudoword was disrupted (i.e., disruption block 4). In the rule blocks, the nonadjacent dependencies *tep X lut* and *sot X mip* were each presented 24 times. In addition to the items containing the nonadjacent dependency rules, each rule block contained 12 filler trials with an $f_1–X–f_2$ structure where f_1 does not predict f_2 (f_1 and f_2 are taken from a set of 24 one-syllable nonwords, not including *tep*, *sot*, *lut*, or *mip*, and *X* refers to the same set of 24 two-syllable nonwords used in the *a–X–b* structure). In the disruption block, the occurrence of *lut* and *mip* was no longer predicted by the *a–X–b* rule: in 24 out of 30 trials, *lut* and *mip* still occurred in the *b* position, but one of the one-syllable fillers *f* occurred in the *a* position (i.e., *f–X–b* structure). The remaining six trials were entire filler items (i.e., $f_1–X–f_2$ structure).

Children performed a word-monitoring task in which they tracked the occurrence of one of the two predictable nonwords (i.e., the *b* element in the *a–X–b* structure). Half of the participants were assigned to *lut* as a target and half to *mip*. Children were instructed to press a green button when they heard the target nonword and to press a red button when they did not hear the target nonword (Lammertink *et al.*, 2019; López-Barroso *et al.*, 2016). As in the SRT task, learning in the A-NADL task is reflected by slower RTs to input in the disruption block than to rule-governed input in the surrounding rule blocks.

In both the SRT and A-NADL tasks, accuracy and RTs to each trial were recorded. As explained, learning is evidenced by shorter RTs to structured input as compared to random input and, therefore, the individual measures of learning used in the regression analysis are difference scores (SRT: normalized RT in disruption block 6 minus mean normalized RT in sequence blocks 5 and 7, A-NADL: normalized RT in disruption block 4 minus mean normalized RT in rule blocks 3 and 5).

General procedure

As mentioned in Section “Participants”, children in the present study were tested as part of a larger project investigating statistical learning and its relation to language in children with and without dyslexia and DLD (van Witteloostuijn *et al.*, 2019; 2021; Lammertink *et al.*, 2019; Lammertink *et al.*, 2020a, 2020b). The test battery was administered one-on-one by an experimenter in the child’s home or school. It took approximately 3 hr to complete and was divided into three testing sessions. Importantly, each testing session consisted of one statistical learning measure, combined with a range of background and language measures. The orders between and

within sessions were counterbalanced and children were randomly assigned to one out of six testing orders.

The CELF word structure, CELF recalling sentences, PPVT, and digit span tasks were dictated by the experimenter. Instructions (SRT, A-NADL) and auditory stimuli (A-NADL, NWR-S) in the statistical learning and NWR-S tasks were pre-recorded by a native Dutch speaker and were played over Sennheiser HD 201 headphones. PPVT images were shown on a Windows Surface 3 tablet. The SRT and A-NADL tasks were programmed and administered through E-prime 2.0 and displayed on the same tablet (Psychology Software Tools, 2012; Schneider et al., 2012). Accuracy and RTs in the SRT task were logged using a Trust Wired GXT 540 gamepad controller, while responses to the A-NADL task were logged through an external button box. Verbal responses in the CELF word structure, CELF recalling sentences, and NWR-S tasks were recorded using an Olympus DP-211 voice recorder.

Scoring and analysis

The following sections provide details of our method of scoring and analyses regarding group comparisons, the error exploration, and the regression model. All analyses are performed in *R* software (*R* Development Core Team, 2008); raw (summary level) data files and *R* Markdown and HTML files containing all analyses reported in the present study can be found on our Open Science Framework (OSF) project page (<https://osf.io/kjctf/>).

Group comparisons

Individual *t*-tests were run on children's raw scores on our outcome measures (CELF word structure and CELF recalling sentences; research question 1) and raw scores on the tests assessing phonological memory (digit span forward, digit span backward, and NWR-S) in order to investigate whether a difference in performance is observed between participants with and without dyslexia. As mentioned, investigations of group differences on the statistical learning measures (SRT and A-NADL) were already reported in detail elsewhere (van Witteloostuijn et al., 2019), and are thus not reanalyzed here.

Error explorations

Children's performance on the CELF word structure and CELF recalling sentences were examined in more detail through error analyses to explore whether children with dyslexia make qualitatively different errors than their TD peers (research question 2). Since items of the CELF word structure are already divided into categories, we inspected the total number of errors (and proportion of answers correct) per category (see Section "Measures of grammatical performance"). To explore potential differences between children with and without dyslexia in their performance on the CELF word structure categories, individual binomial generalized linear mixed-effects (GLMER) models were built for each error category using the *lme4* package for *R* (version 1.1.13; Bates et al., 2014). These models were run on the proportions of errors in a category with group as the predictor (orthogonally contrast-coded

such that the TD group was coded as $-1/2$ and the dyslexia group as $+1/2$) and included a random intercept per subject.

For ease of error analysis, responses to the CELF recalling sentences were recoded as either correct or incorrect (instead of 0, 1, 2, or 3; see Section “Measures of grammatical performance”). Besides scoring the accuracy on the sentence level, we categorized errors according to a predetermined scoring schedule including errors pertaining to the inflectional morphology of verbs (subject–verb agreement, tense, overgeneralization, and lexical errors) and nouns (plural, article choice, and lexical errors), and errors regarding the referential use of pronouns (demonstratives). These error categories combined will be referred to as “specific errors”. The remaining errors (i.e., errors that could not be categorized under specific error categories) were deemed “unspecific errors”, which included omissions, additions, replacements, and displacements of words that we did not analyze further (e.g., uttering a word in a different position in the sentence or switching two words). As in the CELF word structure analysis, Poisson GLMER models were run to explore potential group differences in the number of errors in specific error categories. All models included a random intercept per subject and a random intercept per item. Please note that the model for verb overgeneralizations failed to converge, and therefore the random intercept for an item was removed from this model.

Furthermore, the effects of syntactic complexity, sentence length (number of words), and group (dyslexia vs. TD) on sentence accuracy (i.e., an overall score on the sentence of 0 [incorrect] or 1 [correct]) were explored using a separate GLMER model. 19 sentences were marked as “syntactically complex”; these consisted of passive sentences ($N = 6$) and sentences containing a subordinate clause ($N = 13$). Ninety-five percent confidence intervals (CIs) were computed through Wald’s approximation for CIs and raw sentence length (i.e., number of words in target sentence) was centered and scaled by standard deviation. The categorical predictors included in the model were sentence complexity and group, which were orthogonally contrast-coded. Sentence complexity was coded into two contrasts such that the first contrasted simple (coded as $-2/3$) and complex sentences (passive and subordinate, coded as $+1/3$) and the second contrasted the two complex sentence types (i.e., passive coded as $-1/2$ and subordinate coded as $+1/2$). The coding of group was identical to the coding reported for the regression model: the TD group was coded as $-1/2$ and the dyslexia group as $+1/2$. The random effect structure of the model contained by-subject intercepts and by-subject random slopes for sentence length, sentence complexity, and the interaction between sentence length and sentence complexity.

Since testing on the CELF recalling sentences was halted after five consecutive 0 scores (see Section “Measures of grammatical performance”), it is important to note that testing was discontinued after a similar number of sentences in both participant groups of children with and without dyslexia (dyslexia: 29.5 [$SD = 2.8$], TD: after 30.2 [$SD = 2.0$]). Although a subset of children (i.e., children over the age of 8) did not complete sentences 1 through 8, we disregard this in our error analyses since children were individually matched on age. Of all 2,320 sentences administered to our 100 participants, 10 sentences resulted in null responses that were categorized as missing data and were excluded from analyses (dyslexia: $N = 6$, TD: $N = 4$).

