Spoken Language Samples of Australian Children in Conversation, Narration, and Exposition

Marleen F. Westerveld

Kath Vidler

RUNNING HEAD: Language Samples of Australian Children

Correspondence to

Marleen Westerveld, PhD

Senior Lecturer Speech Pathology

School of Allied Health Sciences

Griffith Institute for Educational Research

Griffith University

Gold Coast, QLD 4222

Phone: 07 5678 7658

M.westerveld@griffith.edu.au

Keywords: school-age children, story retelling, language sample analysis, narratives, conversation, expository

This article was published as:

Westerveld, M. F., & Vidler, K. (2016). Spoken language samples of Australian children in conversation, narration and exposition. *International Journal of Speech-Language Pathology, 18*(3), 288-298. doi: 10.3109/17549507.2016.1159332

Abstract

Purpose: Language sample analysis is a powerful clinical tool for identifying and describing the oral language difficulties of children with language impairment. In response to a lack of existing Australian normative data, the current study collected spoken language samples from 127 children attending the first three years of schooling (YOS). This dataset was compared with New Zealand (NZ) databases of language samples to determine whether clinicians can use overseas databases for appraising language performance of Australian children.

Method: Children participated in several oral language tasks: conversation, personal narratives, story retelling, and exposition (YOS3 only).

Result: Analyses of the spoken language samples revealed a developmental trend of increasing syntactic complexity, semantic diversity, and verbal productivity. Discourse genre had a significant impact on children's language production, with the expository task yielding the syntactically most complex language from the YOS 3 children. Comparisons between the Australian and NZ datasets revealed some differences in performance with the Australian children showing better syntactic complexity.

Conclusion: The Australian dataset of language samples provides clinicians with useful information regarding young school-age children's performance on a range of discourse tasks deemed important for classroom participation. Care should be taken when using the NZ database for diagnostic purposes.

Spoken Language Samples of Australian Children in Conversation, Narration, and Exposition

Introduction

Young school-age children's oral language proficiency is critical to their classroom participation and academic success (Australian Curriculum Assessment and Reporting Authority [ACARA], 2012; Catts & Hogan, 2003; Milosky, 1987). At school, children are expected to participate in a wide range of tasks involving different discourse genres including conversation, narratives; both fictional and personal, and exposition. It is therefore surprising that, to the authors' knowledge, no published norms exist on which to base our expectations of typical oral language performance for young school-age children in Australia. An enhanced understanding of children's typical performance on oral language tasks that are relevant to the Australian Curriculum will assist both educators and speech-language pathologists in identifying children who show weaknesses in spoken language performance compared to their peers. Detailed analysis of these children's spoken language samples will assist in intervention planning and in monitoring of intervention effectiveness. In response to the lack of existing Australian data, the current small-scale study investigated the spoken language skills of young school-age children attending the first three years of their primary school education in Queensland, Australia.

Discourse Genres

This study focuses on three discourse genres that are frequently encountered during the early school years: oral narratives, conversation, and expository discourse (ACARA, 2012; Hughes, McGillivray, & Schmidek, 1997; Westerveld & Moran, 2011). The first genre, oral narratives, has typically been investigated in fictional story re/telling conditions (e.g. Heilmann, Miller, Nockerts, & Dunaway, 2010; Justice, Bowles, Pence, & Gosse, 2010; Justice et al., 2006; Petersen, Gillam, & Gillam, 2008; Westerveld & Gillon, 2010). The

importance of oral narrative ability to academic success is well established (see Boudreau, 2008) and many of the reading materials used during the early school years utilise a narrative format (see Snyder & Caccamise, 2010). Despite this attention to fictional narratives, the importance of personal event narratives should not be underestimated. Personal event narratives can be defined as recounts of past events that have been experienced by the speaker and competence in personal narrative ability has been linked to socio-emotional wellbeing and identity development (Reed & Spicer, 2003). During the first year of school, children are expected to "retell events and experiences with peers and known adults" and "share a personal experience, interest or discovery with peers in a semi-formal situation " (ACARA, 2012, pp. 24-26). Moreover, personal event narratives are important to early writing development as children rely on past personal experiences for early writing tasks, such as "journaling".

The second discourse genre this study focuses on is conversation. "Interacting with others" is a critical academic skill that features prominently in the Australian Curriculum. As an example, in their second year of schooling, students are expected to "engage in conversations and discussions, using active listening behaviours, showing interest, and contributing ideas, information and questions" (ACARA, 2012, p. 31). In young school-age children, conversation elicited in an interview format has been found to be an effective way of eliciting relatively complex language samples in a short period of time. During the interview, examiners ask a series of open-ended questions about school and home events (Evans & Craig, 1992). Moreover, results from Heilmann, Debrock, and Riley-Tillman's (2013) recent study investigating young school-age children's performance on a range of language measures elicited during interviews, confirmed the clinical feasibility of using this context for language sample analysis.

The third genre relates to expository discourse. This is a less researched context, especially for young school-age children (Westerveld & Moran, 2011). As Westerveld and Moran argued, expository discourse may well be regarded as too difficult for this age group. However, results from their research into the expository skills of 6- and 7-year-old children from New Zealand showed the effectiveness of a task tapping procedural expository discourse in eliciting complex language samples. At school, there is an increase in the use of expository materials as children move from learning to read to learning through reading (see Snyder & Caccamise, 2010). This is reflected in the Australian Curriculum (ACARA, 2012), where in their third year of schooling, students are expected to a) engage with texts in which the primary purpose is to inform (p. 35), b) listen for specific purposes and information (p. 38), and c) prepare and give oral presentations, using more formal speech and specific vocabulary about content area topics (p. 39).

Schooling is considered a critical factor in the development of advanced linguistic skills (Nippold, 2004). As mentioned previously, at school children are not only exposed to advanced language, they also practise their skills on a regular basis (ACARA, 2012). As a result, we would expect children's language skills to show a developmental increase in complexity and semantic diversity. The current study therefore investigates whether the Australian dataset of language samples is sensitive to year of schooling and/or age.

