

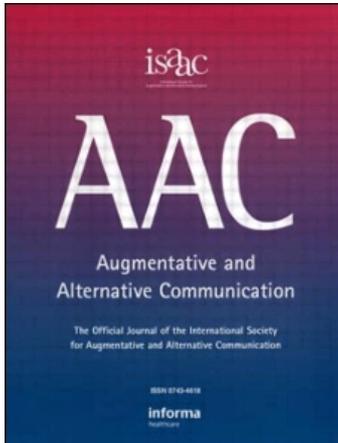
This article was downloaded by: [University of New Hampshire]

On: 22 November 2008

Access details: Access Details: [subscription number 791270723]

Publisher Informa Healthcare

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Augmentative and Alternative Communication

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t713692248>

The Vocabulary of Beginning Writers: Implications for Children with Complex Communication Needs

Sally A. Clendon ^a; Karen A. Erickson ^b

^a Massey University, Auckland, New Zealand ^b University of North Carolina at Chapel Hill, North Carolina, USA

Online Publication Date: 01 January 2008

To cite this Article Clendon, Sally A. and Erickson, Karen A. (2008) 'The Vocabulary of Beginning Writers: Implications for Children with Complex Communication Needs', *Augmentative and Alternative Communication*, 24:4, 281 — 293

To link to this Article: DOI: 10.1080/07434610802463999

URL: <http://dx.doi.org/10.1080/07434610802463999>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The Vocabulary of Beginning Writers: Implications for Children with Complex Communication Needs

SALLY A. CLENDON^{a,*} and KAREN A. ERICKSON^b

^aMassey University, Auckland, New Zealand and ^bUniversity of North Carolina at Chapel Hill, North Carolina, USA

One of the greatest challenges facing augmentative and alternative communication (AAC) professionals is providing children with complex communication needs with access to the vocabulary that they need in order to develop mature language and literacy abilities. The purpose of this study was to analyze the vocabulary used by typically developing early-elementary children in the United States and New Zealand when they write about self-selected topics, in order to inform practices with children with complex communication needs. The children's writing samples were compared across school ages and countries. The school age comparisons highlighted the relatively restricted range of vocabulary used by children in the earliest stages of writing development, and the country comparisons revealed differences in core vocabulary. The findings of this study hopefully will assist AAC professionals as they engage in selecting, prioritizing, and organizing vocabulary to support written language development in children with complex communication needs.

Keywords: Augmentative and Alternative Communication; Language Development; Vocabulary; Writing

INTRODUCTION

The aim of this investigation was to contribute to the body of information available to assist educators and speech-language pathologists as they engage in decision making about teaching and supporting children with complex communication needs who use augmentative and alternative communication (AAC). The investigation explored vocabulary use in typical written language development and its potential application to planning vocabulary sets and language representation systems for children with complex communication needs.

Why Study Written Language?

Several researchers in the field of AAC have analyzed the vocabulary used by typically developing children and then endeavored to relate their studies' findings to children with complex communication needs. Most of these studies have

analyzed spoken vocabulary samples (e.g., Beukelman, Jones, & Rowan, 1989; Fallon, Light, & Paige, 2001). Few studies (Clendon, Sturm, & Cali, 2003; McGinnis & Beukelman, 1989) have analyzed the vocabulary that typically developing children use when they write. Obstacles faced by beginning writers (see Gombert, 1992; Kroll, 1981; McCutchen, 2000; Scardamalia, 1981) resemble those experienced by children who use AAC; therefore a better understanding of written language development for typically developing children will inform practice with children with complex communication needs. For example, both groups of children confront the challenge of taking language that is inside their heads and translating it into an expressive form, using an instrument that is not second nature to them (e.g., a pencil or a communication device). The cognitive, memory, and physical demands of such a process have an influence on the quantity and quality of the language produced. By studying the vocabulary used in written language samples of

*Corresponding author. Speech Language Therapy Programme, School of Education at Albany, Massey University, College of Education, PO Box 102 904, North Shore Mail Centre, Auckland, New Zealand. Tel: +64 9 414 0800. Fax: +64 9 443 9717. E-mail: s.clendon@massey.ac.nz

typically developing writers, we can select and organize vocabulary in ways that may decrease the demands of writing for children with complex communication needs and, consequently, increase the quantity and quality of the writing that they produce.

A Theoretical Model of Writing

The model of writing proposed by Flower and Hayes (1981) describes the process that skilled adult writers use when they compose text. However, this model has also been modified to account for beginning and developing writing (Berninger, Fuller, & Whitaker, 1996). There are four primary cognitive processes in the model: planning, translating, reviewing, and monitoring.

Planning involves setting goals, generating ideas, and organizing thoughts relative to a written text. Young writers are far less sophisticated than older writers when planning (MacArthur, Harris, & Graham, 1994) and often plan while they write rather than before (Graham, Harris, MacArthur, & Schwartz, 1991). *Translating* involves conveying ideas, sensory impressions, and oral language into written language that adheres to print conventions. Effective translation depends upon oral language skills including vocabulary (Abbott & Berninger, 1993). The physical act of generating text – text production – is a component of translating for beginning writers who are typically developing (Berninger et al., 1992) and is influenced by graphomotor skills (Singer & Bashir, 2004). Text production may continue to be a component of translating for individuals who use AAC even when they attain relatively sophisticated levels of writing. Especially for individuals who use scanning and other indirect selection techniques, text production requires conscious attention to the tool itself, which, according to verbal efficiency theory (Perfetti, 1985) would limit the cognitive resources available to other writing processes.

Reviewing is composed of revision (i.e., reorganizing existing text) and evaluation (i.e., appraising the degree to which a text fulfills the writer's plan). Beginning writers first engage in reviewing after completing a text, but more sophisticated writers revise as they write (Berninger, Whitaker, Feng, Swanson, & Abbott, 1996). The final cognitive process, *monitoring*, refers to the writer's ability to attend to and adjust the application of all of the component processes while composing a given text.

