Elicited and Spontaneous Determiner Phrase Production in French-Speaking Children With Developmental Language Disorder

Phaedra Royle

Université de Montréal, Montréal, QC, CANADA

Centre for Research on Language Mind and Brain, Montréal, QC, CANADA

International Laboratory for Brain, Music and Sound Research, Montréal, QC, CANADA

Lena Reising

Rheinisch-Westfälische Technische Hochschule, Aachen, GERMANY

Authors' Note

Correspondence concerning this article should be addressed to Phaedra Royle, Université de Montréal, Faculté de médecine, École d'orthophonie et d'audiologie, C.P. 6128, Succursale Centre-Ville, Montréal, QC, Canada, H3C 3J7. Email: phaedra.royle@umontreal.ca

Acknowledgments

This research was funded in part by an "*Établissement de nouveaux professeurs-chercheur*s" grant from the Fond de recherche du Québec – Société et culture to Royle (FQRSC, 2006-NP-104790).

We thank the participants and their families and Judith Boisvert, Josée Bouffard, Nicolas Bourguignon, Isabelle Delorme, Layal Ghazlé, Isabelle Stine, and Jacinthe Toupin for their assistance in data collection and analysis.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

Abstract

We contrast elicitation and spontaneous speech data in French-speaking children with developmental language disorder and controls, with a focus on determiner phrase and gender agreement. Eight French-speaking children with developmental language disorder and age-matched or language-matched controls were compared on an elicitation task for complex noun-phrases with one or two adjectives (e.g., *la petite maison verte* 'the small green house') and a spontaneous speech sample of 200 utterances containing determiner phrases. Elicitation and spontaneous speech data revealed different profiles in French children with developmental language disorder compared to controls: elicitation tasks revealed specific difficulties with adjective agreement as well as high levels of global error, while spontaneous speech revealed mostly omission and substitution errors, often on determiners. Ultimately, both approaches to evaluating language abilities are complementary, but elicitation tasks might be the most useful tool for rapid identification of difficulties with determiner phrases and agreement in young French-speaking children.

Keywords: Developmental language disorder, French, determiner phrase, gender agreement, elicitation, spontaneous speech

Developmental language disorder (DLD; Bishop, Snowling, Thompson, Greenhalgh, & CATALISE-2 consortium, 2017; in Québec, *trouble développemental du langage*) is found in approximately 7% of the population in Québec, Canada (Ordre des orthophonistes et audiologistes du Québec, 2014). DLD describes children who exhibit persistent language difficulties in the absence of obvious neurological impairment, social deprivation, or low cognitive skills (Leonard, 2014; Ordre des orthophonistes et audiologistes du Québec, 2004). This disorder can impair the lexicon, phonology, morphology, syntax, morphosyntax—for example, accusative clitics (Grüter, 2005), gender agreement (Roulet-Amiot & Jakubowicz, 2006), and verb inflection (Royle, St-Denis, Mazzocca, & Marquis, 2017)—and pragmatics in the expressive or receptive spheres (Leonard, 2014). However, French-speaking children with DLD are somewhat difficult to identify because in the preschool years they show almost no omission or substitution errors in spontaneous speech (Elin T. Thordardottir & Namazi, 2007), but can have reduced syntactic structure repertoires as compared to age-matched controls (e.g., in French noun-phrases; Royle & Stine, 2013). The relative lack of French indicators for this disorder has led to a search for specific markers for DLD in French. Our focus is on grammatical gender agreement and noun phrase structures, or *determiner phrases* (DPs) because they are acquired early and reveal language production difficulties in children with DLD.

French Agreement and Acquisition of the Determiner Phrase

French exhibits subject-verb, number, and person agreement as well as gender agreement, which is especially prevalent in the DP. Masculine is the default gender but French contains equal numbers of masculine and feminine nouns. Many elements in French, such as determiners, relative pronouns, pronoun clitics, adjectives, and other structures agree with the noun with

which they are co-referential, either in gender, in number, or both. However, agreement processes are not always morphologically transparent. For example, pronouns and determiners have plural forms that are underspecified for gender (e.g., *les* def.pl 'the', the accusative clitic *les* 'them', or *leur* dat.cl.pl 'to them') and many adjectives and past participles have invariable forms: compare variable *brun/e* [bʁœ̃/bʁyn] 'brown.m/f' to invariable *jaune* [ʒɔn] 'yellow,' two of our stimulus items.

Longitudinal and cross-sectional corpora showed that typically developing French children produce their first variable feminine adjectives between ages 1;8 and 2;4, and all children show strong mastery of determiner agreement with mean scores of 96% (Valois & Royle, 2009; Valois, Royle, Sutton, & Bourdua-Roy, 2009). Few variable adjectives, however, were found in these corpora—only eight different lemmas were found in the first, and five in the second. Children can produce agreement in elicitation tasks between ages 4 and 7, depending on the structure. Determiner agreement is reliably elicited before age 5, while adjective agreement becomes stable around age 6 or later (Roulet-Amiot & Jakubowicz, 2006). This last phenomenon might be linked to item-based lexical learning (Royle & Valois, 2010) as some young typically developing children (aged 3;0 and 3;4) are unable to comprehend specific adjectives' feminine forms when asked to provide, e.g., a green frog, *Donne moi la grenouille verte* [vɛʁt] 'Give me the.f frog green.f'; response, *Il n'y en a pas* 'There are none', even though they understand the masculine form *vert* [vɛʁ] (Royle & Valois, 2010). Brain imaging data show that neurotypical French adults and children aged 7 or 8 years process agreement in similar ways (Fromont, Royle, & Steinhauer, 2015).

Gender Agreement and Determiner Phrases in Children With Developmental Language Disorder

Some researchers (i.e., Rice & Oetting, 1993; van der Lely, 1998) have proposed that DP agreement is a relative strength in English-speaking children with DLD. However, children with DLD have been found to show agreement difficulties in many languages including Dutch, English, Icelandic, Italian, Swedish (see Leonard, 2016, for a short review), and French (Gopnik, 1990; Parisse & Maillart, 2007; Pizzioli & Schelstraete, 2008). Further, determiner substitutions and omissions have been found in Spanish (Bedore & Leonard, 2001) and Swedish (Hansson, Nettelbladt, & Leonard, 2003). Difficulties with genitive (e.g., *hundens mat* 'the dog's food'), plural inflection, indefinite articles (e.g., *en hund* 'a dog'), and determiner + adjective + noun structures (e.g., *dom smutsiga flickorna* 'the dirty girls') are also observed in Swedish (Leonard, Salameh, & Hansson, 2001). In English, children with DLD have difficulties with DP syntax, including comprehending specificity marked by definite and indefinite determiners or genitive structures such as *that/those bear's/bears' balloon/s* (Ramos, 2000). Russian-speaking children with DLD up to age 10 have been shown to exhibit persistent gender agreement errors on adjectives (Tribushinina & Dubinkina, 2012), while Rakhlin, Kornilov, and Grigorenko (2014) observed difficulties in children with DLD up to age 15 in judging feminine Russian gender, as opposed to masculine gender, which is the default.

However, a coherent picture of determiner adjective and preposition use in the noun phrase has not emerged from the literature on French. For example, in a 20-minute spontaneous-speech sample, Le Normand, Leonard, and McGregor (1993) observed no differences in determiner omissions between a mean length of utterance (MLU) matched group (aged 3–3;6) and French-speaking children with language impairment (aged 4–6 years) on a small number of

utterances (*M* = 29). On the contrary, Parisse and Maillart (2007) observed a higher percentage of syntactic structure errors with determiners and prepositions in four French-speaking children with DLD aged 3–7 years as compared to MLU-matched children, based on short spontaneous-speech samples collected over two to four 20-minutes encounters. However, no error description is presented.

