Determining a Zulu Core Vocabulary for Children who use Augmentative and Alternative Communication

Jocelyn Mngomezulu, Kerstin M. Tönsing, Shakila Dada, and Nomadlozi B. Bokaba

University of Pretoria

Author Note

Jocelyn Mngomezulu, Kerstin M. Tönsing, and Shakila Dada, Centre for Augmentative and Alternative Communication, University of Pretoria; Nomadlozi B. Bokaba, Department of African Languages, University of Pretoria.

This study was based on a Master's thesis completed by the first author under supervision of the second and third authors. The financial assistance of the Oppenheimer Memorial Trust (OMT) and the National Research Foundation (NRF) of South Africa (grant no. TTK150617119597) towards this research is herewith acknowledged. Opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to the OMT or the NRF. The authors would like to thank the children who participated in the study, the school staff who provided assistance with logistical arrangements, and those who acted as research assistants.

Correspondence regarding this article should be addressed to Jocelyn R Mngomezulu; jocelyn.mngomezulu@rocketmail.com

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Abstract

Vocabulary selection is an important aspect to consider when designing augmentative and alternative communication (AAC) systems for children who have not yet developed conventional literacy skills. AAC team members have used core vocabulary lists (representing words most commonly and frequently used by speakers of a natural language) as a resource to assist in this process. To date, there are no core vocabulary lists for Zulu. Therefore, the aim of this study was to identify the vocabulary most frequently and commonly used by Zulu-speaking preschool children, in order to inform vocabulary selection for peers who use AAC. Communication samples from 6 Zulu-speaking participants without disabilities were collected during regular preschool activities. Analyses were conducted both by orthographic words and by morphological analysis of formatives. Due to the linguistic and orthographic structure of Zulu, an analysis by formatives was found to be more useful to determine a core vocabulary. The number of different formatives used, frequency of use, and commonality of use among the participants were identified. A total of 213 core formatives were identified; core formatives related to language structure were used more frequently than those that related to lexical content. The characteristics of this Zulu core vocabulary were consistent with those of core vocabularies established in other languages. Implications for the design of Zulu AAC systems are discussed.

Keywords: Augmentative and alternative communication; Core vocabulary; Preschool; Vocabulary selection; Zulu

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

When children in need of aided augmentative and alternative communication (AAC) have not yet developed conventional literacy skills, providing them with a method of expressing a variety of novel self-generated messages is often challenging (Light & McNaughton, 2012; Smith, 2015). Aided communication systems used by these children typically contain preselected words represented by graphic symbols that the child can access in order to communicate. Vocabulary selection is therefore one of the first steps in designing aided AAC systems for this population. Vocabulary should be meaningful, functional across all communication contexts, relevant to the child, motivating, age-appropriate, and facilitate the language development of the child (Boenisch & Soto, 2015; Trembath, Balandin, & Togher, 2007). At the same time, the vocabulary should also be manageable by the child: Accessing a large number of words stored within an AAC system may pose significant demands on memory and metalinguistic skills, while also requiring physical effort (Thistle & Wilkinson, 2013; Wilkinson & Hennig, 2007). A process of choosing vocabulary from a pool of all possibilities (Yorkston, Dowden, Honsinger, Marriner, & Smith, 1988) must ensue to narrow down the array to a manageable but useful size.

Various methods and sources have been suggested in the literature to facilitate vocabulary selection. These include environmental inventories, informant lists, and the use of existing vocabulary lists to guide selection (Fried-Oken & More, 1992; Yorkston et al., 1988). Regarding the latter, various core vocabulary lists have been generated based on the words most frequently and commonly used by specific groups of individuals, including school- and preschool-aged children (e.g., Beukelman, Jones, & Rowan, 1989; Boenisch & Soto, 2015; Fallon, Light, & Kramer Paige, 2001; Marvin, Beukelman, & Bilyeu, 1994; Robillard, Mayer-Crittenden, Minor-Corriveau, & Bélanger, 2014; Trembath et al., 2007). These core vocabulary lists have been advocated as useful vocabulary sets to include in AAC systems because they

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

consist of a small number of words that are frequently used across various communication contexts and partners and are important for building novel sentences (Baker & Chang, 2006; Deckers, Van Zaalen, Van Balkom, & Verhoeven, 2017; Snodgrass, Stoner, & Angell, 2013; Witkowski & Baker, 2012).