Vocabulary in Writing

Since the late 1800s, the study of vocabulary has interested professionals from fields such as educa-

tion, textbook publishing, and psychology (Johnson, 2000), but few research studies have examined the relationship between vocabulary and writing. Some studies (Farr, Kelleher, Lee, & Beverstock, 1989; Fitzgerald, 1938; Hillerich, 1978; Lorenz, 1931; Rinsland, 1945; Shapiro & Gunderson, 1988; Smith & Ingersoll, 1984) have analyzed the vocabulary used by children below the fourth-grade level in their writing, and focused on the compilation of frequency based word lists for a variety of purposes. Fitzgerald (1938), for example, created a list of 692 high-frequency words used in 1,256 letters written about life outside of school. More recently, Farr et al. (1989) collected a corpus of over 3 million words from more than 21,000 children's writing samples obtained from national and state-wide assessments of writing performance. The data from these samples were used to develop a variety of vocabulary lists.

Rinsland (1945) conducted the largest analysis of the vocabulary used in children's writing. This researcher contacted the administrators of 1,500 elementary schools throughout the United States and requested children's writing samples, including personal notes, stories, poems, compositions, exam papers, reports, and observations. A total of 708 schools responded. The writing samples produced a corpus of 6,012,359 words, written by just over 100,000 children. These words were hand tabulated, recorded on paper, entered in ledgers, and checked before being analyzed to determine the frequency of individual words used at each grade level. Rinsland identified 25,632 different words from the total corpus. The total number of different words ranged from 5,099 in Grade 1 to 17,930 in Grade 8. The total number of words written ranged from 353,874 in Grade 1 to 1,088,343 in Grade 8. As would be expected, the older children wrote more words and used a greater number of different words than the younger children.

Additional studies that examine age-related differences in the vocabulary of beginning writers are required; many studies were carried out more than 30 years ago. Since then, writing instructional practices and the expectations that educators and parents have for children, have undoubtedly changed. Another limitation of the existing research is that only one study (Bear, 1939) has included children below the third-grade level. In at least one of the studies, the reason for this was a concern that children in or below third-grade "may jabber away with ease, fluency and exuberance," but most will "write only under considerable duress" (Hunt, 1964, p. 4). However, it is the written language used by children in the earlier grades that may be most relevant to children with complex communication needs.

Two additional concerns in the existing research are that (a) few studies have collected writing samples about self-selected topics, and (b) no studies have compared the written language generated by children in different countries. All of these limitations were taken into consideration and addressed in the current investigation.

Vocabulary Selection and Language Representation for Children with Complex Communication Needs

Over the past 20 years, significant advances have been made in the AAC technologies that are available to children with complex communication needs. Many children now have access to systems considered far superior to their predecessors. These systems have greater memory capacity, produce more intelligible voice output, and are able to interface with a variety of mainstream technologies, including the Internet. Despite these advances, some problems that have faced the AAC field since its inception continue to challenge clinicians. One of these is identifying the most effective method for selecting vocabulary and representing language on a child's AAC system.

Vocabulary selection requires a comprehensive understanding of the individual child and his or her evolving language needs (Marvin, Beukelman, & Bilyeu, 1994). The vocabulary selected for a child must meet several criteria. It must be "reinforcing and dynamic, responsive and functional, meeting today's needs and tomorrow's goals" (Fried-Oken & More, 1992, p. 52). It must also be appropriate to the child's gender, background, and personality (Light, 1988).

Understanding the processes underlying written language acquisition for typically developing children is critically important during the vocabulary selection process, particularly in order to support children with complex communication needs in becoming writers. Knowing the sequence in which typically developing children acquire different language forms and the types and number of words that children use at different developmental stages may guide clinicians in determining which vocabulary should be included in their AAC systems (Fried-Oken & More, 1992).

Studies have examined the vocabulary use patterns of typically developing children for the purpose of informing vocabulary selection for children with complex communication needs (Ball, Marvin, Beukelman, Lasker, & Rupp, 1999; Banajee, Dicarolo, & Stricklin, 2003; Beukelman et al., 1989; Clendon et al., 2003; Fallon et al., 2001; Fried-Oken & More, 1992; Marvin

et al., 1994; McGinnis & Beukelman, 1989). Two of these studies have documented and analyzed children's writing samples. McGinnis and Beukelman (1989) collected writing samples from 70 second-, 79 third-, 80 fourth-, 68 fifth-, and 77 sixth-grade children who all attended the same elementary school in Lincoln, Nebraska. The writing samples were taken from letter writing activities, science project assignments, and/or language arts assignments. Children across all of the grade levels wrote letters. Science project assignments, however, were only written by fourth-grade children, and language arts assignments were only written by third-, fifth-, and sixth-grade children. The researchers selected these written tasks because they believed that they provided an accurate representation of the typical writing demands placed on a mainstream student. The writing samples were transcribed into computer format and analyzed using the Systematic Analysis of Language Transcripts (SALT) software program (Miller & Chapman, 1984). The results of the analysis were used to assist an AAC team when making decisions about vocabulary representation on a sixth-grade child's augmented writing system.

When the 374 letter writing samples were combined, 161 words occurred at least once per 1000 words. Further analysis revealed that these 161 words accounted for 70% of the total sample; a significant portion of the composite sample was represented by a limited number of words. The commonality (consistency) with which the children in different grades used the 161 most frequent words was also examined. Extensive overlap was evident: 92 out of the 161 words were used in at least four out of the five grade levels (McGinnis & Beukelman, 1989).

McGinnis and Beukelman (1989) also compared the vocabulary used in the letter writing task and in the science project and language arts assignments. The authors acknowledged that these comparisons were limited by the fact that not all of the writing tasks (letter writing, language arts assignments, science project assignments) were completed across all of the grade levels. As might be expected, the results identified greater similarities between the vocabularies used in the letter writing and the language arts assignments (62%), than those used in the letter writing and the science project assignments (37%). In addition, the science project assignments appeared to include a greater number of grammatical words than the other two assignment types.

In the second study, Clendon et al. (2003) collected a total of 482 writing samples from 114 kindergarten and first-grade children. The children were recruited from 10 classrooms in three

school districts in the Southeastern United States. The children generated between 2 and 6 writing samples each. The writing samples were analyzed using the Child Language Analysis (CLAN) (MacWhinney, 1995) software program. The frequencies of words and multiword sequences were computed. Differences were examined across topics and genres.

