Journal of Child Language (2022), 1–24
doi:10.1017/S0305000922000484

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

What Characterizes the Productive Morphosyntax of Norwegian Children with Developmental Language Disorder?

Gisela HÅKANSSON[1,2], Emily Wanda WILLIAMS[4], Jannicke KARLSEN[3,4] and Janne von Koss TORKILDSEN[4,*]

[1]Lund University, Sweden
[2]Linnaeus University, Sweden
[3]Statped, Department of Speech and Language Disorders, Norway
[4]Department of Special Needs Education, University of Oslo, Norway
*Corresponding author. Janne von Koss Torkildsen, Department of Special Needs Education, University of Oslo, Postboks 1140 Blindern, 0318 Oslo, Norway. E-mail: janneto@isp.uio.no

(Received 18 December 2020; revised 29 November 2021; accepted 01 September 2022)

Abstract

Little is known about the productive morphosyntax of Norwegian children with developmental language disorder (DLD). The current study examined morphosyntax in Norwegian-speaking children with DLD (n =19) and a control group that was pairwise matched for age, gender, and intelligence quotient (IQ; n = 19). The children's sentence repetitions were studied through the lens of Processability Theory. The group differences were largest for grammatical structures at the latest developmental stage of the processability hierarchy. The Norwegian subordinate clause word order, belonging to the latest stage of the processability hierarchy, stood out as particularly challenging for children with DLD. Only 2 children with DLD but 16 children in the control group produced a subordinate clause with subordinate clause word order. Categorization of children's errors revealed that children with DLD made more errors of all types (addition, omission, substitution, inflection and word order) but especially errors of omission and inflection.

Keywords: developmental language disorder; Processability Theory; error analysis; morphosyntax; sentence repetition

Introduction

Children with developmental language disorder (DLD) are known to score below age expectations on tests measuring language ability, despite having no obvious biomedical conditions that could explain their language problems (Bishop et al., 2017). Morphosyntax is consistently reported as an area of difficulty in DLD (Ebbels, 2014; Leonard, 2014). However, exactly what aspects of morphosyntax are most affected differs substantially

between languages (Leonard, 2014). Thus, it is critical to examine morphosyntax in DLD in different languages to be able to design targeted assessments and interventions, as well as to understand the breadth of variability in how the disorder can manifest. Previous studies of Norwegian children with DLD have only examined tense morphology (Simonsen & Bjerkan, 1998). The present study aims to fill a gap in the cross-linguistic research literature by investigating productive morphosyntax more broadly in Norwegian children with DLD.

In this paper we use the term DLD in accordance with the Catalise consortiums for English and Norwegian (Bishop et al., 2017; Kristoffersen et al., 2021), but note that many of the studies we review below have used other terms for the same or a similar group of children, such as specific language impairment (SLI).

Morphology in Children with Developmental Language Disorder

Many studies of the language production of children with DLD have focused on morphology, particularly verb morphology. For example, multiple studies have reported problems with verb morphology in English-speaking children with DLD (Rice & Wexler, 1996; van der Lely & Ullmann, 2001). However, when looking behind the accuracy scores to see what the children are doing instead of producing the correct forms, different patterns emerge. Focusing on finiteness, Rice and colleagues (e.g., Rice & Wexler, 1996) found that children with DLD behaved like much younger children and did not use finite verbs. These researchers suggested that English-speaking children with DLD have an extended period of infinitive verb use (optimal infinitive stage) compared with typically developing (TD) children. Other studies of English have scrutinized irregular and regular forms of the past tense and found that children with DLD have more problems with the regular past forms. Thus, van der Lely and Ullmann (2001, p. 202) concluded that the children had a "significant deficit in regular past tense formation."

Most studies examining morphology in DLD have analyzed the language production of English-speaking children. When children with DLD are compared across languages, it is clear that, although there is a general pattern of lower accuracy compared with other children of the same age, children with DLD speaking different languages display problems with different structures (Håkansson, 2017; Leonard, 2014). For example, in contrast to English, Spanish-speaking children with DLD do not have clear problems with tense markings. Instead, many of the verb-related difficulties involve the subjunctive. Children with DLD use significantly fewer verbs in the subjunctive form in spontaneous speech (Sanz-Torrent, Serrat, Andreu & Serra, 2008) and make more errors when the subjunctive is elicited (Castilla-Earls, Pérez-Leroux, Restrepo, Gaille & Chen, 2018). For agglutinating languages, such as Hungarian and Kannada, where each affix marks a morphological feature, children with DLD may produce so-called near-misses; i.e., forms that differ from the target form by a single feature (Lukács, Leonard, Kas & Pléh, 2010; Tiwari, Karanth & Rajashekar, 2017). For Cantonese, where temporal relations are marked using aspect rather than tense, aspectual markers appear to be particularly vulnerable (Fletcher, Leonard, Stokes & Wong, 2005).

Syntax in Children with Developmental Language Disorder

Grammatical difficulties in children with DLD are often described as problems with "structural complexity." What is considered a complex linguistic structure is often

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

determined from the point of view of linguistic theory, and it ranges from the morphological level, with complex morphemes reflecting more than one feature (Jakubowics & Nash, 2001), to the syntactic level, which includes coordinated or subordinated clause structures (Frizelle, Thompson, Duta & Bishop, 2019). In the domain of syntax, problems with subordination have been observed in children with DLD in different languages, and these difficulties have been explained in terms of movement operations (Jensen de López, Sundahl Olsen & Chondrogianni, 2014; Marinis & Armon-Lotem, 2015; Novogrodsky & Friedmann, 2006). Some studies have suggested that the difficulties are related to the number of operations (e.g., Riches, Loucas, Baird, Charman & Simonoff, 2010). Other studies focused on the distance between the moved argument and the gap, finding that a longer distance entails a higher complexity (e.g., Riches et al., 2010; van der Lely, 1998). Finally, the universal markedness hierarchy for relative clauses in the world's languages (Keenan & Comrie, 1977) is suggested to reflect the order of difficulty in processing different types of relatives (Friedmann, Belletti & Rizzi, 2009). For example, subject relative clauses are usually better mastered than object relative clauses, but language-specific properties may change this order. Object relatives can be comprehended as easily as subject relatives if there is salient case morphology which makes the construction clearer (see Guasti, Stavrakaki & Arioso, 2012, comparing Italian and Greek).

Children's accuracy when producing different grammatical structures is not the only aspect to consider; information about what children do when encountering problems, and the types of errors they make, can also provide informative directions for intervention (Frizelle & Fletcher, 2014). Interestingly, there are cross-linguistic differences in how children react when dealing with relative clauses that are too difficult, especially in how they resolve object relative clauses. In some languages, such as English and Hebrew, children with DLD tend to convert object relatives into subject relatives (Meir, Walters & Armon-Lotem, 2015; Novogrodsky & Friedmann, 2006), avoiding movement across arguments, whereas Russian-speaking children resort to simple Subject-Verb-Object (SVO) sentences (Meir et al., 2015), and Danish-speaking children use passives (Jensen de López et al., 2014). These differences may be due to frequency, i.e., that types of relative clauses which are more common in a given language are assumed to be easier to process than less common types (Novogrodsky & Friedmann, 2006).

In Norwegian, relative clauses emerge early in the production of TD children (Ribu, Simonsen, Løver, Strand & Kristoffersen, 2019). However, their internal word order causes problems, because the sentence adverbial in subordinate clauses is placed in front of the verb rather than after the verb as in main clauses (Ringstad & Kush, 2021; Westergaard & Bentzen, 2007). As we will discuss in more detail below, Norwegian-speaking children tend to overgeneralize the main clause word order when they start using relative clauses.