The Effects of Discourse Genre on Spoken Language Performance

Apart from year of schooling, we would expect significant effects for discourse genre on children's spoken language performance (Leadholm & Miller, 1992; Nippold, Hesketh, Duthie, & Mansfield, 2005; Westerveld, Gillon, & Miller, 2004). For example Westerveld, Gillon, and Miller (2004) investigated the effects of elicitation context on 4- to 7-year-old children's spoken language performance. Consistent with the notion that "complex thought drives the development of complex language" (Nippold et al., 2005, p. 1 058), it was found

that the narrative contexts of story retelling and personal event narratives yielded syntactically more complex language samples (in mean length of utterance; MLU) than the conversational context (interview format) across all age groups. Few studies, however, have compared school-age children's performance in expository generation to conversation or personal narrative generation. One exception is the work by Nippold et al. (2005) who investigated syntactic development in children, adolescents, and adults (ages 7 to 49 years) across conversational and expository discourse. Results indicated greater syntactic complexity in the expository context compared to the conversational context, as measured in MLU and clausal density. The current study is the first study to compare children's spoken language performance across three discourse generation genres. An increased understanding of the effects of discourse genre on young school-age children's language generation skills will not only improve our knowledge of children's syntactic proficiency during the early school years, it will likely influence language sample analysis practices. Clinicians may decide to use the most challenging discourse context relevant to the school curriculum and reveal a child's strengths and weaknesses in linguistic performance to complement and confirm standardised test results.

Using Overseas Databases

One way to describe a child's strengths and weaknesses in linguistic performance is by comparing the child's performance to a database of language samples produced by typically developing speakers (Miller, Andriacchi, Nockerts, Westerveld, & Gillon, 2012). In the last two decades research efforts around the world have focused on creating such databases (e.g. Gillam & Pearson, 2004; Miller, 1989 - 1997; Schneider, Dubé, & Hayward, 2009; Westerveld et al., 2004). A continued topic of debate is whether we can safely adopt overseas databases containing samples from English-speaking children when appraising language samples of Australian children (Westerveld, 2011a). In a brief review of existing

cross-cultural research, Westerveld (2011a) expressed some confidence in Australian clinicians using databases from the US or New Zealand (NZ), but urged that further research was needed. In response, Westerveld and Heilmann (2012) compared spoken language samples produced in a story retelling task from typically developing NZ children to samples from the US. Results indicated that spoken language performance on measures of verbal productivity, semantic diversity and story quality did not differ as a result of geographic location. To extend this research, part two of the current study compares conversational, personal narrative, and story retelling samples from Australian children to their age-matched peers from NZ.

The Current Study

This study evaluated spoken language samples of young Australian school-age children in conversational, narrative, and expository contexts. The following questions were asked:

Study one:

1. What level of performance can we expect from young school-age children in Australia in conversation, story retelling, personal narrative, and exposition?
2. Is the Australian dataset of language samples sensitive to year of schooling and age for measures of syntactic complexity, semantic diversity, or mazing behaviour?
3. Is syntactic performance sensitive to discourse context (conversation, personal narrative, exposition)?

Study two:

4. Are there differences in conversational, story retelling, or personal narrative performance between school-age children from Australia and NZ?

Study One: Performance of the Australian children by YOS, Age, and Context

Method

Ethics approval for this project was granted by the University's Human Ethics Committee (PES/31/12/HREC) and the Department of Education and Training, Queensland Government (550/27/1258). A total of 20 speech-language pathologists working for the Department of Education and Training agreed to assist in the project by inviting schools to participate, organising the distribution of information sheets and consent forms, and administering the tasks. For a full description of the recruitment process see Westerveld and Vidler (2015). A total of 127 children participated in the study (64 boys; 63 girls), attending year of schooling (YOS) 1 (n = 44), YOS 2 (41), and YOS 3 (n = 42). These children spoke English as their first language, were progressing normally at school, and had no history of speech and/ or language impairment. Children were from Australian (85.2%), Aboriginal and Torres Strait Islander (4.0%), Pacific Island (0.8%), other (3.2%), or non-specified (6.4%) ethnic backgrounds, as indicated by their parents on the project consent forms. Children attended schools across Queensland (regional: 55; city: 72); these schools represented the full range of socio-economic areas, as indicated by their decile rankings (1 – 10), based on the schools' postcodes and Socio-Economic Indexes for Areas data (SEIFA; Australian Bureau of Statistics, 2011). Distribution of students was as follows: deciles 1 – 3: 38 students; deciles 4 – 7: 57 students; deciles 8 – 10: 32 students.

Procedure. All children were seen individually in a quiet room at the child's school. All sessions were audio-recorded using digital voice recorders. The sessions lasted ~40 mins and included a warm-up task and a range of oral language tasks. Table I provides an overview of the tasks reported in the current study and the number of children who participated in the tasks.

Insert Table I here

Conversation. In this task, children were asked to bring something from the classroom to discuss with the examiner. The child was then asked a series of generic questions about school, after-school activities, and family (see Westerveld et al., 2004, for a copy of the protocol). The aim was to engage the child in conversation for at least 5 mins and elicit 50 utterances. Throughout this task, the examiner was encouraged to listen and follow the child's lead, maintain the child's pace, and use open-ended prompts when needed to encourage the child to continue talking (Miller, 1981).

Story retelling. This task required the children to listen to an English translation of the story "Ko au na galo" (Ana Gets Lost; Swan, 1992). Following the first exposure to the story, children were asked eight comprehension questions, yielding a story comprehension score (SC). Children were provided with the correct information if they provided no answer or if their answer was clearly incorrect. After the second exposure to the story, children were asked to retell the story into the microphone, so that 'other children can listen to your story next time'. This task has been used extensively in the past with young school-age children from NZ and Australia (e.g. Lennox & Westerveld, 2014; Westerveld & Gillon, 2010; Westerveld et al., 2004).

***Personal narratives*.** Personal event narratives were elicited using the conversational map technique developed by Peterson and McCabe (1983). Examiners related a brief scripted personal experience related to a photo prompt to encourage the child to share one of his or her personal event narratives by asking "Did anything like that ever happen to you?" If the child responded "yes", a follow-up question was asked "Can you tell me about it?" To encourage the child to continue talking, the examiner was allowed to use nonverbal encouragement and/or or neutral prompts (uhuh, tell me more). Although we used the same protocol as described in Westerveld et al. (2004), some of the photo prompts were updated

(please contact the author for a copy of the photos and protocol). The aim was to elicit at least 50 child utterances.

