Australian Journal of Educational & Developmental Psychology. Vol 6, pp 1-11

An Error Analysis of Elementary School Children's Number Production Abilities

Sheri-Lynn Skwarchuk[1] & Paul Betts
University of Winnipeg

ABSTRACT

Translating numerals into number words is a tacit task requiring linguistic and mathematical knowledge. This project expanded on previous number production models by examining developmental differences in children's number naming errors. Ninety-six children from grades one, three, five, and seven translated a random set of numerals into number words. Their errors were coded according to location and type. The data were analysed by grade using ANOVAs followed by Tukey post hoc tests. Results suggest that prerequisite linguistic conventions are needed to start and sustain the number production process, and may have implications for curricular recommendations concerning number naming.

INTRODUCTION

Although counting is recognized as a core mathematical competency (National Council for the Teaching of Mathematics, 2000; 2004), little research has concentrated on the process through which children learn to produce numbers (and particularly large numbers) in the cardinal number sequence. Most studies have concentrated on preschoolers' and early school-aged children's use of problem solving to produce numbers in the decades, developing understanding of numeracy and basic "how to count" principles (e.g., Fuson, 1988; Baroody & Price, 1983). Research is emerging on the number production abilities of older children (Bell & Burns, 1981; Fuson, Richards, & Briars, 1982; Power & Dal Martello, 1990; Seron & Fayol, 1994). Once children acquire the decade sequence, researchers have assumed that the remaining large cardinal number words are generated (as opposed to memorized) through the use of a set of conventional rules (Ginsburg, 1977; Baroody, 1987; Fuson, 1990; Hurford, 1975; Kamii, 1981; McCloskey, Sokol & Goodman, 1986; Power & Longuet-Higgins, 1987).

Culminating ideas from various cognitive models and research on number production, the ability to produce spoken number words from their arabic form seems to be dependent on at least four factors: (1) appropriate vocabulary knowledge (e.g., Anglin, 1993), (2) knowledge of a set of implicit compound rules needed to order the vocabulary words correctly within the number word compound (e.g., Hurford, 1975; McCloskey et al., 1986), (3) consideration of some prerequisite linguistic conventions for

[1] Contact
Sheri-Lynn Skwarchuk
Email: s.skwarchuk@uwinnipeg.ca
Paul Betts
Email: p.betts@uwinnipeg.ca
Faculty of Education and Extended Learning
University of Winnipeg
515 Portage Avenue, Winnipeg, Manitoba, Canada, R3B-2E9
Phone: 204-786-9164
Fax: 204-772-7980

translating a numeral into words [and although these conventions have not been formally described together, they have been alluded to in work by Baroody (1987), Kamii, (1981), Fuson, (1990), and McCloskey et al., (1986)], and (4) skills to orchestrate the above knowledge. Producing number words in written form is further complicated since children must also coordinate their knowledge of the alphabetic writing system with the numeration system (Kamii, 1981; McCloskey et al., 1986). These factors are described in detail below.

With regard to vocabulary knowledge, it has been proposed that children do not memorize all of the number words (e.g., Fuson, 1988; Baroody, 1987; Ginsburg, 1977; Seron & Deloche, 1987). Instead, it is speculated that they memorize a minimum of 31 basic number words (i.e., the number words *one* through *nine*, *ten* through *nineteen*, *twenty*, *thirty*…*ninety*, and *hundred*, *thousand*, *million*, and *billion*), and they are able to construct all or almost all additional number words through problem solving.

Most researchers agree that a number production model also must address the coordination of known vocabulary words with a syntactic frame (e.g., Anglin, 1993; Campbell & Clark, 1992; Dehaene, 1992; Seron & Fayol, 1994; McCloskey, et al., 1986; McCloskey, Sokol, Goodman-Schulman, & Caramazza, 1990; McCloskey & Caramazza, 1987). McCloskey et al., (1986) proposed a model based on research with adult aphasics. In this work, number words are represented using five discrete classes in semantic memory: ones, teens, tens, hundreds and larger numeral components known as multipliers (e.g., *thousand*, *million*). The large multi-digit number words are constructed by creating a base syntactic structure of ones, teens, tens, and hundreds components using the information provided in each segment of three digits. Depending on the position of those three digits within a numeral, a multiplier word is inserted into the syntactic component. Meaning depends on the exact ordering of the words within the compound. For example, *one hundred three* communicates a different meaning than *three hundred one*. Further, knowledge of the conventional number word syntax ensures the correct translation of 462,000 as *four hundred sixty-two thousand* and not *four thousand sixty-two hundred* or some other plausible syntactically-generated alternative. Hurford (1975) attempted to document the specific syntactic structures that are relevant to generating cardinal number words. Power and Longuet-Higgins (1987) obtained evidence that the construction of number words is dependent on knowledge of a syntactic structure. With the input of a few vocabulary words and the syntactic structure, numerals could be generated using a computer program.