In addition to the notion of structural complexity, which is discussed above, developmental complexity must also be considered. Developmental complexity refers to the order in which linguistic structures emerge. Structural and developmental complexity may be related, but do not necessarily coincide (Theodorou, Kambanaros & Grohmann, 2017). In this paper we take a developmental perspective on children with DLD, following the suggestion by Leonard (2014) that they may develop language at a slower pace:

> Children with SLI seem to differ from their peers primarily in their slower pace of language development and their greater vulnerability to the more challenging details of the language they are learning. These vulnerabilities seem to fall in the areas of

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

> grammatical computation and phonological short-term memory, but may well represent low points on an ability continuum rather than markers of a separable condition. (Leonard, 2014, p. 4)

Morphosyntax in Norwegian Children with Developmental Language Disorder

Norwegian, the language in focus in this study, is a Germanic language that is closely related to Danish and Swedish. It has a simple verbal morphology and a more complex nominal morphology. Norwegian is a verb-second language (V2), with the verb always taking the second position in main clause declaratives. Subordinate clauses differ from main clauses in that the sentential adverb is placed in front of the finite verb, instead of after the finite verb as in main clauses. In the stage model of typical development LARSP (Language Assessment Remediation and Screening Procedure; Ribu et al., 2019) the profile for Norwegian describes two stages of relative clauses: emergence at Stage III (2;0-2;6), with overgeneralization of main clause word order, and subordinate clause word order with preverbal placement of adverbs occurring at Stage VII (4;6-5;6) (see also Westergaard & Bentzen, 2007, p. 282). Overgeneralization of main clause word order in subordinate clauses is also documented in Swedish-speaking children (Håkansson, 1989, 2017). This is in accordance with the claim by Brandt, Diessel, and Tomasello (2008) that children first use relative clauses that are similar to main clauses.

The details of Norwegian grammar that are most relevant for this study are described in the sections about Processability Theory below. Regarding children with DLD, there is a paucity of studies on productive morphosyntax (but see e.g., Helland, Biringer, Helland & Heimann, 2009, for a study on parental reports of language and communication in Norwegian children with DLD). The one previous study that has specifically examined productive morphosyntax in Norwegian investigated eight 6- to 8-year-old children with DLD and TD controls on an elicitation task targeting past tense conjugations of verbs (Simonsen & Bjerkan, 1998). In contrast to the English-speaking children in the studies by Rice and Wexler (1996), the Norwegian children rarely used infinitives, and in contrast to the findings by van der Lely and Ullmann (2001), the Norwegian children had few problems with common regular patterns. On the contrary, they tended to overgeneralize regular past tense endings from one of the conjugation classes to verbs from other conjugation classes, as well as to irregular verbs.

A Theory of Staged Development: Processability Theory

In this study, we apply a psycholinguistic theory of language acquisition, Processability Theory (PT; Pienemann, 1998), to investigate whether Norwegian children with DLD are at the "low points on an ability continuum" (Leonard, 2014, p. 4) in comparison to children with TD. The choice of a developmental theory implies that we regard the children as language learners in the process of developing grammar when they encounter constraints in the developmental process (Thomas, 2011, p. 33). The focus is on developmental markers, and on what the children can produce instead of what they cannot do. PT suggests five major stages of increasing complexity for the development of morphosyntax. The stages are implicationally ordered and the processing of one stage implies the acquisition of all earlier stages (Pienemann, 1998).

The cross-linguistic validity of the PT stages is supported by empirical data from L2 acquisition of 19 different languages (Håkansson, 2017), agrammatism in

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

Table 1. Processing Procedures, Feature Unification and Examples of Norwegian Grammar According to Processability Theory

Stage	Processing procedures	Feature unification	Outcome: Norwegian grammar
5	Subordinate clause procedure	Main and subclause	Subclause word order
4	Sentence procedure	Between phrases	Subject-verb inversion
3	Phrasal procedure	Within phrases	NP agreement
2	Category procedure	No (local morph.)	Tense
1	Words	No	Invariant words

English-speaking patients with aphasia (Dyson, Håkansson & Ballard, 2022) and Swedish-speaking children with DLD (Håkansson, 2001). For Swedish, the developmental hierarchy has been further supported in a number of longitudinal studies of monolingual and bilingual children with and without DLD (Håkansson, 2017; Håkansson, Salameh & Nettelbladt, 2003; Salameh, Håkansson & Nettelbladt, 2004) and clinical implications have been suggested (Nettelbladt, Håkansson & Salameh, 2007).

The processing routines in PT relate to Levelt's (1989) model of speech production, but while the procedures needed for speech are automated in adults, they are acquired stage-by-stage in children, from annotating of the lexicon, to morphological agreement and subordinate clauses. To operationalize the morphosyntactic stages, the grammatical formalism of Lexical Functional Grammar (LFG; Bresnan, 2001) is used. LFG is a generative grammar, with unification of features as a prominent characteristic (see Bresnan, 2001; Levelt, 1989). The concept of unification of features between words, phrases and clauses is used by PT to describe the development of morphosyntax as a gradual expansion of feature unification. Table 1 illustrates the hierarchy of processing procedures, and thus, the corresponding PT stages, where the scope of feature unification is exemplified by Norwegian grammar for each stage.

In the first stage, words are learned, and the lexicon is annotated with forms, meanings, and syntactic categories. In Stage 2, when the lexicon is categorized, local/lexical affixes can be added – for example, tense markings. In Stage 3, feature unification takes place between words in a phrase – for example, as NP agreement (between article, adjective and noun). The next stage, Stage 4, involves assigning the sentence functions through feature unification between subject and verb. Different languages have different means for this. One example is head marking (agreement marking on verbs), another is dependent marking (case marking on nouns) and a third is word order (e.g., V2). In English, subject-verb agreement emerges at Stage 4, and in Norwegian, subject-verb inversion emerges at this stage. Finally, Stage 5 entails the categorization of main and subordinate clauses, and marking them by different features. This distinction does not occur in all languages, but when it does, it surfaces, for example, in different word orders or verb modes in main and subordinate clauses. In English, there is a difference in word order between main clause direct questions and subordinate clause indirect questions, whereas in Spanish the subjunctive mood is used in subordinate clauses. In Norwegian the adverbial shifts place from the main clause placement after the finite verb to the subordinate clause placement in front of the finite verb. Thus, with PT as an analytic lens, seemingly unrelated phenomena can be connected as parts of a developmental sequence (Dyson & Håkansson, 2017).

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

Processability Theory Methodology

The PT stages are hypothesized to apply to the development of any language and all types of learners. A criterion of EMERGENCE is used to decide the developmental stage. Instead of measuring the distance to the target, the focus is on how the individual proceeds from one stage to another. However, it is important to find distributional evidence for systematic and productive use in obligatory contexts, and to avoid memorized formulas. For example, for inflectional morphology to be regarded as productive, there has to be lexical variation with different verbs in the same tense (spiste 'eat-PAST'; bygde 'built-PAST'; springer 'run-PRES'; legger 'put-PRES') and/or morphological variation with the same verb in different tenses (kjøper 'buy-PRES'; kjøpte 'buy-PAST'). For syntactic development this variation is not necessary, as children rarely rely on memorized chunks for syntax, and thus one occurrence in a sample can be sufficient to show that the structure is processable (Pienemann, 1998, p.133).

The analyses are presented by means of implicational scales (De Camp, 1971). There are implicational relationships between items, where the presence of X implies the presence of Y; and thus they reveal developmental stages in language acquisition (Pienemann, 1998). This makes it possible to unearth development in cross-sectional studies. To determine that the development is systematic, it is desirable to ensure that there are only a few gaps in the data, or in other words, that the presence of one structure always implies the presence of another. A gap in the data would indicate that it is possible to reach a certain stage without completing previous stages. A scalability coefficient of above 0.90 is considered strong enough for the developmental stages to be reliable (Hatch & Lazaraton, 1991, p. 204).

The Current Study

In this study, we examine morphosyntactic aspects of the language production of Norwegian children with and without DLD, aiming to identify the structures that represent key challenges for these children. Our analyses are based on a subtest from a commonly used clinical test of language ability for Norwegian, the Clinical Evaluation of Language Fundamentals IV (CELF-IV; Semel, Wiig & Secord, 2003) – Recalling sentences. From this task, we extract structures that are compatible with the stages of PT and analyze the children's production. In addition, we analyze how much and in what way the children's responses differ from the target sentences. The overarching study used data from a larger project examining language function and quality of life in four groups of children: those with cochlear implants, hearing aids, DLD, and TD (see Torkildsen, Hitchins, Myhrum & Wie, 2019). Only data from children with DLD and TD were used in the current study. We focused on children of primary school age (7;5-13 years), as the study examined a broad range of morphological and syntactic structures, some of which have been shown to emerge late in development even for TD children (Ribu et al., 2019; Ringstad & Kush, 2021; Westergaard & Bentzen, 2007, p. 282).

This study addresses the three following research questions:

1) Will the latest stage (5) in the PT hierarchy pose the largest difficulties for the children with DLD compared with children in the TD control group?
2) Will there be an implicational order between the stages so that children will only use a structure from a higher stage if structures from the lower stages are present?