Gopnik (1990) observed errors on determiners and adjectives in a corpus of 70 utterances for a 9-year-old French-English bilingual child with DLD. When describing pictures, he alternated *le* and *les* for plural targets (**le/les marshmallows* 'the.sg/pl marshmallows'), omitted determiners (*Après tire fleur* 'After [he] throws [a/the] flower'), and made gender errors on determiners and adjectives in spontaneous speech (*I1 prend *un *gros respiration* 'He takes a.m big.m breath', breath is feminine in French). In contrast, using 45-minute spontaneous speech samples—the number of utterances is not specified—Paradis and Crago (2001) observed no difficulties in spontaneous production of determiners and prepositions among 10 French-speaking children with DLD (average age: 7;6) and their age-matched controls, and among their MLU-matched controls (average age: 3;3). Pizzioli and Schelstraete (2008) computed determiner omissions in a sentence-production task focusing on verb argument structure with participants aged 8;1 to 13;0, controls matched for age, and controls matched on sentence production abilities (aged 5;6–6;4). They found that omissions were higher in the DLD group—5–17%, depending on conditions, exact numbers are not provided—than language-matched (2.5%) and age-matched control groups (no errors). Roulet-Amiot and Jakubowicz (2006) elicited adjective production in French children with DLD aged 6;10 to 12;6 and control groups aged 4 and 6 years. In their sentence completion task (e.g., Here the elephant is wearing __ [a green shirt]), children with

DLD showed higher error rates (*M* = 27%), mostly on feminine adjectives, than both control groups (4-year-olds: *M* = 19%; 6-year-olds: *M* = 5%).

These authors also evaluated "comprehension" errors using a semantic categorization task (e.g., Can you eat that?) with auditory presentation of grammatically correct and incorrect DPs (e.g., *un *nouvelle balai* 'a.m new.f broom.m). Despite globally higher categorization error rates in participants with DLD (i.e., 9.5% versus 5.8% in 6-year-olds and 4.1% in adults), no ungrammaticality or slowing effects were found in children with DLD faced with ungrammatical structures. The authors concluded that these children are sensitive to gender agreement during language comprehension and that it is only the extra processing load involved in producing the structures that causes errors. However, it is unclear whether the task was appropriate for this question; at least for these types of structures, the comprehension task could easily be carried out without agreement checking. Furthermore, the absence of reaction time differences between grammatical and ungrammatical structures could point to an *insensitivity* to gender agreement errors in children with DLD. Finally, as far as we can tell, the authors did not control stimuli for any psycholinguistic factors such as age of acquisition or word-internal morphological structure.

What can we conclude from the mixed results presented above? It appears that French-speaking children can exhibit at least some level of error on agreement in determiners and adjective production and do not always resolve these issues, as shown by studies including children aged 6–13 years. The difficulties appear to be more marked on elicitation tasks, but are occasionally also observed in longer spontaneous speech corpora, and this might depend on their age. We believe that elicited agreement is a fruitful domain for study in French child language, as spontaneous production in young children does not systematically distinguish preschool French-speaking children with DLD from their peers on morphosyntactic abilities. For example,

Elin T. Thordardottir and Namazi, (2007) found virtually no differences on a number of grammatical morphemes they studied. Spontaneous speech corpora might therefore over- or under-evaluate their language abilities (Eisenbeiss, 2011; Royle & Elin Thordardottir, 2008; Steel, Rose, Eadie, & Thornton, 2013). Further, because of low adjective variety in production even when children appear to master them (Royle & Valois, 2010), elicitation might be a more efficient approach to study them. We thus used the DP elicitation paradigm developed by Royle and Valois (2010) as well as spontaneous corpus analyses to establish the usefulness of either approach for the identification of agreement difficulties in French and other difficulties that might arise in the DP structure (e.g., syntax, determiner production, and omission of obligatory elements).

Current Study

This study compared two approaches to the investigation of noun-phrase structures in French. The first, an elicitation task using puzzles, was used to evaluate the production and comprehension of adjective and determiner agreement, as well as syntactic structure—mainly Adj-Noun or Noun-Adj order—in simple (e.g., *La maison brune* 'The.f. house brown.f.') and slightly more complex noun-phrases (e.g., *La grande maison brune* 'The.f big.f. brown.f house'). The second approach was used to evaluate noun-phrase production in the spontaneous speech corpus. This approach allowed for the comparison of similar structures across both tasks, but also other types of DP structures not elicited with the puzzle props, which emerged in the corpus.

The major research question was whether children with DLD have difficulties producing DPs with appropriate agreement and other syntactic and morphosyntactic structures that are common in them, especially word-order and prepositional structures. A secondary question was

to compare elicitation and spontaneous speech tasks for the identification of difficulties in DP structures in children with DLD.

Because younger French-speaking children with DLD have not systematically shown obvious difficulties producing adjective agreement or other morphosyntactic structures in spontaneous speech, but have shown more difficulties in controlled elicitation contexts, we expected elicitation tasks to highlight agreement processing difficulties on adjectives and determiners, and difficulties with DP structure, in children with DLD. We expected spontaneous speech corpora to reveal difficulties less saliently, more specifically on adjectives because adjectives are rare and not varied in spontaneous speech corpora for typically developing French-speaking children (see Valois & Royle, 2009, and Valois et al., 2009, for transversal and longitudinal corpus analyses), and because determiner errors (omissions and commissions) were not consistently found in the studies reviewed above. However, with more variety in structures produced (e.g., determiners, prepositions, adjectives, complex nouns phrases with expansions), the corpus has the potential to provide us with interesting data that go beyond the constrained set of items used in controlled experiments. We thus expected the corpus to reveal other difficulties children with DLD might have with DP structures.

Method

Participants

Participants were residents of the province of Québec, Canada, had French-speaking parents, and were exposed to French at least 80% of the time at home or in daycare. They had no history of autism, neurological disorder, hearing loss, learning disabilities, or other medical conditions that could impair language development. Participant characteristics are presented in

Table 1. Parents signed a consent form for their child's participation. All children gave oral consent to participate.

Insert Table 1 here

Children with Developmental Language Disorder. Nine monolingual French-speaking children from Montréal and the Sherbrooke region (three girls and six boys, two of which were twins: D5 and D6) were recruited through specialized services for children with language disorders (i.e., a research hospital and a summer camp). They had been labelled as having *dysphasie* (the term for DLD at the time) by a certified speech-language pathologist, between 9 months and up to 3 years prior to testing. They had maintained their diagnosis as determined by response to intervention. Based on parental questionnaires, all children were monolingual and raised in monolingual environments. One child was excluded because he was significantly older than the controls. The remaining eight were between 5 and 6 years of age (M = 5;7 years, SD = 0;4) and had a mean length of utterance in words (MLUw) of 4.4 (SD = 0.97).

Control groups. In total, 25 control participants were included in this study. Two of our four puzzles (puzzles 3 and 4, see below) had A and B versions. Control children only saw one of the boards for these two puzzles (only A or B) in order to mitigate potential fatigue. We realized that task fatigue was not an issue and all but two children with DLD were presented with the complete set of stimuli (D2 and D4 saw only the A versions). Thus, control matching for puzzle analyses necessitated double pairings. Each child with DLD was matched with two typically developing peers, having done either A or B versions of tasks 3 and 4, on biological sex, parental education, non-verbal memory IQ, and either (a) age: the AGE group, matched for

age within two months always to the advantage of children with DLD, or (b) MLUw: the MLU group based on a sample of 100 utterances (a subset of the 200 utterances used for the corpus analysis). All control children were from a larger pool (N = 151) of children recruited from the greater Montréal area. Because two controls were needed for every pairing on the puzzle tasks, but only one for the spontaneous-speech sample, half of the children matched for the task were randomly selected for the spontaneous speech sample comparisons (details are provided in Table 1). As matching child-by-child made it somewhat difficult to find perfectly matching pairs on all measures, even among our large sample of controls, we reused data from some control children (e.g., N114) as they matched with more than one child with DLD. In total, 12 AGE controls (8 in spontaneous speech, marked with ✧ in Table 1), and 13 MLU controls (8 in spontaneous speech) were chosen. No significant differences were found between DLD and control groups on parental education or MLUw scores (see Elin Thordardottir et al., 2005, for similar findings). AGE-matched controls did not differ with the DLD group on age, but the MLU-matched group did. Significant differences were observed between DLD and both AGE-matched and MLU-matched control groups on receptive vocabulary scores, except for the MLU-matched sub-group in the corpus analysis who nonetheless showed a trend for higher scores than the DLD group (p = .076). Memory IQ scores were not significantly different in groups matched for the puzzle comparisons, but were significantly higher in both MLU- and AGE-matched groups for the corpus comparison.[1, 2]

[1]This might be due to the fact that the (presumably) non-verbal Leiter screen loads heavily on linguistic abilities: it uses cards with visual cues such as a crib, a plate, a boot, etc., that children must reorganize in specific sequences. See also Miller and Gilbert (2008) for a comparison of two non-verbal intelligence tasks and DLD.

