

Comparison of Measures of Morphosyntactic Complexity
in French-Speaking School-Aged Children

Catherine Mimeau, Vickie Plourde, Andrée-Anne Ouellet, and Ginette Dionne

Université Laval, Quebec City, Quebec, Canada

Author Note

Catherine Mimeau, École de psychologie, Université Laval, Quebec City, Quebec, Canada; Vickie Plourde, École de psychologie, Université Laval, Quebec City, Quebec, Canada; Andrée-Anne Ouellet, École de psychologie, Université Laval, Quebec City, Quebec, Canada; Ginette Dionne, École de psychologie, Université Laval, Quebec City, Quebec, Canada.

Correspondence concerning this article should be addressed to Catherine Mimeau, École de psychologie, Université Laval, Pavillon Félix-Antoine-Savard, 2325 rue des Bibliothèques, Quebec City, Quebec, G1V 0A6, Canada. Phone number: 1-418-656-2131, ext. 3083. E-mail: catherine.mimeau@psy.ulaval.ca

Comparison of Measures of Morphosyntactic Complexity
in French-Speaking School-Aged Children

This study examined the validity and reliability of different measures of morphosyntactic complexity, including the Morphosyntactic Complexity Scale (MSCS), a novel adaptation of the Developmental Sentence Scoring, in French-speaking school-aged children. Seventy-three Quebec children from kindergarten to Grade 3 completed a definition task and a narration task. Mean length of utterance (MLU), clause density, and MSCS global score, average frequency scores and average complexity scores were calculated from the transcripts of the two contexts. MLU, clause density, and MSCS global score were correlated with vocabulary knowledge and narrative skills, and they increased as a function of school level, suggesting that they are valid measures of morphosyntactic complexity. Moreover, the three scores were correlated across contexts, suggesting that they are also reliable measures. However, no MSCS average frequency or average complexity score was found to be both valid and reliable. These findings will guide researchers and practitioners who desire to assess the language skills of French-speaking school-aged children.

The development of different aspects of language in young English-speaking children is well documented (e.g., Hoff, 2009). However, less is known about the oral language skills of school-aged children speaking a language other than English. This disparity may arise from the fact that tools that are needed for adequate language assessment are not always easily adaptable from one age group to another or from one language to another. As a consequence, in developmental research, language assessments of non-English-speaking children are sometimes limited to a unique aspect, most often vocabulary, leading to conclusions that may not be representative of the whole range of children's language skills. Thus, this study examines the validity and reliability of different measures of morphosyntax, including a novel measure adapted from the Developmental Sentence Scoring (DSS; Lee, 1974). The objective is to provide a broader pool of methods for assessing French-speaking school-aged children.

Existing Measures of Morphosyntactic Complexity

When measuring morphosyntactic skills in children, analysis of spontaneous speech is often preferred to elicited production tests (e.g., the Wug Test; Berko, 1958), which may underestimate children's knowledge (Dever, 1972). The traditional way of analysing morphosyntactic complexity from spontaneous speech is to calculate the mean length of utterance (MLU; Brown, 1973) by dividing the total number of words or morphemes by the total number of utterances. The popularity of MLU comes from its easiness to compute and to adapt across languages. Nevertheless, this measure has been widely criticized for its lack of validity and reliability, especially in older children.

For instance, numerous studies have shown that MLU is positively correlated with age in toddlers, but that this association sharply decreases after age 3 years (Klee & Fitzgerald, 1985; Parrisé & Le Normand, 2006; Rondal, Ghiotto, Bredart & Bachelet, 1987; Scarborough, Wyckoff & Davidson, 1986), even though morphosyntax development is ongoing until adulthood

(Martinot, 2005; Nippold, Hesketh, Duthie & Mansfield, 2005). These findings and others from research examining the association between MLU and more detailed morphosyntactic cues (Blake, Quartaro & Onorati, 1993; Scarborough, Rescorla, Tager-Flusberg, Fowler & Sudhalter, 1991) suggest that MLU becomes less sensitive to individual differences as children get older. Furthermore, MLU was found to be rather unstable from one measuring time to another (Bornstein, Hahn & Haynes, 2004; Chabon, Kent-Udolf & Egolf, 1982). Therefore, MLU may not be well suited to assess morphosyntactic complexity in school-aged children.

Alternatively, morphosyntactic complexity in older children can be assessed with clause density (Scott & Stokes, 1995), which can be calculated by dividing the total number of clauses (independent and dependent) by the number of independent clauses to express the proportion of embedded clauses produced. As MLU, this measure is easy to compute and to adapt to different languages. Additionally, it is deemed to increase with age up to adolescence. Therefore, clause density, even though less popular than MLU, is frequently chosen when assessing morphosyntactic complexity in school-aged children.

Another recognised measure of morphosyntactic complexity in English is the DSS (Lee, 1974). This measure is based on the typical developmental sequence of English morphosyntax, that is, score is calculated by assigning less points to early-acquired items, such as the conjunction *and*, and more points to later-acquired items, such as the conjunction *while*. One point is also given for each morphosyntactically correct utterance. Then, total score is divided by the total number of utterances to express a mean complexity score. A study evaluating the DSS's validity indicated that score was associated with age in children up to 7 years old (Kemper, Rice & Chen, 1995). Along with the fact that the DSS informs on many different components of morphosyntax, this finding suggests that this measure may be a suitable alternative to other methods when assessing school-aged children. However, the complexity of the DSS makes it

time-consuming to use and difficult to adapt to a language other than English, and so no similar measure exists for children speaking French (but see Maillart, Parrisé, & Tommerdahl, 2012, and Parrisé & Le Normand, 2006, for detailed qualitative measures of morphosyntax in French; see also Toronto, 1976, and Miyata et al., 2013, for adaptations of the DSS to Spanish and Japanese, respectively).

Finally, when assessing morphosyntactic complexity from spontaneous speech, it is important to bear in mind that the context in which speech is produced may have an influence on language. For instance, MLU is typically higher when children tell a story compared with when they answer questions within a conversation (Southwood & Russell, 2004). In addition, children generally use more complex verb forms in conversation than in narration (Wagner, Nettelbladt, Sahlén & Nilholm, 2000). Therefore, when assessing morphosyntactic complexity from spontaneous speech, the choice of a context must be made conscientiously.

The Present Study

Despite the availability of several methods to measure English-speaking children's language (see also the Index of Productive Syntax [Scarborough, 1990] and the Language Assessment Remediation and Screening Procedure [Crystal, Fletcher & Garman, 1976]), there are few means to measure morphosyntax from spontaneous speech in French. Moreover, the validity and reliability of these methods remain poorly documented, for they have been mostly examined in English-speaking children some decades ago. Hence, the objective of the present study is to examine the validity and reliability of MLU, clause density, and the Morphosyntactic Complexity Scale (MSCS; an adaptation of the DSS) to assess morphosyntactic complexity in French-speaking school-aged children.