However, current cognitive models do not clearly articulate the orchestration of appropriate vocabulary and a syntactic structure to produce a number word. Additional rudimentary knowledge about the workings of the number system and the conventions for translating a numeral into words are necessary in parsing a numeral, which we refer to as prerequisite linguistic conventions. These prerequisite linguistic conventions refer to a set of principles related to the general knowledge of the cardinal number system, and are needed to initiate and maintain the translation of arabic numerals into number words. For example, when children are preparing to construct a number word from its arabic form, they must have implicit knowledge, at a minimum, that the symbols 0-9 are used to identify numerical quantities (Kamii, 1981), and their digit value can be translated linguistically (i.e., 1 is *one*). Children must consider that symbols within a numeral are organized from left to right so that both the digit value and the positional value are relevant to number meaning (Kamii, 1981; Kibel, 1992). When producing number words, children must give consideration to digits grouped into sets of three (Fuson, 1990). Finally, children must be aware of the rules that exist when interpreting the role of a zero in a numeral (Cathcart, Pothier & Vance, 2000; Baroody, Gannon, Berent & Ginsburg, 1984). The prerequisite linguistic conventions provide preliminary information, distinct from vocabulary, syntactic knowledge, and place value knowledge, which is used to understand the accepted conventions for reading numerals.

Understanding the factors relevant in number word production is important from a theoretical perspective. Previously developed cognitive models have focused on describing the end-point of numerical development (e.g., Campbell & Clark, 1992; Dehaene 1992; McCloskey et al., 1986; Seron & Deloche, 1987), and responses from brain-injured patients are used as supportive evidence (e.g., McCloskey et. al, 1986; Seron & Deloche, 1987). A developmental approach may provide insight into the process of acquiring number knowledge, in addition to clarifying the complete model of numerical processing in normal adult development.

One method of understanding the factors involved in the process of translating numerals into their number word form is to analyze errors. Ginsburg (1977) interviewed children and analyzed errors to provide evidence that children learn the cardinal sequence informally, first reading numerals on how they look or sound. Seron and Fayol (1994) and Power and Dal Martello (1990) examined children's errors in translating number words into numerals. The authors of both studies reported a high occurrence of errors involving the addition of extra digits, usually zeros to the target numeral, which they interpreted as reflecting problems in understanding the syntactic structure of the number word. They also reported the rare occurrence of errors involving the substitution of the correct digit for an incorrect digit while still maintaining correct syntactic structure. The studies provide some support that (at least) knowledge of both vocabulary words and a compound structure are required to produce number words. However, sample size was low, the studies were conducted in non-English languages, and the focus was on translating the reverse process of number words into numerals, which may not involve the same cognitive mechanisms.

Baroody et al. (1984) examined the number production errors produced by children in a numeral to number word translation task. Most of the children's errors were plausible number word substitutions. However, they reported that in some errors, children either substituted known vocabulary words (e.g., *seventeen* for 47) or they "made up" terms (e.g., *sixteen-twenty* for 162) that "resembled" the target numeral. Although these errors could reflect deficits in vocabulary knowledge (e.g., not knowing the word *forty* to produce 47), the errors could also represent a lack of knowledge concerning the prerequisite linguistic conventions (e.g., incorrectly parsing a three digit numeral into groups of two). Baroody et al. (1984) also reported that numbers involving zeros created particular difficulty for students, and attributed some of these errors to children's rudimentary interpretation of a zero as meaning "nothing." Further research is needed to clarify the issue of needing a specialized linguistic rule to deal with naming numerals containing zero(s), and other conventions that do not easily map onto vocabulary or syntactical concerns as previous reported.

The purpose of the current project was to expand on previous research and develop further understanding of the number production process by examining over 1800 number naming errors for a large set of randomly selected numbers into the billions. We attempt to document the qualitative differences in children's number knowledge with grade. Children were asked to verbalize Arabic numerals. Their errors on this task were transcribed and coded. In terms of a developmental pattern, we predict that younger children who are still learning to coordinate the basics of number production will produce more errors affecting all syntactic sections of the number word, reflecting deficits in knowledge of prerequisite linguistic conventions compared to their older, more capable peers. However, it is hypothesized that the other categories of errors reflecting deficit knowledge in vocabulary and syntactic structure will be dependent on the knowledge of each individual child, and thus, will be distributed fairly evenly across each grade.

METHOD

Participants

Twenty-four children from each of grades one, three, five, and seven, and representing equal numbers of boys and girls within each grade participated. Mean ages of the children were 6 years, 6 months in grade one; 8 years, 8 months in grade three; 10 years, 8 months in grade five; and 12 years, 8 months in grade seven. All children spoke English as their first language and were instructed in English at school. None of the children had any language impediments or learning disabilities as reported by their teachers. Four children were replaced as they did not meet these criteria.