The children used a total of 11,673 words. Of these words, 1590 were unique. As was observed in McGinnis & Beukelman's (1989) study, a significant proportion of the writing samples (70%) were represented by a relatively small number of words ($n=140$). Of the top 40 words, 27 were grammatical words (e.g., pronouns, articles, prepositions, auxiliaries, conjunctions) and 17 were content words (e.g., nouns, verbs, adjectives, and adverbs). The frequency with which particular words and multiword sequences were used varied across genres and topics. The degree of variation was greater for particular words and sequences. For instance, the sequence once upon a time was only used in the story genre. In contrast, the sequence I think was used in the plan, procedure, description, report, and opinion genres.

Perhaps the most important finding from the vocabulary studies conducted to date is that only a limited set of words is needed to represent significant portions of the spoken and written language samples produced by typically developing children. This is encouraging for the field of AAC as it suggests that typically developing children share a core vocabulary comprised of grammatical words and content words, and that this core vocabulary can be used to inform decision-making for children with complex communication needs. However, the lack of commonality observed across different environmental contexts (Marvin et al., 1994) and across different activity contexts (Banajee et al., 2003; McGinnis & Beukelman, 1989) indicates that the core vocabularies identified in these studies were unable to address all of the vocabulary needs of children with complex communication needs. Studies documenting the vocabulary that children use in other environments and in other types of language activities are needed. In addition, it seems possible that variations may exist in the core vocabularies generated by different groups of children. For instance, it is well-recognized that children in different English-speaking countries use different vocabulary words to refer to the same language concepts; for instance children in the United States use the word *swimsuit*, whereas children in New Zealand use the word *togs*. What remains unclear, however, is whether there are differences in the frequency of use patterns for the vocabulary words

that are used by both groups of children. The current investigation addresses this important issue.

METHODS

Research Question

This study was part of a broader investigation that examined the vocabulary, semantics, and syntax of beginning writers. The study was designed to explore the translation process for beginning writers and its application for children with complex communication needs who use AAC. The specific research question addressed in this study was: Are there school age and country-related differences in the vocabulary words used by typically developing beginning writers who reside in the United States and New Zealand when they compose about self-selected topics, and if so, what are these differences?

Participants

Children

The children were recruited from four schools in North Carolina, USA, and three schools in New Zealand. The children in North Carolina were in kindergarten, first-, second-, and third-grade classrooms. The children in New Zealand were in Year 1–4 classrooms. Teachers were asked to send an information letter and consent form home to the parents of all of the children in their classrooms. The information letter described the study and asked the parents to contact the researchers if they had any questions. The parents signed the consent form and returned it to their child's teacher if they were willing for their child to participate. The average consent form return rates were 42% in North Carolina and 50% in New Zealand. In order to participate in the study, the children had to be fluent in English and had to have no history of cognitive, speech, language, and/or hearing problems (as evidenced by school records). A total of 125 children in North Carolina and 113 children in New Zealand participated in the study.

The children represented a variety of ethnic backgrounds. In North Carolina, 72% of the children were United States European, 11% were African American, 3% were Hispanic, 7% were Asian, and 6% were classified as Other. In New Zealand, 70% of the children were New Zealand European, 12% were New Zealand Maori, and 18% were classified as Other. In both North

Carolina and in New Zealand, the Other classification included children from European countries and children who represented multiple ethnic backgrounds.

The children were grouped according to their school age, defined as the length of time that the children had been at school. Children who had been at school for less than one year were assigned a school age of 1; between 1 and 2 years a school age of 2; between 2 and 3 years a school age of 3, and between 3 and 4 years a school age of 4. The creation of this variable was important, as it enabled cross-country comparisons to be made. Most children in the United States start school at the beginning of the school year during which they turn 5 years old. In contrast, most children in New Zealand start school on their fifth birthday. The school age variable was created in order to establish equivalence of grade level in the United States and year level in New Zealand. The number of children within each of the school age categories and their mean chronological ages are depicted in Table 1.

Teachers

The school principals at the participating schools identified teachers who provided regular (at least three times per week) opportunities for children to write about self-selected topics. These teachers were approached and invited to participate in the study. Seventeen teachers in North Carolina and 13 teachers in New Zealand took part. Two of the teachers in New Zealand co-taught within the same class. The teachers in North Carolina had between 2 and 24 years of experience ($M=8$ years). The New Zealand teachers had between 6 and 30 years of experience ($M=15$ years). The North Carolina teachers had bachelor's degrees ($n=11$) or master's degrees ($n=6$). The teachers

in New Zealand had teaching certificates ($n=2$), teaching diplomas ($n=2$), bachelor's degrees ($n=7$), or post-graduate diplomas ($n=2$).

Schools

The four schools in North Carolina were all public elementary schools located in three school districts in the Central Piedmont region. School A was in a university town. Schools B, C, and D were within 13 miles of the university town. Schools B and C were situated in semi-rural settings, and School D was in a small town. When the study was conducted, the number of children enrolled in the schools ranged from between 400 and 700. Under North Carolina's ABC accountability program, all of the schools had been designated as Honor Schools of Excellence, meaning that at least 90% of their children were performing at or above grade level. In the 2004–2005 academic year, Schools A, B, and D exhibited high student growth, and School C exhibited expected student growth (Education First, 2005).

According to the 2000 Census in the United States, the median household income for the census tracts surrounding the four North Carolina schools ranged from US \$40,424 to US \$54,897. The median household income for the tracts surrounding the four schools all exceeded the state average of US \$39,184. The median household income for the United States was US \$41,994 (United States Census Bureau, 2000).

The three schools in New Zealand were located in suburbs of two small North Island cities. Schools E and G were situated in one city and School F was in the other city. Both cities were adjacent to a large North Island city. All schools were public primary schools. School E differed from the other two schools in that it was a

TABLE 1 Number of children (n) and their Mean Chronological Age (CA) for each School Age (SA) at each School.

School	School age							
	SA1		SA2		SA3		SA4	
	n	CA	n	CA	n	CA	n	CA
A	0		18	6;7	16	7;7	8	8;7
B	0		8	6;6	12	7;6	11	8;6
C	0		7	6;4	0		0	
D	14	5;7	25	6;7	6	7;7	0	
Total NC	14	5;7	58	6;7	34	7;6	19	8;6
E	17	5;6	5	6;7	0		0	
F	6	5;11	30	6;6	12	7;4	0	
G	9	5;5	14	6;6	13	7;8	7	8;4
Total NZ	32	5;7	49	6;6	25	7;6	7	8;4

Note: NC = North Carolina; NZ = New Zealand.