To date, there is little empirical evidence regarding the effect on a child's communication and language skills of including core vocabulary in his or her AAC system. One reason may be that proposed benefits of the inclusion of core words are increased grammaticality (correct use of grammar) and greater generativity (the ability to generate novel messages) in message production, but such skills are at a higher level of linguistic proficiency and may take a significant amount of time and training to develop. In addition, populations in need of AAC are small in number and heterogeneous. Thus, the level of skill targeted, the interaction between the vocabulary selected and other characteristics of the system, and the heterogeneity of the population in need of AAC, are all factors that complicate experimentally controlled studies on the effect of the inclusion of core vocabulary on long-term language and communication outcomes (Lund & Light, 2007). However, Soto and Clarke (2017; 2018) have recently conducted conversation-based interventions with children and adolescents who used AAC systems that included core and fringe vocabulary. In the first study, participants with receptive language skills at a 6-year-old level were able to learn to produce grammatically correct novel utterances. Skills were maintained and generalized to conversations with familiar partners. In the second study, four adolescents with motor speech disorders increased their use of various linguistic targets following intervention.

Despite the dearth of evidence on the impact of including core vocabulary in a child's AAC system, core vocabulary lists have been used to design AAC systems. Various

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

commercially available vocabulary sets and language systems for speech generating devices (SGDs), for example, are based on a core-fringe vocabulary arrangement in order to provide easy access to core words, while also containing words that are more specific to certain situations and contexts (fringe vocabulary). AAC team members reportedly include core vocabulary in the AAC systems they design or customize (Dada, Murphy, & Tönsing, 2017; Lund, Quach, Weissling, McKelvey, & Dietz, 2016; Thistle & Wilkinson, 2015), although they also use various other sources to inform vocabulary selection. One may therefore argue that theory and practice support the use of core vocabulary lists to inform AAC system design.

Core vocabulary is not the same across languages (Liu & Sloane, 2006; Shin & Hill, 2016), because it typically reflects structural aspects of the language. These aspects can differ significantly among languages of different linguistic typologies (Lee, Kim, & Park, 2005; Robillard et al., 2014). For example, of the top 10 most frequently used Korean words determined by Shin and Hill (2016), eight cannot be translated into English because they serve a grammatical function that is expressed differently in English. Language-specific studies are therefore needed, and vocabulary frequency lists based on the analysis of conversational language samples have been proposed as the most valid source for core vocabulary (Liu & Sloane, 2006; Madrid & Antonio, 2016; Shin & Hill, 2016). In addition to obtaining primary data, the linguistic unit of analysis (items from the transcripts that are counted for frequency) of core vocabulary studies in a particular language may also need to be adjusted when compared to previous studies in other languages. In many English studies, for example, orthographic words have been counted, strictly defined by the orthographic space. For example, a transcribed sentence like “You can use the one I used” from a speech sample consists of seven written (orthographic) words, each separated by a space (called the orthographic space). In such an

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

analysis, only words that are spelled in an identical fashion are counted together. This sentence would therefore consist of seven unique words because words like *use* and *used* are counted as separate words. In some English studies (e.g., Boenisch & Soto, 2015), however, lexemes were counted. While a lexeme count in English still relies heavily on the orthographic space, certain inflections of words are counted together under what is termed the *lexeme*, or base form, of the word. Past tense forms of verbs, for example, are counted under the imperfect tense form. Therefore the word *used* would be counted together with the word *use.* The sentence above would therefore consist of seven words but contain only six unique lexemes. Because English is an analytic language, meaning that it has relatively few bound morphemes and many words consisting only of one morpheme, both ways of counting core vocabulary seem to have merit. However, languages with different linguistic structures and different orthographic conventions may warrant a different unit of analysis. Morpheme-rich (also termed synthetic) languages with a high morpheme-to-word ratio may require a morphological level analysis, since meaning in these languages is primarily constructed by building up words from different morphemes (Kosch, 2006). In the preceding example, a morphological level analysis would yield a total number of eight morphemes -- because the word *used* consists of the morphemes *use* and *–(e)d*) -- but only seven different morphemes (because the morpheme *use* appears twice).