Regression analysis

We set up a linear regression model to examine whether a range of predictors contribute to individual differences in performance on our outcome measures of grammar (research question 3). This was done using the *lm* function included in *R*, which modeled grammatical performance by a number of control predictors (age, gender, SES, nonverbal reasoning, vocabulary, and attention) and predictors relevant to research question 3 (phonological memory: digit span forward, digit span backward, and NWR-S, statistical learning: SRT, and A-NADL). Group membership (dyslexia vs. TD) was added as a predictor in order to assess interactions between group and other predictors (research question 3a). The significance of predictors to both grammatical measures combined (CELF word structure and recalling sentences) was determined through the *Manova* function in the *car* package for *R* (version 2.1.5; Fox et al., 2012). The effects of phonological memory and of statistical learning on grammar performance were investigated by comparing the full model to models from which all measures assessing phonological memory (digit span forward, digit span backward, and NWR-S) and both statistical learning measures (SRT and A-NADL) were removed. To compare the contribution of predictors to CELF word structure versus CELF recalling sentences (research question 3b), we computed 95% CIs using the profile method (*confint* function in *R*) and examined the overlap of 95% CIs of individual predictors for the two measures of grammar performance. Importantly, raw scores on continuous outcome variables (CELF word structure and CELF recalling sentences) and predictors (age, SES, nonverbal reasoning, attention, PPVT, digit span forward, digit span backward, NWR-S, SRT, and A-NADL) were centered and scaled. The categorical predictors, that is, gender and group, were orthogonally contrast-coded: females were coded as $-1/2$ and males as $+1/2$, and, similarly, the TD group was coded as $-1/2$ and the dyslexia group was coded as $+1/2$.

Results

We present the results regarding our three research questions: the group comparisons pertaining to research question 1 are described in Section “Group comparisons”, followed by the error exploration in Section “Error exploration” (research question 2), and the regression analysis in Section “Regression analysis” (research question 3). While the analyses related to research questions 1 and 3 can be viewed as confirmatory, the analyses related to research question 2 are exploratory in nature. Additionally, our regression analysis provides us with some exploratory findings that may be of interest and are reported separately following recommendations by Wagenmakers et al. (2012).

Group comparisons

Table 2 presents the mean (and *SD*) scores on the two measures of grammar (CELF word structure and CELF recalling sentences), and the phonological memory (digit span forward, digit span backward, NWR-S) and statistical learning (SRT, A-NADL) measures included as predictors in our regression model that is discussed in Section “Regression analysis”. In order to answer our first research question, we examined

Table 2. Mean (and SD) scores on measures of grammar, phonological skills, and statistical learning: raw scores, standardized scores, and group comparisons

	Dyslexia (<i>N</i> = 50)		TD (<i>N</i> = 50)		Group comparisons
	Raw	Standardized	Raw	Standardized	(<i>t</i> (98), <i>p</i>)
CELF WS	26.9 (1.8)	N/A	27.9 (1.3)	N/A	3.08, .0027
CELF RS	55.2 (10.3)	8.5 (2.2)	59.7 (12.9)	9.9 (2.9)	1.96, .053
DS forward	7.3 (1.5)	7.7 (2.6)	8.9 (1.5)	10.7 (2.9)	5.36, 5.5×10^{-7}
DS backward	4.2 (1.1)	9.0 (2.5)	4.5 (1.5)	10.0 (3.2)	3.96, .00014
NWR-S	7.3 (2.7)	N/A	9.7 (3.3)	N/A	1.26, .21
SRT	0.29 (0.28)	N/A	0.27 (0.27)	N/A	
A-NADL	0.15 (0.33)	N/A	0.17 (0.28)	N/A	

Note: Raw scores: CELF word structure (WS) = number of items correct out of 30, CELF recalling sentences (RS) = total score on administered sentences, digit span (DS) = number of items answered correctly out of 16 (forward) or 14 (backward), NWR-S = number of nonwords repeated correctly out of 22, SRT = difference in normalized RTs (RT disruption - RT sequence), A-NADL = difference in normalized RTs (RT disruption - RT rule). Standardized scores represent norm scores (norm = 10). No standardized scores are available for the CELF word structure, NWR-S, SRT, and A-NADL tasks.

group effects on children's grammatical performance as measured by the CELF word structure and recalling sentences subtests (see Table 2 for group comparison statistics). Results reveal that participants with dyslexia achieved significantly lower scores on the CELF word structure. The children with dyslexia also achieved lower scores on the CELF recalling sentences, although this effect did not reach significance. Out of 50 children with dyslexia, 9 received a norm score of 6 (i.e., 10th percentile) or lower on the CELF recalling sentences, while 7 out of 50 TD children received a norm score of 6 or lower. No norm scores are available for the CELF word structure subtest (see Section "Measures of grammatical performance"). Together, these results suggest subtle difficulties in the area of grammar in the group of children with dyslexia.

Furthermore, the children with dyslexia performed significantly worse than the TD children on the digit span forward task and the NWR-S, which both assess phonological processing and short-term memory. No evidence of such a difference between participant groups was found for the digit span backward that targets phonological processing and working memory. Finally, as previously published in Van Witteloostuijn *et al.* (2019), although evidence of learning was found for both statistical learning measures when looking at children with and without dyslexia together, no evidence of a group difference emerged for either the SRT ($p = .61$) or the A-NADL ($p = .87$) task. The correlations between the tasks used to assess grammar, phonological memory, and statistical learning are provided in the appendix.

Error exploration

Word structure

Overall, the children with dyslexia made an average of 3.1 errors out of 30 potential errors (range: 0–10 errors) and the TD children made 2.1 errors (range: 0–5 errors);

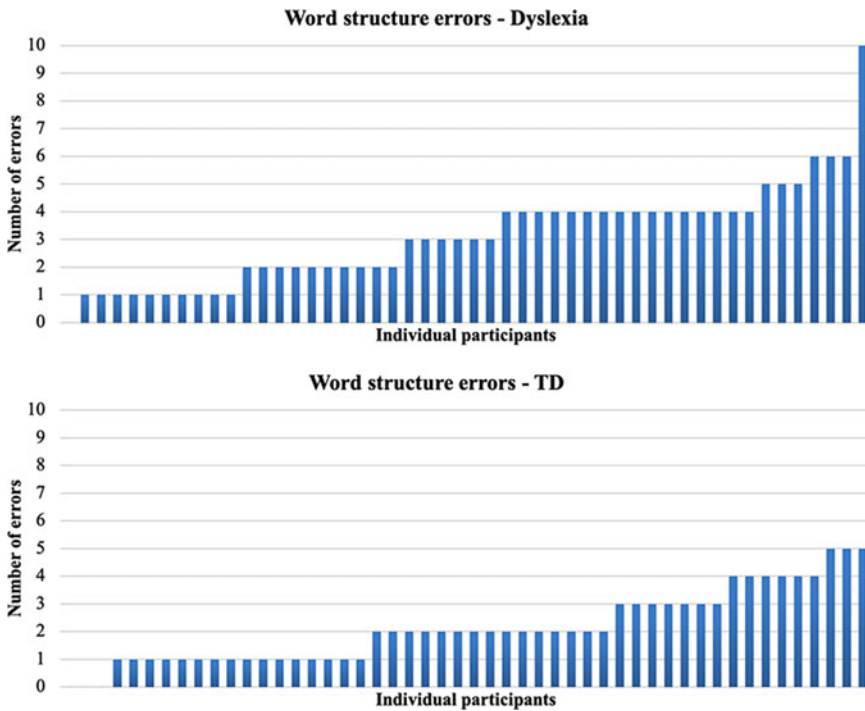


Figure 1. Histogram showing the distribution of performance on the CELF word structure subtest; children with dyslexia are presented in the top graph, TD children are presented in the bottom graph. Each bar represents the number of errors (out of 30 test items) of an individual participant.

see Figure 1). Table 3 presents the children's performance in the CELF word structure task per category. Performance on regular plurals and past tense formation was found to be at the ceiling both in participants with and without dyslexia. Moreover, on categories eliciting demonstrative and personal pronouns, the children with and without dyslexia achieved comparable levels of accuracy. The other categories may inform us about different error patterns in children with dyslexia as compared to their TD peers, as participants with dyslexia achieve numerically lower accuracy levels. The difference in accuracy levels between participants with and without dyslexia reached marginal significance, given the multitude of exploratory binomial GLMER models, on irregular plurals and compound verbs (*estimate* = .94, *SE* = .46, $z(100) = 2.04$, $p = .041$ and *estimate* = .78, *SE* = .30, $z(100) = 2.62$, $p = .0089$, respectively). Differences between the children with and without dyslexia did not reach significance on diminutives or comparative superlatives (*estimate* = .75, *SE* = .47, $z(100) = 1.62$, $p = .11$ and *estimate* = 1.88, *SE* = 2.07, $z(100) = .91$, $p = .36$, respectively).