Catholic school. When the study was conducted, the number of children enrolled in the schools ranged from between 200 and 400.

According to the 2001 Census in New Zealand, the median household income for the areas surrounding the three New Zealand schools ranged from NZ \$35,195 to NZ \$66,789. The median household income for the city that Schools E and G were in was NZ \$45,667 and for the city that School F was in was NZ \$44,320. The median household income for New Zealand was NZ \$39,588 (Statistics New Zealand, 2001). School characteristics are summarized in Table 2.

Procedures

Ethical approval for the study was obtained from the Biomedical Institutional Review Board at the University of North Carolina at Chapel Hill. Teachers were asked to provide their children with opportunities to write about self-selected topics at least three times per week across a 6-week period. Self-selected writing is a common instructional component of classrooms in New Zealand and North Carolina, and involves children learning written language through multiple opportunities to write extended text about student-selected topics (Scott, 2005). To be considered for inclusion in the study, the samples had to be produced without explicit writing prompts, topic suggestions, or guidance regarding form. No restrictions were put into place regarding writing instruction that occurred outside of the self-selected writing sessions. All writing samples produced during self-selected writing sessions across the 6-week period were photocopied for analysis. The classroom teachers were asked to provide translations of any handwriting or spelling attempts that were unclear or unconventional (i.e., any words that could not easily be read by an unfamiliar conventionally literate

person). The teachers were also asked to complete bi-weekly classroom context logs with descriptions of classroom, school, community, national, and international events that may have influenced children's writing topics, vocabulary, etc.

Analysis of the Writing Samples

Sample characteristics

A total of 2,721 writing samples were analyzed. The children produced between 1 and 33 writing samples. The number of writing samples produced decreased as school age increased. The North Carolina children generated an average of 22 writing samples in School Age 1, 13 in School Age 2, 10 in School Age 3, and 5 in School Age 4. The New Zealand children generated an average of 13, 12, 8, and 3 writing samples in School Ages 1, 2, 3, and 4, respectively. All of the children in New Zealand and the children in School Ages 1 and 4 in North Carolina produced a different writing sample each day. However, the children in School Ages 2 and 3 in North Carolina often worked on the same writing sample over multiple days, adding more content to a self-selected sample when asked to write again on a subsequent day. The children were not directed to work on the same sample over multiple days.

Language analysis

The vocabulary was analyzed using the *Child Language Analysis (CLAN)* (MacWhinney, 2006) program. The *Freq* command in CLAN was used to generate word frequency counts and two summary measures: total number of words and total number of different words. A common unit of measurement, the writing session, was established, because the children produced different

TABLE 2 School characteristics.

Country	School	Location	Type	Enrollment ^{a,b}	Median household income
North Carolina	A	Small university town	Public	450	US \$47,063 ^{c,d}
	B	Semi-rural	Public	600	US \$54,897 ^c
	C	Semi-rural	Public	700	US \$54,897 ^c
	D	Small town	Public	400	US \$40,424 ^c
New Zealand	E	Suburban	Integrated ^f (Catholic)	200	NZ \$35,195 ^{e,g}
	F	Suburban	Public	350	NZ \$44,320 ^{e,h}
	G	Suburban	Public	400	NZ \$66,789 ^{e,h}

^aStatistics were obtained from school districts' websites (US) and the Ministry of Education website (NZ). ^bFigures are presented as approximates in order to maintain the anonymity of the participating schools. ^cStatistics were retrieved from the 2000 Census for the tract that each school is geographically located within. ^dThe 2000 Census was conducted prior to the construction of this school and its surrounding neighborhood so this figure should be interpreted with caution. ^eStatistics were retrieved from the 2001 Census. ^fIn intergrated schools, teachers' salaries and learning materials are publicly funded, however their land and buildings are not. These schools teach the New Zealand curriculum but incorporate their own special character (e.g., philosophy or religious belief) (Ministry of Education, 2001). ^gStatistics were retrieved for the city that this school is geographically located within. ^hStatistics were retrieved for the area unit that these schools are geographically located within.

numbers of writing samples, and because some of them children produced one sample each day while others worked on the same sample over multiple days. To calculate the total number of words, the total scores for each child were divided by the total number of writing sessions that he or she took to generate the writing samples that were included in the study. To calculate the number of different words produced per writing session and take into consideration the likelihood that children who generated one multi-day sample would have reused vocabulary if they had written several samples (i.e., one per day) over the same time-frame, each child's writing samples were divided into equal portions by the number of writing sessions. The number of different words-score was then calculated for each portion of the writing sample. This ensured that the children who produced writing samples over multiple days received appropriate credit for their use of different words. It also accounted for the differences in the number of samples produced by individual children. The total number of words and total number of different words were calculated for each writing sample. Mean scores were generated for the set of writing samples produced by each child. These scores were used to calculate means for the four school ages within and across countries. Lastly, the *Cooccur* command was used to find and document the frequency of multiword sequences. The teachers did not always provide translations for the children's spelling attempts. Whenever the researcher was unsure of a word, it was marked with an X and ignored in the analyses.

Data analysis

Descriptive statistics and graphs for total number of words and total number of different words in each school age and country were generated and compared. Formal statistical tests were not used because the variables failed to meet the assumptions of normality and homogeneity of variance.

Core vocabulary lists were generated for the entire writing sample population as well as for each school age, country, and school age within

each country. The most frequently occurring content words (nouns, verbs, adjectives, and adverbs) were identified, as well as the most frequently occurring grammatical words (pronouns, articles, prepositions, auxiliaries, conjunctions). The grammatical words were obtained from lists provided by MacWhinney (1995). Some words in English can be used as both content words and grammatical words. For instance, the word *one* can be used to mean a number (content word) or pronoun (grammatical word). Due to the large volume of writing samples included in the vocabulary analyses, it was not feasible to identify the true meaning of every word. The decision was made to treat all words that could take either form as grammatical words.