Zulu, one of the 11 official languages of South Africa, is the language most frequently spoken as a home language in the country (10 million home language speakers, or 27% of the population; Statistics South Africa, 2012). A further 15 million people report speaking Zulu as an additional language. A total of 95% of all Zulu home language speakers live in South Africa (Lewis, Simons, & Fennig, 2016), with the greatest concentration in KwaZulu-Natal, where 77.8% of the population are Zulu home language speakers (Statistics South Africa, 2012). No

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

prevalence figures are available regarding persons in need of AAC in South Africa. Prevalence studies from Canada, the United Kingdom, and Australia suggest prevalence rates between 1.2 and 1.5% (Beukelman & Mirenda, 2013). By extrapolation, this would suggest that about 120,000 to 150,000 persons from a Zulu home language background are in need of AAC.

Zulu is an Nguni language, falling in the same linguistic grouping as Xhosa, Swati and the Ndebele languages. Like other African languages, Zulu is mainly synthetic agglutinative, relying not on word order and auxiliaries but rather on a series of morphemes (also called formatives) including stems, roots, and affixes to create meaning (Kosch, 2006). Synthetic languages have a higher morpheme-to-word ratio than analytic languages (English being an example of a more analytic language). Zulu is also primarily *agglutinative* (a term derived from the Latin word for *glue*), meaning that morphemes remain relatively constant in their phonological form and are merely "glued" together to form a word. This also means that morphemes remain identifiable within words (Kosch, 2006). An example of agglutination in English is the plural marker '-s', which can be added to various nouns without the noun changing form (e.g., *cat–cats, house–houses*). However, unlike English, Zulu has very few free morphemes or morphemes that can stand on their own. Although identifiable in words, the morphemes are only fully meaningful when together. The agglutinative morphology of Zulu is written conjunctively (Kosch, 2006; Prinsloo & De Schryver, 2001), that is, there is often no orthographic space between morphemes. In rough comparison with English, therefore, many orthographic words in Zulu translate to two or more English words. For example, *Ngiyababona* translates to "*I see them*."

Many core vocabulary studies in English have relied on the orthographic space to define the linguistic units (words) that are counted to determine the most frequent vocabulary items.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Due to its conjunctive orthography and agglutinating nature, this may not be the most useful way of establishing a core vocabulary in Zulu. An alternative is to count morphemes, or what is termed in Zulu as *izakhi zamabizo* (Nyembezi, 1982, p. 43), translated as "formative." Like morphemes, formatives describe the smallest units of the language that perform a grammatical function (structure formatives) or carry lexical meaning (content formatives) (Kosch, 2006). The orthographic word *Ngiyababona,* for example, consists of five formatives, namely *ngi, ya, ba, bon,* and *a.* The structure formative *ba* (an object concord) translates to *them*, and the content formative *bon* (a verb root) translates to *see*. The structure formative *ya* is a present tense morpheme that is required in Zulu but has no direct translation into English. Most content formatives and some structure formatives can be given an approximate English translation. However, neither content nor structure formatives can be used alone (Paulos & Msimang, 1998).

A handful of studies employing word frequency counts in Zulu and other languages in this language family have been conducted within the fields of corpus linguistics (e.g., Allwood et al., 2010; Allwood & Hendrikse, 2003; Prinsloo & De Schryver, 2001) and natural language processing (e.g., Pretorius & Bosch, 2003; Prinsloo & De Schryver, 2001; Spiegler, van der Spuy, & Flach, 2010a). The characteristics of the Zulu language and Zulu orthography described above have presented a number of challenges in these studies. Allwood et al., for example, when reporting on word frequency counts based on orthographic space for Xhosa (a language closely related to Zulu), note that code switches such as the English words *and* and *so* appear within the top 12 most frequently appearing orthographic words.