Expository generation. The Favorite Game or Sport task (FGS) developed and described in full by Nippold et al. (2005) was administered to all children in YOS3. The examiner followed a script and asked the child what their favourite game or sport was, followed by several scripted prompts, including "I am not too familiar with the game of ....." The child was given as much time as needed to provide the explanation. The examiner was allowed to use neutral responses as required to encourage the child to continue.

Transcription and coding. All spoken language samples were digitally recorded and sent to the first author's research lab for transcription and coding by trained research assistants using standard coding conventions for Systematic Analysis of Language Transcripts – New Zealand (SALT-NZ; Miller, Gillon, & Westerveld, 2012). Utterances were segmented into communication units (C-units). Although C-units contain a main clause with all its subordinate clauses (Loban, 1976), following SALT-NZ conventions, elliptical phrases were also considered a C-unit. Sentence fragments were counted as separate C-units when the final intonation contour of the utterance indicated that a complete thought had been spoken. Finally, only complete and intelligible C-units (C&I) were used for analysis. Reformulations, repetitions, and disfluencies were placed in parentheses and considered mazes.

The conversation and personal event narrative samples were cut after the first 50 C&I utterances to enable comparisons across age groups. Unfortunately, a large number of conversational samples did not contain 50 utterances. Closer inspection revealed examiner error (i.e. not engaging the child in conversation for the stipulated 5 mins). It was decided to discard all language samples containing less than 40 utterances (24 in total). See Table I for the number of samples available by YOS group. The Ana Gets Lost story retelling samples were transcribed in full. Once all language samples were transcribed, transcripts were

analysed using SALT-NZ; An Excel data file was produced summarising each dependent measure (see below) for each transcript.

Measures. A range of language measures can be used to describe children's language performance (see Miller, Andriacchi, et al., 2012). To answer the research questions it was decided to focus on measures that are known to be sensitive to year of schooling (or age) and language ability.

*Total number of utterances (UTT).*Total number of utterances in C-units was computed automatically using SALT-NZ. Verbal productivity as measured in number of utterances differentiates school-age children with language-learning disabilities from their age-matched peers with typical development in narrative and expository contexts (Scott & Windsor, 2000).

Number of different words (NDW). NDW was computed automatically using SALT-NZ. NDW as a measure of semantic diversity, successfully differentiated between children with specific language impairment and their age-equivalent peers (Watkins, Kelly, Harbers, & Hollis, 1995).

Mean length of utterance in morphemes (MLU-M). Mean length of C-unit in morphemes was computed automatically using SALT-NZ. MLU-M as a measure of grammatical ability is sensitive to language ability (Scott & Windsor, 2000) and is considered an efficient measure of syntactic development (Nippold, 2007).

Percent maze words (PcMzWds). PcMzWds was computed automatically using SALT-NZ. Mazing is a measure of verbal fluency that is particularly sensitive to task complexity for children with language-learning difficulties (Maclachlan & Chapman, 1988).

To allow for comparison with previous research investigating children's performance on the story retelling task involving the Ana gets Lost story (Westerveld & Gillon, 2010), the transcripts were analysed using the Profile of Oral Narrative Ability (PONA, see Westerveld & Gillon, 2010). Apart from MLU, NDW, and UTT, the PONA includes:

Grammatical accuracy (GA). GA was calculated as the percentage of grammatically accurate utterances (Fey, Catts, Proctor-Williams, Tomblin, & Zhang, 2004). GA is particular sensitive to language ability status (Scott & Windsor, 2000).

Story quality (SQ). SQ was calculated as the total score on the story quality rubric (minimum score 8, maximum score 40), which covers six text structure elements: introduction, main character/s, supporting character/s, conflict, resolution, and conclusion as well as a measure of holistic coherence and a measure of "theme" (Westerveld & Gillon, 2010). This measure is sensitive to age (Westerveld & Gillon, 2010) and successfully differentiates between children with reading difficulties and their age-matched peers with typical development (Westerveld, Gillon, & Moran, 2008).

Story comprehension (SC). SC (maximum score 8) was calculated as the number of questions answered correctly. This measure is sensitive to age (Westerveld & Gillon, 2010) and was successful in differentiating between children with reading difficulties and their age-matched peers with typical development (Westerveld et al., 2008).

Reliability. First, to ensure accuracy and completeness of transcription and coding, the first author checked all transcripts for spelling, utterance segmentation and maze coding and listened to the sound files if the transcripts contained unintelligible segments or utterance segmentation issues. Disagreements were resolved through discussions and corrections were made if needed. Second, 15% of the transcripts and their corresponding sound files were randomly selected. An independent examiner (a research assistant proficient in transcription and SALT coding) listened to the sound files and checked the transcripts for transcription accuracy (percentage of words agreement = 98.5%; total words in agreement = 15 107). There were 14 instances of disagreement in utterance segmentation (agreement = 99.4%; total CU's in agreement = 2 436). For the AGL stories, the first author listened to all sound files, checked all transcripts for accuracy and completeness of transcription and coding, and made

corrections when needed, until 100% agreement was reached. Finally, 15% of the AGL stories were randomly selected and scored independently by a second rater on the quality rubric. Krippendorff alpha coefficients (Krippendorff, 1980) were calculated to document agreement between the two raters. Krippendorff's alpha using ordinal scaling was 0.908 indicating excellent agreement.

Results Study One

The data file generated with SALT-NZ was imported into SPSS (Pasw, 2012). SPSS was then used to analyse the data.