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

3) Are there differences between children with DLD and TD children in terms of how many and what types of errors they make?

Methods

Participants

Thirty-eight children aged 7;5–13;0 participated in the current study, including 19 children with DLD (11 girls and 8 boys) and 19 TD children, who were pairwise matched with the children with DLD on age (within 6 months), gender, and nonverbal abilities (within 1 SD). All the children with DLD were recruited from the educational and psychological counseling service in municipalities across Norway. This service is responsible for assessing and counseling children with developmental difficulties.

The inclusion criteria for the DLD group were as follows:

1. Referral to the educational and psychological counseling service for language difficulties.
2. Independent confirmation of DLD status via scores of 1 *SD* or more below the normative mean on at least two out of the following five standardized assessments:
   a) The British Picture Vocabulary Scale II (BPVS-II; Dunn, Dunn, Whetton & Berley, 1997; Norwegian version by Lyster, Horn & Rygvold, 2010)
   b) The Children's Test of Nonword Repetition (Gathercole, Willis, Baddeley & Emslie, 1994), Norwegian version by Furnes and Samuelsson (2009)
   c) Concepts & Following Directions from CELF-IV (Semel et al., 2003)
   d) Number Repetition 1–Backward from CELF-IV (Semel et al., 2003)
   e) The General Communication Composite (GCC) from the Children's Communication Checklist (CCC-2; Bishop, 2003; Bishop, 2011, Norwegian version by Helland et al., 2009)
3. Nonverbal abilities of at least 70 as measured by the Raven's Coloured Progressive Matrices (CPM) for children aged 5;9–8;11 or Raven's Standard Progressive Matrices (standard version or plus version) for children aged 9;0–12;11 (Raven & Court, 2003).
4. Norwegian as their first language.
5. No diagnosis of other developmental disorders, such as autism or ADHD.
6. Presence of otoacoustic emissions, indicating normal inner ear function and normal hearing with hearing thresholds better than 30 dB (Engdahl, Tambs, Borchgrevink & Hoffman, 2005).

Initially, 34 children were recruited to the DLD group from the educational and psychological counseling service. Five children were excluded from the study for the following reasons: language scores that were better than the criteria specified in (2) above (two children), nonverbal ability scores on Raven's matrices below 70 (two children), and a lack of an audio recording of the relevant CELF-IV subtests (one child). In addition, two children were excluded because we did not succeed in finding a control child who was matched on age, gender and nonverbal IQ, and who satisfied the inclusion criteria for control children as specified below. Finally, eight children with DLD were excluded because they had not completed sentences 5–13 on the Recalling Sentences subtest due to the stopping rule on the test, which entails that administration

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

of the subtest is stopped after four consecutive 0-point responses according to the scoring protocol.

Children in the TD control group were tested using the same assessments as the children with DLD. Criteria for inclusion in the TD group were that they did not score below 1 *SD* from the mean on more than one of the five standardized language tests described in criterion 2 above and did not have a history of language difficulties. In addition, they met inclusion criteria 3–6 above. Parents of all participating children were asked to fill out a background questionnaire regarding child characteristics (e.g., diagnoses and special education services) and information about the parents (e.g., languages spoken in the home).

Descriptive information regarding participants is given in Table 2. No statistically significant difference was found between children in the DLD and control groups on nonverbal abilities measured with Raven's matrices, but there was a moderate difference in effect size (see Table 2).

Methodological Considerations

Sentence repetition has the advantage of giving control of which structures are elicited, and thus, the researcher can formulate precise hypotheses about grammatical proficiency. Also referred to as SENTENCE RECALL or ELICITED IMITATION, sentence repetition has been used in studies of language development since the 1960s, and it is a commonly used method in studies of DLD (Komeili, Marinis, Tavakoli & Kazemi, 2020; Marinis & Armon-Lotem, 2015). The underlying assumption is that children will interpret the content of an utterance and reconstruct it according to the present state of their grammar (the Regeneration Hypothesis; Potter & Lombardi, 1990). The role of memory in sentence repetition is complex and intertwined with the normal mechanisms of spontaneous speech production (Lombardi & Potter, 1992). In immediate sentence recall, the child regenerates the utterance not from a surface representation in short term memory, but from a conceptual representation that probes deeper into working memory and underlying semantic and syntactic knowledge stored in long term memory (Lombardi & Potter, 1992; Potter & Lombardi, 1990). Indeed, many studies have found that performance on sentence repetition tasks reflects the child's general language ability, and speakers perform at the same level in repetition tasks as in other tasks of language production (Devescovi & Caselli, 2007; Håkansson & Hansson, 2000; Marinis & Armon-Lotem, 2015; Potter & Lombardi, 1990; Riches et al., 2010).

Materials

The Recalling Sentences, Formulated Sentences, and Concepts & Following Directions subtests and Backwards Number Repetition subtests from CELF-IV (Semel et al., 2003) measure children's ability to repeat sentences and numbers, and follow verbal instructions. In the Recalling sentences subtest, the child was asked to repeat 18 sentences with increasing complexity as accurately as possible. Following the manual, each repetition was awarded 0–3 points based on the number of grammatical and other substitution and omission errors. The administration was stopped after four consecutive utterances with a score of 0. In this study, we report percentile scores for this test according to the Norwegian manual (see Table 2), and detailed analyses of repetitions to answer research questions 1-3.

In the Concepts & Following Directions subtests from CELF-IV, children were presented with series of pictures and asked to point to them in the order presented by

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

Table 2. Descriptive Characteristics of Children with DLD (n = 19) and Matched Controls (TD; n = 19) and Independent Samples *t*-Tests of Group Differences

	Children with DLD			Children with TD			*t*-test (*df*)		Effect size
	M	*SD*	Range	*M*	*SD*	Range	*t* (42)	*p*	*d*
Age	124.25	17.77	91.63–152.80	124.62	18.73	89.49–154.74	0.62	.951	0.03
Nonverbal IQ	89.21	11.58	75–120	95.79	11.09	85–125	1.789	.082	0.58
Receptive vocabulary	76.63	18.81	16–100	101.21	13.13	79–122	4.67	<.001	1.51
Nonword repetition	17.47	5.31	7–27	23.89	2.21	21–27	4.87	<.001	1.58
CELF-IV Concepts	7.05	2.20	2-12	9.95	2.10	7-15	4.16	<.001	1.35
CELF-IV Num Rep. Bw.	7.84	2.32	3-12	9.47	3.47	1-15	1.71	.098	0.55
CCC-2 GCC	23.35	21.26	1-63	88.00	10.50	57-101	11.51	<.001	3.86
CELF-IV Rec. sen.	5.58	3.34	1–12	11.05	2.12	8–15	6.03	<.001	1.96

Note: CELF-IV Concepts = subtest Concept and Following Directions, CELF-IV Num Rep. Bw. = subtest CELF-IV Number Repetition Backwards, CELF-IV Rec. sen. = subtest Recalling sentences. CCC-2 GCC = Children's Communication Checklist General Communication Composite. Age is given in months. Scores for Nonverbal IQ and Receptive vocabulary are standard scores. Scores for CELF-IV subtests are scaled scores. Scores for the nonword repetition task are raw scores (no Norwegian norms exist for this test). Scores for the Children's Communication Checklist are percentile scores. Effect sizes are Cohen's *d*.

the test (e.g., *"Point at the house in the upper row"* or *"After pointing at the shoe, point at the fish"*). Each task received a score of 0 points or 1 point, and the maximum score was 38.

In the Backwards Number Repetition subtest from CELF-IV, the child was asked to repeat a random sequence of numbers with increasing length in a reverse order. The sequences were orally presented. Each response received a score of 0 or 1 points, and the maximum score is 14.

The BPVS-II (Dunn et al., 1997; Lyster et al., 2010) is a measure of receptive vocabulary. The children were asked to match a word presented orally by the test instructor with one out of four pictures. The stop criterion was 8 out of 12 incorrect answers, and the maximum score was 144.

The children's test of nonword repetition (Gathercole et al., 1994; Norwegian version by Furnes & Samuelsson, 2009) measures the ability to correctly repeat nonwords. Twenty-eight nonwords consisting of two, three, four, or five syllables were presented to the children through earphones, and the children were asked to repeat them as accurately as possible. There were no stop criteria.