A *valid* method is expected to generate scores that are moderately correlated with other measures of language. Such correlations would suggest that while being related to other aspects

of language, the method assesses a different aspect, which, in this case, is thought to be morphosyntax (see Question 1). A valid method should also generate scores that increase as a function of school level, which would indicate that it is sensitive to age-related growth in morphosyntactic complexity (see Question 2a).

In addition, a *reliable* method is expected to generate scores that are correlated across contexts. This inter-contextual correlation would imply that children who speak with an above-average morphosyntax while performing a certain task also speak with an above-average morphosyntax while performing a different task (see Question 3a).

Furthermore, an adequate measure of morphosyntax is expected to generate scores that increase as a function of school level and that are correlated across contexts even when other measures of language are controlled for. This would ensure that the validity and reliability criteria are met because of actual variations in morphosyntax and not only because of variations in other aspects of language captured by the measure (see Questions 2b and 3b).

Finally, once a method is considered as valid and reliable, it is important to determine in which context it can be used. A context that maximises scores is preferable for it reduces any possible underestimation of children's capabilities (see Question 4).

Therefore, the specific questions the study addresses are the following:

1. How are MLU, clause density, and MSCS scores associated with other measures of language?
2. a) Do MLU, clause density, and MSCS scores increase as a function of school level?
b) Do they increase over and beyond other measures of language?
3. a) Are MLU, clause density, and MSCS scores stable across contexts?
b) Are they stable over and beyond other measures of language?
4. How do MLU, clause density, and MSCS scores vary as a function of context?

Method

Participants

Seventy-three students (35 boys and 38 girls) from kindergarten to Grade 3 were recruited from five public primary schools in Quebec City, Canada. To participate in the study, students had to have French as their first language and to have typical language development according to their teacher¹. Furthermore, to create somewhat homogeneous age groups, all students were tested no more than four months before or five months after their birthday. One participant was removed from the study because her parents did not provide her date of birth and other essential information. Mean age of all the remaining participants was 7.55 years ($SD = 1.17$). Kindergarten students ($n = 19$) were on average 6.04 years old ($SD = 0.19$), Grade 1 students ($n = 17$) were on average 7.04 years old ($SD = 0.16$), Grade 2 students ($n = 18$) were on average 8.08 years old ($SD = 0.20$), and Grade 3 students ($n = 18$) were on average 9.08 years old ($SD = 0.17$).

Procedure

Participants were assessed at their school. Kindergarten students completed the Vocabulary subtest of the French version of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III; Wechsler, 2002), and Grade 1 to Grade 3 students completed the Vocabulary subtest of the French version of the Wechsler Intelligence Scale for Children (WISC-III; Wechsler, 1991). All participants also completed the French version of the Edmonton Narrative Norms Instrument (ENNI; Schneider, Dubé & Hayward, 2005). Half of the participants completed the WPPSI/WISC first, and the other half completed the ENNI first. Participants' answers to the WPPSI/WISC and the ENNI were recorded and transcribed by four trained assistants and the first author, who also revised all of the transcripts. MLU, clause density, and MSCS scores were then calculated from these transcripts.

Materials

Contexts. The Vocabulary subtest of the WPPSI/WISC assesses children's vocabulary knowledge by asking them to define a list of words. The WPPSI, which is intended for children between 2 ½ and 7 years of age, consists of a list of 25 words, and the WISC, which is intended for children between 6 and 16 years of age, consists of a list of 30 words. The task ends after the children have defined all of the words or after four consecutive scores of 0. Score is determined by the accuracy of the definitions provided: 0 point is assigned to an incorrect answer, 1 point is assigned to a partially correct answer, and 2 points are assigned to a completely correct answer². The total score is calculated by adding the individual scores of every word (maximum = 43 for the WPPSI and 60 for the WISC), and it is then normalised for age (range: 1-19). Inter-rater reliability is good for these measures, with intra-class correlation coefficients of .97 for the WPPSI and .98 for the WISC (Wechsler, 1991, 2002).

The ENNI assesses 4- to 9-year-old children's narrative skills. In this task, children are asked to tell a practice story and two series of three stories from pictures that are shown to them. The practice story has five pictures and each series includes one 5-, one 8-, and one 13-picture story. The task ends after the children have told the seven stories or after 20 minutes. Score is determined by the mention of essential units of the story grammar (e.g., setting, initiating event, outcome) with every unit being worth 1 or 2 points. Total scores are calculated from the 5-picture story (maximum = 14) and the 13-picture story (maximum = 40) of the first series, and they are then normalised for age ($M = 10$; $SD = 3$). Inter-rater reliability is good for these measures, with Cohen's kappa coefficients of .92 for both stories (Schneider et al., 2005). The normalised scores of the two stories were averaged to a mean score in the present study.

Measures of morphosyntactic complexity. MLU was calculated by dividing the total number of words or morphemes by the total number of utterances. A morpheme was defined as any word or audible ending indicating feminine, plural, tense, or person. For instance, the

utterance *L'horloge a une petite aiguille pour les secondes* (The clock has a small hand for the seconds) includes 14 morphemes (nine words, three audible indications of feminine or plural, and one indication of tense and one indication of person in the verb *a* [has]). An utterance was defined as a sentence containing minimally a subject and a verb, and possibly one or more coordinate and/or subordinate clauses. A sequence of words that was not a sentence but that was separated from the rest of speech by pauses of at least one second was also considered as an utterance. However, hesitations or reflection pauses within a sentence did not divide this sentence into more than one utterance (Rondal, 1997).

Clause density was calculated by dividing the total number of independent and dependent clauses by the number of independent clauses. A dependent clause was defined as a clause that is embedded in an independent clause. Thus, relative clauses, noun clauses, and adverbial clauses were counted as dependent clauses. However, non-embedded clauses, coordinate clauses, and utterances with no inflected verb were counted as independent clauses. As an example, the utterance *Ça veut dire que c'est vieux* (It means that it's old) would receive a score of 2 because it has one independent clause (*Ça veut dire* [It means]) and one dependent clause (*que c'est vieux* [that it's old]).