Performance of the Australian children. To answer research question one, we investigated the children's spoken language performance in the following contexts: conversation (YOS1, YOS2, YOS3), personal narrative (YOS1, YOS2, YOS3), and story retelling (YOS1, YOS2). Table II displays the results. As shown in Table II, there was considerable variability in performance (with large standard deviations) on some measures. However, the mean scores were close to the median scores indicating distributions were generally around the centre. We subsequently calculated skewness statistics to determine the shape of the distribution (i.e. if scores clustered at the low end or at the high end). Values of > 2 standard deviations of skewness are considered significant (Tabachnick & Fidell, 1996). As indicated in Table II, skewed distributions were found in the story retell condition, with UTT and NDW clustering at the low end of the scale for YOS1. For YOS2, PcMzWds clustered at the low end of the scale, whereas GA clustered at the high end of the scale. Closer inspection showed one outlier in the story retelling condition (this child produced 29 utterances; 126 different words). When excluding this case, the distributions for YOS1 were no longer skewed.

Sensitivity of the Australian dataset to year of schooling and age. To answer research question two, Multivariate Analyses of Variance (MANOVA) were used to

investigate overall effects for year-level. Effect sizes were calculated as eta squared (η^2) to document the relative proportion of the variation in each measure that was associated with the children's year of schooling (Lakens, 2013). For example if $\eta^2 = 0.13$, 13% of the variance can be explained by the children's year of schooling. These effect sizes can be interpreted as follows: 0.0099 = small effect size; 0.0588 = medium effect size; 0.1379 = large effect size (Richardson, 2011). To investigate sensitivity to age, a series of linear regression analyses were performed, with age in months as the independent variable and the language measures as the dependent variables.

Insert Table II here

Conversation. There was a significant overall effect for year of schooling level for NDW $F(2, 99) = 3.850$, $p = 0.025$, $\eta^2 = 0.072$, and MLU-M: $F(2, 99) = 3.391$, $p = 0.038$, $\eta^2 = 0.064$, but not for PcMzWds ($p = 0.871$, $\eta^2 = 0.003$). For NDW, post-hoc tests (Tukey) revealed that the YOS3 children significantly outperformed the YOS1 children ($p = 0.018$); there were no significant differences between YOS1 and YOS2 ($p = 0.389$) or between YOS2 and YOS3 ($p = 0.389$). For MLU-M, the YOS3 children significantly outperformed the YOS1 children ($p = 0.045$); there were no significant differences between YOS1 and YOS2 ($p = 0.121$) or YOS2 and YOS3 ($p = 0.915$).

Results from linear regression analyses indicated there were no significant effects for age for MLU-M ($p = 0.072$), or PcMzWds ($p = 0.781$); There was a significant effect for age for NDW ($p = 0.026$).

Personal narratives. There was a significant overall effect for year of schooling level for NDW $F(2, 121) = 8.784$, $p < 0.000$, $\eta^2 = 0.127$, and MLU-M: $F(2, 121) = 5.048$, $p = 0.008$, $\eta^2 = 0.077$, but not for PcMzWds ($p = 0.278$, $\eta^2 = 0.021$). For NDW, post-hoc tests (Tukey) revealed that the YOS2 children significantly outperformed the YOS1 children ($p = 0.004$); there were no significant differences between YOS2 and YOS3. The YOS3 children

performed significantly better than the YOS1 children ($p = 0.001$). For MLU-M, the YOS3 children significantly outperformed the YOS1 children ($p = 0.007$); there were no significant differences between YOS1 and YOS2 ($p = 0.095$) or YOS2 and YOS3 ($p = 0.591$).

Results from linear regression analyses revealed significant effects for age for MLU-M ($p = 0.008$) and NDW ($p < 0.00$), but not for PcMzWds ($p = 0.898$).

Story retelling. There was a significant effect for year level for NDW: $F(1, 82) = 6.495$, $p = 0.013$, $\eta^2 = 0.073$; and GA $F(1, 82) = 5.405$, $p = 0.023$, $\eta^2 = 0.062$; and SQ $F(1, 82) = 7.759$, $p = 0.007$ with the YOS2 children performing better than the YOS1 children. There were no significant differences in performance between the two year groups on UTT ($p = 0.123$, $\eta^2 = 0.029$), MLU-M ($p = 0.121$, $\eta^2 = 0.029$), PcMzWds ($p = 0.811$, $\eta^2 = 0.001$), or SC ($p = 0.072$, $\eta^2 = 0.039$).

Results from regression analyses showed significant effects for age for MLU-M ($p = 0.044$), NDW ($p < 0.000$), UTT ($p < 0.000$), GA ($p = 0.006$), SC ($p = 0.015$), and SQ ($p < 0.000$), but not for PcMzWds ($p = 0.711$).

Effects of elicitation context on spontaneous language measures for the YOS3 children. To answer research question three, we compared the performance of the YOS3 children across the three elicitation contexts. Because the language samples were of different length, only those measures that are independent of sample length were included in the present analysis: MLU-M and PcMzWds. Table III displays the results. MANOVA showed significant effects for context for MLU-M (with a large effect size as shown in Table III), but not PcMzWds ($p = 0.064$). Post-hoc analyses (Tukey) revealed significant differences in MLU-M between conversation and expository and between personal narrative and expository with the expository context yielding the longest MLU-M. There were no statistically significant differences for MLU-M between conversation and personal narratives.

Insert Table III here

Study Two: Comparing the Australian dataset to the NZ database

Method

New Zealand database. The NZ database (Westerveld et al., 2004) was originally created in 2000 and contains language samples from 268 children aged between 4.5 and 7.6 years. These children were randomly selected from schools in major urban areas in NZ. All children spoke English as their first language and had no history of sensory deficits or neurological disorder as indicated by their parents on the consent form. Children came from a range of socio-economic areas and the ethnic make-up of the group reflected the largest cultural groups in NZ at the time (see Westerveld et al., 2004, for further details). Unfortunately, no data regarding the children's year of schooling were collected. Therefore the following comparisons are based on chronological age of the participants.

Procedure. The procedures used in collecting the data have been described in detail elsewhere (Westerveld et al., 2004). In summary, the administration and elicitation procedures were similar to the ones used in the present study, except for a) the photos used in the personal narrative condition were updated for the Australian study, and b) the Ana gets Lost story was shown on a computer screen with an Australian voice-over (as opposed to a more traditional book and accompanying tape). Previous research investigating the effects of showing the story on a computer screen versus the traditional book on children's performance revealed no significant differences (Westerveld, 2011b).