Closer inspection of the error pattern on the irregular plurals (4 items) revealed that the children with dyslexia made most errors on the item *ei* – *eieren* [ɛi – ɛiərə(n)] (“egg – eggs”; 12 errors), followed by *ship* – *schepen* [sxɪp – sxe:pə(n)] (“ship – ships”; 5 errors), and fewest errors were made on *koe* – *koeien* [ku –

Table 3. Word structure: number of errors and accuracy (in percentage) in the different categories per group

	Dyslexia (N = 50) # Errors (% accuracy)	TD (N = 50) # Errors (% accuracy)
Regular plural	3 (98.5)	1 (99.5)
Past tense	3 (98.5)	2 (99)
Demonstratives	32 (68)	34 (66)
Personal pronouns	26 (87)	23 (88.5)
Irregular plural	17 (91.6)	7 (96.5)
Diminutives	18 (92.8)	9 (96.4)
Compound verbs	46 (54)	28 (72)
Comparative superlatives	10 (96)	2 (99)

kujə(n)] (“cow – cows”; 2 errors) and *glas – glazen* [ɣlas – ɣla:zə(n)] (“glass – glasses”; 2 errors). The errors in the TD participants were distributed more equally (*koe – koeien*: one error, *ei – eieren*: two errors, *schip – schepen*: three errors; *glas – glazen*: 0 errors). Generally speaking, errors on irregular plurals were cases of overgeneralization: participants applied the regular plural rules (add/ə(n)/or /s/) to irregular nouns (i.e., *ei – *eien* *[eiə(n)], *schip – *schippen* *[sxɪpə(n)], *glas – *glassen* [ɣlasə(n)], *koe – *koes* *[kus], instead of applying the required more complex suffix (/əɾə(n)/ as in *ei – eieren*) or alteration of the noun stem (i.e., vowel lengthening as in *schip – schepen* [sxɪp – sxɛ:p-ə(n)] and *glas – glazen* [ɣlas – ɣla:z-ə(n)], or stem alteration as in *koe – koeien* [ku – kuj-ə(n)]).

Regarding compound verbs, the majority of errors (dyslexia: 40 out of 46; TD: 27 out of 28) was made on the item *wassen af* (“[they are] washing the dishes”) and only a few errors were made on the item *speelt gitaar* (“[he/she] plays the guitar”). These errors were cases in which the child failed to separate the two verbal elements, such as *zij zijn aan het afwassen* (“they are washing the dishes”; this is not ungrammatical, but maybe an avoidance strategy), and/or cases in which the infinitive form of the verb was used (i.e., **zij gitaar spelen* [“she guitar plays”], or **zij afwassen* [“they washing the dishes”]).

Recalling sentences

In our GLMER model predicting children’s performance on the CELF recalling sentences task, we found a significant effect of sentence length: accuracy was lower for longer sentences than for shorter sentences (odds ratio estimate = a factor of 1.5 per standard deviation, 95% CI = [1.41 ... 1.60], $p = 3.5 \times 10^{-35}$). There was also a significant effect of sentence complexity; accuracy was lower for sentences that contained a complex syntactical structure as compared to simple sentences (odds ratio estimate = 58, 95% CI = [13 ... 261], $p = 1.1 \times 10^{-7}$). These two predictors were found to significantly interact with one another, indicating that the effect of length on performance is stronger in simple sentences than in complex sentences

(estimated odds ratio = 1.35, 95% CI = [1.19 ... 1.55], $p = 7.5 \times 10^{-6}$). Furthermore, accuracy was significantly higher on sentences containing subordinate clauses as compared to passive sentences (odds ratio estimate = 23, 95% CI = [5 ... 119], $p = .00015$) overall and the effect of length was found to be significantly stronger for subordinate clauses than for passive sentences (odds ratio estimate = 1.53, 95% CI = [1.31 ... 1.79], $p = 7.9 \times 10^{-8}$). The effect of group in interaction with sentence complexity ($p = .81$) or sentence length ($p = .87$) is nonsignificant, as are the three-way interactions with group. Thus, performance on the CELF recalling sentences task is influenced by sentence length and sentence complexity in children, and we find no evidence of a difference in performance between children with and without dyslexia regarding the effects of sentence length and sentence complexity.

A total number of 4,762 errors (dyslexia: $N = 2,525$, TD: $N = 2,237$) on the CELF recalling sentences task were classified into predetermined categories (see Section "Error explorations"). The largest proportion of total errors (approximately 74%) were classified as nonspecific: omissions of words (dyslexia: 35.5%, TD: 36.4%), replacements, displacements, and switches of words (dyslexia: 26.9%, TD: 25.6%), and additions of words (dyslexia: 11.1%, TD: 11.9%). Nonspecific errors involved function words slightly more often than content words (dyslexia: 51.2% and 48.8% of nonspecific errors, respectively, TD: 51.9% and 48.1% of nonspecific errors, respectively).

The remaining 26% of errors (dyslexia: $N = 671$, TD: $N = 585$) were labeled as specific errors, divided into errors pertaining to nouns (plurals, article choice, and lexical errors), verbs (subject–verb agreement, tense, overgeneralization, and lexical errors), and demonstrative pronouns (see Section "Error explorations"). Table 4 presents a summary of these results. The children with dyslexia made an average of 13.4 specific errors (range: 1–24 errors) and the TD children made 11.7 errors (range: 1–26 errors; see Figure 2). The largest proportion of specific errors (approximately 55%) was classified as lexical errors, both in the children with dyslexia (verbs: $N = 192$, nouns: $N = 168$) and in the TD children (verbs: $N = 192$, nouns: $N = 139$). Regarding nouns, the children made very few pluralization errors (dyslexia: $N = 6$, TD: $N = 7$). More errors were made concerning article choice: both the choice between indefinite and definite articles (*een* vs. *de/het*; dyslexia: $N = 36$, TD: $N = 30$) and between the two definite articles (*de* vs. *het*; dyslexia: $N = 67$, TD: $N = 41$). Exploratory Poisson GLMER models suggest that children with dyslexia may make more errors regarding the choice between the two definite articles in Dutch (*estimate* = .53, *SE* = .24, $z(2,320) = 2.24$, $p = .025$), an effect that is considered marginally significant given the number of exploratory models. No evidence of a difference between groups was found for errors concerning the choice between indefinite and definite articles (*estimate* = .22, *SE* = .26, $z(2,320) = .87$, $p = .39$). Second, regarding verbal morphology, the children made a small number of subject–verb agreement errors (dyslexia: $N = 15$, TD: $N = 23$) and overgeneralization errors (dyslexia: $N = 13$, TD: $N = 4$), whereas tense errors were more frequent (dyslexia: $N = 121$, TD: $N = 107$). No evidence of a difference in performance between children with and without dyslexia was found regarding the number of subject–verb agreement errors (*estimate* = $-.33$, *SE* = .39, $z(2,320) = -.83$, $p = .40$) and tense errors (*estimate* = .16, *SE* = .18, $z(2,320) = .90$, $p = .37$). The children with dyslexia were found to produce more verb overgeneralization errors, which is again considered marginally significant given the number of exploratory tests

Table 4. Recalling sentences: number of errors in the categories covering verb, noun, and pronoun errors per group

	Dyslexia (N = 50)	TD (N = 50)	
	# Errors	# Errors	
Verbs			
Subject-verb agreement	15	23	38
Tense	121	107	228
Overgeneralization	13	4	17
Lexical	192	192	384
Nouns			
Plural	6	7	13
Article choice	36	30	66
Article choice (definite)	67	41	108
Lexical	168	139	307
Pronouns			
Demonstrative	53	42	95
Total	671	585	1,256

(*estimate* = 1.24, *SE* = .57, $z(2,320) = 2.17$, $p = .030$). These are instances where children apply the regular Dutch past tense rule (i.e., add /te/) to irregular verbs, such as *koop* – **koopte* (correct: *kocht*; “buy – *bayed – bought”). However, please note the low number of errors in this category overall. Lastly, no evidence of a difference in performance between children with and without dyslexia was found regarding the incorrect use of the demonstrative pronoun (dyslexia: $N = 53$, TD: $N = 42$; *estimate* = .25, *SE* = .21, $z(2,320) = 1.19$, $p = .23$).

