Expository language skills of young school-aged children

Marleen F. Westerveld
School of Education at Albany
Massey University
Auckland, New Zealand

Catherine A. Moran
Department of Communication Disorders
University of Canterbury
Christchurch, New Zealand

Address correspondence to:
Dr Marleen Westerveld
Senior Lecturer
Speech and Language Therapy
School of Education at Albany
Massey University
Private Bag 102904
North Shore Mail Centre
Auckland
New Zealand
Phone: + 64 9 4140800
Email: m.westerveld@gmail.com

Abstract

Purpose: This research investigated the expository language skills of young school-aged children with the ultimate aim of obtaining normative data for clinical practice. Specifically, this study examined (a) the level of performance of 6- and 7-year-old children with typical development, and (b) age-related differences between young and older school-aged children.

Method: Expository discourse was elicited in two groups of children, using the favorite game or sport (FGS) task. Performance of the younger age-group (*n* = 61), age 6;0 to 7;11 years, was compared to that of a group of 20 eleven-year-old children from an earlier study. Samples were analyzed on measures of verbal productivity, syntactic complexity, grammatical accuracy, and mazing behavior.

Results: The FGS task was effective in eliciting text-level discourse in young school-aged children. These children produced discourse that resulted in a fairly normal distribution across some of the language production measures. Age-related differences were observed on measures of verbal productivity, mazing behavior, and grammatical accuracy but not on syntactic complexity.

Conclusions: The findings suggest that expository discourse sampling may be a useful addition to a language assessment protocol even for very young school-aged children.

Keywords: Children, expository discourse, language sample analysis, assessment, school-age.

As children progress through the early school years, they are expected to use different text-level discourse genres depending on the situation. An advanced form of discourse that children are exposed to at school is expository discourse, which consists of a monologue providing factual descriptions or explanations of events. Examples include explanations of the rules of a game or procedural descriptions. Expository discourse in young school-aged children has received relatively little attention. One explanation is that expository discourse may be regarded as too difficult for this age group. Considering the possibility that “complex thought drives the development of complex language” (Nippold, Hesketh, Duthie, & Mansfield, 2005, p. 1058), however, this discourse genre may be particularly informative in revealing children’s ability to use complex language structures. The current study aims to address this issue by investigating young school-aged children’s ability to produce expository discourse.

Nippold (2004) identified factors that influence the development of more advanced linguistic skills. Apart from cognitive readiness, schooling is considered critical. At school children are exposed to decontextualized language that contains advanced linguistic structures, for example when reading non-fiction books or when listening to teachers’ explanations of games or experiments. Besides being *exposed* to expository discourse, children need the opportunity to *use* the more advanced linguistic structures characteristic of this discourse type. In New Zealand, children start school (Year 1) on their fifth birthday, and by the time they reach Year 2 of schooling they are expected to use a variety of sentence structures and to form and express ideas and information with reasonable clarity (Levels Two and Three English; Ministry of Education, 2007). Examples of activities requiring competence in expository discourse include sharing of news events and explaining the rules of a game. Although normative data have been collected, both nationally and internationally, describing young school-aged children’s spoken language performance in conversational, personal narrative, and story retelling contexts (e.g., Schneider, Dubé, & Hayward, 2009; Westerveld, Gillon, & Miller, 2004), to the authors' knowledge there are currently no such data available on which to base our expectations of this age group's performance in an expository context. The present investigation therefore examines young school-aged children's ability to produce expository discourse as they progress through Years 2 and 3 of their regular primary school. From a clinical perspective, obtaining normative data of children's spoken language performance in a potentially challenging discourse context will aid in the diagnosis of (specific) language impairment in young school-aged children whose spoken language abilities may seem sufficient in less challenging discourse situations, such as conversation. Furthermore, early detection of language weaknesses in expository discourse may help guide our intervention practices to ensure children possess the required language skills to participate in these important social interactions with their peers and teachers (e.g., Fujiki, Brinton, Morgan, & Hart, 1999).

Existing research into expository discourse of school-aged children has mainly focused on children over the age of eight (Berman & Verhoeven, 2002; Nippold, Hesketh, et al., 2005; Scott & Windsor, 2000). In general these studies have shown that expository discourse improves with age. For example, Nippold, Hesketh, et al. (2005) conducted a cross-sectional study of syntactic development in six age groups (8, 11, 13, 17, 25, and 44 years) and found that with increasing age, there was a gradual increase in (a) overall length of the expository language samples, (b) utterance length, and (c) clausal density. Closer inspection of the performance of the primary school-aged children, however, showed there was no significant improvement in performance from 8 to 11 years of age. This is somewhat surprising as one might expect significantly better performance of the 11-year-olds compared to the 8-year-olds because of a marked increase in exposure to expository discourse, especially in the later primary school years (see Snyder & Caccamise, 2010). As Snyder and

Caccamise summarize, children in first grade are mainly exposed to narrative reading materials, with exposure to expository type text estimated at less than four minutes a day. This balance changes around third or fourth grade, when the emphasis is no longer on 'learning to read' but on 'learning new information through reading' (see Chall, 1996). Although Nippold, Hesketh, et al. (2005) questioned whether the conservative nature of the Tukey test might explain the lack of differences in performance between the 8- and the 11-year-old children in their study, another possible explanation pertains to the age groups they used. It seems feasible that by eight years of age children have already started the transition from 'learning to read' to 'learning through reading', resulting in increased exposure to more advanced reading materials that incorporate informational text. This may well explain the lack of statistically significant differences in performance between 8- and 11-year-old children on measures of syntactic complexity in Nippold, Hesketh, et al.'s (2005) study. To check this hypothesis, the current study will compare the performance of 6- and 7-year-old children, who are in the early stages of reading development, to that of a group of 11-year-old children.