RESULTS

Total Words and Different Words

The writing samples contained a total of 85,759 words, including 5,724 different words. When the writing samples written by the children in North Carolina and New Zealand were combined, the mean number of words written per writing session increased as a function of school age (see Table 3). The children in School Age 1 wrote an average of 11.95 words per session. This increased to 24.96 at School Age 2, 46.08 at School Age 3, and 79.30 at School Age 4. The mean number of different words per session also increased from 8.74 at School Age 1 to 17.92 at School Age 2, 31.72 at School Age 3, and 50.13 at School Age 4.

Similar patterns were evident when the writing samples written by the children in North Carolina and New Zealand were separated (see Table 3). In North Carolina, the mean number of words per writing session increased as follows: 13.21 words per session in School Age 1; 25.57 in School Age 2; 47.00 in School Age 3; and 79.36 in School Age 4. In New Zealand, the same increase was observed, with children in School Ages 1, 2, 3, and 4 writing an average of 10.59, 26.08, 49.07, and 91.99 words per session, respectively. The mean number of different words per session in North Carolina was 8.34 in School Age 1, 17.31 in

TABLE 3 Mean total words and mean different words per writing session for each school age for North Carolina (NC) and New Zealand (NZ) and for the combined sample.

School Age	Mean total words			Mean different words		
	NC	NZ	Countries combined	NC	NZ	Countries combined
1	13.21	10.59	11.95	8.34	8.92	8.74
2	25.57	26.08	24.96	17.31	18.63	17.92
3	47.00	49.07	46.08	30.84	32.92	31.72
4	79.36	91.99	79.30	47.44	57.41	50.13

School Age 2, 30.84 in School Age 3, and 47.44 in School Age 4. In New Zealand, the children in School Age 1 wrote 8.92 different words per session. This increased to 18.63 in School Age 2, 32.92 in School Age 3, and 57.41 in School Age 4. For total number of words and total number of different words, the differences between the two countries were minimal across the first three school ages. In School Age 4, however, the children in New Zealand produced noticeably more words and different words than the children in North Carolina.

Most Frequently Occurring Words

The cumulative frequency analyses revealed that the most frequently occurring 163 words accounted for 70% of the total words used, and the most frequently occurring 39 words accounted for 50% of the total words used. Many of the most frequently occurring words were grammatical words. The ratio of grammatical words to content words was 35:15 (70%) within the 50 most frequently occurring words, 56:44 (56%) within the 100 most frequently occurring words, and 68:82 (45.33%) within the 150 most frequently occurring words. Only three of the most frequently occurring content words (Halloween, Fall, Christmas) appeared to directly correspond to an event or theme that was documented by the teachers in their context logs. Tables delineating the most frequently occurring words are provided on the Centre for Literacy and Disability Studies website: <http://www.med.unc.edu/ahs/clds/resources.html>.

When the countries were combined and the school ages were compared, some interesting patterns were evident. Seven of the top 10 words overall were represented in the top 10 words in all of the school age groups. The three words that were not represented consistently were: is (ranked 11 in School Age 4), we (ranked 14 in School Age 1), and was (ranked 22 in School Age 1). Some words received similar rankings across the four school age groups and others did not. Of the top 50 words overall, the words that had the most disparate rankings were then, one, they, when, so, and said, used very infrequently by children in School Age 1; and am and love, used very infrequently by children in School Age 4.

For the next analysis, the school ages were combined, and the writing samples generated by all of the children in North Carolina were compared to those generated by all of the children in New Zealand. This analysis revealed more similarities than it did differences. For instance, 9 of the top 10 words overall were represented in the top 10 words in both countries. In addition,

15 of the top 20 words overall were represented in the top 20 words in both countries. Many of the words received similar rankings, but some did not. For instance, the word see was ranked 14 in North Carolina and 90 in New Zealand.

When the school ages from the two countries were examined separately, there were a number of disparities. The word he, for example, was used much more frequently by children in School Age 1 in New Zealand (rank = 32) than it was by children in the same school age group in North Carolina (rank = 190). In addition, the word is was used much more frequently by children in School Age 4 in North Carolina (rank = 11) than it was by children in the same school age group in New Zealand (rank = 108).

Most Frequently Occurring Multiword Sequences

Significant variability was evident in the use of multiword sequences. The children in School Ages 1–4 in North Carolina used 19 two-word sequences at a frequency of greater than 5 per 1000, but the children in School Ages 1–4 in New Zealand used 7 two-word sequences. The children in School Age 1 in North Carolina tended to use multiword sequences as simple sentence frames when generating their stories (e.g., I see, I like). The vocabulary analyses indicated that they used these multiword sequences upwards of 887 times per 1000 sequences. In contrast, the most frequently occurring two-word sequence in New Zealand, I am, was only used 25 times per 1000 sequences. Multiword sequences were also used less frequently by children in School Ages 2–4 in both countries. Tables delineating the most frequently occurring multiword sequences are provided on the Centre for Literacy and Disability Studies website: <http://www.med.unc.edu/ahs/clds/resources.html>.

The patterns of use of two-word sequences were very different in the two countries with the school ages combined. Nine of the two-word sequences and 16 of the three-word sequences that appeared in the top 25 rankings for North Carolina did not appear in the top 25 rankings for New Zealand. The sequence I see illustrates the extreme variability. The ranking for I see was 2 in North Carolina and 768 in New Zealand. As might be expected, the patterns of use for I see across the school ages within the two countries were also very different. In North Carolina, the rankings were as follows: 1 in School Age 1, 107 in School Age 2, and 5,186 in School Age 3. It was not used at all by children in School Age 4. The pattern for the children in New Zealand was very different. The children in School Ages 3 and 4 in New Zealand did not use the sequence I see at all, and

the rankings for children in School Ages 1 and 2 were 162 and 3004, respectively.

Other examples of sequences that demonstrated extensive variability were in the and it was fun. The ranking of the two-word sequence in the was 25 or less for all school ages in New Zealand and for School Ages 2, 3, and 4 in North Carolina. In contrast, the same sequence obtained a ranking of 587 for School Age 1 in North Carolina. The three-word sequence it was fun obtained rankings ranging from 8 to 582 in North Carolina and from 18 to 338 in New Zealand.

DISCUSSION

This study identified a number of school-age and country-related differences in the vocabulary words used when beginning writers compose about self-selected topics. In this section, some of the key findings will be discussed, along with their implications for teaching and supporting written language development in children with complex communication needs.