The Ukwabelana Corpus (Spiegler et al., 2010a; Spiegler, Van der Spuy, & Flach, 2010b) developed a system of analysis that may be more useful for vocabulary frequency counts in Zulu. In their endeavour to create an automated formative tagging system (to be used, for example, in

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

automated spell checkers), they developed 207 tags to identify different formatives (Spiegler et al., 2010b). The first author, with input from the fourth author of this paper revised this tagging system for use in the current study. The result was a tagging system that allowed formatives to be identified and counted within orthographic words.

In light of the absence of any Zulu core vocabulary lists that can be consulted as a resource for AAC vocabulary selection, the aim of this study was to determine such a list, and in particular a list that would be appropriate for preschoolers in need of AAC who have not yet developed conventional literacy skills. In light of the linguistic structure of Zulu, the researchers specifically set out to determine (a) the parameter of a core vocabulary based on orthographic words; (b) the parameters of a core vocabulary based on formatives. By comparing the results of the two analyses, the implications that different linguistic units would likely have on system design and system possibilities could also be explored.

Method

Participants

The participants were three females and three males aged between 5;1 and 5;9 (years;months). By age 5, children (regardless of language background) typically use a mature grammar with a variety of morphological and syntactic forms, and typically do not make grammatical errors (Owens, 2016). This has also been confirmed in a longitudinal study of language development of isiZulu children (Suzman, 1990). The participants in the current study met the following selection criteria: (a) the language used most frequently in the home and preschool setting was Zulu; (b) they attended the preschool for a minimum of 5 months prior to data collection and attend regularly for at least two days each week; and (c) their parents had no concerns about the development of their children.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Recruitment of participants commenced after the study had been approved by the Ethics Committee of the Faulty of Humanities at the University of Pretoria. The first author approached the principals of preschools where Zulu was the language of instruction and requested permission to recruit participant and three principals provided permission. All three preschools were situated in northern KwaZulu Natal in a rural area of South Africa. The schools were all attached to primary schools, and shared facilities with them. At one of the sites the children spent break times together with children from higher grades in a shared playground, whereas at the other two sites the children had different break times to other classes. All of the preschools used Zulu as the language of instruction, with English being used occasionally in rhymes and songs. Class sizes ranged from 14 to 39. In each instance, one adult (teacher) was responsible for the class. All of the preschools used the Grade R[1] (Reception Year) curriculum proposed in the National Curriculum and Assessment Policy Statement (CAPS; Department of Basic Education, 2018).

The first author then asked a teacher at each of the three preschools to identify two children in her class (one boy and one girl) who, in her opinion, had age-appropriate speech and language skills for possible inclusion in the study. Teachers were provided with information letters and consent forms, which they sent to the children's parents. Parents of all six children provided consent and also completed a background questionnaire to obtain background information about the child.

The first author then met with the potential participants at each preschool, explained the study procedures to potential participants in child-friendly language, and ensured that the children were aware of their rights, including the right to refuse to take part or withdraw at any time. The children were then given the opportunity to provide or deny assent to participate. Five

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

children assented to participate, while one did not. The teacher at this school therefore identified another child, and obtained parental consent. The researcher explained the study to the child as she had done for the other children. This child did provide assent to participate. On each day of data collection, assent was once again obtained from the participants before fitting the recording equipment. All six participants provided ongoing assent and completed the study.

Research Design

A quantitative descriptive observational design was used, allowing researchers to capture the spoken language that the participants used in their natural environment (preschool settings) via audio recordings.

Materials

Portable digital voice recorders with lapel microphones were used to record the speech of participating children. The recorders were contained in small zipped pouches that children wore strapped around their waist. Microphones were attached to the collars of the children's clothing.

Procedures

Data were collected using the procedure outlined in Trembath et al. (2007). This involved recording the interactions of the children throughout the preschool morning on consecutive mornings at a single site until 1,500 orthographic words had been collected for each participant. No attempt was made to control the activities that the children engaged in, but data were collected while the children engaged in a variety of both teacher-led and child-led activities, in accordance with the preschool's own typical schedule. The recorders were turned on in the morning and continued recording until the school day was over in the early/late afternoon (exact times varied between preschools). A research assistant was always present on site (but not within sight) to assist the children with adjusting the recording equipment throughout the day. On

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

average it took 4 hr 10 min (range: 2 hr 0 min - 6hr 30 min) over 2 days (range: 1–3 days) to collect the quota of 1,500 orthographic words from each participant. The variability in recording time to reach the desired number of words might be explained by a variation in talkativeness and/or also a variation in activities that the different participants were engaged in during the days when recording took place. Although teachers were required to nominate children who they perceived as having typical speech and language development, this did not mean that children were equally talkative. In addition, free play activities typically resulted in more speech being produced by participants, while teacher-led activities often required more listening with only limited and/or rote recitation responses from children. Similar variations in recording time were noted in other studies of preschoolers (e.g., Trembath et al., 2007).