Materials

The sample of number stimuli included 120 randomly selected numbers ranging from 1 to 1,000,000,000,000, divided equally across five number series (i.e., decades, hundreds, thousands, millions and billions). There were also 36 special case numbers which included the basic number vocabulary words (*one, two, three…ten, twenty, thirty…hundred, thousand, million*), as well as the number word equivalents for numbers occurring during certain transition points (e.g., 999; 9,999; 99,999 and 100, 1000, 10,000). All numbers were presented to children in ascending order according to number series (i.e., decades were always presented first, then hundreds), but the numbers were randomized within each series for each participant.

Procedure

Children were interviewed individually and responses were audio-taped. They completed four tasks, but only the task in which an error analysis was conducted is described [consult Skwarchuk and Anglin (2002) for details regarding the other tasks]. On the number production task, 120 randomly selected numbers and 36 special case numbers were presented in arabic form one at a time on a laptop computer screen. Children were asked to name each number and then to count forward or backward from them by one. Both naming and counting trials were included in the study to ensure the results for the project could be generalized to both the processes of number naming and counting. Children continued the task until they either completed all numbers, or until they made eight consecutive incorrect attempts through errors or omissions. Incorrect number naming and counting attempts were identified, transcribed and coded. Preliminary analyses were conducted on both the naming and the counting data. However, since most of the erroneous responses in the counting trials were identical to the naming trials, with the exception that the last digit(s) were increased or decreased by one, only the naming trial findings were analysed in detail and are reported in this paper.

Error Coding

A coding scheme for children's incorrect responses during the number naming task was developed to determine: 1) the location(s) of the compound number word affected by the error(s), and 2) a categorization of the types of errors made. An infrequent number of children's incorrect responses were also categorized as "did not know-no attempt" when they did not know the answer and did not attempt it, or were classified as "incomprehensible" when it was not possible to decipher the response. These error categories were not considered in the analyses.

The location of an error was coded by determining the specific position(s) in the number word compound affected by the error. Errors were classified as occurring in the units, decades, hundreds, thousands, millions, or billions sections of a number word. Incorrect attempts in translating the first, second and third digits from the right were considered to affect the unit, decade and hundred sections, respectively. Errors occurring while translating into words the fourth to sixth digits, the seventh to ninth digits, and the tenth to twelfth digits from the right, were classified as affecting the thousands, millions, and billions sections of a numeral, respectively. Responses could be coded as having more than one error, and they could affect one, several or all sections of a numeral.

After identifying the location of an error, the nature of the error was classified according to three categories: 1) *prerequisite linguistic*, 2) *vocabulary*, and 3) *compound construction*. We examined responses for prerequisite linguistic errors first since we assumed children would need this implicit knowledge to help them get started in producing a response. In the absence of knowledge of prerequisite linguistic conventions, it was assumed that any relevant vocabulary and/or compound construction knowledge would not likely improve response quality.

Prerequisite Linguistic Errors. These errors were coded when children displayed any inadequacy in understanding the prerequisite linguistic conventions for the numeral to number word translation process. These errors were coded when children:

1) read the digits according to their digit value, without considering positional value information consistent with place value conventions (e.g., "*four, five, six*" for 456);
2) added/omitted number words in relation to the number of digits in the numeral (e.g., "*one one hundred*" for 100);
3) interpreted zeros incorrectly by saying them aloud (e.g., "*four hundred zero six*" for 406) or omitting the number series represented by the zero (e.g., "*forty-six*" for 406);
4) did not read the digits in order of magnitude starting with the leftmost digit and moving to the rightmost digit (e.g., "*fourteen*" for 41); and/or

5) did not organize the digits conventionally in groups of three, using the comma or space as an indication of the appropriate grouping (e.g., "*fourteen nine, twenty-three two*" for 149,232).

Vocabulary Errors. If children did not commit prerequisite linguistic errors, their response was then examined for the correct vocabulary words. A vocabulary error was coded when children did not express their knowledge of all appropriate vocabulary words needed. A substitution error was coded when children used one incorrect basic number word in place of another basic number word (e.g., "*one hundred million*" for 100,000). We consider this error a vocabulary problem, because, by our operational definition of number vocabulary knowledge, the child has not communicated knowledge of "million" as having seven to nine digits. A vocabulary omission code was recorded when children did not provide a required vocabulary word in their response, and they did not add any other words in its place (e.g., *thousand* is missing in the thousands location in the response "*two, one hundred one*" for 2,101). This example suggests that children have the basic knowledge to construct number words, including knowledge that digits are grouped into three and that each digit is associated with a name. However, inability to produce the vocabulary word "*thousand*" compromised a correct response.