Maximizing Efficiency with Core Vocabulary

Consistent with other studies that have examined the vocabulary of typically developing children (e.g., Beukelman et al., 1989; McGinnis & Beukelman, 1989), this study found that a relatively small core vocabulary was needed to represent significant portions of the language samples. For example, in this study, the top 163 words accounted for 70% of the total words used and the top 39 words accounted for 50% of the total words used. In McGinnis and Beukelman's study (1989) which examined the writing samples generated by children in Grades 2–6 in letter-writing activities, science project assignments, and/or language arts assignments, 161 words accounted for 70% of the total sample and 46 words accounted for 50% of the total sample. In Clendon et al.'s (2003) study, which examined the writing samples produced by children in kindergarten and first-grade on self-selected topics, 140 words accounted for 70% of the total sample and 39 words accounted for 50% of the total sample.

The identification of a core vocabulary has important implications for helping children with complex communication needs to become effective writers. For instance, if only 39 words account for 50% of what children write, then AAC teams can focus on teaching these words in their reading and spelling instruction, and on providing fast and easy access to these words so that children with complex communication needs

can focus their energies on spelling or otherwise generating less frequent or familiar words.

Comparison of the most frequently occurring words in this study and in the previous two studies revealed that there were a number of words that occurred with high frequency in one study that did not occur with the same high frequency in the other study. McGinnis and Beukelman (1989) provided frequency information for vocabulary generated in letter writing activities and language arts assignments. The overlap between the top 50 words used in this study and the top 50 words used in these two writing activities was 62% for letter writing activities and 56% for language arts assignments. A total of 25 words (50%) appeared on all three lists. Clendon et al. (2003) documented the top 20 grammatical words and the top 20 content words used by the participants in their study. The overlap between the studies was 75% for grammatical words and 65% for content words. The differences across the studies are likely related to the contexts of the writing assignments, the instruction teachers provided outside of the self-selected writing sessions, and the nature of the reading materials the children encountered. Given that none of the studies attempted to control any of these external factors, the proportion of overlap suggests that there is, in fact, a set of core words children use when writing.

The current study adds to the research base by exploring the differences in the vocabulary used across four school age groups and two countries. The implication of the findings of the three studies for children with complex communication needs is that access to the most frequently occurring words overall, will allow efficient access to a large portion of the vocabulary required to write about the topics they choose. If AAC teams develop core vocabularies that change over time, taking into consideration a child's school age, country, and the language contexts within which the child communicates, then it is likely that children with complex communication needs will be much closer to being able to use their AAC systems to write in an efficient manner.

Maximizing Efficiency with Multiword Sequences

The findings of this study suggest that some multiword sequences may be used with sufficiently high frequency to warrant their storage as whole units in AAC systems. The children in School Ages 1–4 in North Carolina used 19 two-word sequences that occurred at a frequency of greater than 5 per 1,000, and the children in School Ages 1–4 in New Zealand used 7. This finding, while highlighting the frequent usage of

some multiword sequences, also draws attention to the fact that the usage patterns for multiword sequences were different for the children in the two countries. The children in School Age 1 in North Carolina used a relatively large number of frequently occurring multiword sequences as simple sentence frames when generating their stories (e.g., I see, I like). The children in School Age 1 in New Zealand and the children in School Ages 2–4 in both countries, however, did not. These findings again suggest the importance of taking into consideration country and school age when making decisions about how best to store vocabulary in a child's AAC system. For younger children in North Carolina, the findings suggest that access to frequently occurring two-word sequences would support efficient written language development, while children in New Zealand of any age might receive limited benefit.

Cross-Cultural Differences in Language Use

The current study explored differences in the written language produced by children in two different countries given the same opportunities to write about self-selected topics. The language analyses carried out identified some differences between the countries in the children's use of high frequency words and sequences. One example of this was the verb see. When the school ages were combined and the writing samples generated by the children in North Carolina were compared to those generated by the children in New Zealand, the verb see was ranked 14 in North Carolina, but 90 in New Zealand.

This finding is particularly informative for New Zealand-based AAC teams. There are very few, if any, AAC systems that are manufactured in New Zealand. Most of the systems that children receive are imported from other countries, particularly from the United States. When introducing one of these AAC systems to a child in New Zealand, it is likely that AAC teams will be concerned about the appropriateness of the vocabulary prestored in these systems. They may be concerned, for instance, that an AAC system manufactured in the United States says the word *sweater*, when children in New Zealand say *jumper*. AAC teams will benefit from knowing that the differences observed in the fringe vocabulary used by children in the two countries are only part of the problem. This study suggests that there are important modifications to be made to some of the high frequency words as well; not because the children in the two countries use different lexemes, but more because the children exhibit different patterns of use. Dealing with country-related differences in core vocabulary in addition

to fringe vocabulary will lead to greater efficiency and precision for children with complex communication needs who use their AAC systems to write.

School Age Differences in Vocabulary Use

The analyses revealed differences in the vocabulary used by the children in the upper School Age groups in comparison to the children who were in School Age 1. Words such as *so*, *when*, and *then*, for example, were used much more frequently by children in School Ages 2, 3, and 4, than they were by children in School Age 1. This is likely indicative of the children's advancing language skills and their use of more sophisticated grammatical structures such as subordinate and relative clauses. The implication of these findings for the AAC field is the need to take into consideration the cognitive, linguistic, memory, and physical demands inherent in learning a new expressive form and the impact that these demands can have on the quantity and quality of the language that children produce. Harpin (1976) stated that:

When children begin the process of learning to write there is a big gap between their general language competence and their performance. The effort involved in learning a new skill is considerable and attention is, naturally enough, on the mechanics of the business. What is drawn on from those oral language resources is sharply restricted. As the act of writing becomes habitual, so more opportunity is available to bring oral competence and written performance into harmony. (p. 52)

The act of translation as it is described in Flower and Hayes' (1981) model is challenging for beginning writers, just as it is for children who use AAC. This study examined the vocabulary use of beginning writers in order to provide a window into a comparable learning scenario. The degree of similarity between the experiences of learning to write using a pencil, and learning to write using an AAC system merit further consideration. It seems probable that when some beginning writers engage in the process of translation that they find themselves restricted by their spelling capabilities. They may substitute more descriptive and elegant adjectives, for example, for more simplistic forms. Obviously, this is influenced by the writing instruction they receive and the type of writing product that is valued in their classroom. Nevertheless, this scenario is similar to that observed in children

with complex communication needs, where what they are able to express in both their face-to-face and written communication is sharply restricted by the vocabulary available to them, and the manner in which that vocabulary is organized.