[2]Note that one child with DLD (D9) was non-compliant during the IQ task and was an outlier that significantly reduced group scores. Excluding this child and her controls for the comparison

Materials

We used four puzzles containing images varying in size or colour to elicit targeted adjectives and nouns (see Royle & Valois, 2010, for details). Participants had to name the pieces they wanted to manipulate. The first puzzle tested colour adjective vocabulary: six colour dots with no gender information, half variable and half invariable, usually produced in the default masculine, although this was not scored for target production. The second depicted eight items: four nouns (half masculine) varying in size, which forced appropriate use of gender-marked determiners and adjectives *small* and *big*, both variable. The third contained four nouns that varied in colour, with 24 items over two versions, A and B, half with variable adjectives (e.g., *the brown duck* vs. *the white duck*) forcing the use of gender-marked determiners and adjectives. The fourth puzzle tested DPs with both colour and size characteristics and contained 48 items over two versions, A and B, using all adjectives and nouns from previous puzzles, all forcing the use of gender-marked determiners and adjectives. Each A and B version of the second and third puzzles contained half of the stimulus items to minimize task fatigue. Appendix A includes the full set of items for version A and Appendix B presents examples of visual stimuli. Images were printed on the backgrounds of puzzle boards and on insert wood pieces. All adjectives and nouns used were early-acquired and high-frequency in oral French based on oral language frequencies in New, Pallier, Ferrand, and Matos (2001) and age-of-acquisition norms from Trudeau, Frank, and Poulin-Dubois (1999; see Royle & Valois, 2010, for details).

Procedure

show that there are significant differences in memory IQ between AGE-matched and DLD groups (p = .037), but that the comparison between MLU-matched and DLD groups is no longer significant but shows a trend (p = .056).

All participants took part in two 1.5-hour experimental sessions within 2 weeks, during which they completed a hearing screening, the *Échelle de vocabulaire en images Peabody*, a receptive vocabulary assessment (Dunn, Thériault-Whalen, & Dunn, 1993), the Leiter Memory Screen (Associated Pairs and Forward Memory subtests; Roid & Miller, 1996), the puzzles, and an audio-video recorded spontaneous-speech sample of the child playing with a standard set of toys (house, circus-train, plasticine, bus). The language samples were used to evaluate structures of interest and to establish MLU scores. Non-directive speech was used to interact with the children while they played. A hearing threshold in both ears above 20 at 500 Hz, or 15 at 1000, 2000, and 4000 Hz resulted in the child being excluded from the experiment. All tasks were run by trained researchers most often in a sound-proof recording lab equipped with an observation post for parents. This took place at the *Centre de recherche Marie-Enfant,* Montréal, Québec. Some children with DLD were tested at their day camp in Sherbrooke, Québec, over 2 days in a quiet room.

During the puzzle barrier tasks, the experimenter installed a screen between herself and the child to avoid pointing. Puzzle boards were presented sequentially in ascending order of difficulty. Two practice stimuli were presented to the child before each puzzle, with target models provided orally (e.g., “If you wanted this piece, you could say ‘give me the pink horse.’ I am behind the screen and cannot see the puzzle so you must tell me with words which piece you want”). The child was then encouraged to ask the experimenter for the piece he or she wanted to put on the puzzle. If he or she successfully performed the first puzzle, the experimenter moved on to the second one, and so on. During the second meeting, a comprehension task was run using the same materials and procedure while reversing experimenter and child roles. All procedures were approved by the *Centre de recherche CHU Ste-Justine* Institutional Review Board (#2167)

and the *Université de Montréal Comité d'éthique à la recherche en santé* (#14-034-CERES-P(2)) committees.

Analyses

Puzzles. Target responses and error types were noted. Responses were scored in three ways. The first scored the full appropriate (target) response, which in the first puzzle was simply naming the colour—extra information, such as producing a determiner, was not scored although it was noted. For the other three puzzles, a point was given when participants produced the full DP including the determiner, the adjective(s) and the noun, as well as appropriate gender and syntactic structure. A second score was given for target feminine adjectives only (correct colour or size and gender) irrespective of the DP structure. A third qualitative analysis was performed on error types when the target DP was not produced. In these, only errors observed at least five times in each group were statistically analyzed.

Colour and size adjectives were analyzed separately because of their different properties, linguistic distributions, and age of acquisition. An important difference between these types is that colour adjectives are both variable and invariable in French, while size adjectives are always variable. Size adjectives are more frequent and are acquired slightly earlier, but are less varied, than colour adjectives (Royle & Valois, 2010). Colour adjectives are canonically post-nominal while size ones are pre-nominal (a typologically unusual position). Their different positions in the DP are considered to be linked to movement properties and features in the grammar of French (Valois, 1991).

Due to the small sample size, target and non-target responses were entered into non-parametric Kruskal-Wallis analyses with group as a between-subjects factor (DLD, AGE, MLU).

Post hoc Mann-Whitney tests with Bonferroni corrections were used to directly compare groups when relevant. Analyses on effects of gender (masculine vs. feminine) or variability (variable vs. invariable) were run when adjectives in the puzzle bore these properties. Because comprehension levels were globally high in all children—the range in children with DLD was 96.9%–98.4%, depending on the puzzle—these data were not analyzed further.

Spontaneous speech. For each child, 200 utterances were examined to provide a reliable picture of his or her linguistic behaviour. Utterances were coded using the Systematic Analysis of Language Transcripts program (Miller & Chapman, 1984–2002; adapted for French by Elin T. Thordardottir, 2005). The first transcription was conducted by the research assistant who had administered the task. Each was checked by a second native speaker, then coded for morphosyntax and checked again. All disagreements were resolved by consensus or coded as incomprehensible (< 1%). The entire DLD corpus was re-transcribed by the first author. Inter-transcriber agreement was 96.6% for words and 97.9% for morphemes.

Between 71 and 171 DPs were produced by each child (M = 127, SD = 29.7), with a total of 3049 DPs included in the analysis. DPs containing unintelligible elements were excluded from the analysis (DPs per child: M = 3.2, SD = 2.3, Min. = 0, Max. = 10). Preposition use was considered when it occurred DP-initially in the form of complex contracted determiners (e.g., *du* and *de la* 'of-the.m/f'). Due to the small sample size and heterogeneous error distributions, Kruskal-Wallis analyses with group as a between-subjects factor (DLD, AGE, MLU) were used. Post hoc Mann-Whitney tests, with Bonferroni corrections for multiple comparisons, were applied to compare groups when relevant.

Results

Puzzles: Target Responses on Full Determiner Phrases and Variable Feminine Adjectives

Puzzles 1 and 2: Colour naming and size determiner phrases. Global results for target responses on each puzzle are reported in Table 2. Analyses reveal no differences among groups for both colour naming ($H = 3.79$, $p > .10$) and size DP production ($H = 1.16$, $p > .10$). As can be seen in Table 2, all groups showed high target response levels, with more variability in the MLU group.

Insert table 2 here

Analyses on target adjectives with feminine forms in size DPs (i.e., *grande/grosse,* 'big', or *petite,* 'small', $n = 4$) again revealed no group differences ($H = 0.8$, $p > .10$), with similar median and mean responses for the three groups. Table 3 provides details on target feminine adjective responses for tasks 2 and 4.

Insert Table 3 here

Puzzle 3: Colour determiner phrases. Because two children with DLD did not complete both A and B versions of this puzzle, we tested whether both versions were equivalent in controls. A *t*-test assuming unequal variance comparing the mean responses on version A ($M = 4.81$, $SD = 0.75$) and B ($M = 4.75$, $SD = 1.66$) of the control subgroups showed no significant differences, $t(14.39) = 0.12$, $p > .10$.