Compound Construction Errors. A compound construction error was coded when all correct vocabulary words were cited somewhere within a response, but the words were not organized correctly according to McCloskey et al.'s (1986) model. The words *one* through *nine* are referred to as units. There is no consistent syntactic structure for the teen number words; these numbers are assumed to be memorized. For the remaining decade number words (i.e., 20-99), the syntactic structure is *tens-units* where *tens* is a number word from the list twenty, thirty…ninety. The use of *tens-units* below includes the appropriate adjustment for the numbers one through nineteen. For hundreds number words (i.e., 100-999), the syntactic structure is *units-hundred tens-units*. The syntactic structure for larger numbers uses multiplier words (e.g., thousand, million, billion) to modify each group of three digits (starting from the right) to reflect the magnitude of the number in question. For example, the syntactic structure for the numbers between 10,000,000 and 99,999,999 is *tens-units million, units-hundred tens-units thousand, units-hundred tens-units*.

Compound construction errors were coded when children produced the number words in the wrong conventional sequence (e.g., "*one thousand hundred*" for 100,000). Compound construction errors differ from prerequisite linguistic errors and vocabulary errors in the sense that all digits seem to be appropriately accounted for and all required vocabulary words are present, respectively, somewhere in the number word compound. However, there is some (usually minor) incorrect adjustment in the ordering of the vocabulary words. Compound construction errors can take the form of an omission or substitution. For example, the response of "*three <u>hundred thousand</u> twenty six ____, one hundred ninety-two*" for 326,192 contains both substitution (i.e., hundred thousand instead of hundred) and omission (i.e., missing thousand) compound construction errors.

Error Analysis

Before conducting analyses on error location, the data were converted into mean proportion scores by dividing the number of correctly identified location sections by the total number of possible locations. This proportioning controlled for the varying numbers of errors made by individual children. Analyses were conducted at each number series (e.g., decades, hundreds), because number words at each increasing number series involve the coordination of many additional words. Based on previous research (Baroody, et. al, 1984) we assumed that children would learn the cardinal number sequence in order, starting with smaller numbers and adding onto them with development. Once children produced eight consecutive errors within a particular number series, we assumed that they would not be able to produce any additional number words. Admittedly, it is possible that children may have known some larger numbers before smaller ones, and this methodological choice may have underestimated children's knowledge in places. However, testing time constraints and participant fatigue required us to make some decisions about the number of trials in the study. Young children took a long time to produce the large numbers, and they became easily overwhelmed when they had to continue to produce responses for the increasingly large numbers. Thus, we did not test all of the children on all of the number words. For the location analyses only, all children were included in the analyses even if they were not exposed to all numbers. If a child was not exposed to a particular number, an error was recorded for all sections of the number word.

Before conducting analyses on error type, mean proportion scores were obtained by dividing the number of errors made within a specific error category by the total number of error types coded. Proportion scores were used to balance the unequal numbers of errors and numbers to which children were exposed at each number level. Only children who had been exposed to at least half of the numbers for a number series were included in an analysis because it could not be assumed that children would make the same kind of error on subsequent trials. The types of errors children made while naming numbers were analyzed within each number series. A separate error type category was occasionally used when both a vocabulary and compound construction error occurred simultaneously within the same location of a number word, and the exact source of the error could not be determined.

The reliability of the codes was found to be within acceptable limits of standard reliability conventions. Two raters coded 25% of the transcripts. Percent agreement was used to determine inter-rater reliability. An 86% agreement was found for error location and an 80% agreement was found for error type. A Kappa value of 0.82 was computed for the specific error type categories. Disagreements were resolved through discussion.

RESULTS

The location of children's errors within a number word compound and the types of errors children produced were analysed as a function of grade using ANOVAs and followed up with Tukey post hoc tests. Since number word compounds become increasingly complex with each number series, analyses were conducted within each number series. Because preliminary analyses did not reveal any effects for gender, analyses were conducted across gender.

Children's mean proportion scores representing the number of locations correctly identified for a particular number series as a function of grade are presented in Table 1. A breakdown of children's errors according to the four error type categories as a function of grade and number series are presented in Table 2. A grade by location ANOVA at the decades revealed a grade effect, $F(1, 92) = 7.41$, $p < .008$, and post hoc tests indicated that grade one children produced more errors than their older counterparts. There was no significant location effect, $F(1, 92) = .16$, $p < .688$, suggesting that when errors were made (particularly by the grade one students), they tended to be distributed across all portions of the number word. A grade by error type ANOVA at the decades revealed a grade effect, $F(1, 91) = 22.08$, $p < .001$, a type effect, $F(3, 273) = 2.89$, $p < .036$ and a grade by type effect, $F(3, 273) = 8.20$, $p < .001$. Post hoc tests revealed that a significant amount of the errors made by the grade one children were attributed to both their lack of knowledge of the prerequisite linguistic rules and appropriate vocabulary terms compared to the other error categories.