Two trained research assistants transcribed the speech of the participants from the recordings. They were trained to distinguish the participant's voice from incidentally recorded speech by adults and peers by noting differences in pitch and volume. They transcribed the participants' speech verbatim using conventional Zulu orthography into the System for Analyzing Language Transcripts (SALT) program (Miller & Iglesias, 2012) loaded onto a Windows-based computer. In order to counteract reactivity by the participants, the first 20 min of each child's recording were not transcribed. In addition, references to the recording equipment or process were omitted from the analysis. A transcript file was created for each participant. The research assistants followed a set of transcription rules based on Trembath et al. (2007). These included (a) transcribing vocalizations that indicated agreement, disagreement, or interrogation in orthographically consistent forms and counting them as words; (b) transcribing words used in chants and rhymes, as well as those uttered in ritual games as a single word so as not to skew frequency counts of those repeated words; and (c) transcribing unintelligible words and portions

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

of orthographic words in a way that allowed omitting them from the analysis. Additionally, as in Trembath et al.'s study the codes CN (child name), TN (teacher name), and additionally PN (place name) were employed for confidentiality while preserving the use of these proper nouns by participants. Additional conventions prescribed by the SALT program were also followed.

A third independent transcriber transcribed a randomly selected 20% of the total recording (based on recording time) from each participant (Ayres & Ledford, 2014) for a second time. Percentage agreement was calculated by dividing the number of agreements by the sum of the number of agreements and disagreements (the latter referring to omissions, additions and differences in transcription), multiplied by 100. Percentage agreement was found to be 81% (range: 79% – 83%). Because the recordings were made during the school day while participants spent time in group settings in the classroom and the outside playground, the microphones also recorded ambient noise and some of the comments from peers and teachers. This reduced the intelligibility of the recorded speech of the target participant at times and resulted in interrater agreement that was relatively low when compared to some other studies (e.g., Boenisch & Soto, 2015, 94%; Trembath et al., 2007, 97.8%); however, Robillard et al. (2014) had lower inter-rater agreement (70%), also ascribed to the ambient noise in the classroom and playground.

Data tagging and analysis. From the participant-specific transcript files, a composite file was created. Orthographic word frequency counts as well as counts of total number of different orthographic words were performed on the transcriptions using the SALT program in order to compare this method of analysis to the formative tagging method. A frequency count for each orthographic word was calculated by dividing the total number of occurrence of the word by the total number of words multiplied by 100. Additionally, each word ranking above 0.05% in frequency was given a commonality score between 1 and 6. A commonality score of 6 meant

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

that all six participants had used the word. Core vocabulary was defined as all words occurring with a frequency of 0.05% or more that had a commonality score of 3 or more. These criteria, although somewhat arbitrary, have been used in other core vocabulary studies (Beukelman et al., 1989; Robillard et al., 2014; Trembath et al., 2007).

The first author then segmented and tagged the transcriptions using the formative tagging system adapted from Spiegler et al. (2010a; 2010b). An independent person trained in the formative tagging system then tagged a randomly selected 20% of each participant's transcription. Percentage agreement was calculated by dividing the number of agreements by the sum of the number of agreements and disagreements. A disagreement was counted when a different formative tag was used, when a tag was omitted, and when the raters had segmented words differently. Percentage of agreement when compared to the tagging of the first author was 94%. Because this agreement was relatively high, the first author's tags were used in the analyses. The same analyses that had been conducted on the untagged transcript (i.e., using the orthographic word as unit of analysis) were then conducted on the tagged transcript to determine the total number of formatives (TNF), total number of different formatives (TNDF), frequency counts, and commonality scores.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Results

Analysis by Orthographic Words

Prior to tagging, the sample of 9,000 orthographic words yielded 3,203 different words. Of these, 238 were used with a frequency of 0.05% and above and had a commonality score of 3 or above. These 238 words accounted for 51.9% of the total sample.