The MSCS (see Appendix A) follows the same design as Lee's (1974) DSS: Every time selected morphosyntactic items were encountered, they received points, later-acquired items more so than early-acquired ones (e.g., *moi* [me] received 1 point and *eux* [them] received 5 points). The selection of morphosyntactic items (articles, personal and impersonal pronouns, prepositions and adverbs, verb tenses, clause types, and relations) and their score (range: 1-9) were established from the typical developmental sequence of French morphosyntax proposed by Rondal for European children between 2 and 6 years old (1978, pp. 176–178). Several aspects of this sequence were also validated more recently in English-speaking and French-speaking

European and Quebec children (Bassano, 2000; Bassano, Maillochon, & Mottet, 2008; Girouard, Ricard, & Gouin Décarie, 1997; Schmitz & Müller, 2008; Strik, 2007; Thordardottir, 2005; Trudeau & Sutton, 2011; Vasilyeva, Waterfall, & Huttenlocher, 2008). In addition to this graded scoring, one point was given for each morphosyntactically correct utterance.

Different scores can be computed from the MSCS. The *global score* is the total number of points divided by the total number of utterances. A *subscore* can be computed for a given category (e.g., Articles) by dividing the total number of points in the category by the total number of utterances. Two other scores can also be calculated from the total number of items in a given category (e.g., the total number of articles produced): (a) the *average frequency* of a category, that is, the total number of items in the category divided by the total number of utterances, and (b) the *average complexity* of a category, that is, the total number of points in the category divided by the total number of items in the category. In fact, the multiplication of these two scores results in the subscore of a category.

To make sure MLU, clause density, and MSCS scores were not artificially inflated, a maximum of two coordinate clauses were kept for each utterance. Any additional coordinate clause was considered as a new utterance (Lee, 1974). Moreover, groups of words and expressions considered as a single unit (e.g., *parce que* [because]) were counted as only one word (Thordardottir, 2005). Finally, utterances of only one morpheme were not included in the calculation of the scores (Rondal, 1997), and nor were utterances that were unintelligible or that contained an unintelligible segment, repeated utterances (Lee, 1974), and utterances not related to the task.

All morphosyntactic scores were calculated by four different raters, except clause density, which was calculated by a single rater. Scores for six participants were calculated from the WPPSI/WISC and from the ENNI by an additional external rater for clause density and by the

study's four raters for the other measures. Intra-class correlation coefficients were above .85 for all scores except complexity of prepositions and adverbs, for which it was .76.

Results

Descriptive Statistics

The WPPSI/WISC elicited 51.01 utterances on average ($SD = 31.95$), and the ENNI, 67.75 ($SD = 15.29$). The mean normalised score of vocabulary, computed from the WPPSI/WISC, was 11.94 ($SD = 2.72$), and the mean normalised score of narration, computed from the ENNI, was 9.71 ($SD = 2.79$). Means and standard deviations for MLU, clause density, and MSCS scores calculated from the WPPSI/WISC and the ENNI are presented as a function of school level in Table 1. Since the correlations between MLU in words and MLU in morphemes were exceptionally strong ($r = .99, p < .001$, for the WPPSI/WISC, and $r = .995, p < .001$, for the ENNI; see Parisse & Le Normand, 2006, for similar results), only MLU in words was used in the analyses. Moreover, since the six MSCS subscores are composites of average frequency and average complexity scores, only the latter were used in the analyses (in addition to the global score). Furthermore, clause density was log-transformed for further analyses because it was not normally distributed.

Correlations between MLU, clause density, and MSCS scores calculated from the WPPSI/WISC and the ENNI are presented in Table 2. For scores calculated from the WPPSI/WISC (below the diagonal) as well as for those calculated from the ENNI (above the diagonal), most correlations between MLU, clause density, MSCS global score, and average frequency scores were significant and positive. However, regarding average complexity scores, there were fewer significant correlations, and no clear pattern emerged.

How Are MLU, Clause Density, and MSCS Scores Associated with Other Measures of Language?

To investigate the validity of MLU, clause density, and the MSCS, correlations were performed between these measures calculated from the WPPSI/WISC and the ENNI and vocabulary knowledge and narrative skills. The results are presented in Table 3. MLU, clause density, MSCS global score, frequency of articles, and frequency of relations, when calculated from either the WPPSI/WISC or the ENNI, were all significantly and positively correlated with both vocabulary knowledge and narrative skills (except for frequency of relations calculated from the ENNI, which was not significantly correlated with vocabulary knowledge). These consistent moderate associations suggest that these measures could be considered as valid. Still, the next analyses provide further evidence regarding the question of validity.

Do MLU, Clause Density, and MSCS Scores Increase as a Function of School Level?

To further investigate the validity of MLU, clause density, and MSCS global score, three repeated measures analyses of variance (ANOVAs) were conducted with context (WPPSI/WISC, ENNI) as the within-subjects variable and school level (kindergarten, Grade 1, Grade 2, Grade 3) as the between-subjects variable. Sex was also included as a between-subjects variable for MLU and MSCS global score, as *t*-tests revealed that girls scored higher than boys on these measures ($ps < .05$ for both measures in the WPPSI/WISC and the ENNI).

Moreover, to further investigate the validity of average frequency and average complexity scores, two repeated measures multivariate analyses of variance (MANOVAs) were conducted with context (WPPSI/WISC, ENNI) as the within-subjects variable and school level (kindergarten, Grade 1, Grade 2, Grade 3) as the between-subjects variable. Sex was also included as a between-subjects variable in both analyses, as *t*-tests revealed that girls scored higher than boys on at least one measure of frequency and one measure of complexity ($ps < .05$ for frequency of articles, prepositions and adverbs, conjugated verbs³, and relations, and

complexity of prepositions and adverbs in the WPPSI/WISC, and for complexity of verb tenses in the ENNI).

MLU. The ANOVA performed with MLU as the dependent variable revealed a significant main effect of school level, $F(3, 67) = 4.88, p = .004, \eta_p^2 = .18$. Post hoc comparison tests performed with a Sidak adjustment indicated that Grade 3 students had higher MLUs than kindergarten ($p = .003$) and Grade 1 ($p = .04$) students.

Clause density. The ANOVA performed with clause density as the dependent variable revealed a significant main effect of school level, $F(3, 68) = 5.17, p = .003, \eta_p^2 = .19$. Post hoc comparison tests performed with a Sidak adjustment indicated that Grade 2 ($p = .02$) and Grade 3 ($p = .02$) students produced denser clauses than kindergarten students.

MSCS global score. The ANOVA performed with MSCS global score as the dependent variable revealed a significant main effect of school level, $F(3, 67) = 2.85, p = .04, \eta_p^2 = .11$. However, post hoc comparison tests performed with a Sidak adjustment indicated no significant difference between any school levels. Still, as shown in Table 1, MSCS global scores generally increase from kindergarten to Grade 3.