Following this, total target DP production scores were computed for 12 items by doubling the score on six items in order to include participants (DLD and controls) who did only task A. Analyses revealed a significant effect of group on target responses ($H = 8.74$, $p = .013$; see Table

2). Post hoc comparisons show that differences between the AGE and DLD groups reached significance with an adjusted alpha of .025 ($U = 52$, $z = 2.05$, $p = .020$, $d = 0.95$). No differences were found between the MLU and DLD groups, $U = 24$, $z = 0.79$, $p > .10$ (see Table 2). Because there were few feminine variable items in this puzzle (three if the participant did both A and B versions), we did not further analyze these items.

Puzzle 4: Size and colour determiner phrases. Again, we checked whether both versions of the puzzle were equivalent. A *t*-test assuming equal variance comparing controls' mean responses on version A ($M = 8.00$, $SD = 2.50$) and B ($M = 9.58$, $SD = 2.64$) showed no significant differences, $t(26) = 1.62$, $p > .10$.

Following this, total target production scores on 24 DPs were computed for analyses (see Table 2) by doubling the score on 12 items in order to include participants (DLD and controls) who only did task A. Analyses reveal a significant effect of group on target responses ($H = 7.27$, $p = .003$). Post hoc comparisons, with an adjusted alpha of .025, revealed a significant difference between the AGE and DLD groups ($U = 53$, $z = 2.15$, $p = .002$, $d = 1.22$), but not between MLU and DLD groups ($U = 0.21$, $z = 2.57$, $p > .10$). A comparison on feminine variable adjective targets ($n = 18$) transformed to a score over nine in order to include participants who did only task A, again revealed a significant effect of group on target responses ($H = 13.21$, $p = .001$; see Table 3). Post hoc tests revealed a significant difference between the DLD and AGE groups ($U = 61.5$, $z = 3.05$, $p = .001$, $d = 2.13$). The comparison between the DLD and MLU groups did not reach significance ($U = 46.5$, $z = 1.47$, $p = .07$).

Qualitative analyses of non-target responses on puzzles. In this section, we provide a detailed analysis of errors and non-target responses found in the elicitation task as they were expected to reveal different production strategies in the three groups. These responses were collated from the three puzzles involving DPs (puzzles 2, 3, and 4). The two children with DLD who did not perform both versions of puzzles 3 and 4 had their non-target responses on these tasks multiplied by two, while errors for matched-control pairs of children on A and B versions were added together to provide representative numbers of errors for a child who had done both versions. A total of 611 errors or non-target productions were observed across groups (DLD = 274, AGE = 87, MLU = 250). Only errors observed more than five times in a given group were considered for the analyses. These were gender errors on size or colour adjectives (e.g., *La grenouille *vert* 'the.f frog green.m') or determiners (e.g., **La petit bateau blanc* 'the.f white.m small.m, a.m boat'), determiner omissions (e.g., __ *canard rouge* '__ duck red'), and splitting the DP (e.g., *Le petit bateau blanc* 'the small white boat' → *Le blanc petit, un bateau* 'the.m white.m small.m, a.m boat'.[3,4]

Table 4 presents detailed information about observations. Kruskal-Wallis analyses with Group as a between-subjects factor (DLD, AGE, MLU) for gender errors on colour adjectives revealed a significant main effect of Group ($H = 10.39$, $p = .006$). Children with DLD produced significantly more gender errors ($n = 62$) on colour adjectives than the AGE group ($n = 20$; $U = 4$, $z = 2.89$, $p = .002$, $d = 2.15$), but only showed a trend for a difference with the MLU group (n

[3]Responses that were observed at least five times, but linked to fewer than three children in either group were word order errors (e.g., *Le bateau blanc → Le blanc bateau* 'the.m white boat'), and other errors (e.g., copula or conjoined DP structures: *Une maison rouge (elle) est petite* 'A red house (it) is small', or *Le bateau blanc et gros* 'The boat white and big').
[4]Adjective commissions (e.g., *Purple* for *brown* or *mommy* for *big*) and determiner commissions (e.g., *La grenouille grise → *Une grenouille *gris* 'A/one.f frog grey.m') were observed more than five times, but were rare in controls and thus not analyzed further.

= 43; $U = 16$, $z = 1.63$, $p = .052$, $d = 0.83$). Of interest was that one child (D6) tended to overuse the non-default feminine form for *green* (*verte* [vɛ(ʁ)t]), which is observed nowhere else in the 151 control participants aged 3–6 years (unpublished data). A trend for an effect of Group was found for gender errors on size adjectives ($H = 5.14$, $p = .075$). Because children with DLD and the AGE group showed highly different patterns, we performed post hoc comparisons on these two groups. Differences were observed between DLD and AGE groups in the production of gender errors on size adjectives, but these did not reach our alpha criterion of .025 ($n = 46$ vs. 5; $U = 15.5$, $z = 1.68$, $p = .047$, $d = 1.26$). No other differences between groups on error patterns were observed.

Insert table 4 here

Spontaneous speech corpus. The children's spontaneous speech corpus was analyzed to compare similar DP structures across both tasks, but also other types not elicited with the puzzles. Based on 200 utterances per participant, we analyzed all DP contexts with determiners, adjectives, and other complex DP structures. All groups showed high levels of DP production, ranging from 74 in one MLU-matched child to 165 in one child with DLD[5]. Total numbers of DP structures produced by each group were 939 in the DLD group ($M = 130.8$, $SD = 71.6$), 1073 in the AGE group ($M = 134.1$, $SD = 20.2$), and 1046 in the MLU group ($M = 117.4$, $SD = 34.2$). A total of 302 analyzable errors were observed on DP structures. Children with DLD produced half of these (i.e., 155), AGE controls produced 63, and MLU controls 84. Four error types—

[5]Note that numbers include proper noun phrases as these were occasionally introduced by the preposition *à* used in genitive structures as in *à Maman* 'Mommy's'. These usually accounted for less than 10% of the corpus.

agreement errors on adjectives or determiners, substitutions, omissions, and overregularizations—reached the five-item-per-group criterion. All other error types including word order errors, added elements, and other errors were negligible ($n \leq 5$ in a given group). Determiner and adjective agreement errors involved using an inappropriate gendered adjective or determiner in a determiner phrase (*la *vert* 'the.f green.m'). Substitutions involved definiteness errors (1a), number errors (1b), lexical-semantic errors (1c), and preposition substitutions (1d). Most omissions involved determiners but some involved nouns (e.g., *six* __ 'six (years old)').

(1) a. *des* *chevaux* for *les* *chevaux*
det.indef horse.pl det.def horse.pl
'some horses' 'the horses'

b. *le* *traces* for *les* *traces*
the.m.sg [lœ] traces the.pl [lɛ] traces
'the traces' 'the traces'

c. *un lion* for *un tigre*
'a lion' 'a tiger'

d. *sa trompe dedans l'épan* [epã] for *la* *trompe sur l'éléphant*,
its trunk inside the-elephant the.def.f trunk on the-elephant
'his trunk inside the elephant' 'the trunk on the elephant'

Four overregularization patterns were observed in the corpus. Three were linked to the non-application of obligatory morphophonological processes in contraction (2a), elision (2b), and liaison (2c). French-speaking children master contraction in elicitation by age 5 and some

children at younger ages (approx. 3;01; Béchara, 2015). Overregularizations of contractions, elisions, or liaisons were rare but more common in participants with DLD. A fourth type of overregularization observed more frequently in controls was the misuse of irregular plural or singular nouns (2d). We checked all potential error types in the corpus outside the DP and only one other significant difference in patterns between groups was observed on verb number agreement, which was more common in children with DLD, but still quite rare (e.g., *Les fleurs *est belles,* 'the.pl flowers *is beautiful;' DLD: $M = 2.25$, $Mdn = 2$, Range = 2–8; AGE: $M = 0.38$, $Mdn = 0$, Range = 0–2; $z = -2.23$, $p < .05$).