Limitations of the Study

In studies that attempt to quantify language skills, the effects of context must be taken into consideration (Nelson, 1988) and the findings must always be interpreted with a degree of caution. In the current study some of the less obvious contextual influences became more apparent when the multiword sequences were examined. For instance, the two-word sequence *is for* received a ranking of 1 in School Age 3 in North Carolina. This was because the children in this school age group wrote a large number of alphabet stories (e.g., A is for ants, B is for bananas, ...). The teachers in two of the participating classrooms had read alphabet stories to the children in their classrooms as part of the language arts curriculum. Reportedly, the children had enjoyed the stories and when it came time for self-selected writing, a number of them chose to write their own. While this form of writing was self-selected, it is clear that the context heavily influenced the selection.

The fact that context has had an influence on the findings of this study reduces their generalizability. Perhaps the need to minimize contextual influences has led previous researchers to ask children to write on controlled topics. However, asking children to write about self-selected topics was important in the current study because the purpose was to elicit a broader range of vocabulary and language structures. In addition, when children write about self-selected topics, they are not restricted by their background knowledge in the way that they are when they write under more researcher or teacher-directed conditions.

Beyond this teacher-reported influence, the country-related differences may have been related to instructional differences more broadly. For instance, the high prevalence of simple sentence frames (e.g., *I see*, *I like*) observed in the writing samples generated by children in School Age 1 in North Carolina may be indicative of the type of reading books that children are currently exposed to in United States classrooms. No specific data were collected to support this hypothesis, so it is difficult to determine the exact nature of the instructional differences between the two countries and their potential influence on the study's findings.

Other limitations were inherent in this study. The small number of New Zealand children in School Age 4 ($n = 7$) may have resulted in some distorted findings. For instance, some of the top 25 three-word sequences used by the children in School Age 4 in New Zealand may have seemed surprising, as they were not sequences that one would expect to occur very frequently in written language. Examples included the sequences *breathing hot flaming* and *flaming hot fire*. These sequences may have received a high ranking because there were so few New Zealand children particularly in School Age 4. In fact, these sequences were only used twice, but because of the small sample size, they received a high ranking.

The amount of time that the children had to write each day was not controlled. This facilitated the recruitment process, as more teachers were willing to participate when they learned that the researcher was not asking them to make any changes to their typical classroom practice. However, the fact that some children generated one sample per day and other children generated one sample over multiple days was also a limitation of the study. The vocabulary measures were weighted by number of writing days in order to reduce the impact that this limitation had on the study's findings, but the weighting procedure did not eradicate the limitation altogether.

Another limitation was the presence of unintelligible words (spelling attempts for which the teachers did not provide translations) in the children's writing samples. Whenever the researcher was unsure of a word, it was marked with an X and ignored in the analyses. Of the total words, 1.2% were marked as unintelligible and deleted from the overall corpus.

Future Directions

Such a large sample of children's writing lends itself to a myriad of further analyses. Relationships between the word frequencies observed in this study and in large corpora of written English should be explored. In addition, detailed analyses of multiword sequences might include examining frequently occurring sequences with variable slots (e.g., "I like ___") and ranking possible slot fillers according to their frequency of use so that they can be stored efficiently in children's AAC systems. The identification and examination of larger constructions such as "*I like ___ because ___*" would also be of interest. Comparison of the types of vocabulary that children use across different types of text (e.g., different genres and topics), and investigation of school age and country differences in the types of texts

children produce, would permit examination of the relationship between text type and vocabulary use.

CONCLUSION

One of the most important goals for speech-language pathologists, teachers, and other professionals who work with children with complex communication needs is to help these children to develop the ability to communicate both efficiently and precisely in face-to-face and written forms. In order to achieve this goal, members of AAC teams must make thoughtful decisions about vocabulary selection and language representation. The findings of the current study provide research-based evidence that can be used to help guide AAC teams as they make these important decisions.

This study analyzed 2721 writing samples written by 125 children in North Carolina and 113 children in New Zealand. The findings identified school-age and cross-cultural differences in the words used by typically developing children when composing about self-selected topics. The school-age comparisons highlighted the restricted vocabulary use in the earliest stages of writing development. The country comparisons revealed differences in core vocabulary.

The field of AAC is exciting because it is constantly evolving. New technologies are being developed at rapid rates. In order to maximize the capabilities of these technologies, however, children with complex communication needs must have access to language that is organized thoughtfully and logically, and that supports the development of both efficient and precise communication. The findings of this study highlight the need to carefully consider a child's school age and country when making decisions about vocabulary selection and language representation. Increased understanding of these issues will help AAC teams to support children with complex communication needs to overcome the challenges inherent in learning to use AAC to compose and to achieve the goal of engaging in effective communication.

Acknowledgements

This paper is based on a doctoral dissertation that the first author completed in 2006 for the PhD in Speech and Hearing Sciences degree program at the University of North Carolina at Chapel Hill. The dissertation was completed under the direction of Dr Karen Erickson and was entitled "The Language of Beginning Writers: Implications for

Children who use Augmentative and Alternative Communication."