The choice of elicitation task is an important consideration. Expository discourse can be elicited in different conditions, such as asking the child to explain or describe a procedure (Masterson & Kamhi, 1991), to provide a summary of a short descriptive film (Scott & Windsor, 2000), or to discuss the issue of interpersonal conflict after watching a three minute video (Berman & Verhoeven, 2002). For the present study we decided to use the task developed by Nippold, Hesketh, et al. (2005) to ensure the elicitation procedure would be effective in eliciting expository language in young school-aged children. This 'favorite game or sport task' (FGS) requires the child to speak from his or her own experience, thus allowing the child to draw from domain-specific topic knowledge (see Nippold, 2010), and was used successfully with 8-year-old children (Nippold, Hesketh, et al., 2005). Utilizing the same task and elicitation protocol as the one used in previous research will not only extend the research by Nippold and colleagues to a younger age group and population, but will also allow for comparisons between the studies.

In summary, the current study aimed to replicate and extend previous research by Nippold, Hesketh, et al. (2005) into expository discourse. First, we investigated the spoken language performance of a reasonably large group of young school-aged children to determine the effectiveness of the FGS task in eliciting expository discourse in 6- and 7-year-old children, with the aim of creating norms for this age-group. The expository discourse samples were analyzed on a range of spoken language measures to gain insight into children's syntactic ability, verbal productivity, and mazing behavior. Second, we compared the performance of these young school-aged children to that of a group of 11-year-old children from an earlier study (Nippold, Moran, Mansfield, & Gillon, 2005) to determine the effects of age on measures of verbal productivity, syntactic ability, and verbal fluency. More specific, in the present study we addressed the following questions:

1. Do 6- and 7-year-old children with typical development produce expository discourse in response to the FGS task?
2. What level of performance can be expected of this age-group on a range of language production measures?
3. Does the performance of 6- and 7- year-old children differ from that of 11-year-old children?

Based on the results from previous research, we expected that the FGS task would be effective in eliciting text-level discourse in young school-aged children. It was anticipated that performance of this group of children could be used for normative purposes, despite some variability in language performance. Finally, we expected that the group of 11-year-old children would significantly outperform their younger peers on all language production measures.

Method

Participants

The 6- and 7-year-old participants were recruited from three primary schools located in suburban Auckland, New Zealand (NZ). The schools were awarded mid socio-economic status based on the Ministry of Education ranking system. Each school in NZ is allocated a decile ranking, using Census information from its students' addresses, based on socio-economic indicators (household income, occupation, household crowding, parents' educational qualifications, and income support). These deciles are typically used by the government to determine the level of funding to state schools; the lower the school's decile, the more funding the school receives. Teachers were asked to hand out information sheets to parents of children who (a) were between the ages of 6;0 and 7;11 (years;months); (b) had no known history of hearing disorder, neurological disorder, or speech-language therapy; (c) spoke English as their first language; and (d) were progressing normally at school. A total of 65 consent forms were returned. To verify children demonstrated typical receptive vocabulary skills, all children were assessed on the Peabody Picture Vocabulary Test – Third Edition (PPVT-III; Dunn & Dunn, 1997); only children whose standard scores fell between 76 and 124 were included in the study; four children were excluded based on standard scores > 124. The group's mean score on the PPVT-III was 104.8 (*SD* = 10.9, range = 80 – 122). Thus a total of 61 children, 36 girls and 25 boys, participated from NZ European (72%), Maori (15%), Pasifika (9%), and "other" (4%) ethnic backgrounds.

In the 11-year-old age group, 20 children, between 11;0 and 11;11 (*M* = 11;4) years of age, participated. These children were originally recruited as part of a larger study comparing expository discourse to conversation across ages and cultures (Nippold, Moran, et al., 2005).

In that study, children, adolescents, and adults (age 8, 11, 17, and 25 years) participated. The 11-year-old participants were recruited from primary schools in suburban Christchurch, NZ, which ranged from mid-low socioeconomic status to mid-high socioeconomic status as measured by the Ministry of Education ranking system. The ethnicity of the group was representative of the schools with 75% NZ European, 12% Maori, 7% Pasifika, and 6% other ethnicity. All participants spoke English as a first language and demonstrated PPVT-III scores (Dunn & Dunn, 1997) within the 76 – 124 range (M = 101.6, SD = 11.1, range = 85 - 122). An independent samples t test revealed no significant differences between the two age groups on the PPVT – III, $t(79) = 1.17$, $p = .26$.

Procedure

For the younger age-group, three undergraduate speech-language therapy students conducted the assessments over a four-week period under close supervision of the first author (a certified speech-language therapist). The students were trained extensively to ensure they adhered to the elicitation protocol, and debriefing sessions were held on a daily basis. All children were seen individually in a quiet room in their school environment on three separate occasions. All sessions were taped using a digital voice recorder (Olympus DM-1). Care was taken to prevent the same examiner from assessing the same child more than twice. During the first session the examiner documented verbal agreement, rapport was established, and the PPVT-III was administered as well as a story retelling task; the second session consisted of a personal narrative and a story retelling task. The favorite game or sport (FGS) task, the focus of this study, was administered during the third and final session.

During this final session, the children were first given an expository retell task, not related to the present study. Second, the FGS task was administered (Nippold, Hesketh, et al., 2005). In this task, the examiner carefully followed a script. In summary, the child was asked what his or her favorite game or sport was and why. The examiner then asked the child to explain the game or sport, using the pragmatically felicitous prompt “I am not too familiar with the game of [..]”. Finally, the child was asked what a player should do to win a game of [..]. For specific information regarding the protocol, see Nippold, Hesketh, et al. (2005). The child was allowed as much time as necessary to finish the explanation. The examiner made sure to show interest in the child’s explanation and only used neutral responses as needed to encourage the child to continue.

For the older age-group, similar elicitation conditions were used. The child was seen individually by a speech-language therapy student in a quiet room at the child’s school. The examiner first administered the PPVT-III and then engaged the child in a 5-minute conversation, followed by the FGS task as described above and in Nippold, Hesketh, et al. (2005).