Regression analysis

Regression analysis: confirmatory findings

In order to answer research question 3, we performed a linear regression analysis to investigate the effects of phonological memory (digit span forward, digit span backward, and NWR-S) and statistical learning (SRT and A-NADL) on children’s performance on the CELF word structure and recalling sentences subtests. MANOVA results show that NWR-S (Wilk’s $\lambda = .88$, $F[2,80] = 5.28$, $p = .0070$) scores significantly affect grammar performance (CELF word structure and recalling sentences combined). The effects of the digit span forward (Wilk’s $\lambda = .96$, $F[2,80] = 1.89$, $p = .16$) and backward (Wilk’s $\lambda = .99$, $F[2,80] = .22$, $p = .81$) tasks do not reach significance. Importantly, when we compare the full model to a model where the phonological memory measures (digit span forward, digit span backward, and NWR-S) are removed, this results in a significant decrease in the fit of the model ($F[12,162] = 2.80$, $p = .0017$). The model provides no evidence of an effect of statistical learning on explaining individual differences in performance on the CELF

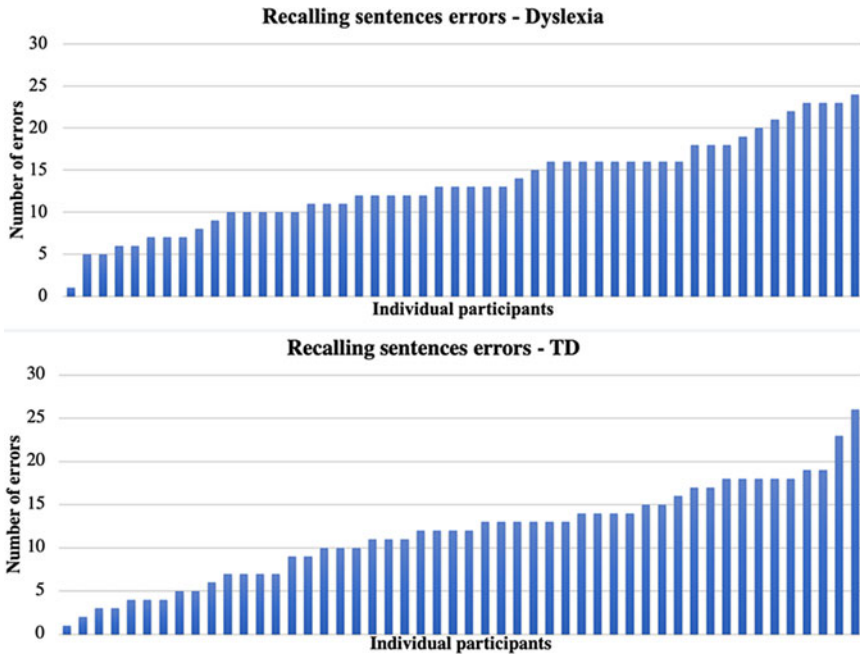


Figure 2. Histogram showing the distribution of performance on the CELF recalling sentences subtest; children with dyslexia are presented in the top graph, TD children are presented in the bottom graph. Each bar represents the number of specific errors of an individual participant.

word structure and recalling sentences (SRT: Wilk's $\lambda = .99$, $F[2,80] = .36$, $p = .70$; A-NADL: Wilk's $\lambda = .99$, $F[2,80] = .29$, $p = .75$). Comparing the full model to a reduced model where the statistical learning measures (SRT and A-NADL) are removed does not reveal a significant difference in fit between the models ($F[8,162] = 1.04$, $p = .41$). Thus, we can conclude that phonological memory skills contribute to the grammatical performance of children with and without dyslexia; and we find no evidence for or against the hypothesis that statistical learning contributes to children's grammatical performance.

Regarding potential differences between children with and without dyslexia (research question 3a), the model shows a significant interaction between phonological processing and short-term memory, as measured by the NWR-S, and group (NWR-S * Group: Wilk's $\lambda = .92$, $F[2,80] = 3.51$, $p = .035$). To follow-up on this interaction, the correlation between nonword repetition and grammatical performance is significant in both groups (TD: $r(48) = .58$, $p = 1.0 \times 10^{-5}$; dyslexia: $r(48) = .51$, $p = .00019$). No other interactions with group are found to be significant. As for potential differences in the contribution to CELF word structure versus recalling sentences performance (research question 3b), we cannot conclude that there is a difference in effect of the NWR-S due to overlapping 95% CIs (effect NWR-S on CELF word structure: *estimate* = $+.20$, 95% CI [$-.05 \dots +.46$], $p = +.11$; effect NWR-S on CELF recalling sentences: *estimate* = $.28$, 95% CI [$+.09 \dots +.47$], $p = .0044$).

Regression analysis: Exploratory findings

The regression model provides us with some exploratory findings that may be of interest. First, nonverbal reasoning (Raven; Wilk's $\lambda = .87$, $F[2,80] = 6.06$, $p = .0035$) and vocabulary (PPVT; Wilk's $\lambda = .81$, $F[2,80] = 9.28$, $p = .00024$) are found to contribute to CELF word structure and recalling sentences performance combined. Second, there is no evidence that the effect of group membership contributes to children's grammatical performance over and beyond the other predictors included in the model (Wilk's $\lambda = .95$, $F[2,80] = 2.31$, $p = .11$).

In addition to these findings provided by the model, we wish to further explore the effects of phonological memory and statistical learning on grammar performance. First, the results described above suggest that the effect of phonological memory is largely carried by the NWR-S. This is corroborated by a further analysis: removal of digit span forward and backward does not significantly affect the fit of the model ($F[12,162] = .56$, $p = .81$). Similarly, we wish to explore the effect of the A-NADL task on its own, since this statistical learning task is considered to model aspects of grammar acquisition. Removing the SRT task from the model does not result in a significant decrease in fit ($F[4,162] = .90$, $p = .46$), and the effect of the A-NADL task on its own remains nonsignificant (Wilk's $\lambda = .99$, $F[2,80] = .31$, $p = .73$). Please note that the effect of the SRT and A-NADL tasks combined also did not reach significance (Section "Regression analysis: confirmatory findings").

Discussion

The goal of this study was to examine the performance of Dutch-speaking school-aged children with and without dyslexia on standardized measures of inflectional morphology and syntax. We investigated whether phonological memory and statistical learning ability contributed to children's grammatical performance, in order to shed light on the underlying causes of the linguistic difficulties associated with dyslexia. Here, we first discuss the findings concerning group and error pattern analyses of tasks assessing inflectional morphology and syntax (research questions 1 and 2), followed by a discussion of the contributions of phonological memory and statistical learning to children's grammatical performance (research question 3).

Grammatical performance in children with dyslexia

In line with previous studies examining the performance of children with dyslexia on (standardized) tests of grammar, children with dyslexia in the present study achieved slightly lower scores on the CELF word structure subtest that assesses inflectional morphology (see also Joanisse *et al.*, 2000) and the CELF recalling sentences, targeting both morphology and syntax (see also Carroll & Myers, 2010). When investigating the effects of a range of predictors on grammatical performance, results showed that group membership (i.e., having a diagnosis of dyslexia or not) did not contribute to individual differences in grammar over and beyond other contributors to performance (e.g., vocabulary, nonverbal reasoning, nonword repetition, and digit span tasks). Together, these findings agree with earlier findings that showed that difficulties in the area of grammar in individuals with dyslexia exist, but they are subject to substantial individual variation (Rispen *et al.*, 2004) and that the grammatical

problems appear to be subtle (Rispen & Been, 2007), at least in 8- to 11-year old children.