Average frequency. The MANOVA performed with average frequency scores as the dependent variables revealed a significant main effect of school level, Wilks' $\Lambda = .62, F(15, 174) = 2.22, p = .01, \eta_p^2 = .15$. Univariate tests of between-subjects effects showed that only frequency of articles varied as a function of school level, $F(3, 67) = 4.13, p = .01, \eta_p^2 = .16$. Post hoc comparison tests performed with a Sidak adjustment indicated that Grade 3 students produced more articles than kindergarten students ($p = .01$).

Average complexity. The MANOVA performed with average complexity scores as the dependent variables revealed a significant main effect of school level, Wilks' $\Lambda = .46, F(18,$

176) = 3.10, $p < .001$, $\eta_p^2 = .23$. Univariate tests of between-subjects effects showed that complexity of pronouns, $F(3, 67) = 6.40$, $p = .001$, $\eta_p^2 = .22$, complexity of verb tenses, $F(3, 67) = 2.82$, $p = .045$, $\eta_p^2 = .11$, complexity of clause types, $F(3, 67) = 3.72$, $p = .02$, $\eta_p^2 = .14$, and complexity of relations, $F(3, 67) = 4.70$, $p = .005$, $\eta_p^2 = .17$, varied as a function of school level. Post hoc comparison tests performed with a Sidak adjustment indicated that Grade 3 students produced more complex clause types than kindergarten ($p = .01$) and Grade 2 ($p = .01$) students, and more complex relations than kindergarten students ($p = .01$). Nonetheless, they also indicated that kindergarten students produced more complex pronouns than Grade 2 ($p < .001$) and Grade 3 ($p = .01$) students, and more complex verb tenses than Grade 2 students ($p = .04$), that is, that complexity of pronouns and complexity of verb tenses decreased as a function of school level.

Interactions with context. Interactions between school level and context need to be considered, as a measure for which no main effect of school level was found could have increased as a function of school level only in one of the contexts. Nevertheless, no interaction showing an increase as a function of school level was found ($ps > .16$)⁴.

Controlling for vocabulary knowledge and narrative skills. To ensure that MLU, clause density, and MSCS scores are valid measures of morphosyntax and not of general language, the previous ANOVAs and MANOVAs were conducted again with vocabulary knowledge and narrative skills as covariates. The analyses revealed a significant main effect of school level for MLU, $F(3, 65) = 3.58$, $p = .02$, $\eta_p^2 = .14$, clause density, $F(3, 66) = 3.21$, $p = .03$, $\eta_p^2 = .13$, average frequency, Wilks' $\Lambda = .66$, $F(15, 169) = 1.80$, $p = .04$, $\eta_p^2 = .13$, and average complexity, Wilks' $\Lambda = .47$, $F(18, 170) = 2.90$, $p < .001$, $\eta_p^2 = .22$, but not for MSCS global score, $F(3, 65) = 2.32$, $p = .08$, $\eta_p^2 = .10$ (although the effect was almost significant). Univariate

tests of between-subjects effects and post hoc comparison tests performed with a Sidak adjustment yielded results close to those obtained without the control.

Overall, as indicated by their increase as a function of school level, MLU, clause density, MSCS global score, frequency of articles, complexity of clause types, and complexity of relations could be considered as valid measures. However, when taking away the effect of vocabulary knowledge and narrative skills to reveal which measures assess morphosyntactic complexity specifically, only MLU, clause density, complexity of clause types, and complexity of relations remain. Furthermore, when considering the results from the correlations performed previously, only MLU and clause density meet both criteria to be considered as valid measures of morphosyntactic complexity.

Are MLU, Clause Density, and MSCS Scores Stable across Contexts?

To investigate the reliability of MLU, clause density, and the MSCS, correlations were performed between these measures calculated from the WPPSI/WISC and the ENNI. The results are presented in Table 2 (on the diagonal). MLU, clause density, MSCS global score, frequency of pronouns, frequency of conjugated verbs, and frequency of relations were significantly and positively correlated across contexts. All of these inter-contextual correlations remained significant when controlling for vocabulary knowledge and narrative skills (MLU: $r = .59, p < .001$; clause density: $r = .33, p = .01$; MSCS global score: $r = .56, p < .001$; frequency of pronouns: $r = .43, p < .001$; frequency of conjugated verbs: $r = .55, p < .001$; frequency of relations: $r = .46, p < .001$), suggesting that these measures are reliable.

Taken with our previous findings, these correlations suggest that only MLU and clause density meet all the criteria for validity and reliability. However, MSCS global score could also be considered as an adequate measure of morphosyntactic complexity, as it only marginally

failed to increase as a function of school level when vocabulary knowledge and narrative skills were controlled for.

How Do MLU, Clause Density, and MSCS Scores Vary as a Function of Context?

To investigate which context is better suited to use MLU, clause density, and MSCS global score, the effect of context was examined in the repeated measures ANOVAs conducted previously with these measures. A marginally significant main effect of context was found for MLU, $F(1, 67) = 3.44, p = .07, \eta_p^2 = .05$, with higher MLUs in the ENNI than in the WPPSI/WISC. In addition, a significant main effect of context was found for clause density, $F(1, 68) = 181.25, p < .001, \eta_p^2 = .73$, with denser clauses in the WPPSI/WISC than in the ENNI. However, no significant main effect of context was found for MSCS global score, $F(1, 67) = 1.44, p = .23, \eta_p^2 = .02$, suggesting that children have comparable scores in the WPPSI/WISC and in the ENNI.

Discussion

The objective of this study was to examine the validity and reliability of MLU, clause density, and the MSCS to assess morphosyntactic complexity in French-speaking school-aged children. Whereas MLU, clause density, and MSCS global score proved to be valid and reliable methods, MSCS average frequency and average complexity scores failed to (a) be associated with other measures of language, (b) increase as a function of school level, and/or (c) be stable across contexts. Furthermore, while the MSCS generated similar global scores when calculated from a definition and a narration task, MLU was found to generate higher scores when calculated from a narration task, and clause density was found to generate higher scores when calculated from a definition task.

MLU

Despite the numerous criticisms it has received (e.g., Chabon et al., 1982; Klee & Fitzgerald, 1985), MLU seems to be an adequate method to assess morphosyntactic complexity in French-speaking school-aged children, at least between kindergarten and Grade 3. Indeed, regardless of the context from which it was calculated, it was moderately correlated with vocabulary knowledge and narrative skills. Furthermore, children entering school produced on average shorter utterances than did children in higher school levels, and MLUs calculated from the definition task were associated with MLUs calculated from the narration task. These results remained significant even when controlling for vocabulary knowledge and narrative skills to ensure that morphosyntactic skills specifically rather than general language skills be assessed.