A grade by location ANOVA for the hundreds revealed a grade effect, $F(1, 92) = 9.34$, $p < .003$, a location effect, $F(2, 184) = 9.57$, $p < .001$ and a grade by location effect, $F(2, 184) = 5.91$, $p < .003$. Post hoc tests revealed that grade one students produced more errors than their older counterparts, and that more errors were produced at the hundreds section than the decades and the units sections in grade one. A grade by error type ANOVA revealed a grade effect $F(3, 89) = 24.88$, $p < .001$, a type effect $F(3, 267) = 18.43$, $p < .001$, and a grade by type effect $F(9, 267) = 16.87$, $p < .001$. Post hoc analyses revealed that grade one children produced more prerequisite linguistic errors and vocabulary errors compared to all other error types combined.

At the thousands level, a grade by location ANOVA revealed a grade effect, $F(1, 92) = 19.32$, $p < .001$, a location effect, $F(3, 276) = 25.92$, $p < .001$, and a grade by location effect, $F(3, 276) = 3.12$, $p < .026$. Post hoc analyses (excluding grade one students due to low sample sizes) indicated that grades three and five children produced greater proportions of errors in the thousands portion of the number word, than they did in the units, decades, and hundreds sections. With respect to the error category analyses at the thousands level, a grade by error type ANOVA revealed a grade effect, $F(2, 68) = 13.93$, $p < .001$, a type effect, $F(3, 204) = 6.69$, $p < .001$, and a grade by type interaction, $F(6, 204) = 2.81$, $p < .012$. The interaction was likely due to error type category differences in grade three, but not at the other

Table 1: Children's Mean Proportion Scores (and Standard Deviations) Representing the Number of Locations Correctly Identified Within Each Number Series for Each Grade.

	Grade 1	Grade 3	Grade 5	Grade 7	Total
DECADE SERIES					
Units	.863 (.250)	1.0 (.000)	1.0 (.000)	1.0 (.000)	.970 (.140)
Decades	.868 (.250)	1.0 (.000)	1.0 (.000)	1.0 (.000)	.998 (.140)
HUNDREDS SERIES					
Units	.434 (.389)	.979 (.068)	1.0 (.000)	1.0 (.000)	.858 (.312)
Decades	.286 (.420)	.972 (.078)	1.0 (.000)	1.0 (.000)	.815 (.373)
Hundreds	.269 (.436)	.951 (.203)	1.0 (.000)	1.0 (.000)	.805 (.396)
THOUSANDS SERIES					
Units	.127 (.303)	.649 (.312)	.899 (.185)	.974 (.110)	.662 (.410)
Decades	.106 (.264)	.646 (.312)	.929 (.135)	.976 (.110)	.664 (.410)
Hundreds	.078 (.224)	.589 (.345)	.913 (.141)	.977 (.111)	.639 (.420)
Thousands	.082 (.205)	.418 (.327)	.792 (.287)	.937 (.173)	.557 (.417)
MILLIONS SERIES					
Units	.024 (.119)	.229 (.317)	.594 (.395)	.865 (.286)	.428 (.438)
Decades	.024 (.119)	.229 (.317)	.637 (.364)	.920 (.212)	.453 (.440)
Hundreds	.024 (.119)	.205 (.310)	.632 (.366)	.922 (.212)	.446 (.442)
Thousands	.007 (.034)	.089 (.281)	.519 (.529)	.813 (.472)	.357 (.489)
Millions	.005 (.026)	.137 (.250)	.509 (.452)	.847 (.331)	.375 (.446)
BILLIONS SERIES					
Units	.000 (.000)	.042 (.163)	.370 (.448)	.698 (.417)	.277 (.421)
Decades	.000 (.000)	.042 (.163)	.370 (.448)	.719 (.411)	.283 (.425)
Hundreds	.000 (.000)	.042 (.163)	.370 (.448)	.717 (.410)	.282 (.425)
Thousands	.000 (.000)	.040 (.164)	.347 (.475)	.701 (.462)	.272 (.435)
Millions	.000 (.000)	.038 (.163)	.349 (.481)	.701 (.462)	.272 (.437)
Billions	.000 (.000)	.024 (.102)	.306 (.461)	.660 (.468)	.247 (.422)

grades. In grade three, prerequisite linguistic, vocabulary, and compound construction errors were coded more often than the vocabulary/compound construction combined category.

A grade by location ANOVA for the millions series revealed a grade effect, $F(1, 92) = 146.12$, $p < .001$, a location effect, $F(4, 368) = 14.58$, $p < .001$, and a grade by location effect, $F(4, 368) = 2.39$, $p < .05$. Post hoc tests (excluding grade one students due to low sample sizes) indicated that grade three and grade five children produced more errors in the millions section, followed by the thousands section, compared to the proportion of errors made on the units, decades, and hundreds sections combined. For the error type analyses, a grade by type ANOVA revealed a grade effect, $F(2, 49) = 4.83$, $p < .012$, but no type effect or interaction. Error types were equally distributed across error categories for the eldest three grades.