Results Study Two

Due to differences in age ranges between the two datasets, it was decided to limit comparisons to those age groups for which at least 20 samples were available. As a result, for the conversation and personal narrative contexts, comparisons were made for the 6- and 7-year-old age groups only. For the story retelling context, comparisons could be made for the

5-, 6- and 7-year-age groups. The following subset of the NZ database was used: a) conversation*:* 154 samples, b) personal narratives: 137 samples, c) story retelling: 176 samples. Because (within those age groups) the Australian children were significantly older than the NZ children, analyses of covariance were used (controlling for age) to determine the effects of "country". The results for the conversation and personal narrative contexts are displayed in Table IV.

Conversation. There were no significant differences on any of the language measures for the 6-year-olds. In contrast, the Australian 7-year-olds performed significantly better than their NZ peers on measures of semantic diversity (NDW) and grammatical complexity in MLU-M (with medium effect sizes). There were no group differences in mazing behaviour.

Personal narratives. For both age groups, the Australian children outperformed the NZ children on measures of semantic diversity and grammatical complexity (with small to medium effect sizes). There were no differences between the groups on mazing behaviour.

Insert Table IV here

Story retelling. For the 5-year-old age groups, there was a significant difference in performance on grammatical complexity (MLU-M) with the Australian children outperforming the NZ children. In contrast, the NZ children showed better performance in story comprehension. There were no group differences on any of the other measures (see Table V). For the 6-year-olds, significant group differences were found on MLU-M and grammatical accuracy, with medium effect sizes. Although the Australian children used significantly longer utterances, the NZ children showed higher grammatical accuracy. No significant differences were found on any of the other measures. For the 7-year-old age groups, similar results were found. The Australian children used significantly longer utterances, but showed significantly lower grammatical accuracy than the NZ children (with a large effect size). There were no group differences on any of the other language measures.

Insert Table V here

General Discussion

This study investigated the spoken language performance of 127 children attending the first three years of their primary schooling in Queensland, Australia. All children produced spoken language samples in conversation, personal narrative, and story retelling contexts. The YOS3 children also produced an expository language sample.

First, we described the children's performance by year of schooling (YOS) on measures of semantic diversity, grammatical complexity, and mazing behaviour. To determine the potential usefulness of these data for local norming purposes, distribution statistics were calculated. Consistent with previous research, considerable variability in performance was observed (Justice et al., 2006; Westerveld & Gillon, 2010). However, median scores were close to mean scores and only some measures showed skewed distributions. It is therefore suggested that the current results may be used to determine a child's relative performance to his or her typically developing peers, as long as the child is similar to the participants in the current study (i.e. attending the first three years of school and speaking English as the main language).

The next question asked if the dataset of language samples was sensitive to year of schooling for measures of syntactic complexity, semantic diversity, and mazing behaviour. As expected, and consistent with previous research there was a clear developmental trend of increasing mean length of utterance and higher semantic diversity with age (Westerveld et al., 2004). In contrast, mazing behaviour was relatively stable over time, which is consistent with findings from Leadholm and Miller (1992) and Heilmann, Miller, and Nockerts (2010). Despite the lack of sensitivity of this measure to developmental change, as Heilmann, Miller, and Nockerts (2010) point out this measure may still be useful in clinical practice for describing disfluent behaviour, including the frequency of repetitions and revisions at sound-,

word-, and phrase-levels. Mazing behaviour is susceptible to the formulation load of the task (Leadholm & Miller, 1992), and a higher than expected level of mazing behaviour may reveal processing difficulties in children with language impairment (Thordardottir & Weismer, 2002).

The third question compared the performance of the YOS3 children across three generation contexts. As shown in Table III, the expository context yielded significantly longer utterances (with large effect sizes) than the personal narrative or conversational contexts. In contrast, no significant context effects were found for mazing behaviour. These results are consistent with those from Nippold et al. (2005) who investigated syntactic performance of 8-year-old children from the US across conversational and expository discourse. The results add to our existing knowledge by revealing higher MLU-M in expository generation compared to personal narrative generation in young school-age children. It was somewhat surprising to find no significant differences in syntactic complexity between the personal narrative and conversation contexts. Closer inspection of the results from our previous research with NZ children that used the same elicitation tasks (Westerveld et al., 2004), revealed that 5- and 7-year-old children (but not the 6-year-olds) produced significantly longer utterances in personal event narration compared to conversation. However, the YOS3 students in the current study were on average 7;10 (range 7;4 to 8;4), which may explain the differences in findings. Taken together, the findings confirm the usefulness of the expository elicitation context (the Favourite Game or Sport task; Nippold et al., 2005) for young school-age children (Westerveld & Moran, 2011). Although further research is needed investigating the performance of children with identified language impairment on this task, clinicians may add this task to their battery of language assessments to reveal a child's strengths and weaknesses in linguistic performance in a challenging discourse context that is relevant to the curriculum.

The final aim of this study was to compare the language samples of the Australian children to those of their NZ peers, contained in the NZ database (Miller, Gillon, et al., 2012). Children were divided into age groups and analyses corrected for age differences where needed. As shown in Tables IV and V, the measure that seemed particularly sensitive to geographic origin (across all ages and conditions) was MLU-M, with the Australian children outperforming their NZ peers. However, effect sizes were considered medium, explaining between 0.7% and 8.1% of the variability. In addition, the Australian children showed stronger performance in semantic diversity than their NZ counterparts in conversation and personal event narration (but not in story retelling), with small to medium effect sizes. Finally, in the story retelling condition, the NZ children (6-and 7-year-olds) showed better grammatical accuracy and the NZ 5-year-olds showed significantly better story comprehension.