(2) a. *de les tirex* [dœletˢirɛks] for *des tirex* [detˢirɛks]

'some t-rexes'

b. *le nenfant* [lœnãfã] for *l'enfant* [lãfã]

'the child'

c. *l'zami l'dinosaure* [lzamildᶻnozɔʁ] for *l'ami du dinosaure* [lamidydᶻnozɔʁ]

'the friend the dinosaur' 'the friend of the dinosaur'

d. *un chevaux* [ʃœvo] for *un cheval* [ʃœval]

'a horses' 'a horse'

We compared groups on the four identified error types with Kruskal-Wallis tests, and when relevant, post hoc Mann-Whitney tests with an adjusted alpha of .025. Descriptive information is provided in Table 5. Significant group effects were found for substitution ($H = 8.85$, $p = .012$) and omission errors only ($H = 10.58$, $p = .005$). Follow-up comparisons revealed significant differences in substitution errors produced by DLD and AGE groups ($U = 9.5$, $z =$

2.31, $p = .010$, $d = 1.56$), as well as with the MLU group ($U = 6.5$, $z = 2.63$, $p < .01$, $d = 1.77$). Omissions (e.g., _ *porte* '(the) door') were significantly more common in the DLD group than the AGE group ($U = 2.0$, $z = 3.10$, $p = .001$, $d = 1.65$), but not in the MLU group ($U = 16.5$, $z = 1.58$, $p = .057$, $d = 0.70$).

Insert table 5 here

Based on the assumption that the narrow focus of the puzzles might have provided poorer information about adjective inventories and gender-marking abilities than the corpus, we reviewed each child's spontaneous production of these items and compared these to elicitation patterns (Table 6). Table 6 illustrates that only two children with DLD (D8 and D9) spontaneously produced four or more feminine variable adjectives; the other six did not exhibit productive use of these forms (Hiriarteborde, 1973). One (D8) showed normal scores on elicitation tasks, but another (D9) did not. The global inventory of feminine adjectives spontaneously produced by the DLD group was small: seven lemmas, four of which were tested in the tasks. We also noted that only D4 showed any signs of difficulties with gender marking on determiners in spontaneous speech. Most other errors observed on determiners were substitutions and omissions.

Insert table 6 about here

In sum, we observed a number of characteristics of DP production in two contexts: (a) a controlled barrier task with puzzles eliciting size and colour adjectives in simple and slightly

more complex DPs and (b) in a spontaneous speech play context. The puzzle tasks revealed difficulties producing target configurations for DP structures involving colour adjectives with little difficulties with size adjectives in children with DLD compared to age-matched controls, but few differences with MLU-matched controls. A finer analysis on error types revealed similar patterns on gender-marked colour adjectives: children with DLD produced more errors than age-matched controls, while trends for differences were observed on size adjectives. An analysis of corpus errors on similar DP structures revealed that gender errors were not a feature distinguishing groups, but rather that omission and substitution errors in the DP were salient features of spontaneous productions by children with DLD.

Discussion

Although our task was run on a small group of French-speaking children, some interesting and important patterns emerged. Our elicitation tasks revealed that globally, the DP is well mastered in these children. Any lexical-semantic errors on colour adjectives or in the corpus appear to cut across participant groups and thus do not appear to be the root cause for observed differences between them. Robust comprehension of these structures was also noted. However, specific difficulties were found when children with DLD were asked to produce DPs with variable adjectives, mostly feminine colour adjectives. Determiner errors were only salient in some children: D4 omitted a quarter of his determiners and made one definiteness error, and D7 omitted two and made six gender errors on them, most often using feminine for the masculine, which is highly unusual, even in clinical observations. Finally, D3 omitted or substituted a smaller number of determiners. Children with DLD were also more likely to produce many more errors than age-matched controls, especially on adjectives, whether size or colour. Note that

these size adjectives are acquired at age 3 in French (Royle & Valois, 2010; Valois & Royle, 2009; Valois, et al., 2009).

Globally, our task results are consistent with cross-linguistic data from Swedish and Russian, as well as previous French studies of DP production. As mentioned above, older Russian children with DLD have difficulties producing gender-marked adjectives in an antonym elicitation task (e.g., *open* vs. *closed*; Tribushinina & Dubinkina, 2012). Gender difficulties on Swedish adjectives in elicited DP probes also have been observed, but were less salient than determiner-omission and substitution errors (Leonard et al., 2001).

Our results on adjective production, however, are not as clear-cut as those provided in Roulet-Amiot and Jakubowicz (2006) who showed important difficulties with variable adjectives in children with DLD. This could be explained by their use of non-contrasting pictures where adjectives were not pragmatically salient. It is also possible that some of their items were much harder than those in our task. For example, they used derived adjectives and did not strictly control for age-of-acquisition norms. Therefore, we cannot be sure the task was not tapping into lexical-semantic knowledge. Regarding determiner errors, in the sentence elicitation task reported above, Pizzioli and Schelstraete (2008) observed determiner omissions in older French-speaking children with DLD. However, the tasks they used were quite different from ours.

We checked whether a spontaneous speech corpus would be more instructive in identifying agreement or DP-structure difficulties in children with DLD. Spontaneous speech analyses highlighted other differences between French-speaking children with DLD and their peers, especially omission and substitution errors, which were more common in the DLD group than both control groups. Although agreement errors are found on determiners and adjectives, they were relatively rare in our study, but adjectives were also under-represented in the corpus.

In particular, and as found with younger children, the corpus provides little information about adjective agreement abilities, and might even over-evaluate them if, for example, we were to score total correct adjective production or even masculine variable adjective production (Table 6, columns 5 and 6). Our data were not coherent with previous studies of determiner production in spontaneous speech showing little difficulties in determiner production (e.g., low rates of omission in obligatory contexts) in French-speaking children with DLD (Le Normand et al., 1993; Paradis & Crago, 2001). One reason for this disparity may be that in Le Normand et al.'s (1993) study, only 29 DP contexts on average were analyzed per child. The Paradis and Crago (2001) study, however, does not appear to have this issue as a 45-minute sample of spontaneous speech was taken, yet no information was provided on the number of utterances produced or analyzed.

Contrasting elicitation and spontaneous-speech samples, we can affirm that both approaches were useful in evaluating adjective agreement, DP structure, and lexical-semantic abilities, and also revealed significant difficulties with determiners in children with DLD. Only the puzzle tasks revealed gender agreement as an area of weakness in our participants. Because French adjective morphology is opaque (Royle, 2011) and thus difficult for all children, differences between typical and impaired language development were not as robust as what has been found for other languages with productive and transparent morphology. However, we did observe particular patterns only in children with DLD, namely overuse of feminine for masculine forms on determiners and adjectives, which is highly unusual. This behaviour could be a clear marker of language disorders in French, and should be investigated further. Royle and Stine (2013) suggested that some errors observed in the corpus are worth pursuing in further elicitation studies, although they do not stand out in the spontaneous data. These include

morphophonological alternations in elision, liaison, and contraction (see examples in (1) above), which are easily elicited in controlled contexts (Béchara, 2015).

Regarding other errors found in the corpus, only verb number agreement errors stand out as a potential marker for DLD in these children, and these are quite rare. It is well known that recording and coding a spontaneous speech corpus demands more time and effort to implement than a short elicitation task. Thus, although some interesting features of DLD speech can be extracted from the corpus, we believe elicitation tasks hold the potential to rapidly identify children with language production difficulties in an ecological way. While not discounting the usefulness of spontaneous speech samples for providing global measures of lexical richness, syntactic and morphosyntactic abilities, we acknowledge the puzzles' utility to quickly tap into agreement and short sentence structure difficulties in French children, and their potential to help identify children with language development delay or disorders in their pre-school years. It is also relevant from a clinical point of view, as well as from developmental and theoretical points of view, to distinguish colour and size adjectives as we have done, when evaluating DP production in children with DLD. Remember that we observed that children with DLD did not distinguish themselves from controls on size adjectives, which are acquired early, but showed worse results on colour ones. Grouping these two types of adjectives in a speech corpus analysis or in an elicitation task could water down interesting information about a child's linguistic abilities or would make it more ambiguous where the difficulties were observed.