To explore the nature of the observed (subtle) difficulties within the CELF word structure and recalling sentences subtests, we performed a fine-grained analysis of children's error patterns. Here, we highlight the most important findings. First, regarding the CELF word structure subtest, no evidence of a difference between participant groups was found on the production of diminutives (i.e., producing the correct diminutive suffix on nouns as in *boom-pje*; "tree-DIM" [diminutive marker]; see also Boersma, 2018), comparative superlatives (e.g., *snel, snel-ler, snel-st*; "fast, fast-er, fast-est"), regular and irregular past tense, or pronouns. Please note that accuracy on demonstrative pronouns was low in both participant groups (dyslexia 68%, TD: 66%): children overuse the common demonstrative pronoun *die* ("that") in situations where the neuter pronoun *dat* ("that") is required. The overgeneralization of the common gender in Dutch is a pattern previously described for TD children (see, e.g., Blom et al., 2008), and is therefore not unexpected. Interestingly, participants with dyslexia were found to achieve scores close to ceiling performance on items targeting regular plurals (98.5% accuracy), while accuracy was found to be lower than in their TD peers on items assessing irregular plurals. Errors on irregular plurals were cases of overgeneralization of the regular plural rule. As suggested by Ullman (2001) in his declarative/procedural model of language, the use of irregulars is thought to be supported by the mental lexicon, while the use of regulars depends on the application of structural rules (i.e., grammar). Thus, in the case of irregular plurals, instead of retrieving the correct (irregular) plural form from their lexical memory (e.g., *ei, ei-eren*; "egg-PL" [plural marker]), participants with dyslexia were more likely to incorrectly apply the regular pluralization rule than TD participants (e.g., *ei, ei-*en*; "egg-*PL"). If grammatical problems in children with dyslexia were the consequence of an underlying deficit in statistical learning, one would not expect this pattern of results. Rather, this pattern of findings may suggest a problem with lexical retrieval for language production in individuals with dyslexia, which is in line with previous studies indicating poor performance on tasks assessing lexical retrieval (i.e., rapid automatized naming; e.g., Bexkens et al., 2015). Furthermore, participants with dyslexia were outperformed by TD participants on separable compound verbs. This is an indication of difficulties with the production of the correct verb-second word order in Dutch: the finite verb (i.e., the verb that expresses tense and/or agreement) appears in the second position (*zij wassen af*, "they wash up") and, thus, the production of an infinite verb in the second position is ungrammatical (**zij afwassen*, "they washing up"). Problems related to the verb-second phenomenon have previously been demonstrated in children with DLD and are argued to be the result of underlying processing and working memory deficits (e.g., Blom et al., 2014; de Jong, 1999; Rice & Wexler, 1996; Verhoeven et al., 2011). Moreover, both overregularization and V2 avoidance strategies are known to occur in the language production of younger TD children and have been proposed to be the result of weak memory traces (e.g., Marcus et al., 1992; Wexler, 1994). Thus, likewise, limitations in the retrieval of lexical information may help explain the difficulties with these phenomena in older children with dyslexia.

Second, regarding the CELF recalling sentences subtest, children's accuracy was lower when sentences were longer and when they were syntactically more complex.

There was no evidence that these effects of sentence length and syntactic complexity affected children with and without dyslexia differently. Thus, although CELF recalling sentence performance is influenced by both short-term memory load and syntactic complexity, we find no evidence that this effect is more pronounced in children with dyslexia as previously reported by Robertson and Joannis (2010) for sentence comprehension. As for specific error types, the children with dyslexia made more errors with respect to choosing the correct definite article (common *de* or neuter *het*) and produced more overgeneralization errors regarding past tense (applying the regular morphological rule to irregular verbs, e.g., *koop*-*te; “buy-*ed”). From a statistical learning perspective, one would expect to find more errors in regular past tense formation as compared to the irregular past tense formation in individuals with dyslexia; a pattern that is not reflected in our findings. In contrast, as for the irregular plural errors in the CELF word structure subtest, the irregular past tense error types explained above appear to be lexical in nature. In Dutch, the correct choice between the common and neuter article depends on the lexical knowledge of the noun: since Dutch gender is largely arbitrary, it has to be stored in the mental lexicon for each noun separately (e.g., Blom *et al.*, 2008; Orgassa & Weerman, 2008). Thus, we find errors suggesting difficulties in lexical retrieval, both in the CELF word structure and the CELF recalling sentences subtests.

Importantly, there are a number of limitations that we would like to point out here. First, future research needs to investigate these exploratory findings regarding differences in error patterns between children with and without dyslexia to test whether the findings reported here are reliable and generalizable. Second, the CELF word structure subtest may not have been maximally sensitive to differences in performance in the current sample due to the fact that it is designed to test children between 5 and 8 years of age. Finally, it is worth noting that the results presented here are based on few items (e.g., compound verbs in the CELF word structure) and/or a low number of errors overall (e.g., overgeneralization of the regular past tense in the CELF recalling sentences). Moreover, the items on which we based our error analysis were not designed to provide insights regarding the underlying reasons for (subtle) grammatical difficulties in dyslexia. In order to investigate this further, future studies should a priori distinguish between specific error types that would be expected based on a phonological or statistical learning account of dyslexia. Future studies comparing children with dyslexia not only to a group of age-matched TD children, but also to a group of children with DLD, may further our understanding of the extent of grammatical difficulties in dyslexia and of the overlap with the problems observed in DLD.

Contributions to grammar performance in children with and without dyslexia

The second aim of the present study was to investigate whether phonological memory and statistical learning ability impact children’s grammatical performance. Research question 3 asked if phonological memory and/or statistical learning ability contributed to individual differences in the CELF word structure and/or recalling sentences subtests. Our analyses controlled for children’s age and SES, as well as scores on tasks measuring their nonverbal reasoning, vocabulary, and attention. We conclude that phonological processing and phonological short-term and

working memory contribute to the grammatical performance of children with and without dyslexia, above and beyond other predictors in the model. Thus, the results from our regression analysis support the idea that the grammatical problems observed in dyslexia may be partially explained by an underlying weakness in the area of phonology (e.g., Shankweiler et al., 1995; Joanisse et al., 2000). More specifically, problems with the processing and short-term storage of phonological information, as measured by nonword repetition and digit span tasks, contribute to difficulties in the areas of inflectional morphology and syntax (see also Robertson & Joanisse, 2010). The correct processing and memorization of verbal material is relevant in both the CELF word structure and recalling sentences subtests, since they involve the processing of spoken sentences and either completing (CELF word structure) or repeating (CELF recalling sentences) these sentences. The link between phonological memory and grammar performance in the present study is further supported by the finding that children with dyslexia make more errors than TD children on compound verbs, which has previously been related to a phonological processing and memory limitation in children with DLD (e.g., Blom et al., 2014). Similarly, it is in line with the fact that participants were affected by sentence length in their performance on the CELF recalling sentences. Taken together, these results underline the contribution of phonological processing and phonological memory to grammatical performance, and support the hypothesis that (some of the) grammatical problems observed in dyslexia result from an underlying problem in the area of phonology (e.g., Shankweiler et al., 1995). Note that the exploratory error analyses discussed in Section “Grammatical performance in children with dyslexia” support these findings.

We could not conclude whether or not statistical learning ability, as assessed through SRT and A-NADL tasks, contributes to children’s grammatical performance. It is important to note that no evidence of group differences in statistical learning performance was reported for the present sample (see Van Witteloostuijn et al., 2019). In previous studies, statistical and procedural learning have been shown to be impaired in individuals with dyslexia (e.g., Gabay et al., 2015; Lum et al., 2013) and DLD (e.g., Lammertink et al., 2017; Lum et al., 2014), and have been related to grammatical abilities in TD children (e.g., Clark & Lum, 2017; Kidd, 2012; Kidd & Arciuli, 2016). However, not all published data provide evidence of a statistical or procedural learning impairment in individuals with dyslexia (e.g., Kelly et al., 2002; Rüsseler et al., 2006). Meta-analyses (Schmalz et al., 2017; van Witteloostuijn et al., 2017) show that there is wide variation in the magnitude of group differences across studies, which are likely associated with both participant and task variables (e.g., the type of structure to be learned, the modality in which statistical learning is assessed, task complexity, etc.). These meta-analyses have also raised the issue of a potential publication bias in the field, suggesting that the evidence for differences between TD children and children with dyslexia or DLD is weaker than it may seem.