The discrepancy between these results and those from previous research may arise from the fact that most other studies examined English-speaking children. Indeed, French has a more complex morphology than English (e.g., verbs are inflected differently for all persons), and so French-speaking children may acquire the complex features of their language over a prolonged period of time, which would preserve the association between age and MLU longer. In fact, recent findings indicate that MLU increases significantly in 4 ½- to 5 ½-year-old children speaking French (Thordardottir, Keheyia, Lessard, Sutton, & Trudeau, 2010), and in 4- to 8-year-old children speaking Italian (Venuti et al., 2011), another morphologically complex language, whereas there is no association with age in English-speaking children above 3 years old (Rondal et al., 1987). Although only MLU in words was used in the present study, its exceptionally strong correlations with MLU in morphemes ($r_s \geq .99$) allow us to make this supposition.

Clause Density

As is the case for MLU, clause density seems to be an adequate method to assess morphosyntactic complexity in French-speaking school-aged children, at least between kindergarten and Grade 3. Indeed, regardless of the context from which it was calculated, it was

moderately correlated with vocabulary knowledge and narrative skills. Furthermore, children in higher school levels produced on average denser clauses than did children just entering school, and clause densities calculated from the definition task were associated with clause densities calculated from the narration task. These results remained significant even when controlling for vocabulary knowledge and narrative skills to ensure that morphosyntactic skills specifically rather than general language skills be assessed.

These findings are in agreement with what was expected. Indeed, some studies have revealed that adults produce more relative clauses than children (Martinot, 2005; Nippold et al., 2005), supporting the idea that morphosyntactic development persists well beyond early childhood. Our results indicate that clause density can capture such age-related variations.

The MSCS

Similarly to MLU and clause density, MSCS global score could be an adequate method to assess morphosyntactic complexity in French-speaking school-aged children, at least between kindergarten and Grade 3. In fact, this score was strongly correlated with MLU in both contexts (see Table 2). Regardless of the context from which it was calculated, MSCS global score was correlated with vocabulary knowledge and narrative skills. However, the correlations were more modest than for MLU and clause density. Furthermore, children entering school had on average lower MSCS global scores than did children in higher school levels, but the effect of school level disappeared when vocabulary knowledge and narrative skills were taken into account. Finally, MSCS global scores calculated from the definition task were associated with MSCS global scores calculated from the narration task, even when controlling for vocabulary knowledge and narrative skills. Therefore, even though the validity of MSCS global score is somewhat less definite than that of MLU and clause density, and considering the fact that this score takes a long time to

compute, it could be used when the objective is to get a detailed portrait of morphosyntactic complexity.

The validity and reliability of MSCS global score in the present study support previous findings indicating that DSS global score is a valid measure of morphosyntax in English-speaking school-aged children (Kemper et al., 1995). Hence, the adaption of the method to French-speaking children proved to be successful, as both the DSS and the MSCS yield global scores that increase as a function of age or school level in school-aged children.

Nevertheless, results of the present study question the validity and reliability of MSCS average frequency and average complexity scores. Indeed, out of the 12 scores examined, only frequency of articles, frequency of conjugated verbs, and frequency of relations were correlated with other language measures. Furthermore, the only relevant differences between children entering school and those in higher school levels were that the former produced fewer articles, simpler clause types, and simpler relations than the latter. Finally, only frequency of pronouns, frequency of conjugated verbs, and frequency of relations were correlated across contexts. Still, these results are novel, since, to our knowledge, no one has previously examined the validity and reliability of the MSCS or a similar measure separately for its different categories. If a specific category had been shown to be valid and reliable, it would have been useful for researchers and practitioners to know, as this category could have been used alone to assess morphosyntactic skills, thus reducing the lengthy computation process required by the MSCS.

It should be noted that this inconclusive finding might result from the small sample size of the study. Indeed, the groups comprised only between 17 and 19 children in each grade, raising the possibility that some insignificant results be due to a lack of statistical power. Further research should therefore examine the validity and reliability of MSCS average frequency and

average complexity scores in a larger sample of children before strong claims can be made as to the psychometric properties of these scores.

Context

The results of this study indicate that both definition and narration tasks are adequate contexts from which to calculate morphosyntactic complexity. Indeed, MLU, clause density, and MSCS global score were correlated with vocabulary knowledge and narrative skills whether they were calculated from one context or the other. In fact, correlations between the morphosyntactic measures calculated from the WPPSI/WISC and vocabulary knowledge (assessed with the WPPSI/WISC) were comparable to those between the morphosyntactic measures calculated from the ENNI and vocabulary knowledge, and the same was true for correlations with narrative skills. In other words, the morphosyntactic measures were not over-correlated with the skill assessed by the context from which they were calculated (*viz.*, vocabulary knowledge or narrative skills), which is desirable, as it suggests that the measures are somewhat independent.

Nonetheless, the morphosyntactic measures acted differently across contexts: While MSCS global scores were equivalent in the definition and the narration tasks, MLU was higher in the narration task, and clause density was higher in the definition task. These results, which point out that morphosyntactic scores should not be compared if calculated from different contexts, are consistent with previous findings. Indeed, longer utterances are usually produced in narrative contexts in comparison with conversational/questioning contexts (Southwood & Russell, 2004), and giving definitions was found to enhance the production of relative clauses (Friedmann, Aram & Novogrodsky, 2011).

Consequently, if morphosyntactic complexity is assessed with MLU, then a narration context should be favoured, as this context maximises the length of utterances. Inversely, if morphosyntactic complexity is assessed with clause density, then a definition context should be

favoured, as this context maximises the density of clauses. As for MSCS global score, a definition or a narration context could be chosen, as both elicit comparable scores.

Conclusion

In summary, the present study showed that MLU and clause density, despite their simplicity, are valid and reliable methods to assess morphosyntactic complexity in French-speaking school-aged children. MSCS global score, although more time-consuming to compute, is also an adequate measure. However, MSCS average frequency and average complexity scores, even though more informative of specific components of morphosyntax, all lack validity and/or reliability. Moreover, definition generation and storytelling both seem to be appropriate contexts from which to calculate MLU, clause density, and MSCS global score in French-speaking school-aged children. Thus, these tasks could allow the computation of two language scores (i.e., a morphosyntax score in addition to a vocabulary or a narration score), which would better represent language than a single score. Further research should examine the validity and reliability of MLU, clause density, and the DSS in individuals beyond 10 years of age and/or speaking a language other than English or Quebec French, including European French, which may slightly differ from Quebec French with regard to morphosyntax. The MSCS could also be tested among children with language impairments. Together with the present study, such investigations would guide researchers and practitioners in the choice of an appropriate method to measure language in a given population.