Analysis by Formatives

The tagged transcripts contained 20,137 formatives (TNF), including 1,110 different formatives (TNDF). With a frequency of at least 0.05% and a commonality score of ≥3 applied, a core vocabulary of 213 different formatives was determined. These 213 formatives accounted for 17,738, or 88%, of the total sample. From Figure 1, it is apparent that the core formatives were used by most participants, with an average commonality score of 5.2 among all core formatives. Within the top 60 formatives, only one was used by five rather than all six participants. All 213 core formatives are provided in the appendix, together with the most frequent orthographic word in which they occurred.

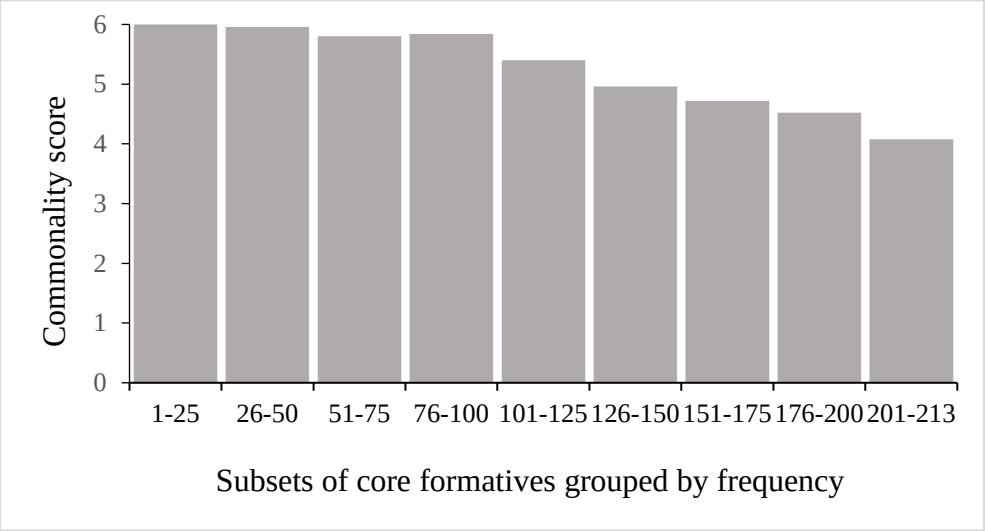

Figure 1. The average commonality scores of subsets of core formatives grouped by relative frequency.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Table 1

Content and Structure Formatives Occurring in the Composite Sample

Need heading	Part of speech	Formative	TNDF[a]	TNF[b]	Frequency in total sample (%)
Content	Substantive	Noun roots	28	773	3.8
		Pronouns	14	1 450	7.2
	Qualificative	Relative stems	4	192	0.9
		Adjective roots	8	215	1.1
		Enumerative roots	1	73	0.4
	Predicative	Verb roots	61	2 511	12.5
	Descriptive	Adverb roots	11	423	2.1
	Conjunction	Conjunctions	5	149	0.7
	Interjection	Interjections	13	593	2.9
			151	6 429	31.6
Structure	Substantive	Noun prefixes	12	1 648	8.2
		Demonstrative formatives	3	354	1.7
		Presentative formatives	2	145	0.7
	Qualificative	Relative concords	1	297	1.5
		Adjective concords	1	111	0.5
		Enumerative concords	1	46	0.2
		Possessive concords	1	394	1.9
	Predicative	Verb auxiliaries	20	1 986	9.9
		Verb concords	13	4 430	21.8
	Descriptive	Adverbial formatives	5	810	4.0
	Interrogative	Anclitic formatives (interrogation)	3	183	0.9
	Phrase	Enclitic formatives (vocative)	3	397	2.0
			66	10 838	53.3
Names		Proper names	3	621	3.1
Total			213	17 738	88.0