The General Communications Composite (GCC) from the Children's Communication Checklist (CCC-2; Bishop, 2011; Norwegian adaption: Helland et al., 2009) includes items from eight different scales (e.g., syntax, semantics, and non-verbal communication), and covers both structural language and pragmatic skills. The checklist was completed by the parents.

Raven's CPM (Raven & Court, 2003) is a multiple choice test of nonverbal ability. There were no stop criteria for this test.

Procedure

Test administration

All tests were administered individually and according to their manuals. The tests included in this study were part of a test battery used in the larger project, and additional tests not reported here included tests of hearing in noise, sustained attention, and reading fluency, as well as parent reports of health-related quality of life. For most children, the test battery was distributed over two days within a two-week period, and each test session lasted between one and two hours.

Transcription

The children's responses to the Recalling Sentences task were recorded and subsequently transcribed by the second author and a research assistant. Transcribers also noted whether the child had phonological difficulties or spoke a regional dialect of Norwegian. Transcriptions from approximately 15% of the children were checked for inter-transcriber reliability. The transcriptions were found to be very similar (less than 5% of words were different between the two transcribers).

The current study focused on productive morphosyntax regardless of phonological or articulatory ability. For this reason, phonological difficulties such as mispronunciation of certain speech sounds were not written into the transcription. Also excluded from analyses were interjections or comments unrelated to the target sentence and repeated sounds or parts of words. If the child attempted to self-correct his or her response, the last version of the utterance was the one used for analysis. Dialectical variations of words, including verb and noun inflection, were transcribed as is.

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

Analytical Approach

Data selection

From the 18 examples in the Recalling Sentences task, we selected the nine sentences that were attempted by all children in the sample for the error analyses (see Table 3). For the PT analyses, we selected five of these nine sentences that provided obligatory contexts for the PT structures, as shown in Table 4. Specifically, the materials consisted of two obligatory contexts for each of the morphological structures and one obligatory context for each of the syntactic structures. The selection was based on the stages from PT. Our adaptation of the PT stages to Norwegian is shown in Table 1 (general examples of how the PT stages relate to Norwegian grammatical structures) and Table 3 (the specific target

Table 3. The Target Sentences for PT and Error Analyses

Sentence (English Translation)	Obligatory Context for PT	Analyses where sentence is included
(1) *Lisa står etter Pelle i den lange køen.* 'Lisa is standing behind Pelle in the long queue.'	Phrasal (NP agreement, PT3)	Error and PT
(2) *Den store brune hunden spiste opp all kattens mat.* 'The large brown dog ate up all of the cat's food.'	Category (tense, PT2) + Phrasal (NP agreement, PT3)	Error and PT
(3) *Jentene og guttene bygde et sandslott på stranden.* 'The girls and boys built a sandcastle on the beach.'	Category (tense, PT2)	Error and PT
(4) *Fordi det er fridag i morgen, kan vi være oppe lenge.* 'Because tomorrow is a holiday, we can stay up late.'	–	Error
(5) *Sommerhuset vårt, som ligger ved havet, trenger å bli malt.* 'Our summer house, which is by the sea, needs to be painted.'	–	Error
(6) *Treneren kunne ikke finne håndballene som laget spilte med i fjor.* 'The coach could not find the handballs that the team played with last year.'	–	Error
(7) *Siden vi skulle bowle senere, sprang vi direkte hjem etter fotballkampen.* 'Since we were going bowling later, we ran straight home after the football match.'	Sentence (inversion, PT4)	Error and PT
(8) *Hvis det ikke slutter å regne før klokken atten, må tenniskampen avlyses.* 'If it doesn't stop raining before six o'clock, the tennis match will need to be cancelled.'	–	Error
(9) *Hunden Peik, som bare er en liten valp, liker å leke med den store blå ballen sin.* 'The dog Peik, who is just a little puppy, likes to play with his big blue ball.'	Subclause (subcl word order, PT5)	Error and PT

PT level refers to the highest PT level of grammatical structures included in the sentence.

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

Table 4. Two Examples of Obligatory Contexts for Morphological PT Stages and One Example of the Syntactical PT Stages, Selected from CELF-IV

PT stage	Processing Procedure	Example from CELF IV Sentence Repetition or Formulated Sentences
2	Category procedure	Den store brune hunden **spiste** opp all kattens mat
		The big-DEF brown-DEF dog-DEF eat-PST all cat-DEF-GEN food
	Past tense	'The big brown dog ate all the cat's food'
	Category procedure	Jentene og guttene **bygde** et sandslott på stranden
	Past tense	Girl-PL.DEF and boy-PL.DEF build-PST a sandcastle on beach-DEF
		'The girls and the boys built a sandcastle on the beach'
3	Phrasal procedure:	Lisa står etter Pelle i **den lange køen**
	NP agreement	Lisa stand behind Pelle in the long-DEF queue-DEF
		'Lisa stands behind Pelle in the long queue'
	Phrasal procedure:	Den **store brune hunden** spiste opp all kattens mat
	NP agreement	The big-DEF brown-DEF dog-DEF eat-PST all cat-DEF-GEN food
		'The big brown dog ate all the cat's food'
4	Sentence procedure:	Siden vi skulle bowle senere, **sprang vi** direkte hjem etter kampen
	Subject-verb inversion	Since we would-PST bowl later, run-PST we directly home after game-DEF
		'Since we were bowling later, we ran directly home after the game'
5	Subordinate clause procedure:	Hunden Peik **som bare er** en liten valp liker å leke med den store blå ballen sin
	ADV before V	Dog-DEF Peik who-REL just is a tiny puppy like to play with the big-DEF blue ball-DEF his
		'The dog Peik, who is just a tiny puppy, likes to play with his big blue ball'

Note., DEF = definite, PST = past tense, REL = relative. Target structures are shown in bold.

sentences in our study in relation to PT). The first and the third authors scored responses for PT stages. The second and the last author scored the errors.

PT analyses (research questions 1 and 2)

Sentences were categorized into PT stages according to the grammatical structure in the sentence (see detailed scoring criteria in Supplementary Materials, Table A1). Following the PT procedure, we scored two examples of a morphological structure, and one example of a syntactic structure as evidence that the stage was processable. The inter-rater reliability for the PT analysis (highest PT level produced by each individual child) was almost perfect (Cohen's kappa = .919, $p < .001$, McHugh, 2012).

Error analyses (research question 3)

Each production from the sub-set of nine sentences from the Recalling Sentences task was analyzed in terms of the total number of errors made. An "error" in this context refers to any deviation from the target sentence, to the exclusion of phonological errors as described above. Dialectical variations of repeated words were accepted as exact repetitions. During a sentence repetition task, the child processes the target sentence and reconstructs his or her response using the grammar and vocabulary available to him or her (Potter & Lombardi, 1990; Vinther, 2002). In addition, children may also use their dialect as a filter through which they process and respond to such as task, and thus we decided to disregard dialectal variations as has been done in similar studies in other languages (Taha, Stojanovik & Pagnamenta, 2021).

Exact repetition

For this scoring method, each production was assigned a score of 1 if it was exactly the same as the target sentence. If the response contained one or more errors, it scored 0. An exact repetition allowed for acceptable dialectical alternatives both in pronunciation and in morphology or vocabulary. Inter-rater reliability for exact repetitions was almost perfect (Cohen's kappa = .977, $p < .001$, McHugh, 2012).

Error types

Each production scored as a non-exact repetition was further analyzed to determine the number of errors contained within the child's response relative to the target sentence. Errors were assigned to the following categories: omission of function words or content words, substitution of content words or function words, addition of words, inflectional changes, and word order changes (see detailed scoring criteria in Supplementary Materials, Section C). The inter-rater reliabilities (Cohen's kappa) for the subcategorization into different error types were all above McHugh's (2012) criterion of .6 for acceptable agreement (addition of words = .779, $p < .001$; word order changes = .894, $p < .001$; omission of function words = .779, $p < .001$; omission of content words = .879, $p < .001$; substitution of function words = .811, $p < .001$; substitution of content words = .790, $p < .001$; inflectional changes = .776, $p < .001$).

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

Results

Research question 1

Group data

Figure 1 gives the proportion of children in each group (DLD and TD) who produced the target structure at each PT level, where 2 is the lowest and 5 is the highest.