Nonetheless, even in the absence of a group effect, associations between statistical learning performance and grammatical performance could have been obtained. However, our data do not provide evidence for (or against) the relationship between statistical learning on the one hand and inflectional morphology and syntax on the other hand. While this may seem surprising, other studies have similarly reported

null results when looking at the relationship between statistical learning and language performance (e.g., West *et al.*, 2017). Recently, the reliability of statistical learning measures has been questioned, especially in child participants (e.g., Arnon, 2019a; West *et al.*, 2017). Therefore, current statistical learning measures may not be suitable to examine the hypothesized relationship with linguistic performance (e.g., Arnon, 2019b). Note, however, that the statistical learning measures used in the present study were reliable at detecting learning in child participants with and without dyslexia overall. In line with concerns relating to reliability, statistical learning measures have been shown to only weakly correlate amongst each other (e.g., Schmalz *et al.*, 2019; Siegelman & Frost, 2015), which may help explain the mixed results regarding the relationship between statistical learning and measures of linguistic performance (i.e., some studies reporting significant correlations and others reporting null findings). Of course, these factors do not exclude the possibility that statistical learning plays an important role in language acquisition and is therefore related to children's grammatical performance, but merely affect our ability to evaluate this link (Arnon, 2019b). More research is needed in order to improve on present methodologies of measuring statistical learning and to more reliably evaluate its relationship to language.

Finally, we would like to return to lexical storage and/or retrieval as potential additional sources of variation in grammatical performance, and of grammatical difficulties in dyslexia. Of course, lexical knowledge, in general, is one of the crucial building blocks of the comprehension and production of language, and lexical knowledge is affected in children with DLD (see McGregor, 2009 for a review). This relationship is also apparent from the present study: children's receptive vocabulary knowledge contributes to their performance on inflectional morphology and syntax. More specifically, however, children with dyslexia were shown to experience difficulties in irregular plurals (CELF word structure), irregular past tense (CELF recalling sentences), and the choice between the common and neuter definite article (CELF recalling sentences). We would like to speculate that, besides phonological processing and memory, the automatic access and retrieval of lexical representations may be impaired in dyslexia (see also Bexkens *et al.*, 2015), while the representations themselves may be unimpaired. A similar line of reasoning has been suggested regarding the retrieval processes of representations of speech sounds and phonology (Boets *et al.*, 2013; Griffiths & Snowling, 2001; Ramus & Szenkovits, 2008; Rispens *et al.*, 2015). If individuals with dyslexia are unable to efficiently retrieve lexical representations from long-term memory (e.g., irregular plural or past tense forms), they are more likely to apply the regular morphological rule instead, resulting in over-generalizations as described in the present study.

In summary, deficits in the area of phonological processing and phonological short-term and working memory, as well as lexical retrieval, are likely to contribute to the linguistic performance of children with dyslexia, not only in the area of literacy skills but also regarding inflectional morphology and syntax. Of course, we cannot rule out the possibility that the observed relationship between short-term and working memory and long-term language knowledge is bidirectional. Linguistic representations (syntactic/morphological representations) stored in long-term memory seem to impact on the performance of verbal working memory tasks, in the sense that long-term linguistic representations are automatically activated,

which is advantageous when recalling sentences as compared to a list of words (Jefferies et al., 2004). Evidence from experiments with children furthermore suggests that the quality and robustness of long-term memory representations, such as lexical and syntactic knowledge, influence performance on tasks involving verbal short-term and/or working memory (e.g., Kidd et al., 2007; Mainela-Arnold & Evans, 2005; Munson et al., 2005), which has led to discussions whether working memory and representations of linguistic knowledge are distinct and separable entities (e.g., MacDonald & Christiansen, 2002; Mainela-Arnold & Evans, 2005). Future work is needed in order to clarify whether the link between short-term and working memory on the one hand, and linguistic performance, on the other hand, is in fact a bidirectional relationship (see also Marshall, 2020; Riches, 2020).

Nonetheless, the observations of the present study fit with suggestions that multiple cognitive deficits may help explain the range of behavioral difficulties associated with dyslexia and other developmental disorders, as well as the comorbidity between different disorders (e.g., Law et al., 2017; Pennington, 2006). Already in 1999, Wolf and Bowers suggested the double deficit hypothesis: impairments in phonology or rapid automatized naming were assumed to cause dyslexia, with more severe problems when both phonological and rapid automatized naming difficulties were present in a single individual. As mentioned previously, there is also emerging evidence that children with a FRdys experience delays in oral language development in early childhood (Snowling & Melby-Lervåg, 2016). More research is needed to increase our understanding of the exact nature of the underlying causes of dyslexia and to shed light on the so-called “risk factors” of developing developmental disorders such as dyslexia (Pennington, 2006). Investigations of multiple sources of variance simultaneously, as attempted in the present study, may shed light on these open questions.

Acknowledgments. We would first and foremost like to thank all children, their parents, and the primary schools (De Bosmark; De Kemp; De Ludgerschool, and De Ieme) that participated in our study. We are also grateful to Imme Lammertink for collaboration on the development of the statistical learning measures. Furthermore, we thank Iris Broedelet, Sascha Couvee, and Darlene Keydeniers for help with testing the control group and Dirk Jan Vet for technical assistance in developing the tasks. This work was supported by the Netherlands Organization for Scientific Research (NWO) through a personal Vidi grant awarded to Judith Rispen (grant number 276-89-005). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

References

- Alloway, T. P., & Gathercole, S. E. (2005). The role of sentence recall in reading and language skills of children with learning difficulties. *Learning and Individual Differences*, *15*(4), 271–282.
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, *80*(2), 606–621.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. American Psychiatric Publishing.
- Arnon, I. (2019a). 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.
- Arnon, I. (2019b). Statistical learning, implicit learning, and first language acquisition: A critical evaluation of two developmental predictions. *Topics in Cognitive Science*, *11*, 504–519.
- Aro, T., Eklund, K., Nurmi, J. E., & Poikkeus, A. M. (2012). Early language and behavioral regulation skills as predictors of social outcomes. *Journal of Speech, Language, and Hearing Research*, *55*, 395–408.

- Baddeley, A.** (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.
- Bannard, C., Lieven, E., & Tomasello, M.** (2009). Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences*, *106*(41), 17284–17289.
- Bar-Shalom, E. G., Crain, S., & Shankweiler, D.** (1993). A comparison of comprehension and production abilities of good and poor readers. *Applied Psycholinguistics*, *14*(2), 197–227.
- Bates, D., Maechler, M., Bolker, B., & Walker, S.** (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1–7.
- Bexkens, A., van den Wildenberg, W. P., & Tijms, J.** (2015). Rapid automatized naming in children with dyslexia: Is inhibitory control involved? *Dyslexia*, *21*(3), 212–234.
- Bishop, D. V. M., & Snowling, M. J.** (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin*, *130*(6), 858–886.
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T.** (2017). CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development. *Journal of Child Psychology & Psychiatry*, *58*, 1068–1080.
- Blom, E., & Boerma, T.** (2019). Reciprocal relationships between lexical and syntactic skills of children with Developmental Language Disorder and the role of executive functions. *Autism & Developmental Language Impairments*, *4*. doi: [10.1177/2396941519863984](https://doi.org/10.1177/2396941519863984).
- Blom, E., Polišenská, D., & Weerman, F.** (2008). Articles, adjectives and age of onset: The acquisition of Dutch grammatical gender. *Second Language Research*, *24*(3), 297–331.
- Blom, E., Vasić, N., & de Jong, J.** (2014). Production and processing of subject–verb agreement in monolingual Dutch children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *57*(3), 952–965.
- Boersma, T. A.** (2018) Variability in the acquisition of alomorphs: The Dutch diminutive and past tense. Doctoral dissertation, University of Amsterdam, The Netherlands.
- Boets, B., de Beek, H. P. O., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., Bulthé, J., Sunaert, S., Wouters, J., & Ghesquière, P.** (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science*, *342*(6163), 1251–1254.
- Bowey, J. A.** (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics*, *22*(3), 441.
- Braams T., & de Vos T.** (2015). *Schoolvaardigheidstoets spelling*. Boom uitgevers Amsterdam.
- Brus B., & Voeten M.** (1972) *Eén Minuut Test. Vorm A en B. Schoolvorderingen voor het lezen, bestemd voor het tweede t/m het vijfde leerjaar van de lagere school*. Berkhout Testmateriaal.
- Cantiani, C., Lorusso, M. L., Perego, P., Molteni, M., & Guasti, M. T.** (2013). Event-related potentials reveal anomalous morphosyntactic processing in developmental dyslexia. *Applied Psycholinguistics*, *34*(6), 1135–1162.
- Carroll, J. M., & Myers, J. M.** (2010). Speech and language difficulties in children with and without a family history of dyslexia. *Scientific Studies of Reading*, *14*(3), 247–265.
- Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E.** (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research*, *48*, 1378–1396.
- Chen, A., Wijnen, F., Koster, C., & Schnack, H.** (2017). Individualized early prediction of familial risk of dyslexia: A study of infant vocabulary development. *Frontiers in Psychology*, *8*, 156.
- Clark, G. M., & Lum, J. A.** (2017). Procedural memory and speed of grammatical processing: Comparison between typically developing children and language impaired children. *Research in Developmental Disabilities*, *71*, 237–247.
- de Jong, J.** (1999). Specific language impairment in Dutch: Inflectional morphology and argument structure. Doctoral dissertation, University of Groningen, The Netherlands.
- Dunn, L. M., & Dunn, L. M.** (2005). *Peabody picture vocabulary test-III-NL* [Dutch version by L. Schlichting]. Harcourt Test Publishers.
- Edwards, J., Beckman, M. E., & Munson, B.** (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, *47*, 421–436.
- 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*, 321–335.