The first author is indebted to the Rotary Foundation for the financial support that they provided through the Ambassadorial Scholarship Program and to Dr Thomas S. and Mrs Caroline H. Royster for the financial support and educational opportunities that they provided through the University of North Carolina Graduate School's Society of Fellows Program. The first author would like to express her sincere thanks to the principals who supported her research, to the teachers and students who participated and to Ms Mary Holman and Ms Caroline Harper for patiently and generously assisting with transcribing the writing samples.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

- Abbott, R., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary and intermediate grade students. *Journal of Educational Psychology, 85*(3), 478–508.
- Ball, L. J., Marvin, C. A., Beukelman, D. R., Lasker, J., & Rupp, D. (1999). Generic talk use by preschool children. *Augmentative and Alternative Communication, 15*, 145–155.
- Banajee, M., Dicarlo, C., & Stricklin, S. B. (2003). Core vocabulary determination for toddlers. *Augmentative and Alternative Communication, 19*, 67–73.
- Bear, M. V. (1939). Children's growth in the use of written language. *Elementary English Review, 16*, 312–319.
- Berninger, V. W., Fuller, F., & Whitaker, D. (1996). A process model of writing development across the life span. *Educational Psychology Review, 8*(3), 193–218.
- Berninger, V. W., Whitaker, D., Feng, Y., Swanson, L., & Abbott, R. (1996). Assessment of planning, translating, and revision in junior high writers. *Journal of School Psychology, 34*, 23–52.
- Berninger, V. W., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing: An Interdisciplinary Journal, 4*, 257–280.
- Beukelman, D. R., Jones, R. S., & Rowan, M. (1989). Frequency of word usage by nondisabled peers in integrated preschool classrooms. *Augmentative and Alternative Communication, 5*, 243–248.
- Cleendon, S., Sturm, J., & Cali, K., (2003, November). *The vocabularies of beginning writers: Implications for students who use AAC*. Seminar presented at the American Speech-Language-Hearing Association Convention, Chicago, IL.
- Education First. (2005). *North Carolina school report cards*. Retrieved 13 June 2007 from: <http://www.ncreportcards.org/src/>
- Fallon, K. A., Light, J. C., & Paige, T. K. (2001). Enhancing vocabulary selection for preschoolers who require augmentative and alternative communication. *American Journal of Speech-Language Pathology, 10*, 81–94.

- Farr, R., Kelleher, C., Lee, K., & Beverstock, C. (1989). *An analysis of the spelling patterns of children in grades two through eight*. Bloomington: Center for Reading and Language Studies, Indiana University.
- Fitzgerald, J. A. (1938). The vocabulary and spelling errors of third-grade children's life-letters. *Elementary School Journal*, 38, 518–527.
- Flower, L., & Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32, 365–387.
- Fried-Oken, M., & More, L. (1992). An initial vocabulary for nonspeaking preschool children based on developmental and environmental language sources. *Augmentative and Alternative Communication*, 8, 41–56.
- Gombert, J. E. (1992). *Metalinguistic development*. IL: University of Chicago Press.
- Graham, S., Harris, K., MacArthur, C., & Schwartz, S. (1991). Writing and writing instruction for students with learning disabilities: Review of a research program. *Learning Disabilities Quarterly*, 14, 89–114.
- Harpin, W. (1976). *The second 'R': Writing development in the junior school* (315 pp.). London: George Allen & Unwin.
- Hillerich, R. L. (1978). *A writing vocabulary of elementary children*. Springfield, IL: Charles C. Thomas.
- Hunt, K. (1964). *Differences in grammatical structures written at three grade levels, the structures to be analyzed by transformational methods*. Tallahassee: Florida State University.
- Johnson, D. D. (2000). Just the right word: Vocabulary and writing. In R. Indrisano & J. R. Squire (Eds.), *Perspectives on writing* (pp. 162–186). Newark, DE: International Reading Association.
- Kroll, B. M. (1981). Developmental relationships between speaking and writing. In B. M. Kroll & R. J. Vann (Eds.), *Exploring speaking-writing relationships: Connections and contrasts* (pp. 32–54). Urbana, IL: National Council of Teachers of English.
- Light, J. (1988). Interaction involving individuals using augmentative and alternative communication systems: State of the art and future directions. *Augmentative and Alternative Communication*, 4, 66–82.
- Lorenz, E. L. (1931). The writing vocabulary of third grade children. *Elementary English Review*, 8, 21–22, 24.
- MacArthur, C., Harris, K., & Graham, S. (1994). Helping students with learning disabilities plan compositions: Instruction in cognitive and metacognitive strategies. In E. Butterfield (Ed.), *Children's writing: Toward a process theory of the development of skilled writing* (pp. 173–198). Greenwich, CT: JAI Press.
- MacWhinney, B. (1995). *The CHILDES project: Tools for analyzing talk* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- MacWhinney, B. (2006). Child Language Analysis (CLAN) (Version 3) [Computer software]. Child Language Data Exchange System. Retrieved 5 January 2006, from: <http://childes.psy.cmu.edu/>
- Marvin, C. A., Beukelman, D. R., & Bilyeu, D. (1994). Vocabulary-use patterns in preschool children: Effects of context and time sampling. *Augmentative and Alternative Communication*, 10, 224–236.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35, 13–23.
- McGinnis, J. S., & Beukelman, D. R. (1989). Vocabulary requirements for writing activities for the academically mainstreamed student with disabilities. *Augmentative and Alternative Communication*, 5, 183–191.
- Miller, J., & Chapman, R. S. (1984). Disorders of communication: Investigating the development of language in mentally retarded children. *American Journal of Mental Deficiency*, 88, 536–545.
- Nelson, N. W. (1988). Reading and writing. In M. A. Nippold (Ed.), *Later language development: Ages nine through nineteen* (pp. 97–125). Boston: College Hill.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Rinsland, H. D. (1945). *A basic vocabulary of elementary school children*. New York: Macmillan.
- Scardamalia, M. (1981). How children cope with the cognitive demands of writing. In C. H. Frederiksen & J. F. Dominic (Eds.), *Writing: The nature, development, and teaching of written communication* (pp. 81–103). Hillsdale, NJ: Lawrence Erlbaum.
- Scott, C. M. (2005). Learning to write. In H. W. Catts, & A. G. Kamhi (Eds.), *Language and reading disabilities* (pp. 233–273). Boston, MA: Pearson.
- Shapiro, J., & Gunderson, L. (1988). A comparison of vocabulary generated by grade 1 students in whole language classrooms and basal reader vocabulary. *Reading Research and Instruction*, 27(2), 40–46.
- Singer, B. D., & Bashir, A.S. (2004). Developmental variations in writing composition. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 559–582). New York: Guilford.
- Smith, C., & Ingersoll, G. (1984). *Written vocabulary of elementary school pupils, ages 6–14*. Bloomington: Indiana University.
- Statistics New Zealand (2001). *Census of population and dwellings*. Retrieved 13 June 2007 from: <http://www.stats.govt.nz/census/default.htm>
- United States Census Bureau (2000). *Decennial census*. Retrieved 13 June 2007 from: http://factfinder.census.gov/home/saff/main.html?_lang=e