As is clear from Figure 1, the structures that were difficult for the children with DLD were the syntactic structures belonging to PT stages 4 and 5: – subject-verb inversion at stage 4 and subordinate clause word order in a relative clause with adverbials at stage 5. However, the proportion of children who produced structures at stage 4 did not differ significantly between children with DLD (M = .53, SD = .51) and TD (M = .74, SD = .45) (t (36) = 1.34, p = .19, d = 0.48). On the other hand, there was a statistically significant difference between the proportions of children with DLD (M = .11, SD = .32) and TD (M = .84, SD = .38) who produced structures at stage 5 (t (36) = 6.56, p < .001, d = 2.08). In fact, only two of the children with DLD (D9, D15) produced the subordinate clause word order. In contrast, most of the TD children (16 out of 19) produced the target subordinate clause word order (see sentence 9 in Table 3).

Research question 2

The results for children's individual development are laid out in implicational scales (see Table 5). This type of scaling is based on the principle that the presence of a rule at a higher stage implies the presence of a rule at a lower stage (i.e., children who are unable to produce stage 3 will not produce stage 4 either). The implicational scale revealed that the stages developed in a predictable sequence. Only five of the 38 children produced

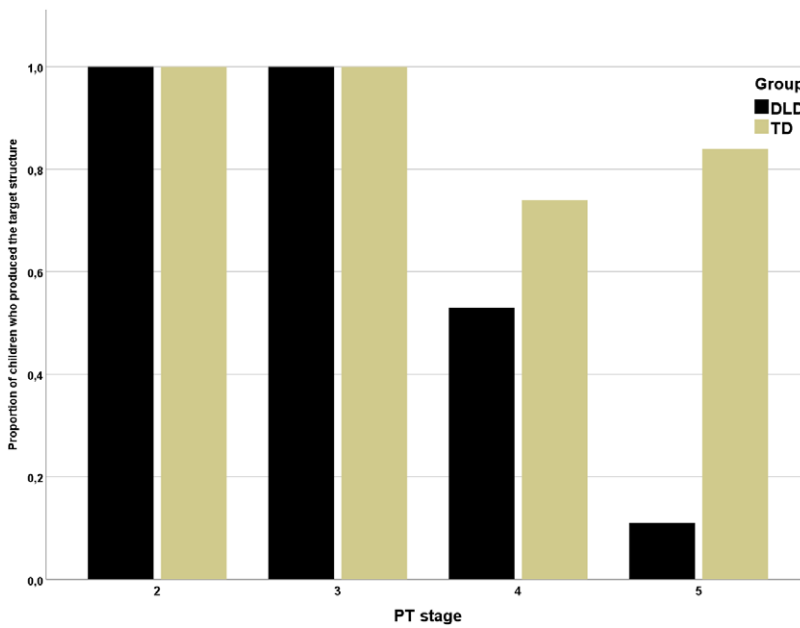

Figure 1. The proportion of children in the DLD and TD groups who produced the target structure at each PT level.

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

Table 5. Implicational Table for Use of PT Stages in the Recalling Sentences Task

ID	stage 2	stage 3	stage 4	stage 5
D1	+	+	–	–
D3	+	+	–	–
D7	+	+	–	–
D8	+	+	–	–
D13	+	+	–	–
D14	+	+	–	–
D17	+	+	–	–
D20	+	+	–	–
T5	+	+	–	–
D2	+	+	+	–
D4	+	+	+	–
D6	+	+	+	–
D10	+	+	+	–
D12	+	+	+	–
D16	+	+	+	–
D19	+	+	+	–
D21	+	+	+	–
D22	+	+	+	–
T9	+	+	+	–
T18	+	+	+	–
D15	+	+	–	+
T2	+	+	–	+
T8	+	+	–	+
T15	+	+	–	+
T20	+	+	–	+
D9	+	+	+	+
T1	+	+	+	+
T3	+	+	+	+
T4	+	+	+	+
T6	+	+	+	+
T7	+	+	+	+
T11	+	+	+	+
T12	+	+	+	+
T13	+	+	+	+
T14	+	+	+	+

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

Table 5. *(Continued)*

ID	stage 2	stage 3	stage 4	stage 5
T16	+	+	+	+
T19	+	+	+	+
T22	+	+	+	+

Note. D[number] refers to individual children in the DLD group. T[number] refers to individual children in the TD group. Scalability: 0.956, which indicates the degree to which the developmental stages follow an implicational order (Hatch & Lazaraton, 1991).

structures at higher stages (stage 5) without also producing structures at the preceding stage (stage 4). This gave a high scalability (0.956), which shows that the implicational order of the stages is robust (Hatch & Lazaraton, 1991). In other words, the presence of a structure from the higher stages implies the presence of structures from the lower stages. This hierarchical pattern of grammatical development is also demonstrated in the children's production of earlier stages when they do not produce later stages. When attempting to repeat the stage 5 structure (subordinate clause with subordinate clause word order), children used structures from stage 2 (e.g., child D12 using present tense: den hunden leke-r med ballen 'the dog-DEF play-PRS with ball-DEF), and/or stage 3 (e.g., child D10 using NP agreement: den stor-e blå ball-en sin 'the big-DEF blue ball-DEF his'). See Supplementary Materials (Supplementary Materials, section B) for transcription and classification of all responses to sentence 9.

The results from the individual data show no qualitative difference between the children with DLD and the children with TD. For example, one of the typically developing children (T21), like most children with DLD, did not produce any structures at stages 4 and 5. One of the children with DLD (D9) behaved like most of the TD children and produced structures at stages 4 and 5. There was no statistically significant relation between age and PT level in either the TD ($r = .18$, $p = .48$) or the DLD group ($r = -.16$, $p = .52$). For the TD-sample this might be due to little variation, as most TD children ($n = 16$) performed at level 5. For the children with DLD, the cause is probably more complex and possibly related to the severity of their language disorder.

Research Question 3

We address research question three through two different methods of scoring: exact repetitions and errors by type.

Exact repetitions

The results from the exact repetition scoring method are presented in Figure 2. An independent-samples t-test was conducted on the sum score for exact repetitions, with a possible score range of 0 (no exact repetitions) to 9 (exact repetition on all of the nine sentences). These results show that children with DLD produced significantly fewer exact repetitions overall (M=2.89, SD=1.94) than their TD counterparts (M=5.89, SD=1.52), (t (36) = -5.30, $p > .001$, $d = 1.72$). Visual investigation of Figure 2 shows that some sentences in particular produced a greater difference in performance between groups than

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

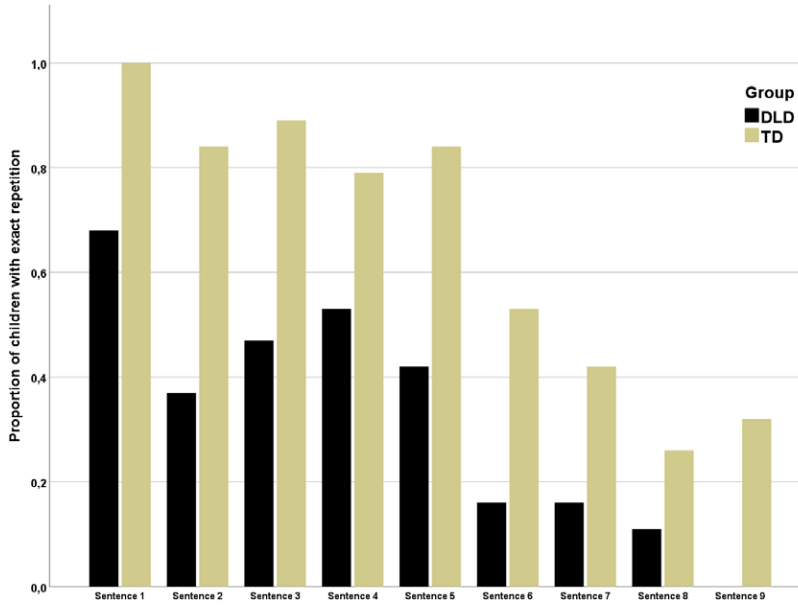

Figure 2. The proportion of children in the DLD and TD groups who produced an exact repetition of each target sentence. For the nine target sentences and their English translations, see Table 3.

others. While it appears that sentences 6, 7 and 8 were equally difficult for children with DLD, there were no exact repetitions of sentence 9 amongst the children with DLD. We do not see the same tendency for the TD children, who in fact produced slightly more exact repetitions for sentence 9 than sentence 8. In both groups we see that the proportion of exact repetitions is highest for the first sentence and steadily declines, with slight variations, towards the last sentence. This can be attributed to the test design itself, whereby the target sentences gradually increase in both length and complexity as the test is administered.