Transcription and Analysis

All audio files were transferred to a computer, using Olympus DSS Player Pro Dictation Module (Version 4.4.0) software, and transcribed while using headphones and Olympus RS28 foot pedals. Transcripts were coded using Systematic Analysis of Language Transcripts, New Zealand Version conventions (SALT-NZ; Miller, Gillon, & Westerveld, 2008). The transcripts from the younger age group were transcribed by two undergraduate speech-language therapy students who were trained by the first author. Utterance segmentation was based on T-units, defined as one main clause with all its subordinate clauses (Hunt, 1970). Identical procedures as those described by Nippold, Hesketh, et al. (2005) were used. That is, only finite clauses (containing a subject and a predicate) were included. Utterances that did not fit this description were coded as fragments (FRG; on the transcripts) and excluded from further analysis. The full transcripts were used, regardless of length. Only complete and intelligible (C&I) T-units were used for analysis (i.e., T-units containing unintelligible segments or unfinished utterances were excluded). All reformulations, repetitions, and disfluencies were placed in parentheses and considered mazes.

To ensure accuracy and completeness of transcription and coding, the following procedure was used. The first author (a) checked all transcripts for spelling, error coding (see below), and/or utterance segmentation errors, (b) listened to the sound files if the transcript contained any unintelligible segments or utterance segmentation issues, (c) made corrections when needed, and (d) categorized each dependent clause as a relative clause (RC), an adverbial clause (AVC), or a nominal clause (NOM). An academic colleague, who is considered an expert in linguistic analysis, assisted with the dependent clause categorization when needed. Second, 20% of the transcripts and their corresponding sound files were randomly selected. An independent examiner (a researcher experienced with language sample analysis) listened to these sound files and checked for transcription accuracy (percent words agreement = 99.4%, range = 97.1 - 100% per transcript). There was one instance of a disagreement in utterance segmentation. There were no disagreements in coding of mazes. Finally, the second author checked all T-units, fragments, and codes (dependent clauses and grammatical errors) and highlighted any disagreements with the initial segmentation or coding. For each type of clause, the percentage of agreement between the two investigators was as follows: 1055 instances of agreement in utterance segmentation (total 1060) = 99.6% agreement; 350 instances of agreement in dependent clause coding (total 358) = 97.8% agreement; 63 instances of agreement in RC coding (total 66) = 95.5% agreement; 199 instances of agreement in AVC coding (total 205) = 97.1% agreement, and 83 instances of agreement in NOM coding (total 87) = 95.5% agreement. In addition, there were 175 instances of agreement in grammatical error coding (total 177) = 98.9% agreement. All disagreements were resolved through discussion so that 100% agreement was attained for all clause types and grammatical errors.

For the 11-year-old group, an independent examiner experienced with language sample analysis checked the coding accuracy of all the transcripts. Agreement for clauses and T-units was as follows: (a) for T-units agreement was 99.5%, (b) for adverbial clauses 96.4%, (c) for nominal clauses 93.8%, and (d) for relative clauses 92.4%. The disagreements were reviewed and any issues were resolved to 100% agreement for all coding of clauses and T-units. For mazes, there was one instance of disagreement on a mazed utterance and that was adjusted.

Language Measures

We selected the measures based on several criteria. First, only language production measures that are known to be sensitive to age and language ability were included. Second, to allow for comparison with previous research, only frequently used measures were selected. Finally, for conceptually related variables that showed a high correlation ($r > .90$), only one measure was included. As a result, number of different words was not included, as it was highly correlated to the overall length of the sample ($r > .90$). As a result, the following measures remained:

Verbal productivity. The child's verbal productivity was measured as:

- Total length of the sample in number of T-units (Total T-units). Total T-units was calculated automatically using SALT. Total T-units only contained complete and intelligible utterances; fragments were excluded (see Appendix B). Previous research into language sample length has found age-related changes for expository discourse in individuals with typical development (Berman & Verhoeven, 2002; Nippold, Hesketh, et al., 2005), and sensitivity of this measure for language ability using a narrative generation task (Fey, Catts, Proctor-Williams, Tomblin, & Zhang, 2004).

- Rate (word per minute: WPM). To calculate WPM, the child's total number of completed words (including those in mazes and fragments) was divided by the total elapsed time in minutes. Research into WPM in conversation, narrative, and expository contexts has shown this measure to be sensitive to age (Heilmann, Miller, & Nockerts, 2010) and language ability (Scott & Windsor, 2000).

Syntactic complexity. Syntactic complexity was measured as:

- Mean length of T-unit (MLTU). Utterance length is a known indicator of advanced language development (e.g., Nippold, 2007). In addition, MLTU has been found to be sensitive to language ability (Scott & Windsor, 2000).
- Clausal density (CD). CD was calculated as the total number of clauses divided by the total number of T-units. Following Nippold, Hesketh, et al.'s (2005) coding process, only finite clauses (independent and subordinate) were identified and coded; non-finite clauses were ignored. The following dependent clauses were coded: (a) Relative clause use (PcRC; percentage of T-units containing a relative clause), (b) Adverbial clause use (PcAVC), and (c) Nominal clause use (PcNOM). Please refer to Appendix A for further information on dependent clause coding. In addition, Appendix B shows examples of clause structures produced by one of the participants. Clausal density is a known indicator of advanced language development (e.g., Nippold, 2007). For example, in a cross-sectional study of syntactic development in children, adolescents, and adults (age 7;8 to 49;9 years), the two best growth indicators were MLTU and production of relative clauses (Nippold, Hesketh, et al., 2005).

Grammatical accuracy. Grammatical accuracy was calculated as the percentage of grammatically accurate T-units (GA) (Fey et al., 2004). Scott and Windsor (2000) found this measure to be particularly sensitive to language ability. To code for grammatical accuracy, all T-units containing an error that rendered them grammatically incorrect (e.g., verb tense error, omitted auxiliary/copula, omitted verb, incorrect conjunction, omitted bound morpheme, or incorrect word order) were considered “not grammatically accurate” (see Fey et al., 2004, for more details). Widely accepted colloquial grammatical constructions (e.g., “There’s cars coming”) were not counted as incorrect.