Steel et al. (2013) emphasized that spontaneous speech data often underestimate children's linguistic competence, but as we have shown, it can also overestimate it because children may opt to use structures they master (Leonard, 2016). Two previous studies of Spanish-speaking children showed that elicitation revealed that those with DLD had difficulties

producing plural nouns, while spontaneous speech did not distinguish them from typically developing children (Bedore & Leonard, 2001, 2005). In our study, difficulties with determiners varied according to task. Table 6 highlights the fact that only D3, D4, and D7 had difficulty with determiners on the puzzles, but most participants with DLD showed determiner errors (i.e., agreement, substitution, or omission) in spontaneous speech. The opposite held for adjective agreement abilities. All participants with DLD, except D8, showed difficulties on puzzle 4, and all produced gender-marking errors on the tasks but only half (i.e., D2, D4, D6, and D8) showed any errors (between one and three) on the small number of feminine adjectives they use in the corpus.

Additionally, spontaneous speech samples often do not provide sufficient samples for contexts that the clinician or researcher wants to investigate (Gerken, 2000). This is definitely the case for gender-marked adjectives and complex DPs, which are rare in the corpus. This could be due to pragmatic factors (e.g., the referent is salient, no extra information is needed to identify it) or to avoidance strategies, and cannot be clarified in this type of data. Thus, we believe it is useful to apply multiple methods to build a complete picture of the child's linguistic abilities. In the case of D8, none of the measures, except error rates on the task, seemed to highlight linguistic difficulties. However, rapid identification of specific difficulties might be better served by elicitation tasks. Eisenbeiss (2011) highlighted the tension between using elicitation experiments, which can be challenging for children, and obtaining insufficient information from spontaneous speech. However, the first three puzzles used here can be used with typically-developing French 3-year-olds without difficulty (Royle & Valois, 2010) and were developed to ecologically reflect other puzzles available to parents or daycares. They thus hold the potential to

provide information on lexical, syntactic, and agreement abilities in young French-speaking children.

Limitations

A limitation of this study was the small sample size for the group with DLD. Data from larger samples are required to confirm the usefulness of these puzzles as a screen for oral language disorders in children aged 5 to 6 years. Furthermore, two of the children were not presented with the B versions of puzzles 3 and 4. Both A and B versions should be used for the same child for clear norms to emerge. It is possible that the easiest puzzles could be useful in identifying younger children with DLD (e.g., between ages 3 and 5), but this has not yet been established. Also, these tasks might prove difficult for second language learners who notoriously find French gender agreement difficult, even if their mother tongue has this feature (see Royle, Bergeron, & Marquis, 2015). Thus, this task would not necessarily be useful in identifying language impairment in contexts where a learner has not yet fully mastered French. Studies on second language mastery of gender agreement would be useful to further deepen our understanding of these issues.

References

Béchara, J. (2015). *Évaluation de la production morphosyntaxique chez un enfant québécois francophone porteur d'implant cochléaire* (Master's Thesis, Université de Montréal, Montréal, Canada). Retrieved from https://papyrus.bib.umontreal.ca/xmlui/handle/1866/13457

Bedore, L. M., & Leonard, L. B. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44,* 905–924. doi:10.1044/1092-4388(2001/072)

Bedore, L. M., & Leonard, L. B. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Applied Psycholinguistics, 26,* 195–225. doi:10.1017.S0142716405050149

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

Dunn, L., Thériault-Whalen, C., & Dunn, L. M. (1993). *Échelle de vocabulaire en images Peabody*. *Adaptation française du Peabody Picture Vocabulary Test Revised*. Toronto, Canada: Psycan.

Eisenbeiss, S. (2011). *CEGS: An elicitation took kit for studies on case marking and its acquisition*. Department of Language and Linguistics, Essex University, Essex, United Kingdom. Retrieved from https://pdfs.semanticscholar.org/719c/481117a88743be0464e89063b9e8f5930aba.pdf

Elin T. Thordardottir. (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. *International Journal of Language and Communication Disorders, 40,* 243–278. doi:10.1080/13682820410001729655

Elin Thordardottir, Gagné, A., Levy, J., Kehayia, E., Lessard, N., Sutton, A., & Trudeau, N. (2005, December). *Systematic language sample analysis in French: Normative data for conversation.* Paper presented at the Emergence of Language Abilities Conference, Lyon, France. Abstract retrieved from http://www.ddl.cnrs.fr/Colloques/ELA2005/abstracts/Ela2005-THO1_P.pdf

Elin T. Thordardottir, & Namazi, M. (2007). Specific language impairment in French-speaking children: Beyond grammatical morphology. *Journal of Speech, Language, and Hearing Research, 50,* 698–715. doi:10.1044/1092-4388(2007/049)

Fromont, L. A., Royle, P., & Steinhauer, K. (2015, October). *Finding agreement: An on-line study of gender processing, in adults and children.* Paper presented at the Seventh Annual Meeting of the Society for the Neurobiology of Language, Chicago, IL.

Gerken, L. (2000). Examining young children's morphosyntactic development through elicited production. In L. Menn & N. Bernstein Ratner (Eds.), *Methods for studying language production* (pp. 45–93). Mahwah, NJ: Lawrence Erlbaum Associates.

Gopnik, M. (1990). Feature blindness: A case study. *Language Acquisition, 1,* 139–164. doi:10.1207/s15327817la0102_1

Grüter, T. (2005). Comprehension and production of French object clitics by child second language learners and children with specific language impairment. *Applied Psycholinguistics, 26,* 363–391. doi:10.1017/S0142716405050216

Hansson, K., Nettelbladt, U., & Leonard, L. B. (2003). Indefinite articles and definite forms in Swedish children with specific language impairment. *First Language, 23,* 343–362. doi:10.1177/01427237030233005

Hiriarteborde, A. (1973). Sur la généralisation de quelques marques grammaticales dans le langage d'enfants de 3 ans 1/2. *Études de Linguistique Appliquée, 9,* 101–124.

Le Normand, M.-T., Leonard, L. B., & McGregor, K. K. (1993). A cross-linguistic study of article use by children with specific language impairment. *International Journal of Language & Communication Disorders, 28,* 153–163. doi:10.3109/13682829309041463

Leonard, L. B. (2014). *Children with specific language impairment* (2nd ed.). Cambridge, MA: MIT Press.

Leonard, L. B. (2016). Noun-related morphosyntactic difficulties in specific language impairment across languages. *First Language, 36,* 3–29. doi:10.1177/0142723715626070

Leonard, L. B., Salameh, E.-K., & Hansson, K. (2001). Noun phrase morphology in Swedish-speaking children with specific language impairment. *Applied Psycholinguistics, 22,* 619–639. doi:10.1017/S0142716401004076

Miller, C. A., & Gilbert, E. (2008). Comparison of performance on two nonverbal intelligence tests by adolescents with and without language impairment. *Journal of Communication Disorders, 41,* 358–371. doi:10.1016/j.jcomdis.2008.02.003

Miller, J., & Chapman, R. (1984–2002). *Systematic analysis of language transcripts: Software for analyzing English and Spanish language transcripts*. Madison, WI: University of Wisconsin, Language Analysis Laboratory.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE^{TM}. *L'Année psychologique, 101,* 447–462. doi:10.3406/psy.2001.1341

Ordre des orthophonistes et audiologistes du Québec. (2004). *Troubles primaires du langage : Dysphasie. Guide et outils cliniques*. Unpublished clinical guidelines.