Total errors and error types

The results from the total error analysis show that children with DLD made on average significantly more sentence repetition errors (M = 28.16, SD =13.8) than the TD children (M =7.05, SD = 5.19) (t (36) = -6.23, $p > .001$, d =2.03). There was also a much larger variation of scores within the DLD group than the TD group. Results from the error type analysis along with significance values and effect sizes are presented in Figure 3. The most common types of errors made by children with DLD were omissions of function words and content words, followed by substitution of content words and additions. The most common types of errors made by TD children were substitution of content words, omission of content words and additions. The children with DLD made significantly more errors within each error type than the TD children. Effect sizes (Cohen's d) for the group comparisons (independent samples T-tests) are all large, but vary from $d = 0.89$ for additions to $d = 1.82$ for omissions of function words (see Figure 3).

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

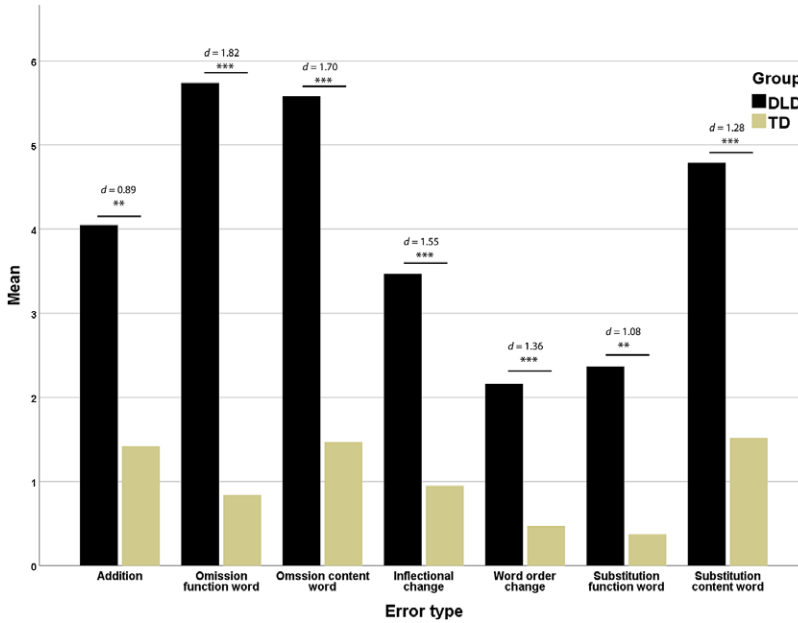

Figure 3. The mean number of errors in each error category for children in the DLD and TD groups. Significance levels for the group difference: *** = p < .001, ** = p < .01. Effect sizes are Cohen's *d*.

Discussion

The theoretical assumption for this study is that grammar is acquired in predictable stages and that children with TD and DLD follow the same stages, but with a different rate of acquisition. The study used a sentence repetition task to examine productive morpho-syntax in Norwegian children with DLD from the point of view of PT and its five stages, applied to Norwegian. In addition, the children's responses were further examined in a detailed error analysis.

The first research question concerned whether children with DLD have more problems with the latest stage (Stage 5) than the earlier stages of PT. Our data show that the PT hierarchy can be applied to Norwegian children with DLD, and that they have more difficulties with stage 5 than stage 4 compared to the children in the TD control group. This suggests that DLD and TD children follow the same morphosyntactic development, with a difference in rate of development. Similar results have been reported from Swedish children with DLD (e.g., Håkansson, 2017).

To capture the language-specific challenges of Norwegian, we aimed to map the developmental sequence of this language according to PT. Data at both the group level and the individual level (where each child with DLD was pairwise matched with a control child for age, nonverbal IQ, and gender) indicated that children with DLD struggle with the higher PT stages (stage 4: sentence procedure; subject-verb inversion; stage 5: subclause procedure; subclause word order; Pienemann, 1998).

In this study, the most advanced PT stage, specifically the difference between main and subordinate clauses in placement of the sentential adverbial, uncovered the challenge of Norwegian subclause word order, which has not been mentioned in previous discussions

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

of language disorder in Norwegian. However, it should be noted that we would not have detected the challenges with this structure without the Recalling Sentences task. Sentence repetition encourages children to use their full linguistic ability, and has been employed to assess language skills in many studies (e.g., Frizelle & Fletcher, 2014; Marinis & Armon-Lotem, 2015; Riches et al., 2010). As Frizelle, Thompson, McDonald, and Bishop (2018, p. 1191) stated, "It may be the case that, unless given a task that obligates the use of complex syntax or where the structures are primed, children at this age will not tap into their full linguistic ability."

The second research question concerned the implicational order between the stages, and whether the children will use a structure from a higher stage only if structures from lower stages are present. The analyses demonstrated that the later stages are underpinned by the earlier stages among most children, both children with DLD and TD. Furthermore, even when the children with DLD experienced great problems in repeating the Stage 5 example (the subordinate clause word order), they managed to use structures from Stages 2, 3 and 4 in their reformulations. It should be noted that the word order in Norwegian subordinate clauses is superficially identical to main clause word order when the subordinate clause does not contain adverbials. Thus, it has been suggested that children in their initial stages of language acquisition assume that the verb placement in main clauses may be extended to subordinate clauses (Ribu et al., 2019; Ringstad & Kush, 2021; Westergaard & Bentzen 2007, p. 282).

The third research question concerned differences between the children with DLD and TD in how many and what types of errors they make in the Recalling Sentences task. Specifically, we applied two scoring methods to the sentence repetition task: exact repetition and total errors by type. The scoring method for total errors by type allowed us to cautiously attempt to identify differences in morphosyntactic production across groups. We found an important difference between the two groups, with the children with DLD giving significantly fewer exact repetitions of the target sentences. Furthermore, their responses were characterized as deviating further from the target sentences than the responses of the TD children on all the measures studied; that is to say, as a group they made significantly more errors of all types than the TD group. This is not surprising, as previous studies on sentence repetition tasks have shown them to reveal a special weakness for children with DLD, and to have good diagnostic accuracy (Komeili et al., 2020; Marinis & Armon-Lotem, 2015; Riches et al., 2010; Taha et al., 2021). Our results suggest that sentence repetition could be a useful tool for identifying Norwegian-speaking children with DLD, and the current study provides some direction as to which grammatical structures may be the most effective in distinguishing between children with DLD and TD.

Regarding types of errors, the children with DLD made the most errors of omission, and their omission errors were almost equally distributed between function words and content words. Conversely, the TD children made almost half the number of omissions of function words than of content words. Thus, the largest effect size was for the difference in omission of function words between groups, although the difference in omission of content words between groups was not much smaller. The omission of function words is interesting as it could indicate a particular weakness in the more grammatical aspects of language, as has been demonstrated for children with DLD in many languages (Håkansson & Hansson, 2000; Jensen de López et al., 2014; Marinis & Armon-Lotem, 2015). Future studies of this kind could examine the contexts in which Norwegian children with DLD omit function words.

Qualitatively, the error profiles of the two groups have many similarities. This adds support to the notion that the difference between children with DLD and TD children is mainly a quantitative, not a qualitative, difference (Leonard, 2014). The results on error profiles show that children with DLD made significantly more inflectional changes than children with TD, which is in line with what was reported for verb inflections in Simonsen and Bjerkan (1998). However, in our study the responses of children with DLD were not characterized by incorrect (non-existing) inflections; rather, they had a greater tendency than the TD children to transform the inflections of verbs and nouns, e.g., by using verbs in a different tense, or nouns in a different number, than in the target sentence.