Verbal fluency. Verbal fluency was calculated as the percentage of maze words (% Maze words). Mazing behavior (i.e., repetitions, false starts, reformulations) (Loban, 1976) has been linked to sentence length and grammatical complexity in studies involving morpho-syntactic development in preschool children (Rispoli & Hadley, 2001). More specific, Rispoli and Hadley found that sentences containing mazes tended to be longer and more complex than fluent ones. Moreover, excessive use of mazing behavior may indicate linguistic vulnerability, especially when the cognitive demands of a task increase (MacLachlan & Chapman, 1988; see also Westerveld & Gillon, 2008).

Results

Preliminary Analyses

The results were analyzed using statistical software (PASW, 2008, version 18). For group comparisons, Eta squared (η^2) values were generated for each of the analyses as an estimate of the effect size. These values document the amount of explained variance in a dependent variable as a function of group. Interpretation of the effect size of η^2 is as follows: small effect size: $\eta^2 < .06$, medium effect size: $\eta^2 = .06 - .15$, large effect size: $\eta^2 > .15$ (Cohen, 1988).

In general, and answering the first research question, we found that the FGS task was effective in eliciting extended discourse in 6- and 7-year-old children with typical development. All children participated in this task and produced on average 17.4 T-units, although two children produced as few as 4 T-units. To determine the appropriateness of treating this group of 6-and 7-year-old children as one age-group, linear regression analyses were performed for each language production measure, using age in months as the independent variable. Age effects were not significant for any of the measures ($df = 59$): Total T-units ($r^2 = .004$, $p = .62$), WPM ($r^2 = .01$, $p = .45$), MLTU ($r^2 = .018$, $p = .30$), CD ($r^2 = .002$, $p = .71$), PcRC ($r^2 = .003$, $p = .69$), PcAVC ($r^2 = .009$, $p = .47$), PcNOM ($r^2 < .001$, $p = .90$), GA ($r^2 < .001$, $p = .90$), or % Maze words ($r^2 = .002$, $p = .76$). Because of these non-significant results for age, the children's performance was analyzed as one group; the results are shown in Table 1.

Insert Table 1 here

Examination of the correlational data between the language measures showed variability in the strength of the associations between different measures. The correlation coefficients are reported in Table 2. As expected, the two verbal productivity measures (Total T-units and WPM) were highly correlated. When investigating the relationship between MLTU and clausal density (CD; see grey shaded area in Table 2), it appeared that MLTU was a good indicator of clausal density ($r = .72$). Adverbial clause use in particular was highly correlated with CD ($r = .86$). The correlation between GA and CD was also significant ($r = -.35$), and GA showed a moderate negative correlation with MLTU ($r = -.44$), indicating that GA decreased with increasing MLTU and increasing CD.

Insert Table 2 here

Distribution Statistics

To investigate the normality of the sample's distribution, and thus the potential of using these data for normative purposes, we performed several calculations. These included mean and median scores, percentile scores, and skewness and kurtosis statistics. As shown in

Table 1, most mean scores (except for MLTU, CD, and GA) were above the median scores, indicating off-center distributions. Five of the measures showed a significantly skewed distribution with scores clustering at the low end of the scale (i.e., skewness statistic > .61), and one measure showed a significantly skewed distribution where scores clustered at the high end of the scale (GA). Closer investigation of the Kurtosis statistics (investigates if the distribution is peaked or flat) showed that 4 of the 9 measures (PcRC, PcAVC, PcNOM, and GA) showed a significantly peaked distribution (i.e., more than two standard errors of kurtosis > 1.21).

In summary, these results show that the performance of the 6- and 7-year-old age group on the expository task was not normally distributed. Performance on measures of sample length (Total T-units) and dependent clause use clustered on the low end of the distribution spectrum (floor effect), whereas performance on the GA measure clustered on the high end of the scale (ceiling effect). The only measures that showed symmetrical and normally shaped distributions were MLTU, % Maze words, and WPM.

Individual Variation

As indicated by the large standard deviations, there was considerable variability in performance on most of the measures. This was particularly noticeable on the use of dependent clauses, i.e., relative, nominal, and adverbial clauses. The range of performance was also large. For example, two children produced only 4 T-units, whereas three children produced more than 40 T-units. To determine if this variability was caused by outliers, closer inspection of the data was performed. Results revealed fewer than 3 outliers per language measure, indicating that the individual variation seen in the results most likely reflects the use of spontaneous language sampling as a context for linguistic analysis.

Dependent Clause Use

To further investigate the children's use of complex language, the percentage of children using different types of dependent clauses, irrespective of the number of utterances contained in the language sample, was calculated. As shown in Figure 1, 49% of the children produced at least one relative clause, 70% of the children used at least one nominal clause, and 75% of the children used at least one adverbial clause. Approximately 11% of the children produced no dependent clauses at all.

Differences in Performance between Age Groups

Table 3 shows the performance of both the 6- and 7-year-old group and the 11-year-old comparison group on the language production measures. To determine whether there were differences in performance between the two age groups, univariate analysis of variance (ANOVA) tests were completed using each of the language sample measures as the dependent variable and group (6- and 7-year-olds versus 11-year-olds) as the between-subjects variable. Results indicated there were significant effects for age group on measures of Total T-units, $F(1,79) = 25.27$, $p < .001$, $\eta^2 = .24$; % Maze words, $F(1,79) = 5.35$, $p < .05$, $\eta^2 = .06$; and GA, $F(1,79) = 16.87$, $p < .001$, $\eta^2 = .18$. The older age group produced longer samples, showed greater verbal fluency, and demonstrated better grammatical accuracy than the younger age group. There were no significant effects for age group for MLTU, $F(1,79) = 1.40$, $p = .24$, $\eta^2 = .02$; CD, $F(1,79) = .44$, $p = .51$, $\eta^2 = .01$; nor for any of the dependent clause measures PcRC, $F(1,79) = 2.37$, $p = .13$, $\eta^2 = .03$; PcAVC, $F(1,79) = 0.004$, $p = .95$, $\eta^2 = <.001$; or PcNOM, $F(1,79) = .08$, $p = .78$, $\eta^2 = .001$.