Acknowledgements

We are grateful to the children who participated in this study. We also thank Marie Gwen Castel-Girard, Haniel Baillargeon-Lemieux, and Laurence Tanguay-Garneau, for their contribution to testing, transcription, and codification, and Natacha Trudeau, for her significant advice on the draft of the paper.

Funding

This work was supported by the Social Sciences and Humanities Research Council of Canada.

Notes

¹ Four participants' parents reported their child had a language impairment and/or had seen a speech therapist. However, the participants were kept in the study because their scores did not substantially change the results of the analyses.

² For the first seven words in the WPPSI, 1 point is assigned to a completely correct answer.

³ Frequency of verb tenses and frequency of clause types are the same measure – frequency of conjugated verbs – as both were calculated by dividing the total number of conjugated verbs by the total number of utterances.

⁴ Interactions between school level and context were found for some average complexity scores, but since they showed a decrease as a function of school level, they will not be discussed further.

References

- Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, 27, 521–559.
doi:10.1017/S0305000900004396
- Bassano, D., Maillochon, I., & Mottet, S. (2008). Noun grammaticalization and determiner use in French children's speech: A gradual development with prosodic and lexical influences. *Journal of Child Language*, 35, 403–438. doi:10.1017/S0305000907008586
- Berko, J. (1958). The child's learning of English morphology. *Word*, 14, 150–177. Retrieved from <http://chilides.talkbank.org>
- Blake, J., Quartaro, G., & Onorati, S. (1993). Evaluating quantitative measures of grammatical complexity in spontaneous speech samples. *Journal of Child Language*, 20, 139–152.
doi:10.1017/S0305000900009168
- Bornstein, M. H., Hahn, C. S., & Haynes, O. M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language*, 24, 267–304. doi:10.1177/0142723704045681
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Chabon, S. S., Kent-Udolf, L., & Egolf, D. B. (1982). The temporal reliability of Brown's mean length of utterance (MLU-M) measure with post-stage V children. *Journal of Speech, Language, and Hearing Research*, 25, 124–128. Retrieved from <http://jslhr.asha.org>
- Crystal, D., Fletcher, P., & Garman, M. (1976). *The grammatical analysis of language disability: A procedure for assessment and remediation*. New York, NY: Elsevier.
- Dever, R. B. (1972). A comparison of the results of a revised version of Berko's test of morphology with the free speech of mentally retarded children. *Journal of Speech, Language, and Hearing Research*, 15, 169–178. Retrieved from <http://jslhr.asha.org>

- Friedmann, N., Aram, D., & Novogrodsky, R. (2011). Definitions as a window to the acquisition of relative clauses. *Applied Psycholinguistics*, *32*, 687-710.
doi:10.1017/S0142716411000026
- Girouard, P. C., Ricard, M., & Gouin Décarie, T. (1997). The acquisition of personal pronouns in French-speaking and English-speaking children. *Journal of Child Language*, *24*, 311–326.
doi:10.1017/S030500099700305X
- Hoff, E. (2009). *Language development* (4th ed.). Belmont, CA: Wadsworth.
- Kemper, S., Rice, K., & Chen, Y. J. (1995). Complexity metrics and growth curves for measuring grammatical development from five to ten. *First Language*, *15*, 151–166.
doi:10.1177/014272379501504402
- Klee, T., & Fitzgerald, M. D. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language*, *12*, 251–269.
doi:10.1017/S0305000900006437
- Lee, L. L. (1974). *Developmental sentence analysis: A grammatical assessment procedure for speech and language clinicians*. Evanston, IL: Northwestern University Press.
- Maillart, C., Parisse, C., & Tommerdahl, J. (2012). F-LARSP 1.0: An adaptation of the LARSP language profile for French. *Clinical Linguistics & Phonetics*, *26*, 188-198.
doi:10.3109/02699206.2011.602459
- Martinot, C. (2005). *Comment parlent les enfants de 6 ans ? Pour une linguistique de l'acquisition* [How do 6 year-old children talk? For a linguistics of acquisition]. Besançon, France: Presses universitaires de Franche-Comté.
- Miyata, S., MacWhinney, B., Otomo, K., Sirai, H., Oshima-Takane, Y., Hirakawa, M., . . . Itoh, K. (2013). Developmental Sentence Scoring for Japanese. *First Language*, *33*, 200–216.
doi:10.1177/0142723713479436

- Nippold, M. A., Hesketh, L. J., Duthie, J. K., & Mansfield, T. C. (2005). Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research, 48*, 1048–1064. doi:10.1044/1092-4388(2005/073)
- Parisse, C., & Le Normand, M. T. (2006). Une méthode pour évaluer la production du langage spontané chez l'enfant de 2 à 4 ans [A method to evaluate spontaneous speech production in the 2- to 4-year-old child]. *Glossa, 97*, 20–41. Retrieved from <http://www.glossa.fr>
- Rondal, J. A. (1978). *Langage et éducation* [Language and education]. Brussels, Belgium: Mardaga.
- Rondal, J. A. (1997). *L'évaluation du langage* [The evaluation of language]. Sprimont, Belgium: Mardaga.
- Rondal, J. A., Ghiotto, M., Bredart, S., & Bachelet, J. F. (1987). Age-relation, reliability and grammatical validity of measures of utterance length. *Journal of Child Language, 14*, 433–446. doi:10.1017/S0305000900010229
- Scarborough, H. S. (1990). Index of Productive Syntax. *Applied Psycholinguistics, 11*, 1–22. doi:10.1017/S0142716400008262
- Scarborough, H. S., Rescorla, L., Tager-Flusberg, H., Fowler, A. E., & Sudhalter, V. (1991). The relation of utterance length to grammatical complexity in normal and language-disordered groups. *Applied Psycholinguistics, 12*, 23–45. doi:10.1017/S014271640000936X
- Scarborough, H. S., Wyckoff, J., & Davidson, R. (1986). A reconsideration of the relation between age and mean utterance length. *Journal of Speech, Language, and Hearing Research, 29*, 394–399. Retrieved from <http://jslhr.asha.org>