Although the differences in performance as a function of country of origin were significant, it is important to highlight the relatively small effect sizes (except for GA and SC). In fact, these effect sizes (i.e. up to 8.1% of the variability) are similar to those reported by Heilmann, Miller, and Nockerts (2010) when comparing conversational samples from children residing in different locations (and states) in the US, namely Wisconsin and California. Despite these modest differences between the datasets, these authors decided to combine them into one database. One option might thus be to combine the Australian and NZ datasets into one database. The alternative would be to keep them separate to preserve 'national integrity' as Heilmann, Miller, and Nockerts (2010) put forward in support of their decision to keep the NZ and US databases separate. A third option may be to combine the NZ and Australian datasets and integrate them into SALT-NZ (Miller, Gillon, et al., 2012), which can provide users with the option of choosing a subset (i.e. AU or NZ) when comparing a language sample from a client with language impairment to one of the databases. This final

option seems the most reasonable one considering that some measures (GA and SC) seem more sensitive to geographic location than others.

Limitations and Future Directions

All participants attended school in Queensland, and the results may not apply to children from other states or territories in Australia. However, as we argued previously, with the introduction of a common Australian curriculum framework (ACARA, 2012), similar levels of performance may be expected of children attending their first three years of primary school education across Australia.

Another limitation is the relatively small sample size per age group or year-of-schooling, which may impact the usefulness of this dataset for providing a benchmark of typical performance. Further debate is clearly needed to address this question of how many language samples are needed to provide enough information for norming purposes, especially considering the inherent variability of spontaneous language (Miller, Heilmann, Nockerts, Andriacchi, & Iglesias, 2006).

We did not control for order effects when comparing the children's performance across the three discourse generation contexts of conversation, personal narratives, and exposition. It may well be that the children produced longer sentences towards the end of the assessment session (i.e. on the expository task) because of a practice effect. Despite this potential practice effect, there were no differences in children's performance in the conversation versus the personal narrative contexts, which was unexpected and may indicate that a practice effect may not transcend across discourse contexts. Future research should control for this potential practice effect by randomizing the order of the tasks.

Unfortunately we did not record the NZ children's year of schooling level. It is therefore not clear if some of the differences in performance between the Australian and NZ children could be attributed to a different schooling system.

Finally, further research is clearly needed to determine the potential of this dataset for diagnostic purposes. Although recent evidence suggests the usefulness of conversational and story retelling databases for correctly classifying children with language impairment (Heilmann, Miller, & Nockerts, 2010), to our knowledge no such data exist for personal narrative or expository contexts.

Conclusion

This study attempted to address an important gap in our current knowledge regarding the typical performance of Australian children in a range of discourse tasks that are relevant to the Australian curriculum (ACARA, 2012). By integrating the results from the current study into the SALT-NZ databases, clinicians will be able to gain insight into the language production abilities of typical Australian children. Although many Australian clinicians report eliciting spontaneous language samples as part of their routine assessment (Westerveld & Claessen, 2014), availability of local data may encourage more clinicians to record and transcribe these language samples. Comparing language samples to the corresponding database will promote a better understanding of a client's strengths and weaknesses in language performance compared to same-age peers from Australia and/or NZ. This detailed analysis of spoken language performance will not only promote effective goal setting, it will also enable detailed progress monitoring following intervention.

Acknowledgments

This project was funded by an Emerging Researcher Grant, awarded to the first author by the School of Arts, Education and Law, Griffith University. In-kind assistance was provided by the Queensland Department of Education and Training. The views expressed in this publication do not necessarily present the views of the Department. The project was also supported by SALT Software, LLC. SALT Software did not participate in the analysis or interpretation stage of the project and did not review the manuscript prior to submission.

Thank you to the speech-language pathologists for volunteering their time to assist with recruitment and collection of the data.

References

Australian Bureau of Statistics. (2011). Census of Population and Housing: Socio-Economic Indexes for Arease (SEIFA), Australia, 2011. http://www.abs.gov.au/ausstats/abs@.nsf/mf/2033.0.55.001/

Australian Curriculum Assessment and Reporting Authority [ACARA]. (2012). The Australian Curriculum - English. www.australiancurriculum.edu.au

Boudreau, D. (2008). Narrative abilities - Advances in research and implications for clinical practice. *Topics in Language Disorders*, *28*, 99-114.

Catts, H. W., & Hogan, T. P. (2003). Language basis of reading disabilities and implications for early identification and remediation. *Reading Psychology, 24*, 223-246.

Evans, J. L., & Craig, H. K. (1992). Language sample collection and analysis: Interview compared to freeplay assessment contexts. *Journal of Speech and Hearing Research*, *35*, 343-353.

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*, 1301-1318.

Gillam, R. B., & Pearson, N. A. (2004). *Test of Narrative Language*. Austin, TX: Pro-ed.

Heilmann, J., DeBrock, L., & Riley-Tillman, T. C. (2013). Stability of measures from children's interviews: The effects of time, sample length, and topic. *American Journal of Speech-Language Pathology, 22*, 463-475. doi: 10.1044/1058-0360(2012/11-0035)

Heilmann, J., Miller, J. F., & Nockerts, A. (2010). Using language sample databases. *Language, Speech, and Hearing Services in Schools, 41*, 84-95. doi: 10.1044/0161-1461(2009/08-0075)

Heilmann, J., Miller, J. F., Nockerts, A., & Dunaway, C. (2010). Properties of the Narrative Scoring Scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology, 19*, 154-166. doi: 10.1044/1058-0360(2009/08-0024)

Hughes, D., McGillivray, L., & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessment*. Eau Claire, WI: Thinking Publications.

Justice, L. M., Bowles, R., Pence, K., & Gosse, C. (2010). A scalable tool for assessing children's language abilities within a narrative context: The NAP (Narrative Assessment Protocol). *Early Childhood Research Quarterly, 25*, 218-234.

Justice, L. M., Bowles, R. P., Kaderavek, J. N., Ukrainetz, T. A., Eisenberg, S. L., & Gillam, R. B. (2006). The index of narrative microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech-Language Pathology, 15*, 177-191.

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Newbury Park, CA: Sage.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in psychology, 4*(November, Article 863), 1-12.

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

Lennox, M., & Westerveld, M. F. (2014). Enhancing emergent literacy performance of Australian students from disadvantaged backgrounds in their first year of school: A preliminary investigation *Journal of Clinical Practice in Speech-Language Pathology, 16*(2), 58-65.

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.