Insert Table 3 here

To further investigate potential age-related changes in syntactic performance, we wanted to establish what percentage of the children in each age group demonstrated dependent clause use. It was found that 80% of the 11-year-old children used at least one relative clause (compared to 49% of the younger children, see Figure 1), 90% of the children used at least one nominal clause (70.5% of the younger children), and 90% used at least one adverbial clause (75.4% of the younger children), irrespective of the total length of the expository discourse sample.

Insert Figure 1 here

Discussion

In this study we investigated the expository discourse of 61 children with typical development, between ages 6;0 and 7;11, using the favorite game or sport task (FGS, see Nippold, Hesketh, et al., 2005). All children were in their second or third year of primary school education. The first question this study addressed was whether the expository sampling context was effective in eliciting extended text-level discourse in young school-aged children. The results confirmed the hypothesis. All children were happy to participate and produced, on average, approximately 17 T-units. Individual variation was large, however, with nearly 9% of the children producing less than 6 T-units. Although these results indicate that the children in this study had enough knowledge about their favorite game or sport to attempt an explanation, they may not have had a clear mental model for this type of discourse to provide a lengthy exposition. It may be assumed that children at this age and stage of schooling will have had limited practice in providing factual explanations (see Snyder & Caccamise, 2010, for a discussion). It is reasonable to expect that, with exposure and practice, children will learn to provide longer, more detailed, and structured explanations about games and/or sports, with the purpose of informing the listener. It is interesting to note, however, that the individual variation in sample length is consistent with the results from Nippold, Hesketh, et al. (2005), who found wide ranges (with relatively large standard deviations) on this measure in all their age groups (8, 11, 13, 17, 25, and 44) in both expository and conversational contexts. These results once more highlight the inherent variability in language production measures derived in spontaneous language sampling situations (see Miller, Heilmann, Nockerts, Andriacchi, & Iglesias, 2006, for a discussion).

Next, we investigated the level of performance of the 6- and 7-year-old age group to determine the potential of obtaining valid normative data of expository discourse performance using the FGS task. Results showed that some language production measures were normally distributed, namely MLTU, % Maze words, and WPM (rate). Most other measures showed skewed distributions with scores that clustered on the low end of the scale (Total T-units, clausal density [CD], and percentage of dependent clause use). The exception was grammatical accuracy (GA) that showed scores clustered on the high end of the scale. These abnormal distributions were not unexpected, and should not preclude the clinical use of these data. As mentioned previously, it has now been well established that there is large variability in performance on spontaneous language tasks, adding to the debate of whether spontaneous language sampling can be standardized (Miller et al., 2006). One important argument in favor is that spontaneous language sampling measures have been shown to be sensitive to language status, with measures such as WPM, MLTU, and grammatical ability showing particular sensitivity (Heilmann et al., 2010). Although further research into expository language in young school-aged children with language impairments is needed, comparing children's individual performance to the normative data of the present study may be clinically useful in describing strengths and weaknesses in expository language performance. Results from this study suggest that several measures may be used to help establish impaired language performance, including MLTU, percent maze words, and WPM. Because of the skewed distributions and comparatively large standard deviations on measures of GA, CD, and sample length, comparing a child's performance on these measures to the percentile values (see Table 1) will be more appropriate and will provide an indication of the child's relative performance compared to his or her peers.

To investigate the complexity of the children's language performance in expository discourse, the syntactic measures were analyzed more closely. Results indicated limited use of dependent clauses in this age group, as measured by the percentage of T-units containing a particular dependent clause. When investigating the number of dependent clauses used, however, irrespective of the length of the transcript, approximately 50% of the children produced at least one relative clause (see Figure 1), 70% used at least one nominal clause, and 75% of the children used at least one adverbial clause.

Several analyses were conducted to explore patterns of relationships between the different language production measures. Consistent with Nippold, Hesketh, et al.'s (2005) work, significant correlations were found between MLTU and clausal density measures. This suggests MLTU is a fairly good predictor of dependent clause use in this age group and will be useful clinically to provide an indication of the child's syntactic ability. Analysis of the correlations between the language production measures also revealed a moderate negative correlation between GA and MLTU (see Table 2), indicating that longer utterances were associated with lower grammatical accuracy. Consistent with a limited capacity working memory model, a trade-off between linguistic behaviors (see Crystal, 1987, for a discussion) in this cognitively challenging context was not unexpected. Faced with the challenge of conveying longer linguistic units, children's grammatical accuracy suffered. These findings therefore support the usefulness of the GA measure to capture a break-down in linguistic performance when constructing longer T-units. Finally, it was noted that there was no correlation between mazing behavior and any of the other language measures. This was somewhat unexpected. One possible explanation is that grammatical accuracy is more susceptible to cognitive demands than mazing behavior. Alternatively, the measure of mazing behavior itself may not have been sensitive enough. To check this, additional analyses were performed, using the mean number of mazes per T-unit, and the percentage of T-units containing mazes as the dependent variables. There were high correlations between these three mazing measures (i.e., $r > .80$). Both the mean number of mazes per T-unit and the percentage of T-units containing mazes showed a significant negative correlation with GA, but not with any of the other measures, indicating sensitivity of these measures of mazing behavior to grammatical accuracy. Future research should investigate which measure of mazing behavior is most sensitive to language status and developmental change.

The third research question addressed age-related changes in performance on the FGS task. It was expected that 11-year-old children, after approximately five years of primary school education would show significantly better performance on measures of verbal productivity, verbal fluency, and grammatical ability than the group of 6- and 7-year-old children. The results partly supported the hypothesis. It was found that the older group produced significantly more T-units, containing a lower percentage of maze words, than the younger group. Moreover, the 11-year-old children demonstrated a higher grammatical accuracy. However, no significant differences were found on measures of syntactic complexity, including MLTU and CD.