- Schmitz, K., & Müller, N. (2008). Strong and clitic pronouns in monolingual and bilingual acquisition of French and Italian. *Bilingualism: Language and Cognition, 11*, 19–41. doi:10.1017/S1366728907003197
- Schneider, P., Dubé, R. V., & Hayward, D. (2005). *The Edmonton Narrative Norms Instrument*. Retrieved from <http://www.rehabresearch.ualberta.ca/enni>
- Scott, C. M., & Stokes, S. L. (1995). Measures of syntax in school-age children and adolescents. *Language, Speech, and Hearing Services in Schools, 26*, 309–319. Retrieved from <http://lshss.asha.org>
- Southwood, F., & Russell, A. F. (2004). Comparison of conversation, freeplay, and story generation as methods of language sample elicitation. *Journal of Speech, Language, and Hearing Research, 47*, 366–376. doi:10.1044/1092-4388(2004/030)
- Strik, N. (2007). L'acquisition des phrases interrogatives chez les enfants francophones [The acquisition of interrogative clauses in French-speaking children]. *Psychologie française, 52*, 27–39. doi:10.1016/j.psfr.2006.07.003
- Thordardottir, E. (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. *International Journal of Language & Communication Disorders, 40*, 243–278. doi:10.1080/13682820410001729655
- Thordardottir, E., Keheyia, E., Lessard, N., Sutton, A., & Trudeau, N. (2010). Typical performance on tests of language knowledge and language processing of French-speaking 5-year-olds. *Canadian Journal of Speech-Language Pathology and Audiology, 34*, 5–16. Retrieved from http://www.caslpa.ca/english/resources/cjslpa_home.asp
- Toronto, A. S., (1976). Developmental assessment of Spanish grammar. *Journal of Speech and Hearing Disorders, 41*, 150–171. doi:10.1044/jshd.4102.150

- Trudeau, N., & Sutton, A. (2011). Expressive vocabulary and early grammar of 16- to 30-month-old children acquiring Quebec French. *First Language, 31*, 480–507.
doi:10.1177/0142723711410828
- Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science, 11*, 84–97. doi:10.1111/j.1467-7687.2007.00656.x
- Venuti, P., de Falco, S., Esposito, G., Zannella, N., Villotti, P., & Bornstein, M. H. (2011, April). *Mean length of utterance (MLU) in children aged 4 to 9 years: A cross-sectional study.* Poster session presented at the meeting of the Society for Research in Child Development, Montreal, Canada.
- Wagner, C. R., Nettelbladt, U., Sahlén, B., & Nilholm, C. (2000). Conversation versus narration in pre-school children with language impairment. *International Journal of Language & Communication Disorders, 35*, 83–93. doi:10.1080/136828200247269
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence* (3rd ed.). San Antonio, TX: Psychological Corporation.

Table 1

Means (and Standard Deviations) for MLU, Clause Density, and MSCS Scores Calculated From the WPPSI/WISC and the ENNI as a Function of School Level

	WPPSI/WISC				ENNI			
	Kindergarten	Grade 1	Grade 2	Grade 3	Kindergarten	Grade 1	Grade 2	Grade 3
MLU in words	8.28 (2.67)	8.48 (2.59)	9.31 (1.64)	10.61 (2.89)	8.61 (2.10)	8.83 (2.27)	9.77 (1.37)	11.32 (2.67)
Clause density	1.38 (0.25)	1.43 (0.21)	1.56 (0.19)	1.58 (0.26)	1.14 (0.10)	1.13 (0.12)	1.21 (0.10)	1.21 (0.08)
MSCS scores								
Global score	20.85 (6.81)	19.91 (6.53)	21.53 (3.78)	24.72 (7.35)	21.54 (5.72)	20.22 (5.07)	22.34 (3.63)	26.01 (6.38)
Average frequency								
Articles	0.93 (0.33)	0.99 (0.27)	1.02 (0.22)	1.10 (0.27)	0.91 (0.31)	1.18 (0.34)	1.21 (0.30)	1.26 (0.41)
(Im)pers. pronouns	0.90 (0.42)	0.73 (0.45)	0.88 (0.39)	1.12 (0.42)	1.46 (0.47)	1.35 (0.44)	1.43 (0.44)	1.68 (0.65)
Prep. and adv.	0.69 (0.24)	0.60 (0.19)	0.64 (0.14)	0.66 (0.28)	0.53 (0.25)	0.41 (0.14)	0.55 (0.17)	0.51 (0.15)
Verb tenses	2.05 (0.58)	2.07 (0.59)	2.11 (0.34)	2.36 (0.58)	1.73 (0.37)	1.72 (0.40)	1.90 (0.30)	2.15 (0.48)
Clause types	2.05 (0.58)	2.07 (0.59)	2.11 (0.34)	2.36 (0.58)	1.73 (0.37)	1.72 (0.40)	1.90 (0.30)	2.15 (0.48)
Relations	0.71 (0.43)	0.71 (0.44)	0.89 (0.30)	1.01 (0.48)	0.73 (0.42)	0.75 (0.46)	0.87 (0.32)	1.06 (0.40)
Average complexity								
Articles	3.04 (0.23)	2.92 (0.17)	2.92 (0.19)	2.86 (0.20)	2.86 (0.10)	2.88 (0.09)	2.85 (0.11)	2.88 (0.13)
(Im)pers. pronouns	3.88 (0.39)	3.63 (0.50)	3.19 (0.45)	3.18 (0.34)	3.04 (0.31)	2.98 (0.22)	3.08 (0.20)	3.20 (0.23)
Prep. and adv.	2.88 (0.57)	2.51 (0.56)	2.91 (0.58)	3.06 (0.49)	3.38 (0.77)	3.32 (0.60)	3.49 (0.71)	3.17 (0.64)
Verb tenses	3.08 (0.12)	3.10 (0.08)	3.16 (0.16)	3.18 (0.20)	4.28 (0.86)	3.60 (0.64)	3.49 (0.44)	3.89 (1.14)
Clause types	1.21 (0.23)	1.26 (0.15)	1.24 (0.12)	1.34 (0.21)	1.24 (0.10)	1.24 (0.10)	1.20 (0.07)	1.32 (0.11)
Relations	4.00 (0.57)	4.23 (0.62)	4.31 (0.46)	4.41 (0.45)	2.86 (0.56)	2.81 (0.39)	3.16 (0.48)	3.32 (0.54)

Note. MLU = mean length of utterance; MSCS = Morphosyntactic Complexity Scale; WPPSI = Vocabulary subtest of the Wechsler Preschool and Primary Scale of Intelligence; WISC = Vocabulary subtest of the Wechsler Intelligence Scale for Children; ENNI = Edmonton Narrative Norms Instrument; (Im)pers. pronouns = personal and impersonal pronouns; Prep. and adv. = prepositions and adverbs.