- Facoetti, A., Paganoni, P., & Lorusso, M. L. (2000). The spatial distribution of visual attention in developmental dyslexia. *Experimental Brain Research*, *132*(4), 531–538.
- Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., . . . Heiberger, R. (2012). *Package 'car'*. R Foundation for Statistical Computing.
- Frizelle, P., & Fletcher, P. (2014). Relative clause constructions in children with specific language impairment. *International Journal of Language and Communication Disorders*, *49*(2), 255–264.
- Gabay, Y., Thiessen, E. D., & Holt, L. L. (2015). Impaired statistical learning in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, *58*(3), 934–945.
- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513–543.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436.
- Griffiths, Y. M., & Snowling, M. J. (2001). Auditory word identification and phonological skills in dyslexic and average readers. *Applied Psycholinguistics*, *22*(3), 419–439.
- Iao, L. S., Ng, L. Y., Wong, A., & Lee, O. T. (2017). Nonadjacent dependency learning in Cantonese-speaking children with and without a history of specific language impairment. *Journal of Speech, Language, and Hearing Research*, *60*(3), 694–700.
- Jefferies, E., Ralph, M. A. L., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language*, *51*(4), 623–643.
- Jiménez-Fernández, G., Vaquero, J. M., Jiménez, L., & Defior, S. (2011). Dyslexic children show deficits in implicit sequence learning, but not in explicit sequence learning or contextual cueing. *Annals of Dyslexia*, *61*(1), 85–110.
- Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology*, *77*(1), 30–60.
- Joanisse, M. F., & Seidenberg, M. S. (1998). Specific language impairment: A deficit in grammar or processing? *Trends in Cognitive Sciences*, *2*(7), 240–247.
- Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences*, *96*(13), 7592–7597.
- Kelly, S. W., Griffiths, S., & Frith, U. (2002). Evidence for implicit sequence learning in dyslexia. *Dyslexia*, *8*(1), 43–52.
- Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184.
- Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, *87*(1), 184–193.
- Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A crosslinguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*(6), 860–897.
- Kidd, E., & Kirjavainen, M. (2011). Investigating the contribution of procedural and declarative memory to the acquisition of past tense morphology: Evidence from Finnish. *Language and Cognitive Processes*, *26*(4-6), 794–829.
- Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, *18*(1), 146–154.
- Kort W., Schittekatte M., & Compaan E. (2008) *Clinical Evaluation of Language Fundamentals-4-NL (CELF-4-NL)*. Pearson.
- Koster, C., Been, P. H., Krikhaar, E. M., Zwartz, F., Diepstra, H. D., & Van Leeuwen, T. H. (2005). Differences at 17 months: Productive language patterns in infants at familial risk for dyslexia and typically developing infants. *Journal of Speech, Language and Hearing Research*, *48*(2), 426–438.
- 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.
- Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2020a). Statistical learning in the visuomotor domain and its relation to grammatical proficiency in children with and without developmental language disorder: A conceptual replication and meta-analysis. *Language Learning and Development*, *16*(4), 426–450.

- Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2020b). Children with developmental language disorder have an auditory verbal statistical learning deficit: Evidence from an online measure. *Language Learning*, *70*(1), 137–178.
- Lammertink, I., Van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J. (2019). Auditory statistical learning in children: Novel insights from an online measure. *Applied Psycholinguistics*, *40*(2), 279–302.
- Law, J. M., Vandermosten, M., Ghesquière, P., & Wouters, J. (2017). Predicting future reading problems based on pre-reading auditory measures: A longitudinal study of children with a familial risk of dyslexia. *Frontiers in Psychology*, *8*, 124.
- Le Clercq, C. M., van der Schroeff, M. P., Rispens, J. E., Ruytjens, L., Goedegebure, A., van Ingen, G., & Franken, M. C. (2017). Shortened nonword repetition task (NWR-S): A simple, quick, and less expensive outcome to identify children with combined specific language and reading impairment. *Journal of Speech, Language, and Hearing Research*, *60*(8), 2241–2248.
- Leikin, M., & Zur Hagit, E. (2006). Morphological processing in adult dyslexia. *Journal of Psycholinguistic Research*, *35*(6), 471–490.
- López-Barroso, D., Cucurell, D., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2016). Attentional effects on rule extraction and consolidation from speech. *Cognition*, *152*, 61–69.
- Lum, J. A., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex*, *51*, 1–10.
- Lum, J. A., Ullman, M. T., & Conti-Ramsden, G. (2013). Procedural learning is impaired in dyslexia: Evidence from a meta-analysis of serial reaction time studies. *Research in Developmental Disabilities*, *34*(10), 3460–3476.
- Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T. K., Laakso, M. L., Leinonen, S., Leppanen, P., Lyytinen, P., Poikkeus, A., Puolakanaaho, A., Richardson, U., & Viholainen, H. (2001). Developmental pathways of children with and without familial risk for dyslexia during the first years of life. *Developmental Neuropsychology*, *20*(2), 535–554.
- MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54.
- Mainela-Arnold, E., & Evans, J. L. (2005). Beyond capacity limitations: Determinants of word recall performance on verbal working memory span tasks in children with SLI. *Journal of Speech, Language & Hearing Research*, *48*(4), 897–909.
- Mann, V. A., Shankweiler, D., & Smith, S. T. (1984). The association between comprehension of spoken sentences and early reading ability: The role of phonetic representation. *Journal of Child Language*, *11*(3), 627–643.
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., Xu, F., & Clahsen, H. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*(4), 1–178.
- Marshall, C. (2020). Investigating the relationship between syntactic and short-term/working memory impairments in children with developmental disorders is not a straightforward endeavour. *First Language*, *40*(4), 491–499.
- McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the “specifics” of specific reading disability and specific language impairment. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(7), 869–874.
- McGregor, K. K. (2009). Semantic deficits associated with developmental language disorders. In R. G. Schwartz (ed.), *The handbook of child language disorders*. Psychology Press.
- Melby-Lervåg, M., Lyster, S. A., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352.
- Metsala, J. L., Stavrinos, D., & Walley, A. C. (2009). Children’s spoken word recognition and contributions to phonological awareness and nonword repetition: A 1-year follow-up. *Applied Psycholinguistics*, *30*(1), 101–121.
- Mintz, T. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition*, *30*, 678–686.
- Mintz, T. (2003). Frequent frames as a cue for grammatical categories in 327 child directed speech. *Cognition*, *90*, 91–117.

- Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, *1*, 31.
- Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*(1), 302–331.
- Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, *48*, 1033–1047.
- Nash, H. M., Hulme, C., Gooch, D., & Snowling, M. J. (2013). Preschool language profiles of children at family risk of dyslexia: Continuities with specific language impairment. *Journal of Child Psychology and Psychiatry*, *54*(9), 958–968.
- Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: Reuniting the developmental disorders? *Trends in Neurosciences*, *30*(4), 135–141.
- Nicolson, R. I., & Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex*, *47*(1), 117–127.
- Orgassa, A., & Weerman, F. (2008). Dutch gender in specific language impairment and second language acquisition. *Second Language Research*, *24*(3), 333–364.
- Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*(2), 385–413.
- Psychology Software Tools, Inc. (2012). E-Prime 2.0.
- Qi, Z., Sanchez Araujo, Y., Georgan, W.C., Gabrieli, J.D., & Arciuli, J. (2019). Hearing matters more than seeing: A cross-modality study of statistical learning and reading ability. *Scientific Studies of Reading*, *23*(1), 101–115.
- R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, *13*(2), 212–218.
- Ramus, F., Marshall, C. R., Rosen, S., & van der Lely, H. K. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain*, *136*(2), 630–645.
- Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *The Quarterly Journal of Experimental Psychology*, *61*(1), 129–141.
- Raven J. (2003). *Raven progressive matrices*. Springer.
- Reggiani, D. (2010). Dyslexia and the acquisition of syntax: Passive and control. Doctoral dissertation, University of Verona, Italy.
- 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.
- Riches, N. (2020). Commentary on special issue: Syntax and verbal short term memory. *First Language*, *40*(4), 500–506.
- Rispens, J., Baker, A., & Duinmeijer, I. (2015). Word recognition and nonword repetition in children with language disorders: The effects of neighborhood density, lexical frequency, and phonotactic probability. *Journal of Speech, Language, and Hearing Research*, *58*(1), 78–92.
- Rispens, J., & Baker, A. (2012). Nonword repetition: The relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research*, *55*(3), 683–694.
- Rispens, J., & Been, P. (2007). Subject-verb agreement and phonological processing in developmental dyslexia and specific language impairment (SLI): A closer look. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists*, *42*(3), 293–305.
- Rispens, J. E., Been, P. H., & Zwarts, F. (2006). Brain responses to subject-verb agreement violations in spoken language in developmental dyslexia: An ERP study. *Dyslexia*, *12*(2), 134–149.
- Rispens, J., de Bree, E., & Kerkhoff, A. (2014). What's in a suffix? The past tense in Dutch children with reading problems. In J. Hoeksema, D. Gilbers, & P. Hendriks (Eds.), *Black book: A festschrift in honor of Frans Zwarts* (pp. 271–281).
- Rispens, J., Roeleven, S., & Koster, C. (2004). Sensitivity to subject-verb agreement in spoken language in children with developmental dyslexia. *Journal of Neurolinguistics*, *17*(5), 333–347.

- Robertson, E. K., & Joanisse, M. F. (2010). Spoken sentence comprehension in children with dyslexia and language impairment: The roles of syntax and working memory. *Applied Psycholinguistics*, *31*(1), 141–165.
- Rüsseler, J., Gerth, I., & Münte, T. F. (2006). Implicit learning is intact in adult developmental dyslexic readers: Evidence from the serial reaction time task and artificial grammar learning. *Journal of Clinical and Experimental Neuropsychology*, *28*(5), 808–827.
- Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development*, *61*(6), 1728–1743.
- Schittekatte, M., Groenvynck, H., Fontaine, J., & Dekker, P. (2007). *TEAch: Test of Everyday Attention for Children*. Pearson.
- Schmalz, X., Altoè, G., & Mulatti, C. (2017). Statistical learning and dyslexia: A systematic review. *Annals of Dyslexia*, *67*(2), 147–162.
- Schmalz, X., Moll, K., Mulatti, C., & Schulte-Körne, G. (2019). Is statistical learning ability related to reading ability, and if so, why? *Scientific Studies of Reading*, *23*(1), 64–76.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime 2.0 reference guide manual*. Psychology Software Tools.
- Shankweiler, D., Crain, S., Katz, L., Fowler, A. E., Liberman, A. M., Brady, S. A., Thornton, R., Lundquist, E., Dreyer, L., Fletcher, J. M., Stuebing, K. K., Shaywitz, S. E., & Shaywitz, B. A. (1995). Cognitive profiles of reading-disabled children: Comparison of language skills in phonology, morphology, and syntax. *Psychological Science*, *6*(3), 149–156.
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120.
- Sigurdardottir, H.M., Danielsdottir, H.B., Gudmundsdottir, M., Hjartarson, K.H., Thorarinsdottir, E.A., & Kristjánsson, Á. (2017). Problems with visual statistical learning in developmental dyslexia. *Scientific Reports*, *7*(1), Article 606.
- Singh, S., Walk, A.M., & Conway, C.M. (2018). Atypical predictive processing during visual statistical learning in children with developmental dyslexia: An event-related potential study. *Annals of Dyslexia*, *68*(2), 1–15.
- Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia*, *7*(1), 37–46.
- Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin*, *142*(5), 498–545.
- Steady, L.M., Compton, D.L., Petscher, Y., Elliott, J.D., Smith, K., Rueckl, J.G., Sawi, O., Frost, S.J., & Pugh, K.R. (2019). Development and prediction of context-dependent vowel pronunciation in elementary readers. *Scientific Studies of Reading*, *23*(1), 49–63.
- Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neurosciences*, *20*(4), 147–152.
- Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience*, *5*(9), 721–728.
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*(2), 105–122.
- Ullman, M.T., Earle, F.S., Walenski, M., & Janacsek, K. (2019). The neurocognition of developmental disorders of language. *Annual Review of Psychology*, *71*(5), 389–417.
- Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*(3), 399–433.
- van Alphen, P., De Bree, E., Gerrits, E., De Jong, J., Wilsenach, C., & Wijnen, F. (2004). Early language development in children with a genetic risk of dyslexia. *Dyslexia*, *10*(4), 265–288.
- Van den Bos, K. P., Spelberg, H., Scheepsmma, A., & De Vries, J. (1994). *De Klepel. Vorm A en B. Een test voor de leesvaardigheid van pseudowoorden. Verantwoording, handleiding, diagnostiek en behandeling*. Berkhout.
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*(1), 2–40.
- Vender, M., Mantione, F., Savazzi, S., Delfitto, D., & Melloni, C. (2017). Inflectional morphology and dyslexia: Italian children's performance in a nonword pluralization task. *Annals of Dyslexia*, *67*(3), 401–426.

- Verhoeven, L., Steenge, J., & van Balkom, H.** (2011). Verb morphology as clinical marker of specific language impairment: Evidence from first and second language learners. *Research in Developmental Disabilities, 32*(3), 1186–1193.
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A.** (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*(6), 632–638.
- Waltzman, D. E., & Cairns, H. S.** (2000). Grammatical knowledge of third grade good and poor readers. *Applied Psycholinguistics, 21*(2), 263–284.
- West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C.** (2017). The procedural learning deficit hypothesis of language learning disorders: We see some problems. *Developmental Science, 21*(2).
- Wexler, K.** (1994). Optional infinitives, head movement, and the economy of derivation. In D. Lightfoot, & N. Hornstein (Eds.), *Verb movement*. Cambridge University Press.
- Wijnen, F.** (2013). Acquisition of linguistic categories: Cross-domain convergences. *Birdsong, Speech, and Language: Exploring the Evolution of Mind and Brain, 157–177*.
- van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J.** (2017). Visual artificial grammar learning in dyslexia: A meta-analysis. *Research in Developmental Disabilities, 70*, 126–137.
- van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J.** (2019). Statistical learning abilities of children with dyslexia across three experimental paradigms. *Plos one, 14*(8), e0220041.
- van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J.** (2021). The contribution of individual differences in statistical learning to reading and spelling performance in children with and without dyslexia. *Dyslexia*, DOI: [10.1002/dys.1678](https://doi.org/10.1002/dys.1678).
- Wolf, M., & Bowers, P. G.** (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*(3), 415–438.

Cite this article: van Witteloostuijn, M., Boersma, P., Wijnen, F., and Rispens, J. (2021). Grammatical performance in children with dyslexia: the contributions of individual differences in phonological memory and statistical learning. *Applied Psycholinguistics 42*, 791–821. <https://doi.org/10.1017/S0142716421000102>