The increase in productivity (as measured by Total T-units) with level of schooling is consistent with previous research into expository discourse in school-aged populations (Berman & Verhoeven, 2002) as well as our own previous cross-sectional research investigating the effects of age on verbal productivity in a fictional narrative retelling context (Westerveld et al., 2004). Interestingly, Nippold, Hesketh, et al. (2005) failed to find effects for age when comparing the 8- to 11-year-old age groups on verbal productivity using the FGS task. It seems feasible that the smaller age difference used in Nippold, Hesketh, et al.'s study explains the difference in findings. As stated previously, the children in the current study were beginning readers and had presumably not made the transition to 'learning through reading', which typically occurs from the age of eight (grade 4). In addition, it is from about grade 4 that children are more exposed to expository type materials at school (see Snyder & Caccamise, 2010, for a discussion), and one might expect children to show accelerated development in verbal productivity during that transitional stage. Although the superior performance in verbal productivity of the 8-year-old group in Nippold, Hesketh, et al.'s (2005) study compared to the 6-and 7-year-olds in the current investigation lends support to this hypothesis, future research using a longitudinal design will help to confirm this assumption.

It was interesting to find significant age-related improvements in verbal fluency (reduced mazing behavior) as well as in grammatical accuracy (GA). Not only did the older children speak more, they were also better at constructing their sentences, as signified by their almost faultless use of grammar (mean of 94.7%) and reduced mazing behavior. In contrast, the lack of a significant impact of age on the children's syntactic complexity was somewhat unexpected. As a group, the 11-year-old children used only slightly longer sentences, with slightly higher clausal density. These results indicate that level of schooling may not affect syntactic complexity as measured by MLTU or clausal density. When investigating age-group performances using the same task, Nippold, Hesketh, et al. (2005) found a significant increase in MLTU from 13 to 17 years of age (but not between 11 and 13 years of age). Perhaps an improvement in syntactic complexity (as measured by MLTU or CD) occurs following primary school education as children get exposed to increasingly more formal language situations and tasks, including written essays and expository textbooks.

A different explanation regarding the lack of sensitivity for age of the MLTU measure relates to a trade-off between linguistic skills within a limited processing capacity model (see Crystal, 1987). It is interesting to note that results from previous research have consistently indicated a small decline in performance on MLTU as well as clausal density around the age of 13 or grade 9 (e.g., Leadholm & Miller, 1992; Loban, 1976; Nippold, Hesketh, et al., 2005), regardless of which elicitation context was used. Rather than attributing this decline to the (lack of) sensitivity of the language production measure (i.e., MLTU), perhaps a stronger explanation is that cognitive resources are allocated to other linguistic processes, such as choice of vocabulary and grammatical accuracy. Consistent with this line of thinking, results from the current project showed a significant increase in verbal productivity as well as a significant improvement in grammatical accuracy and verbal fluency. In other words, despite demonstrating a mean utterance length that is similar to their younger peers, the 11-year-old children's explanations were longer, more fluent, and grammatically more accurate than those of their 6- and 7-year-old peers.

Limitations

There are several limitations to the present research. First, the two age groups were seen by different research groups for different studies, resulting in small procedural differences. The younger children participated in this task during the third assessment session, which could have given them a slight advantage over the older children, whose expository skills were elicited during the first session. That is, the younger children may have felt more at ease to participate in the task which, theoretically, could have inflated their performance. In addition, the older group came from a wider range of socio-economic backgrounds than the younger children and it is not clear if this would have affected the results.

It is not known if the results from this research would generalize to other geographical locations. As mentioned earlier, children in NZ start school at the age of 5, which means that the 6-and 7-year-old children in the current study have had between one and three years of formal primary school education. It is reasonable to expect that exposure to expository material at school may have positively influenced this group's ability to participate in the FGS task. It is interesting to note, however, that previous research comparing expository discourse performance of 11- and 17-year-old children from the US and NZ revealed no significant differences on measures of syntactic complexity (Nippold, Moran, et al., 2005). Similarly, when comparing the oral narrative production skills of 6- and 7-year-old children from the US and NZ in a story retelling context (Westerveld & Heilmann, 2010), no differences were found on measures of syntactic complexity or verbal productivity. Further research, across geographical locations, is needed to investigate the effects of “year of schooling” on young school-aged children’s ability to produce expository discourse.

Summary and Clinical Implications

Results from the present investigation contribute to our expanding knowledge base of children’s spontaneous language use in an expository discourse situation. The findings clearly indicate that the FGS task is successful in eliciting text-level discourse in young school-aged children. Integration of the transcripts into a software program such as Systematic Analysis of Language Transcripts (SALT-NZ; Miller et al., 2008) will potentially provide clinicians with easy access to local normative data for young school-aged children on spoken language performance that is relevant to the academic setting.

Acknowledgments

This research was supported by a grant from the Massey University Research Fund. The technical and financial support from SALT Software, LLC is gratefully acknowledged. Thanks are extended to the children and the children’s class teachers for supporting this research and for sharing their stories.

References

Berman, R. A., & Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: Speech and writing. *Written Language and Literacy, 5*(1), 1-43.

Chall, J. S. (1996). *Stages of reading development*. Fort Worth, TX: Harcourt Brace College Publishers.

Cohen, J. (1988). *Statisticial power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crystal, D. (1987). Towards a "bucket" theory of language disability: Taking account of interaction between linguistic levels. *Clinical Linguistics and Phonetics, 1*, 7-22.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test - III*. Circle Pines, MN: American Guidance Service.

Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, J. B., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research, 47*(6), 1301-1318.

Fujiki, M., Brinton, B., Morgan, M., & Hart, C. H. (1999). Withdrawn and sociable behavior of children with language impairment. *Language, Speech, and Hearing Services in Schools, 30*, 183-195.