Our findings thus suggest that omission of tense, or optional infinitive, a proposed clinical marker for DLD in English (Rice & Wexler, 1996), does not appear to be a major problem in all languages. In English, there is feature unification between the subject and the verb, which means that it is difficult to disentangle tense (PT stage 2) from subject-verb agreement (PT stage 4). In Norwegian, tense only marks time, and there is no feature unification between the verb and the subject in terms of person or number. Lexical morphology appears at an early stage according to PT (stage 2). Earlier empirical studies of Norwegian have reported seemingly conflicting evidence relating to the development of tense marking. A study that adapted the LARSP (Ball, Crystal & Fletcher, 2012) into Norwegian (N-LARSP) by Ribu et al. (2019) described tense marking in TD children as an early emerging structure (2;0–2;6 years). In contrast, Simonsen and Bjerkan (1998) found that six-year-olds still produced a substantial number of inflectional errors for verbs in the small regular conjugation class and for irregular verbs. This discrepancy can likely be explained by differences in methodology. N-LARSP documented the emergence of a structure, while Simonson and Bjerkan aimed to find percentages of correct use of past tense markings for verbs belonging to different regular classes and for irregular verbs. Thus, for structures that emerge early but have a protracted development toward full mastery, different methodological approaches may yield diverging results. It should be noted that the target sentences in our study only contained three past tense forms of main verbs, two verbs from the small regular group and one irregular verb. Simonsen and Bjerkan (1998) studied all Norwegian verb groups, and used a task where children had to generate the past tense forms from the present tense form. Thus, it is possible that the children in our sample would have shown more incorrect inflections if a larger set of different verbs had been included, or if the task had required children to generate past tense conjugations from input in the present tense. In sum, our results show a range of different error types in children with DLD, which underscores the importance of considering a broad range of morphosyntactic phenomena when examining language production.

Limitations and Future Directions

The Recalling Sentences task used in the present study was not designed specifically to capture the PT stages, and contained a relatively low number of obligatory contexts that could be used to evaluate the emergence of the two highest PT stages. It is likely that a task encompassing more examples of all the structures predicted by PT would provide a fuller picture of the development of Norwegian. In addition, all tests were administered according to the manuals. This meant that the stopping rule was followed, with the result that children who did not repeat all the target sentences could not be included in the current study. Therefore, the data analyzed were from the 19 children with DLD who

https://doi.org/10.1017/S0305000922000484 Published online by Cambridge University Press

attempted to repeat all the sentences. An additional eight children with DLD, those with the lowest performance, had to be excluded. Thus, the data do not represent the full possible variability in grammatical skills in the DLD group. Still, the differences between the DLD and TD groups were substantial.

We did not attempt to investigate the possible effects of limitations in short term or working memory on language production in this study. According to Riches (2012), specific errors made by children with DLD on sentence repetition tasks are most likely to reflect difficulties with underlying syntactic and semantic knowledge. Syntactic complexity was seen to affect error rate irrespective of sentence length (Riches, 2012). Our underlying assumption is that the errors made in a sentence repetition task are mainly a reflection of the children's overall linguistic abilities, rather than of a memory-specific impairment.

Clinical Implications

The main finding of this study is that it is possible to use a developmental perspective and analyze children with DLD as language learners in the process of acquiring the morphosyntax of Norwegian. The finding that the stages of grammatical development followed an implicational order has clinical ramifications. It means that, after the assessment of children's stages, a clinician can facilitate children's development by focusing on the next stage in the sequence (see Nettelbladt et al., 2007). From an assessment perspective, we found that the Recalling Sentences task functioned well as a starting point for detailed error analyses, as the obligatory contexts created by this task were quite difficult even for TD primary school children, and encouraged them to show their linguistic potential. A structure where the subordinate clause word order differed from main clause word order appeared to be especially sensitive in distinguishing between children with DLD and TD. Thus, items with subordinate clause word order should be considered in the development of clinical language assessments for Norwegian primary school children.

Acknowledgements. We wish to thank the children who participated. Thanks to the educational and psychological counseling services, Statped, and primary schools who assisted us in recruitment of participants. We are grateful to Ona Bø Wie for collaboration on the larger project that these data belong to, and to Ellen Brinchmann, Åsrun Valberg, Marit Gismarvik, Christiane Haukedal, and the master's students at the Department of Special Needs Education, University of Oslo who helped us with data collection and scoring.

Conflict of Interest. The authors have no conflicts of interest to disclose. The manuscript represents original work, and has not been published or submitted for review elsewhere.

Supplementary Materials. To view supplementary material for this article, please visit http://doi.org/10.1017/S0305000922000484.

References

Ball, M., **Crystal, D.**, & **Fletcher, P.** (2012). *Assessing grammar: The languages of LARSP*. Bristol, United Kingdom: Multilingual Matters.

Bishop, D. (2003). Test for Reception of Grammar II. London: Harcourt Assessment.

Bishop, D. V. (2011). CCC-2. The Children's Communication Checklist. *Manual (Norwegian version)*. Stockholm: Pearson Assessment.

Bishop, D. V., **Snowling, M. J.**, **Thompson, P. A.**, **Greenhalgh, T.**, **Catalise-2 Consortium**, **Adams, C.**, … **Boyle, C.** (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of

problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, **58**(10), 1068–1080.

Brandt, S., **Diessel, H.**, & **Tomasello, M.** (2008). The acquisition of German relative clauses: A case study. *Journal of Child Language*, **35**(2), 325–348.

Bresnan, J. (2001). *Lexical-functional syntax*. Malden, MA: Blackwell

Castilla-Earls, A., **Pérez-Leroux, A. T.**, **Restrepo, M. A.**, **Gaille, D.**, & **Chen, Z.** (2018). The complexity of the Spanish subjunctive in bilingual children with SLI. *Language Acquisition*, **25**(1), 72–84.

De Camp, D. (1971). Implicational scales and sociolinguistic linearity. *Linguistics*, **9**, 30–43.

Devescovi, A., & **Caselli, M. C.** (2007). Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language and Communication Disorders*, **42**(2), 187–208.

Dunn, L. M., **Dunn, L. M.**, **Whetton, C.**, & **Berley, J.** (1997). *The British Picture Vocabulary Scale II*. Windsor, United Kingdom: GL Assessment.

Dyson, B., & **Håkansson, G.** (2017). *Understanding second language processing: A focus on Processability Theory*. Amsterdam: John Benjamins.

Dyson, B., **Håkansson, G.**, & **Ballard, K. J.** (2022). A Developmental Approach to Assessing and Treating Agrammatic Aphasia. *American Journal of Speech-Language Pathology*, **31**(3), 1188 – 1204.

Ebbels, S. (2014). Effectiveness of intervention for grammar in school-aged children with primary language impairments: A review of the evidence. *Child Language Teaching and Therapy*, **30**, 7–40.

Engdahl, B., **Tambs, K.**, **Borchgrevink, H. M.**, & **Hoffman, H. J.** (2005). Otoacoustic emissions in the general adult population of Nord-Trøndelag, Norway: III: Relationships with pure-tone hearing thresholds. *International Journal of Audiology*, **44**, 15–23.

Fletcher, P., **Leonard, L.**, **Stokes, S.**, & **Wong, A. M.-Y.** (2005). The expression of aspect in Cantonese-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **48**, 621–634.

Friedmann, N., **Belletti, A.**, & **Rizzi, L.** (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, **119**, 67–88

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.

Frizelle, P., **Thompson, P.**, **Duta, M.**, & **Bishop, D.** (2019). Assessing children's understanding of complex syntax: A comparison of two methods. *Language Learning*, **69**(2), 255–291.

Frizelle, P., **Thompson, P. A.**, **McDonald, D.**, & **Bishop, D. V.** (2018). Growth in syntactic complexity between four years and adulthood: evidence from a narrative task. *Journal of Child Language*, **45**(5), 1174–1197.

Furnes, B., & **Samuelsson, S.** (2009). Preschool cognitive and language skills predicting kindergarten and grade 1 reading and spelling: A cross-linguistic comparison. *Journal of Research in Reading*, **32**(3), 275–292.

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

Guasti, M. T., **Stavrakaki, S.**, & **Arioso, F.** (2012). Cross-linguistic differences and similarities in the acquisition of relative clauses: Evidence from Greek and Italian. *Lingua*, **122**(6), 700–713.

Håkansson, G. (1989). The acquisition of negative placement in Swedish. *Studia Linguistica*, **43**, 47–58.

Håkansson, G. (2001). Tense morphology and verb-second in Swedish L1 children, L2 children and children with SLI. *Bilingualism: Language and Cognition*, **4**, 85–99.

Håkansson, G. (2017). Typological and developmental considerations on specific language impairment in monolingual and bilingual children: A Processability Theory account. *Language Acquisition*, **24**(3), 265–280.