A grade by location ANOVA for the billions series revealed a grade effect, $F(1, 92) = 5.74$, $p < .019$, a location effect, $F(5, 460) = 2.32$, $p < .042$, and a grade by location interaction, $F(5, 460) = 4.79$, $p < .001$. Post hoc analyses (excluding grade one and three students due to low sample sizes) revealed significant location differences for grade five students only. Greater proportions of the errors made by grade five children were localized in the billions section of the numeral compared to any other section of the numeral. A grade by error type ANOVA did not reveal any significant main effects or interactions.

Table 2: Children's Mean Proportion Scores (and Standard Deviations) Within Each Number Series as a Function of Error Type and Grade.

	Grade 1	Grade 3	Grade 5	Grade 7	Total
DECADE SERIES	**N= 24**	**N=24**	**N=24**	**N=24**	**N=96**
Prerequisite Linguist	.950 (.084)	1.0 (.000)	1.0 (.000)	1.0 (.000)	.988 (.046)
Vocabulary	.946 (.121)	1.0 (.000)	1.0 (.000)	.999 (.004)	.987 (.063)
Compound Construct	.984 (.048)	1.0 (.000)	1.0 (.000)	1.0 (.000)	.996 (.024)
Comb Vocab + CC	.984 (.048)	1.0 (.000)	1.0 (.000)	1.0 (.000)	.996 (.024)
HUNDREDS SERIES	**N=24**	**N=24**	**N=24**	**N=24**	**N=96**
Prerequisite Linguist	.437 (.390)	.994 (.013)	1.0 (.000)	1.0 (.000)	.871 (.299)
Vocabulary	.672 (.388)	.983 (.069)	1.0 (.000)	1.0 (.000)	.922 (.230)
Compound Construct	.694 (.390)	1.0 (.000)	1.0 (.000)	1.0 (.000)	.931 (.223)
Comb Voab + CC	.694 (.390)	1.0 (.000)	1.0 (.000)	1.0 (.000)	.931 (.223)
THOUSANDS SERIES	**N=0**	**N=23**	**N=24**	**N=24**	**N=71**
Prerequisite Linguist	N/A	.766 (.239)	.933 (.107)	.982 (.080)	.895 (.108)
Vocabulary	N/A	.747 (.231)	.942 (.103)	.993 (.016)	.896 (.179)
Compound Construct	N/A	.803 (.222)	.937 (.104)	.993 (.028)	.912 (.162)
Comb Vocab + CC	N/A	.819 (.235)	.957 (.097)	.998 (.009)	.926 (.162)
MILLIONS SERIES	**N=0**	**N=7**	**N=21**	**N=24**	**N=52**
Prerequisite Linguist	N/A	.755 (.202)	.805 (.311)	.945 (.195)	.863 (.257)
Vocabulary	N/A	.729 (.226)	.794 (.312)	.988 (.029)	.875 (.238)
Compound Construct	N/A	.746 (.268)	.783 (.331)	.965 (.136)	.862 (.265)
Comb Vocab + CC	N/A	.804 (.210)	.824 (.318)	.999 (.000)	.902 (.230)
BILLIONS SERIES	**N=0**	**N=0**	**N=11**	**N=19**	**N=30**
Prerequisite Linguist	N/A	N/A	.955 (.151)	.941 (.229)	.946 (.201)
Vocabulary	N/A	N/A	.917 (.179)	.928 (.231)	.924 (.210)
Compound Construct	N/A	N/A	.919 (.259)	.947 (.229)	.937 (.238)
Comb Vocab + CC	N/A	N/A	.955 (.151)	.947 (.229)	.950 (.201)

When children produced errors for billions series numerals, they were equally distributed across all error categories.

DISCUSSION

Our findings extend previous developmental understandings and cognitive models of number knowledge, suggesting an emerging developmental model for number production. Researchers agree that knowledge of both number word vocabularies and syntactic structures (e.g., McCloskey et al., 1986; Seron & Fayol, 1994; Power & Dal Martello, 1990) are important in the developmental process of number production. But evidence from our study indicates that the prerequisite linguistic conventions are also fundamental to number naming. From a developmental perspective, grade one children produced the greatest proportion of prerequisite linguistic errors, and the proportion of these errors substantially decreased with grade. Furthermore, these errors typically affected the entire number word response. Preliminary linguistic errors were coded most frequently when grade one children used unconventional digit groupings (i.e., representing 49% of all prerequisite linguistic codes such as in the example "*ninety-nine-one*" for 991) and right-to-left number reading reversals (i.e., accounting for an additional 26% of the prerequisite linguistic codes such as in the example "*fourteen*" for 41). When these errors are produced, children do not have knowledge of the fundamentals of number production, irrespective of their vocabulary and compounding construction knowledge.