Ordre des orthophonistes et audiologistes du Québec. (2014). *Mémoire de l'Ordre des orthophonistes et audiologistes du Québec présenté à la Commission de la santé et des services sociaux relativement au projet de loi no 10. Loi modifiant l'organisation et la gouvernance du réseau de la santé et des services sociaux notamment par l'abolition des agences régionales*. Retrieved from http://www.ooaq.qc.ca/actualites/OOAQ-memoire-PL10-24oct2014.pdf

Paradis, J., & Crago, M. (2001). The morphosyntax of specific language impairment in French: An extended optional default account. *Language Acquisition, 9,* 269–300. doi:10.1207/S15327817LA0904_01

Parisse, C., & Maillart, C. (2007). Phonology and syntax in French children with SLI: A longitudinal study. *Clinical Linguistics & Phonetics, 21,* 945–951. doi:10.1080/02699200701610634

Pizzioli, F., & Schelstraete, M.-A. (2008). Argument-structure complexity in children with specific language impairment: Evidence from the use of grammatical morphemes in French. *Journal of Speech, Language, and Hearing Research, 51,* 706–721. doi:10.1044/1092-4388(2008/050)

Rakhlin, N., Kornilov, S. A., & Grigorenko, E. L. (2014). Gender and agreement processing in children with developmental language disorder. *Journal of Child Language, 41,* 241–274. doi:10.107/S03050009I200058X

Ramos, E. B. (2000). *Acquisition of noun phrase structure in children with specific language impairment.* (Doctoral Dissertation, University of Massachusetts, Amherst, MA). Retrieved from https://scholarworks.umass.edu/dissertations/AAI9960783

Rice, M. L., & Oetting, J. B. (1993). Morphological deficits of children with SLI. Evaluation of number marking and agreement. *Journal of Speech and Hearing Research, 36,* 1249–1257. doi:10.1044/jshr.3606.1249

Roid, G. H., & Miller, L. J. (1996) *Leiter International Performance Scale-Revised*. Wood Dale, IL: Stoelting.

Roulet-Amiot, L., & Jakubowicz, C. (2006). Production and perception of gender agreement in French SLI. *Advances in Speech Language Pathology, 8,* 335–346. doi:10.1080/14417040601009420

Royle, P. (2011). On the existence of C/Ø alternations in French adjectives: Theoretical and empirical questions. *Proceedings of the 17th International Congress of Phonetic Sciences*, 1730–1733. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Royle/Royle.pdf

Royle, P., Bergeron, E., & Marquis, A. (2015). L'acquisition du genre quand il est déjà acquis : l'expérience d'enfants hispanophones apprenant le français. *Revue canadienne d'orthophonie et d'audiologie, 39*, 238–259.

Royle, P., & Elin T. Thordardottir. (2008). Elicitation of the passé composé in French preschoolers with and without specific language impairment. *Applied Psycholinguistics, 29,* 341–365. doi:10.1017/S0142716408080168

Royle, P., St-Denis, A., Mazzocca, P., & Marquis, A. (2017). Insensitivity to verb conjugation patterns in French children with SLI. *Clinical Linguistics & Phonetics, 32,* 128–147. doi:10.1080/02699206.2017.1328706

Royle, P., & Stine, I. (2013). The French noun phrase in preschool children with SLI: Morphosyntactic and error analyses. *Journal of Child Language, 40,* 945–970. doi:10.1017/S0305000912000414

Royle, P., & Valois, D. (2010). Acquisition of adjectives in Quebec French as revealed by elicitation data. *Journal of French Language Studies, 10,* 313–338. doi:10.1017/S0959269510000013

Steel, G., Rose, M., Eadie, P., & Thornton, R. (2013). Assessment of complement clauses: A comparison between elicitation tasks and language sample data. *International Journal of Speech-Language Pathology, 15,* 286–295. doi:10.3109/17549507.2013.777852

Tribushinina, E., & Dubinkina, E. (2012). Adjective production by Russian-speaking children with specific language impairment. *Clinical Linguistics & Phonetics, 26,* 554–571. doi:10.3109/02699206.2012.666779

Trudeau, N., Frank, H., & Poulin-Dubois, D. (1999). Une adaptation en français québécois du MacArthur Communicative Development Inventory. *Revue d'orthophonie et d'audiologie, 23*, 61–73.

Valois, D. (1991). *The internal syntax of DP*. Los Angeles, CA: UCLA Dissertations in Linguistics.

Valois, D., & Royle, P. (2009). Partitivity, atomization, and noun-drop: A longitudinal study of French child language. *Language Acquisition, 16,* 82–105. doi:10.1080/10489220902773079

Valois, D., Royle, P., Sutton, A., & Bourdua-Roy, E. (2009). L'ellipse du nom en français : le rôle des données de l'acquisition pour la théorie linguistique. *Canadian Journal of Linguistics/Revue canadienne de linguistique, 54,* 339–366. doi:10.1017/S0008413100001286

van der Lely, H. K. J. (1998). SLI in children: Movement, economy, and deficits in the computational-syntactic system. *Language Acquisition, 7,* 161–192. doi:10.1207/s15327817la0702-4_4

Table 1

Participant Characteristics

	Age (months)	MLUw	MLUm	Sex	MEM IQ	EVIP (centile score)	EDUC (years)
DLD group[a]							
1. D2	65	4.02	5.18	F	109	71	22
2. D3	70	3.75	4.91	M	96	87	14
3. D4	60	2.88	3.79	M	90	< 1	mv
4. D5	67	4.39	5.79	M	96	75	11
5. D6	67	5.38	7.17	M	90	30	11
6. D7	69	4.17	5.50	M	106	11	12
7. D8	69	4.64	5.10	F	74	14	11
8. D9	71	6.02	8.06	F	53+	15	14
Mean[c]	*67.25*	*4.41*	*5.69*		*89.25*	*37.9*	*13.57*
SD	*3.49*	*0.97*	*1.35*		*18.16*	*34.23*	*3.95*
AGE group[a,b]							
1. N55✧	63	4.70	6.10	F	106	79	15
2. N50✧	69	4.25	5.37	M	122	97	14
2. N109	69	5.13	6.93	M	90	81	14
3. N42✧	58	4.59	6.47	M	103	50	14.5
4. N20✧	67	4.49	5.92	M	100	99	13.5

4. N114	66	4.92	6.83	M	93	98	14
5. N20✧	67	4.49	5.92	M	100	99	13.5
5. N114	66	4.92	6.83	M	93	98	14
6. N101✧	69	5.27	7.31	M	103	97	14.5
6. N151	72	7.38	9.38	M	103	81	12
7. N127✧	68	3.65	5.57	F	112	64	12
7. N69	67	5.65	7.31	F	125	99	12
8. N9✧	71	4.11	5.77	F	100	66	10.5
8. N8	71	5.28	6.96	F	106	96	13.5
AGE Puzzle Controls (*n* = 14)							
Mean[c]	*66.64*	*4.92*	*6.62*[*]		*104*	*86*[**]	*13.36*
SD	*5.17*	*0.88*	*1.02*		*10.10*	*16.20*	*1.26*
AGE Corpus Controls (*n* = 8)							
Mean[c]	*66.5*	*4.44*	*6.05*		*105.75*[*]	*81.38*[*]	*13.44*
SD	*4.14*	*0.47*	*0.61*		*7.72*	*19.41*	*1.5*
MLU group[a,b]							
1. N63✧	49	3.67	4.74	F	74	80	20
2. N96✧	45	3.87	4.64	M	106	68	10
2. N36	43	3.75	4.61	M	90	77	13.5
3. N103✧	38	2.68	3.53	M	96	52	14.5
4. N41✧	56	4.28	5.71	M	109	50	13
4. N141	46	4.34	5.66	M	106	79	14.5

5. N115✧	49	5.61	7.51	M	125	50	15
5. N129	68	5.14	6.79	M	131	86	14
6. N80✧	46	4.21	5.66	M	122	73	12
6. N96	45	4.11	5.82	M	106	68	10
7. N92✧	43	4.21	5.56	F	143	62	15.5
7. N71	58	4.15	5.19	F	96	66	10
8. N3✧	44	6.15	7.77	F	87	30	14
8. N11	51	6.21	8.12	F	106	95	16
MLU Puzzle Controls (n = 14)							
Mean[c]	*48.64****	*4.46*	*5.81*		*106.93*	*66.86***	*13.71*
SD	*7.65*	*0.99*	*1.32*		*18.50*	*17.08*	*2.72*
MLU Corpus Controls (n = 8)							
Mean[c]	*46.25***	*4.34*	*5.64*		*107.75**	*58.13*	*14.25*
SD	*5.28*	*1.09*	*1.43*		*22.22*	*15.91*	*2.93*

Note. AGE = controls matched on age; DLD = children with developmental language disorder; MLU = controls matched on mean length of utterance. mv = missing value; EDUC = mean parental education; EVIP = *Échelle en images Peabody* [French Peabody Picture Vocabulary Test receptive vocabulary]; F = feminine; M = Masculine; MEM IQ = Leiter memory subtest (Associated Pairs and Forward Memory subtests, $^{+}$ this child was non-compliant on the MEM IQ; however, she appeared to have normal cognitive abilities in all respects); MLUw = mean length of utterances in words; MLUm = mean length of utterances in morphemes.
[a]Numbers refer to children with DLD and their matched controls. [b]Participants marked with a ✧ were included in the spontaneous speech analysis. [c]*T*-tests between DLD and control groups $^{*}p < .05$, $^{***}p < .001$.