Table 2

Correlations Between MLU, Clause Density, and MSCS Scores Calculated From the WPPSI/WISC and the ENNI

Measure	3														
	1	2	a	b						c					
				i	ii	iii	iv	v	vi	i	ii	iii	iv	v	vi
1. MLU in words	0.65***	0.62***	0.94***	0.43***	0.66***	0.46***	0.92***	0.92***	0.87***	-0.07	0.12	0.02	0.08	0.27*	0.27*
2. Clause density	0.63***	0.42***	0.63***	0.31**	0.33**	0.23*	0.65***	0.65***	0.62***	-0.05	0.10	-0.07	0.02	0.20	0.65***
3. MSCS scores															
a) Global score	0.96***	0.69***	0.60***	0.28*	0.71***	0.55***	0.90***	0.90***	0.87***	-0.05	0.23	0.17	0.24*	0.30*	0.27*
b) Av. frequency															
i. Articles	0.77***	0.40***	0.68***	0.19	-0.16	0.19	0.28*	0.28*	0.33**	0.09	-0.08	-0.20	-0.19	-0.01	0.10
ii. (Im)pers. pron.	0.80***	0.53***	0.87***	0.41***	0.40***	0.35**	0.71***	0.71***	0.60***	-0.20	0.28*	0.18	-0.13	0.15	0.15
iii. Prep. and adv.	0.63***	0.24*	0.65***	0.54***	0.49***	0.08	0.35**	0.35**	0.40***	-0.10	0.21	0.42***	0.09	0.01	-0.03
iv. Verb tenses	0.89***	0.62***	0.91***	0.61***	0.73***	0.53***	0.58***	1.00***	0.84***	-0.17	0.12	0.01	-0.03	0.23	0.34**
v. Clause types	0.89***	0.62***	0.91***	0.61***	0.73***	0.53***	1.00***	0.58***	0.84***	-0.17	0.12	0.01	-0.03	0.23	0.34**
vi. Relations	0.87***	0.82***	0.92***	0.54***	0.78***	0.50***	0.80***	0.80***	0.51***	-0.11	0.10	0.13	0.06	0.23	0.06
c) Av. complexity															
i. Articles	-0.14	-0.13	-0.07	-0.13	-0.16	0.12	-0.13	-0.13	-0.05	0.03	-0.12	-0.08	0.27*	0.29*	-0.14
ii. (Im)pers. pron.	-0.31**	-0.30*	-0.24*	-0.10	-0.34**	0.07	-0.25*	-0.25*	-0.33**	0.35**	-0.11	0.04	-0.04	-0.05	0.18
iii. Prep. and adv.	0.35**	0.29*	0.38**	0.14	0.41***	0.18	0.22	0.22	0.38***	0.03	-0.15	0.15	0.05	-0.08	-0.18
iv. Verb tenses	0.50***	0.17	0.46**	0.33**	0.52***	0.32**	0.35**	0.35**	0.38**	-0.19	-0.45***	0.25*	0.14	0.26*	-0.07
v. Clause types	0.21	0.08	0.18	0.04	0.11	-0.02	0.14	0.14	0.17	-0.13	-0.07	-0.16	0.12	0.14	0.20
vi. Relations	0.16	0.58***	0.23*	0.11	0.19	-0.02	0.17	0.17	0.22	-0.04	-0.06	0.08	-0.03	-0.10	0.21

Note. Correlations for the WPPSI/WISC are presented below the diagonal, correlations for the ENNI are presented above the diagonal, and correlations between the WPPSI/WISC and the ENNI are presented in boldface on the diagonal. MLU = mean length of utterance; MSCS = Morphosyntactic Complexity Scale; WPPSI = Vocabulary subtest of the Wechsler Preschool and Primary Scale of Intelligence; WISC = Vocabulary subtest of the Wechsler Intelligence Scale for Children; ENNI = Edmonton Narrative Norms Instrument; Av. frequency = average frequency; (Im)pers. pron. = personal and impersonal pronouns; Prep. and adv. = prepositions and adverbs; Av. complexity = average complexity.

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Table 3

Correlations Between Vocabulary Knowledge and Narrative Skills, and MLU, Clause Density, and MSCS Scores Calculated From the WPPSI/WISC and the ENNI

	Vocabulary knowledge ^a		Narrative skills ^b	
	WPPSI/WISC	ENNI	WPPSI/WISC	ENNI
MLU in words	0.36**	0.30*	0.34**	0.40***
Clause density	0.36**	0.33**	0.25*	0.35**
MSCS scores				
Global score	0.28*	0.23*	0.28*	0.30*
Average frequency				
Articles	0.34**	0.26*	0.27*	0.49***
(Im)personal pronouns	0.12	-0.02	0.19	-0.09
Prepositions and adverbs	0.08	0.00	0.15	0.15
Verb tenses	0.23	0.24*	0.19	0.29*
Clause types	0.23	0.24*	0.19	0.29*
Relations	0.35**	0.19	0.32*	0.32**
Average complexity				
Articles	-0.08	0.03	0.03	0.02
(Im)personal pronouns	-0.17	0.07	-0.12	0.01
Prepositions and adverbs	0.14	-0.13	0.34*	-0.12
Verb tenses	0.12	0.09	0.24*	0.08
Clause types	0.01	0.17	0.04	0.00
Relations	0.19	0.30*	-0.01	0.14

Note. MLU = mean length of utterance; MSCS = Morphosyntactic Complexity Scale; WPPSI = Vocabulary subtest of the Wechsler Preschool and Primary Scale of Intelligence; WISC = Vocabulary subtest of the Wechsler Intelligence Scale for Children; ENNI = Edmonton Narrative Norms Instrument.

^aAssessed with the WPPSI/WISC. ^bAssessed with the ENNI.

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Appendix

Morphosyntactic Complexity Scale (MSCS)

Score	Articles	Personal and Impersonal Pronouns	Prepositions and Adverbs	Verb Tenses	Clause Types	Relations
1		<i>moi</i>	Preposition of possession	Imperative	Imperative	
			<i>pour</i>		Affirmative declarative	
2	<i>un, une</i>	<i>toi, je, tu, il</i>	Adverb of place		Interrogative with interrogative word without subject-verb order inversion	Coordination (except of cause or result)
3	<i>le, la</i>	<i>elle, vous, me, le, la</i>	Preposition of place	Present infinitive		
			<i>avec</i> (comitative)	Present		
4	<i>des, les</i>	<i>nous, on</i>		Compound past		
				Periphrastic future		
5		<i>lui, eux, ils, elles, les, te, soi, se, leur, en, y</i>	<i>avec</i> (instrumental)		Negative declarative	Relative clause
						Noun clause
						Adverbial clause or coordination of cause
						Adverbial clause or coordination of result
6			Adverb of time	Past infinitive		Adverbial clause (except of cause, result or time)
7				Simple future	Interrogative with subject-verb order inversion	
				Imperfect		
8			Preposition of time	Conditional		Adverbial clause of time
				Other tenses		
9					Passive	