Although prerequisite linguistic errors were produced most frequently by our grade one cohort, older children occasionally produced them. In one example, a grade seven student confused the commas in the numerals with decimals, translating each comma as "point" throughout. For example, the number 405,188,139 was read as "*four hundred and five point, one eight eight point, one three nine*." The confusion between the commas and the decimals only became apparent with numbers in the millions section and beyond. In another example, a grade three student labeled each digit with a basic number word, resembling place value conventions. The numbers 6,893 and 964,509 were named as "*six million, eighty nine thousand and three*" and "*ninety six trillion, four hundred million, fifty thousand and nine*," respectively. This same student made no errors for numerals less than 1000, correctly named 1000 and 1,000,000, and reversed the names for 1,000,000,000 and 1,000,000,000,000. The fact that these errors continue to occur in the older grades suggests that children may not be acquiring this implicit knowledge for number construction and, as a result, cannot even get started in producing a plausible response.

Once children seemed to demonstrate an understanding of the prerequisite linguistic principles, other factors related to vocabulary and compound construction knowledge might impede number naming. Vocabulary errors occurred at all grades, they were equally distributed across the omission and substitution error subtypes, but once again, they were highly consistent with the problem solving endeavors of each child. For example, one grade three male produced all of the numbers up to and including the ten thousands series correctly. But his incorrect number naming attempts in the hundred thousands series simply omitted the word *thousand* (e.g., "*six hundred fifty-two, thirteen*" for 652,013). Another grade three student correctly named numbers up to the low thousands, but used the incorrect word *million* for subsequent attempts (e.g., "*eight million, three hundred and thirty-one*" for 8331). Since he did not show knowledge that *thousand* is represented by at least four digits, his response was coded as a vocabulary substitution. These older children's responses resembled close approximations to the desired response, albeit one or two missing vocabulary words compromised their response.

Compound construction errors occurred most frequently in grade three and beyond. Incorrect attempts varied as a function of each child's individual knowledge of the compound structure. When these errors occurred, they most commonly affected the hundred thousands sections of numbers. The compound construction error may have been common for children in the thousands as it is the first time in the number sequence where an increase in the number of digits is not associated with a new number name. Children are required to learn a new compound convention to deal with the discrepancies at the ten and hundred thousand levels.

Furthermore, once children have a basic understanding of how to get a numeral to number word translation process started, there seems to be a strong interplay between knowing vocabulary words and the acceptable syntactic structure. Some children knew appropriate vocabulary words but they could not coordinate this knowledge correctly in the compounding sequence. Other children had knowledge of the syntactic structure but lacked the requisite vocabulary. Although both vocabulary development and compound construction knowledge are important for number production, there are no developmental differences in the acquisition of these factors.

Our findings are consistent with research conducted by McCloskey et al. (1986) and Baroody et al. (1984), and provide confirmation of the factors cited in models on number production. The location data provide evidence that there are developmental differences in children's knowledge of the syntactic frame. Errors produced often affected only the largest portion of the number word, suggesting that children likely learn to add onto their already existing syntactic frame knowledge, piece by piece when producing larger and larger numbers.

Results from this study outline two areas in the theoretical number production model that require further attention. First, it seems apparent that translating numerals into number words and vice versa does not involve the same process, and the relation between these procedures should be examined further. Seron and Fayol (1994) and Power and Dal Martello (1990) report the occurrence of transcription and lexical errors in their studies, which were not observed in the current work. One obvious difference in the number production process across these studies is that vocabulary words are provided in the number word to numeral translation, but they are not provided in the reverse process.

Second, several numbers involving many zeros were included in this study (e.g., 100; 10,000,000; 11,000,204), but strong individual differences within grade suggest that the rules for interpreting zeros is not clearly understood. Some children showed difficulty with numbers containing many zeros. These children often skipped the numeral or provided incorrect verbalizations. Some children found the numbers containing zeros relatively easy, and they were the only numbers that some children received credit for within a particular number series. These individual differences within grade eliminated any significant overall effects.

This study has implications for future research and recommendations to the field of education. Researchers need to explore the manner in which children are exposed to the cardinal number sequence in school, particularly with respect to large number names. Currently, curricular documents and textbooks suggest exposing children to number words from larger and larger number series with each increasing grade (e.g., low thousands in grade two, ten thousands in grade three), with little reflection on how to name and count the numbers. Although we know there is a syntactic structure to organizing number words, children typically receive little guidance on the process, and they are encouraged to use a recursive approach. Future research is needed to explore the effects of explicit instruction of relevant vocabulary terms, compound construction rules, and linguistic conventions in relation to number word production processes. There may be educational benefits to synchronizing the parsing of number naming curricular outcomes with period and place value knowledge (see Cathcart et al., 2000 for clarification), the syntactic frame of the thousands, and relevant prerequisite linguistic conventions. Finally, research should explore the relation between the kinds of errors made and children's overall mathematical and linguistic abilities. Studies involving older children with learning disabilities or mathematical talents may provide further insight into the number production process, and improve the teaching of the cardinal number sequence.