Miller, J. F. (1981). *Assessing language production in children: Experimental procedures*. Austin, TX: Pro-ed.

Miller, J. F. (1989 - 1997). *The Reference Database Project*: Waisman Center, Language Analysis Laboratory, University of Wisconsin, Madison.

Miller, J. F., Andriacchi, K., Nockerts, A., Westerveld, M., & Gillon, G. (2012). *Assessing Language Production using SALT Software. A Clinician's Guide to Language Sample Analysis. New Zealand version*. Middleton, WI: SALT Software.

Miller, J. F., Gillon, G. T., & Westerveld, M. F. (2012). *Systematic Analysis of Language Transcripts (SALT), New Zealand Version 2012 [computer software]*. Madison, WI: SALT Software LLC.

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.

Milosky, L. M. (1987). Narratives in the classroom. *Seminars in Speech and Language, 8*, 329-343.

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., 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.

PASW. (2012). PASW Statistics 20 [Computer Software]. Chicago, IL: IBM SPSS Statistics.

Petersen, D. R., Gillam, S. L., & Gillam, R. R. (2008). Emerging procedures in narrative assessment - The index of narrative complexity. *Topics in Language Disorders, 28*(2), 115-130.

Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. New York: Plenum.

Reed, V. A., & Spicer, L. (2003). The relative importance of selected communication skills for adolescents' interactions with their teachers: High school teachers' opinions. *Language, Speech, and Hearing Services in Schools, 34*, 343-357.

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*(2), 135-147. doi: 10.1016/j.edurev.2010.12.001

Schneider, P., Dubé, R. V., & Hayward, D. (2009). The Edmonton Narrative Norms Instrument. 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*, 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.

Swan, E. (1992). *Ko au na galo (Ana gets lost)*. Wellington, NZ: Learning Media, Ministry of Education.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). NY: Harper Collins.

Thordardottir, E. T., & Weismer, S. E. (2002). Content mazes and filled pauses in narrative language samples of children with specific language impairment. *Brain And Cognition, 48*, 587-592.

Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech and Hearing Research, 38*, 1349-1355.

Westerveld, M.F., & Gillon, G.T. (2010). Profiling oral narrative ability in young school-aged children. *International Journal of Speech-Language Pathology, 12*, 178-189. doi: 10.3109/17549500903194125

Westerveld, M.F., Gillon, G.T., & Moran, C. (2008). A longitudinal investigation of oral narrative skills in children with mixed reading disability. *International Journal of Speech-Language Pathology, 10*, 132-145. doi: 10.1080/14417040701422390

Westerveld, M. (2011a). Sampling and analysis of children's spontaneous language: From research to practice. *ACQuiring Knowledge in Speech, Language and Hearing, 13*, 63-67.

Westerveld, M. F. (2011b). *Children's story retelling and comprehension performance under different task conditions*. Paper presented at the American Speech-Language-Hearing Association Annual Convention, San Diego, CA, US.

Westerveld, M. F., & Claessen, M. (2014). Clinician survey of language sampling practices in Australia. *International Journal of Speech-Language Pathology, 16*, 242-249.

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*, 195-208.

Westerveld, M. F., & Heilmann, J. J. (2012). The effects of geographic location and picture cues on children's story retelling performance. *Asia Pacific Journal of Speech, Language, and Hearing, 15*, 129-143.

Westerveld, M. F., & Moran, C. A. (2011). Expository language skills of young school-age children. *Language, Speech ,and Hearing Services in Schools, 42*, 182-193. doi: 10.1044/0161-1461(2010/10-0044)

Westerveld, M. F., & Vidler, K. (2015). The use of the Renfrew Bus Story with 5- to 8-year-old Australian children. *International Journal of Speech-Language Pathology, 17*, 304 - 313. doi: 10.3109/17549507.2015.1024168

Table I. Overview of number of participants (with mean age and age range) participating in the oral language tasks across year of schooling (YOS)

	YOS1	**YOS2**	**YOS3**
Conversation			
n	37	32	33
Age (years;months)	6;0	6;11	7;11
	5;5 – 6;5	6;5 – 7;7	7;4 – 8;4
Personal narratives			
n	42	40	42
Age	6;0	6;10	7;10
	5;5 – 6;5	6;4 – 7;6	7;4 – 8;4
Story retelling (AGL) *			
n	44	40	Not administered
Age	6;0	6;10	
	5;5 – 6;5	6;4 – 7;5	
Expository [$]			
n	Not administered	Not administered	42
Age			7;10
			7;4 – 8;4

AGL, Ana gets lost (Swan, 1992); * the AGL task is considered too easy for children in YOS3; [$] The expository task was only used with the YOS3 children.

Table II. Mean performance across year of schooling groups in three elicitation contexts (with SD), median, and range.

	YOS1	**YOS2**	**YOS3**
Conversation[$]			
n	37	32	33
NDW	118.8 (19.5)	125.3 (19.9)	132.4 (22)
	120	129	130
	75 - 180	80 - 161	94 - 189
MLU-M	5.86 (1.10)	6.48 (1.32)	6.61 (1.44)
	5.93	6.59	6.16
	3.74 – 9.06	3.8 – 9.84	3.9 – 10.16
PcMzWds	10.1 (4.6)	10.1 (4.8)*	10.6 (4.2)
	9.0	10.5	10.0
	2 - 23	3 - 26	4 – 21
Personal narratives[$]			
n	42	40	42
NDW	120.5 (18.9)	134.8 (19.8)	137 (19.8)
	121.5	136.5	133.5
	68 - 158	92 - 192	94 - 179
MLU-M	6.22 (1.13)	6.8 (1.35)	7.09 (1.37)
	6.23	6.91	7.00
	3.64 – 7.86	4.04 – 9.28	4.92 – 10.22
PcMzWds	8.7 (4.9)	9.9 (4.1)	8.5 (3.8)
	8.0	9.5	8.0
	0 - 22	2 – 18	2 – 17
Story retell			
n	44	40	Not administered
UTT	12.1 (5.6)*	23.95 (5.1)	
	11.5	13.5	
	4 – 36	3 – 25	
NDW	47.9 (20.0)*	58.1 (16.3)	
	44.5	58.0	
	13 – 126	18 – 91	
MLU-M	7.69 (1.45)	8.16 (1.41)	
	7.73	7.85	
	4.23 – 11.13	5.59 – 14.0	
PcMzWds	12.2 (7.1)	12.6 (7.1)*	
	11.0	13.5	
	0 – 29	2 - 37	
GA	75.7 (17.5)	83.6 (12.9)*	
	76.9	83.7	
	33.3 – 100	40 - 100	
SC	6.00 (1.5)	6.55 (1.2)	
	6.00	7.00	
	2 – 8	4 - 8	
SQ	24.1 (6.5)	28.2 (6.7)	