Håkansson, G., & **Hansson, K.** (2000). Comprehension and production of relative clauses: A comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, **27**, 313–33.

Håkansson, G., **Salameh, E.-K.**, & **Nettelbladt, U.** (2003). Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics*, **41**, 255 – 288.

Hatch, E., & **Lazaraton, A.** (1991). *The research manual: Design and statistics for applied linguistics*. Rowley, MA: Newbury House.

Helland, W. A., **Biringer, E.**, **Helland, T.**, & **Heimann, M.** (2009). The usability of a Norwegian adaptation of the Children's Communication Checklist Second Edition (CCC-2) in differentiating between language

impaired and non-language impaired 6- to 12-year-olds. *Scandinavian Journal of Psychology*, **50**(3), 287–292.

Jakubowics, C., & **Nash, L.** (2001). Functional categories and syntactic operations in (ab)normal language acquisition. *Brain and Language*, **77**, 321–339.

Jensen de López, K., **Sundahl Olsen, L.**, & **Chondrogianni, V.** (2014). Annoying Danish Relatives: Comprehension and production of relative clauses by Danish children with and without SLI. *Journal of Child Language*, **41**(1), 51–83.

Keenan, E. L., & **Comrie, B.** (1977). Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry*, **8**(1), 63–99.

Komeili, M., **Marinis, T.**, **Tavakoli, P.**, & **Kazemi, Y.** (2020). Sentence repetition in Farsi-English bilingual children. *Journal of the European Second Language Acquisition*, **4**(1), 1–12.

Kristoffersen, K. E., **Rygvold, A.-L.**, **Klem, M.**, **Valand, S. B.**, **Asbjørnsen, A.**, & **Næss, K.-A.** (2021). Terminologi for vansker med språk hos barn og unge. En konsensusstudie. [Terminology for difficulties with language in children and adolescents. *A consensus study.] Norsk tidsskrift for logopedi [Norwegian Journal of Speech Language Pathology*), **3**, 6–23.

Leonard, L. B. (2014). Specific language impairment across languages. *Child Development Perspectives*, **8**(1), 1–5.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge. Massachusetts: MIT Press.

Lombardi, L., & **Potter, M. C.** (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language*, **31**(6), 713–733.

Lukács, A., **Leonard, L. B.**, **Kas, B.**, & **Pléh, C.** (2010). The use of tense and agreement by Hungarian-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*, **52**, 98–117.

Lyster, S.-A. H., **Horn, E.**, & **Rygvold, A.-L.** (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av British Picture Vocabulary Scale, Second Edition. [Vocabulary and vocabulary development in Norwegian-speaking children and adolescents. Results from a study using the British Picture Vocabulary Scale, Second Edition.] *Spesialpedagogikk*, **9**, 35–43.

Marinis, T., & **Armon-Lotem, S.** (2015). Sentence repetition. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children. Disentangling bilingualism from language impairment* (pp. 95–121). Bristol, United Kingdom: Multilingual Matters.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, **22**(3), 276–282.

Meir, N., **Walters, J.**, & **Armon-Lotem, S.** (2015). Disentangling SLI and bilingualism using sentence repetition tasks: The impact of L1 and L2 properties. *International Journal of Bilingualism*, **20**(4), 421–452.

Nettelbladt, U., **Håkansson, G.**, & **Salameh, E.-K.** (2007). Specifik intervention av fonologi, grammatik och lexikon [Specific intervention for phonology, grammar, and lexicon]. In U. Nettelbladt, & E-K. Salameh (Eds.), Språkutveckling och språkstörning hos barn [*Language development and language disorders in children*] (pp. 289–309). Lund, Sweden: Studentlitteratur.

Novogrodsky, R., & **Friedmann, N.** (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *Advances in Speech-Language Pathology*, **8**(4), 364–375.

Pienemann, M. (1998). *Language processing and second language development: Processability Theory*. Amsterdam: John Benjamins Publishing.

Potter, M. C., & **Lombardi, L.** (1990). Regeneration in the Short-Term Recall of Sentences. *Journal of Memory and Language*, **29**, 633–654.

Raven, J., & **Court, J.** (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Ribu, I. S., **Simonsen, H. G.**, **Løver, M. A.**, **Strand, B-M. S.**, & **Kristoffersen, K. E** (2019). N-LARSP: A developmental language profile for Norwegian. In M.J. Ball, P. Fletcher, D. Crystal. (eds) *Grammatical Profiles: Further languages of LARSP*. (pp. 1–48) Bristol, UK: Multilingual Matters

Rice, M., & **Wexler, K.** (1996). Toward tense as a clinical marker of specific language impairment. *Journal of Speech, Language and Hearing Research*, **39**, 1236–1257.

Riches, N. G. (2012). Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language, & Communication Disorders*, **47**, 499–510

Riches, N. G., **Loucas, T.**, **Baird, G.**, **Charman, T.**, & **Simonoff, E.** (2010). Sentence repetition in adolescents with specific language impairment and autism: An investigation of complex syntax. *International Journal of Language, & Communication Disorders*, **45**, 47–60.

Ringstad, T., & **Kush, D.** (2021). Learning embedded verb placement in Norwegian: Evidence for early overgeneralization. *Language Acquisition*, **28**(4), 411–432.

Salameh, E.-K., **Håkansson, G.**, & **Nettelbladt, U.** (2004). Developmental perspectives on bilingual Swedish-Arabic children with and without language impairment: a longitudinal study. *International Journal of Language, & Communication Disorders*, **1**, 65–71.

Sanz-Torrent, M., **Serrat, E.**, **Andreu, L.**, & **Serra, M.** (2008). Verb morphology in Catalan and Spanish in children with Specific Language Impairment: a developmental study. *Clinical Linguistics, & Phonetics*, **22**(6), 459–474.

Semel, E., **Wiig, E. H.**, & **Secord, W. A.** (2003). Clinical evaluation of language fundamentals (4th ed.). San Antonio, TX: Harcourt Assessment.

Simonsen, H. G., & **Bjerkan, K. M.** (1998). Testing past tense inflection in Norwegian: a diagnostic tool for identifying SLI children. *International Journal of Applied Linguistics*, **8**(2), 251–270.

Taha, J., **Stojanovik, V.**, & **Pagnamenta, E.** (2021). Sentence repetition as a clinical marker of Developmental Language Disorder: Evidence from Arabic. *Journal of Speech, Language, and Hearing Research*, 1–24 (advance online publication).

Theodorou, E., **Kambanaros, M.**, & **Grohmann, K. K.** (2017). Sentence repetition as a tool for screening morphosyntactic abilities of bilectal children with SLI. *Frontiers in Psychology*, **8**, 2104.

Thomas, M. S. C. (2011). Constraints on language development. In P. Fletcher, & J. F. Miller (Eds.), *Developmental theory and language disorders* (pp. 11–34). Amsterdam, Netherlands: John Benjamins.

Tiwari, S., **Karanth, P.**, & **Rajashekar, B.** (2017). Specific language impairment in a morphologically complex agglutinative Indian language – Kannada. *Journal of Communication Disorders*, **66**, 22–39.

Torkildsen, J. v. K., **Hitchins, A.**, **Myhrum, M.**, & **Wie, O. B.** (2019). Speech-in-noise perception in children with cochlear implants, hearing aids, developmental language disorder and typical development: The effects of linguistic and cognitive abilities. *Frontiers in Psychology*, **10**, 2530.

Van der Lely, H. K. (1998). SLI in children: Movement, economy, and deficits in the computational-syntactic system. *Language acquisition*, **7**(2-4), 161–192.

van der Lely, H. K., & **Ullmann, M.** (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, **16**(29), 177–217.

Vinther, T. (2002). Elicited imitation: a brief overview. *International Journal of Applied Linguistics*, **12**(1), 54–73.

Westergaard, M. R., & **Bentzen, K.** (2007). The (non-) effect of input frequency on the acquisition of word order in Norwegian embedded clauses. In I. Gülzow, & N. Gagarina (Eds.), *Frequency effects in language acquisition: defining the limits of frequency as an explanatory concept*. (pp. 271–306) Berlin: Mouton de Gruyter

Cite this article: Håkansson G., Williams E.W., Karlsen J., & Torkildsen J.v.K. (2022). What Characterizes the Productive Morphosyntax of Norwegian Children with Developmental Language Disorder? *Journal of Child Language* 1–24, https://doi.org/10.1017/S0305000922000484