Note: Portions of this research were funded by a Natural Sciences and Engineering Research Council graduate scholarship awarded to the first author. The authors gratefully acknowledge the children, parents and teachers of the schools where the research was conducted. We also thank Jim Clark for his advice on data analysis, Jeremy Anglin for his input when the project was first conceptualized, and Jo-Anne LeFevre, John Kirby and several anonymous reviewers for their comments on earlier versions of this manuscript.

Author Biographical notes:

Dr. Sheri-Lynn Skwarchuk is an assistant professor in the Faculty of Education and Extended Learning at the University of Winnipeg. She is interested in children's mathematical development and predictors of success in school math for pre-school and early school aged children.

Dr. Paul Betts is an assistant professor in the Faculty of Education and Extended Learning at the University of Winnipeg. He recently completed a Ph.D. in mathematics curriculum at the University of Regina. His research interests include mathematical cognition and mathematics teacher development.

REFERENCES

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development, 58* (10, Serial No. 238, 1- 166).

Baroody, A. J. (1987). *Children's mathematical thinking: A developmental framework for preschool, primary and special education teachers*. New York, NY: Teacher's College Press.

Baroody, A. J., Gannon, K. E., Berent, R., & Ginsburg, H. P. (1984). The development of basic formal mathematical abilities. *Acta Paedolgica, 1*, 133-150.

Baroody, A. J. & Price, J. (1983). The development of the number-word sequence in the counting of three-year-olds. *Journal for Research in Mathematics Education*, *14*, 361-368.

Bell, M. S. & Burns, J. F. (1981). Counting and numeration capabilities of primary school children: A preliminary report. *Proceedings of the third annual meeting of the North American chapter of the international group for the psychology of mathematics education*. Minneapolis, MN.

Campbell, J. I. D. & Clark, J. M. (1992). Cognitive number processing: An encoding-complex perspective. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 457-491). Amsterdam, North-Holland: Elsevier Science.

Cathcart, W. G., Pothier, Y. M., & Vance, J. H. (2000). *Learning mathematics in elementary and middle schools*. (3rd Ed., pp. 85-107). Scarborough, ON: Prentice-Hall.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1-42.

Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.

Fuson, K. C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction, 7*, 343-403.

Fuson, K. C., Richards, J., & Briars, D. J. (1982). The acquisition and elaboration of the number word sequence. In C. J. Brainerd (Ed.), *Children's logical and mathematical cognition* (pp.33-92). New York, NY: Springer-Verlag.

Ginsburg, H. (1977). *Children's arithmetic: The learning process*. New York: Litton Educational Publishing Inc.

Hurford, J. R. (1975). *The linguistic theory of numerals*. Cambridge, M.A.: Cambridge University Press.

Kamii, M. (1981). Children's ideas about written number. *Topics in Learning and Learning Disabilities, October*, 47-59.

Kibel, M. (1992). Linking language to action. In T. B. Miles & E. Miles (Eds.) *Dyslexia and mathematics*. London: Routledge.

McCloskey, M. & Caramazza, A. (1987). Cognitive mechanisms in normal and impaired processing. In G. Deloche & X. Seron (Ed.) *Mathematical disabilities: A cognitive neuropsychological perspective* (pp. 201-219). Hillsdale, NJ: Lawrence Erlbaum Associates.

McCloskey, M., Sokol, S. M., Goodman-Schulman, R. A., & Caramazza, A. (1990). Cognitive representations and processes in number production: Evidence from cases of acquired dyscalculia. In A. Caramazza (Ed.) *Cognitive neuropsychology and neurolinguistics: Advances in models of cognitive function and impairment* (pp. 1-32). Hillsdale, NJ: Lawrence Erlbaum Associates.

McCloskey, M., Sokol, S. M., & Goodman, R. A. (1986). Cognitive processes in verbal-number production: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: General, 115*, 307-330.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: The National Council of Teachers of Mathematics, Inc.

National Council of Teachers of Mathematics. (2004). *A family's Guide: Fostering your child's success in school mathematics*. Reston, VA: The National Council of Teachers of Mathematics, Inc.

Power, R. J. D. & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes, 5,* 237-254.

Power, R. J. D. & Longuet-Higgins, H. C. (1987). Learning to count: A computational model of language acquisition. In H. C. Longuet-Higgins (Ed.), *Mental processes: Studies in cognitive science* (pp. 249-280). Cambridge: MIT Press.

Seron, X. & Deloche, G. (1987). The production of counting sequences by aphasics and children: A matter of lexical processing? In G. Deloche & X. Seron (Eds.), *Mathematical disabilities: A cognitive neuropsychological perspective* (pp. 171-200). Hillsdale, NJ: Lawrence Erlbaum Associates.

Seron, X. & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology*, *12*, 281-300.

Skwarchuk, S. & Anglin, J. M. (2002). Children's acquisition of the English cardinal number words: A special case of vocabulary development. *Journal of Educational Psychology, 94*, 107-125.