Heilmann, J. J., Miller, J. F., & Nockerts, A. (2010). Using language sample databases. *Language, Speech, and Hearing Services in Schools, 41*(1), 84-95.

Hunt, K. W. (1970). Syntactic maturity in school children and adults. *Monographs of the Society for Research in Child Development, 35*(Serial No.134, No.1).

Leadholm, B. J., & Miller, J. F. (1992). *Language sample analysis: The Wisconsin guide*. Madison, WI: Wisconsin Department of Public Instruction.

Loban, W. (1976). *Language development: Kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.

MacLachlan, B. G., & Chapman, R. S. (1988). Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of Speech and Hearing Disorders*, *53*, 2-7.

Masterson, J. J., & Kamhi, A. G. (1991). The effects of sampling conditions on sentence production in normal, reading-disabled, and language-disabled children. *Journal of Speech and Hearing Research, 34*, 549-558.

Miller, J., Gillon, G., & Westerveld, M. (2008). SALT NZ: Systematic Analysis of Language Transcripts - New Zealand Version. Madison, WI: SALT Software.

Miller, J. F., Heilmann, J., Nockerts, A., Andriacchi, & Iglesias. (2006, November). *Can language sample analysis be standardized?* Paper presented at the ASHA Convention, Miami Beach, Florida.

Ministry of Education. (2007). *The New Zealand Curriculum for English-medium teaching and learning in years 1 - 13*. Wellington, New Zealand: Learning Media.

Nippold, M. A. (2004). Research on later language developments: International perspectives. In R. A. Berman (Ed.), *Language development across childhood and adolescence: Volume3. Trends in language acquisition research (TILAR)* (pp. 1-8). Amsterdam: John Benjamins.

Nippold, M. A. (2007). *Later language development: School-age children, adolescents, and young adults* (3rd ed.). Austin, TX: Pro-ed.

Nippold, M. A. (2010). Explaining complex matters: How knowledge of a domain drives language. In M. A. Nippold & C. M. Scott (Eds.), *Expository discourse in children, adolescents, and adults* (pp. 41-61). NY: Psychology Press.

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*(5), 1048-1064.

Nippold, M. A., Moran, C. A., Mansfield, T. C., & Gillon, G. T. (2005, July). *Expository discourse development in American and New Zealand youth: A cross-cultural comparison*. Paper presented at the the 10th International Congress for the Study of Child Language, Berlin, Germany.

PASW. (2008, version 18). [Computer Software]. Chicago, IL: IBM SPSS Statistics.

Rispoli, M., & Hadley, P. (2001). The leading-edge: The significance of sentence disruptions in the development of grammar. *Journal of Speech, Language, and Hearing Research, 44*(5), 1131-1143.

Schneider, P., Dubé, R. V., & Hayward, D. (2009). The Edmonton Narrative Norms Instrument. Retrieved from http://www.rehabmed.ualberta.ca/spa/enni

Scott, C. M., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research, 43*(2), 324-339.

Snyder, L., & Caccamise, D. (2010). Comprehension processes for expository text: Building meaning and making sense. In M. A. Nippold & C. M. Scott (Eds.), *Expository discourse in children, adolescents, and adults* (pp. 13-39). NY: Psychology Press.

Westerveld, M. F., & Gillon, G. T. (2008). Oral narrative intervention for children with mixed reading disability. *Child Language Teaching and Therapy, 24*(1), 31-54.

Westerveld, M. F., Gillon, G. T., & Miller, J. F. (2004). Spoken language samples of New Zealand children in conversation and narration. *Advances in Speech-Language Pathology*, *6*(4), 195-208.

Westerveld, M. F., & Heilmann, J. (2010, June). *Narrative analysis: The effects of geographic location on children's spoken language performance*. Paper presented at the Symposium on Research in Child Language Disorders, Madison, Wisconsin.

Table 1

Performance of the 6-And 7-Year-Old Children on the Language Production Measures (N=61)

	Mean (*SD*)	Median	Percentiles				Kurtosis	Skewness
Measure			10	25	50	90		
Total T-units	17.4 (9.8)	16.0	6	10.5	16	32.8	0.53	0.97*
MLTU	8.67 (1.62)	8.71	6.3	7.7	8.7	10.8	0.18	0.03
% Maze words	14.1 (7.2)	13.0	5.0	9.0	13.0	24.8	- 0.54	0.35
PcRC	5.3 (7.2)	0	0	0	0	16.2	2.48*	1.58*
PcNOM	9.0 (9.5)	6.3	0	0	6.3	24.4	1.30*	1.28*
PcAVC	18.8 (16.5)	18.2	0	3	18.2	40.8	1.87*	1.04*
CD	1.33 (0.22)	1.33	1.0	1.18	1.33	1.63	0.91	0.65*
GA	82.0 (13.4)	83.3	64.8	75	92.1	99.3	1.33*	- 1.07*
WPM	65.5 (23.5)	64.2	33.5	50.6	64.2	97.9	- 0.23	0.17

Note. MLTU = mean length of utterance in T-units; PcRC = percent T-units containing a relative clause. NOM = nominal clause; AVC = adverbial clause; CD = clausal density; GA = grammatical accuracy in percent grammatically correct utterances; WMP = words per minute.

* indicates that the performance on that measure shows a significant level of skewness / kurtosis.

Table 2

Correlations Between the Language Production Measures Produced by the 6- and 7-Year-Old Group

	Total T-units	**MLTU**	**%Mzwds**	**PcRC**	**PcNOM**	**PcAVC**	**CD**	**GA**	**WPM**
Total T-units	1.00	.16	-.08	.24	-.11	.06	.08	.17	.83**
MLTU	---	1.00	-.08	.42**	.28*	.60**	.72**	-.44**	.40**
% Maze words	---	---	1.00	-.12	-.03	-.15	-.16	-.22	-.09
PcRC	---	---	---	1.00	-.05	.18	.44**	-.02	.30*
PcNOM	---	---	---	---	1.00	.09	.46**	-.25	.04
PcAVC	---	---	---	---	---	1.00	.86**	-.35**	.20
CD	---	---	---	---	---	---	1.00	-.35**	.27*
GA	---	---	---	---	---	---	---	1.00	.07
WPM	---	---	---	---	---	---	---	---	1.00

Note. MLTU = mean length of utterance in T-units; PcRC = percent T-units containing a relative clause. NOM = nominal clause; AVC = adverbial clause; CD = clausal density; GA = grammatical accuracy in percent grammatically correct utterances; WMP = words per minute. The shaded area highlights the correlations between the syntactic measures.