	24.0	28.0
	12 - 40	12 - 40

NDW, number of different words; MLU-M, mean length of utterance in morphemes; PcMzWds, percent maze words; UTT, utterances; GA, grammatical accuracy in percent grammatically correct utterances; SC, story comprehension, maximum score is 8; SQ, story quality, scores range from 8 – 40; [$] transcripts were cut after the first 50 complete and intelligible utterances. * indicates the data are significantly skewed

Table III. Mean performance across elicitation contexts for YOS3 students only (with SD).

	Conversation	**Personal narratives**	**Expository**	**Effect size (η^2)**
n	33	42	42	
UTT[$]	49.7 (.9)	49.9 (.8)	30 (13.2)	
MLU-M	6.61 (1.44)	7.09 (1.37)	8.80 (1.61)	0.294*
PcMzWds	10.6 (4.2)	8.5 (3.8)	10.3 (4.9)	0.047

[$] transcripts were cut after the first 50 complete and intelligible utterances. MLU-M, mean length of utterance in morphemes; PcMzWds, percent maze words. * $p < .000$

Eta squared (η^2) values represent the amount of explained variance in language sample measures as a function of context.

Table IV. Performance by country and age group in conversation and personal narrative contexts (with SD).

	6-year-olds			7-year-olds		
	Australia	**NZ**	**η^2**	**Australia**	**NZ**	**η^2**
Conversation[$]						
n	33	95		33	59	
Age in months	76.5 (2.4)	76.8 (2.7)		89.5 (3.5)*	87.5 (2.0)*	
NDW	120 (16.5)	115.1 (16.1)	0.025	132.6 (19.7)*	116.9 (17.9)*	0.112
MLU-M	5.95 (1.21)	5.75 (1.00)	0.007	6.70 (1.30)*	5.8 (1.17)*	0.081
PcMzWds	9.91 (4.4)	9.70 (6.1)	0.001	10.3 (4.8)	8.6 (3.6)	0.041
Personal narratives[$]						
n	44	86		46	51	
Age in months	76.5 (2.5)	76.9 (2.8)		89.2 (3.6)*	87.4 (2.0)*	
NDW	126.9 (19.6)*	115.6 (19.9)*	0.070	137.6 (22.6)*	120.9 (16.6)*	0.120
MLU-M	6.61 (1.16)*	6.09 (1.02)*	0.052	7.01 (1.39)*	6.25 (0.87)*	0.071
PcMzWds	9.7 (4.3)	10.5 (6.2)	0.003	8.5 (4.0)	9.7 (3.9)	0.015

NZ, New Zealand; NDW, number of different words; MLU-M, mean length of utterance in morphemes; PcMzWds, percent maze words; [$] transcripts were cut after the first 50 complete and intelligible utterances. * indicates differences are significant $p < .05$. Eta squared (η^2) values represent the amount of explained variance in language sample measures as a function of country.

Table V. Performance by country and age group on the story retelling task (with SD).

	5-year-olds			6-year-olds			7-year-olds		
	AU	**NZ**	**η^2**	**AU**	**NZ**	**η^2**	**AU**	**NZ**	**η^2**
Story retell									
n	20	38		43	85		21	53	
Age in	68.15	65.8		76.4	77.1		86.2	87.4	
months	(2.2)*	(3.0)*		(2.4)	(2.7)		(1.9)	(2.1)	
UTT	11.4	10.5	0.000	12.9	13.0	0.000	14.7	15.7	0.004
	(4.4)	(4.1)		(5.6)	(4.8)		(5.4)	(4.8)	
NDW	45.2	39.4	0.014	52.8	48.7	0.011	60	58.1	0.005
	(16.4)	(13.0)		(19.6)	(15.7)		(17.6)	(14.9)	
MLU-M	7.63	6.86	0.076	7.94	7.30	0.062	8.08	7.38	0.076
	(1.29)*	(1.18)*		(1.66)*	(0.93)*		(1.06)*	(0.93)*	
PcMzWds	11.5	10.9	0.002	12.7	11.7	0.002	12.6	10.0	0.030
	(8.9)	(7.6)		(5.9)	(9.3)		(7.6)	(5.3)	
GA	78.8	87.0	0.048	77.6	86.3	0.070	83.9	92.4	0.137
	(18.6)	(14.3)		(16.2)*	(13.1)*		(11.7)*	(8.8)*	$^{\$}$
SC	5.9	6.6	0.099	6.2	6.5	0.010	6.8	6.3	0.031
	(1.8)*	(1.2)*^		(1.3)	(1.5)$^{\$}$		(1.1)	(1.4)$^{@}$	
SQ	25	21.5	0.033	24.7	24.5	0.000	29.7	27.7	0.018
	(6.4)	(5.9)		(6.6)	(6.5)		(6.9)	(6.9)	

AU, Australia; NZ, New Zealand; NDW, number of different words; MLU-M, mean length of utterance in morphemes; PcMzWds, percent maze words; UTT, utterances; GA, grammatical accuracy in percent grammatically correct utterances; SC, story comprehension, maximum score is 8; SQ, story quality, scores range from 8 – 40. ^ only 37 samples available. $^{\$}$ data for 79 samples available. $^{@}$ data based on 52 samples. * indicates differences are significant $p < .05$. Eta squared (η^2) values represent the amount of explained variance in language sample measures as a function of country. $^{\$}$ $\eta^2 = .137$ indicates that 13.7% of the variance in GA for the7-year-olds can be explained by country.