Table 2

Target Responses for Colour Adjectives (Puzzle 1) and Full Determiner Phrases (Puzzles 2–4)

Task	*H, p*	*M*	Median	*SD*	Range
Puzzle 1 (colour naming, 6 items)					
DLD	3.79, .15	5.50	6	0.76	4–6
AGE		5.75	6	0.46	5–6
MLU		4.63	5	1.60	1–6
Puzzle 2 (size DPs, 8 items)					
DLD	1.16, .56	6.00	6	2.14	2–8
AGE		6.88	8	1.81	3–8
MLU		5.63	6.5	2.88	0–8
Puzzle 3 (colour DPs, 12 items, averaged over 6)[a]					
DLD	8.74, < .05	4.31	4.75	1.41	1–5.5
AGE		5.31	5.5	0.46	4.5–6
MLU		4.19	4.5	0.96	2–5
Puzzle 4 (size and colour DPs, 24 items, averaged over 12)[a]					
DLD	7.27, < .05	6.94	6.5	2.86	2–11
AGE		9.75	10	1.58	7–11.5
MLU		7.25	8	1.87	3.5–9.5

Note. AGE = controls matched on age; DLD = children with developmental language disorder; DP = determiner phrases; MLU = controls matched on mean length of utterance.
[a]Participants with DLD (n = 2) and controls (AGE, n = 2; MLU, n = 2) who only did version A of Puzzle 3 had the opportunity to produce 6 full DP structures, and for Puzzle 4 had the opportunity to produce 12 full DP structures.

Table 3

Target Response Results on Feminine Adjectives in Puzzles 2 and 4

Task	*H, p*	*M*	Median	*SD*	Range
Puzzle 2 (size adjectives, 4 items)					
DLD	0.8, .67	2.50	4	1.69	0–4
AGE		3.10	4	1.46	0–4
MLU		2.38	4	1.85	0–4
Puzzle 4 (colour and size adjectives, 18 items, averaged over 9) [a]					
DLD	13.21, < .01	4.38	4.75	2.60	0–8
AGE		8.50	8.5	0.85	7.5–10[a]
MLU		6.25	6.25	1.41	4.5–8

Note. AGE = controls matched on age; DLD = children with developmental language disorder; MLU = controls matched on mean length of utterance.
[a]Participants with DLD (n = 2) and controls (AGE, n = 2; MLU, n = 2) who only did version A of Puzzle 4 had the opportunity to produce 10 variable feminine adjectives.

Table 4

Descriptive Statistics for Error Productions on Determiner Phrases in Tasks 2, 3, and 4

	H, p	*M*	Median	*SD*	Range	Nb
Colour Adjective Gender						
DLD	10.39, < .01	7.75	8	2.76	3–12	8
AGE		2.50	2	2.07	0–6	6
MLU		5.38	5	2.97	1–11	8
Size Adjective Gender						
DLD	5.14, .078	5.75	5.5	5.57	0–14	5
AGE		0.625	0	1.41	0–4	1
MLU		5.375	5	3.80	0–10	6
Determiner Gender						
DLD	0.10, .616	2.5	1	4.11	0–12	4
AGE		0.625	0.5	0.74	0–2	4
MLU		0.875	0	1.81	0–5	2
Determiner Omission						
DLD	4.29, .117	6.25	0	12.24	0–35	3
AGE		0	0	0.00	0	0
MLU		7.125	2	11.15	0–32	5
Split DP						
DLD	1.22, .054	8.701	2	8.70	0–20	5
AGE		4.625	3	5.83	0–18	6
MLU		11.25	8	11.94	0–35	7

Note. AGE = controls matched on age; DLD = children with developmental language disorder; DP = determiner phrases; MLU = controls matched on mean length of utterance; Nb = number of children exhibiting this pattern, recall that control groups represent pairs of children.

Table 5

Descriptive Statistics for Error Productions on Determiner Phrases in the Spontaneous Speech Corpus

	H, p	*M*	Median	*SD*	Range	Nb
Gender Agreement Errors						
DLD	3.32, .19	3.25	2.5	2.66	0–9	7
AGE		1.75	2	1.39	0–3	6
MLU		1.38	1.5	1.06	0–2	6
Omissions						
DLD	10.58, < 01	9.38	7	6.30	3–21	8
AGE		1.88	1.5	1.36	0–4	7
MLU		5.38	3.5	5.07	0–16	7
Substitutions						
DLD	8.85, .012	4.5	4.5	2.27	1–8	8
AGE		1.63	2	1.30	0–4	6
MLU		1.25	1	1.28	0–3	5
Overregularization						
DLD	0, 1	1.25	1	1.04	0–3	6
AGE		2.0	1	2.56	0–6	5
MLU		2.13	1	3.36	0–10	5

Note. AGE = controls matched on age; DLD = children with developmental language disorder; MLU = controls matched on mean length of utterance; Nb = number of children (out of 8) exhibiting this pattern.

Table 6

Comparison Task Results for Adjective Production, and Adjective or Determiner Errors in Spontaneous Speech in Children with Developmental Language Disorder

	Elicitation			Corpus analysis					
	Tasks (target responses)			Adjective use				Determiner	
	2	3	4	Variable			Total Adjectives	Errors[c]	
	Size	Col.	Siz. + Col.	Masc.	Fem.	Lemmas	Lexemes (Lemmas)	Masc.	Fem.
1. D2	2/8	4/6	4/12	3/3	2/5	4	16 (8)	3 (1)	6 (1)
2. D3*	6/8	9/12	12/24	3/3	2/2	4	6 (5)	4 (0)	0 (0)
3. D4*	6/8	1/6	2/12	9/9	0/3	4	13 (7)	16 (3)	9 (4)
4. D5	6/8	**10/12**	12/24	17/17	1/1	6	29 (12)	7 (0)	4 (2)
5. D6	4/8	**10/12**	14/24	29/29	2/3	7	38 (11)	12 (0)	7 (0)
6. D7*	**8/8**	9/12	18/24	34/35	2/2	10	54 (32)	4 (1)	1 (0)
7. D8	**8/8**	**11/12**	**21/24**	11/12	3/4	5	39 (21)	2 (1)	1 (0)
8. D9	**8/8**	**10/12**	19/24	12/12	8/8	10	36 (20)	2 (0)	5 (0)
Mean	*5.6/8*	*8.6/12*[a]	*14/24*[b]	*0.98*	*0.75*	*6.25*	*29 (14)*	*6.25*	*2.14*
SD	*1.61*	*2.83*	*6.35*			*2.55*	*15.96 (9.2)*	*5.15*	*3.23*
Median	*6*	*9.5*	*14*			*5.50*	*32.50 (11.5)*	*4*	*4.5*

Note. Results are presented as number target/total number or number of different lemmas. Bold numbers indicate that scores on the elicitation tasks are within normal range, as defined by scores for the age-matched controls. Fem. = feminine; Lemma = number of different words produced; Masc. = masculine; Total = Nb of items produced.
Target/total productions: [a]averaged over 12 items; [b]averaged over 24 items; [c]Includes omission, agreement and substitution (e.g., definiteness), errors specific to gender are in parentheses.[*] = indicates children who made determiner errors in the elicitation task.