**p* <.05. ***p* < .001

Table 3.

Group Performance (with Standard Deviations) on the Language Production Measures

Measure		Age 6 /7	Age 11	Effect sizes
		n = 61	*n* = 20	η^2
Total T-units	*M* (*SD*)	17.4 (9.8)	32.4 (15.9)	.24**
	Range	4 - 44	8 - 60	
MLTU	*M* (*SD*)	8.7 (1.62)	9.1 (1.3)	.02
	Range	5.0 – 13.0	6.9 – 12.1	
% Maze words	*M* (*SD*)	14.1 (7.2)	10.1 (4.9)	.06*
	Range	0 - 30	3 - 21	
PcRC	*M* (*SD*)	5.3 (7.2)	8.0 (5.8)	.03
	Range	0 – 31.8	0 – 20.8	
PcNOM	*M* (*SD*)	9.0 (9.5)	9.6 (6.3)	.001
	Range	0 – 41.2	0 - 25	
PcAVC	*M* (*SD*)	18.8 (16.5)	18.6 (10.8)	<.001
	Range	0 - 80	0 - 35	
CD	*M* (*SD*)	1.33 (0.22)	1.36 (0.13)	.01
	Range	1.0 – 2.0	1.1 – 1.53	
GA	*M* (*SD*)	82.0 (13.4)	94.7 (5.3)	.18**
	Range	40.0 - 100	80.0 - 100	

Note. MLTU = mean length of utterance in T-units; PcRC = percent T-units containing a relative clause. NOM = nominal clause; AVC = adverbial clause; CD = clausal density; GA = grammatical accuracy in percent grammatically correct utterances; WMP = words per minute.

* $p < .05$. ** $p < .001$

Appendix A

Dependent Clause Coding

Mirroring Nippold, Hesketh, et al.'s (2005) coding procedure, only finite independent and dependent clauses were coded, i.e., clauses that contain a subject and a predicate. Independent clauses were coded as [IC]. The following three types of dependent clauses were identified and coded:

Adverbial clauses [AVC] begin with a subordinating conjunction. Examples include:

- And if they get the highest number [AVC] when the game's finished [AVC], they win [IC].
- And then once you've done that [AVC] (uhm) we pull out the blue mats and the (o other k) white mat [IC].
- And if you remember that [AVC] and you don't get hit [AVC] you win the game [IC].

Relative clauses [RC] describe a noun and generally immediately follow the noun they describe.

- But we (like) have to hit the person [IC] who's (um) doing that [RC].
- And he brings me to all the games [IC] that I can go to [RC].
- And you've got lines [IC] where you're allowed to go up to [RC].

Nominal clauses name persons, places, things or ideas. These clauses often answer the question 'what'?

- And whoever grabs the ball (um) [NOM] they (um) get to start with the ball in centre [IC].
- And that's [IC] how they lose the game sometimes [Nom].

- And whoever finishes all their beads [NOM] wins [IC].

Appendix B

Expository Discourse Sample Produced by T14, a 7;4-Year-Old Boy.

T14 produced 30 T-units, MLTU = 9.7, and % Maze words = 8%, Clausal Density = 1.43.

+ Topic: game_of_life boardgame

E what is your favorite game or sport?

C the_game_of_life [FRG].

E oh and why is life your favorite game?

C because it's a big board [IC].

C and it has lots of cards [IC] if you do it [AVC].

C and if you're college [AVC], you get to be first [IC].

C but it's a longer track [IC].

C but if you're doing career (uhm) job [AVC], you can't get the college cards [IC].

C but if you go to college part [AVC], you get lots of money and things [IC].

E mhm.

+ question three - how to play the game

C there's only four players [IC].

C and (um) if you jump on the yellow part [AVC], which is change_job [RC], you get another career card [IC].

C but you only pick one of them [IC].

C because (you one of) you just pick two [IC].

C and then you just read it [IC].

C then you just choose it [IC].

C and you just do it [IC].

C and then if you jump on home_card [AVC], you get a new home [IC].

C and Sponge_Bob and Patrick and Squidwort and Mr_Crabs (have to get) one of them have[EW:has] to get all the way to (um) the end of the board [IC] before (um) time runs out [AVC][NGA].

C and that's only (um) 100 minutes [IC].

C but (the um) the game board has (only um) lots of things [IC] that are white [RC].

C and sometimes (um) my sister and I just lift it up a little [IC].

C and we just pretend [IC] that it's in (like) water [Nom].

C so we just like it (like) that way [IC].

+ question four - how to win the game

C (um) you just have to (um) try your best by spinning (um) that number thing [IC].

C but it's not a dice [IC].

C it's just a spinning_top [IC].

C and it (shows) shows you the number [IC] that you should go on [RC].

C and then you just do that number [IC].

C but if you have the bigger number [AVC] and it says stop [AVC], you just have to stop there [IC].

C and you get career_job [IC].

C and you just pick it up [IC].

C and then you will know [IC] what it's like [Nom].

E mhm.

C and you'll be that person [IC].

Note. E = examiner; C = child; [FRG] = fragment; [IC] = independent clause; [AVC] = adverbial clause; [Nom] = nominal clause; [RC] = relative clause; [EW] = error at word-level; [NGA] = not grammatically accurate; ( ) = mazes; _ underscore indicates that the compound word or phrase should be regarded as one word.