[a]Total number of different formatives. [b]Total number of formatives

The core formatives were then divided into structure and content formatives, based on Doke (1939). A total of 61 different structure formatives that are directly related to language structure occurred 10,801 times in the core vocabulary accounting for 60.8% of all core formative occurrences, while 145 different content formatives occurred 6,316 times, accounting for 35.6% of all core formative occurrences. Proper names accounted for the other 3.6% of occurrences. The different types of content and structure formatives, the total number of different

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

types of formatives, as well as the number of different formatives in each category are provided in Table 1. The table also shows the frequency of occurrence of different types of formatives appearing in the composite sample. Regarding content formatives, the core vocabulary contained 28 different noun roots, 61 verb roots, eight adjective roots and 11 adverb roots; however, the top 100 formatives contained only seven noun roots, 19 verb roots, three adverbs or adverb roots, and no adjective roots.

Discussion

At first glance, the Zulu core vocabulary based on an analysis of orthographic words in this study (first analysis) may seem to resemble other core vocabularies, with a greater number of words containing a smaller number of unique words and an even smaller group of words used with a high frequency. However, the coverage of the Zulu orthographic core words (51.9%) is much smaller than that found in English (e.g., Beukelman et al., 1989; Trembath et al., 2007) and French (Robillard et al., 2014) --with Robillard et al. based on speech samples of monolingual French children, where the core vocabulary accounted for around 80% of the total sample. In these three studies, the criteria used to define core vocabulary as well as age and number of participants were comparable to those in the present study, thus, the smaller coverage of the Zulu core vocabulary is likely a result of differences in linguistic typology and orthographic conventions between Zulu and English/French.

The content of the identified core orthographic word list also differs, with conjunctions and interjections ranking highly, as opposed to syntax-related vocabulary typically found in core (Robillard et al., 2014; Trembath et al., 2007). Although other types of words such as verbs, demonstratives, and possessives also appear in the list, the structure of Zulu means that these words contain a cocktail of information specific to the context of the speaker and the object.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Including such words on an AAC system is unlikely to afford the person using the system a level of generativity in producing novel utterances. Take, for example, the word *ngiyababona* (I see them) as referred to in the introduction. By providing this word in its entirety on an AAC system, only its specific message can be expressed; there is no potential for the elements to be recombined for new meanings.

This particular set of words may be useful when selecting vocabulary for AAC systems that are not intended to promote sentence construction and grammatical skills (e.g., activity- or phrase-based vocabulary for a communication board or an SGD with a limited number of messages) but are primarily aimed at enhancing early pragmatic aspects of communication such as turn-taking. Such a list may also be useful for rate enhancement, whereby the most frequently used words can be represented as a whole in the AAC system on a "quick-messages" page and accessed with a single hit rather than compiled from several formatives each time they are used. In contrast, the second analysis of the Zulu speech samples by formatives enabled the identification of a core list that is in many ways more comparable to other core vocabularies established in different studies.

As in previous core vocabulary studies (e.g., Boenisch & Soto, 2015; Robillard et al., 2014; Shin & Hill, 2016; Trembath et al., 2007), participants used a small set of core formatives, which accounted for a large proportion of their conversations. Another finding that was similar to previous studies was that structure formatives (although making up less than 50% of the core vocabulary) were used most frequently in the sample, highlighting that units of speech used for grammatical and syntactical purposes are used often. When relying on informants to generate word lists, these units are often underrepresented because they do not carry meaning by themselves (Morrow, Mirenda, Beukelman, & Yorkston, 1993; Yorkston et al., 1988). This

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

results in AAC systems that are noun- and verb-dominated (Banajee, Dicarlo, & Stricklin, 2003; Dark & Balandin, 2007) and that limit the construction of a variety of syntactical structures through which a range of linguistic meanings can be expressed (Sutton, Soto, & Blockberger, 2002). Generating a list of linguistic units used most frequently in actual conversations as one source to inform vocabulary selection seems the most reliable way of ensuring that those units of speech that provide the grammatical framework for the language are actually represented on AAC systems.

This is further underlined by the finding that the different types of structure formatives reflect the grammatical make-up of the Zulu language. For example, concords appear frequently. This is a particular requirement of the Zulu language, where concords are used to bring about agreement between the noun and other parts of speech in the sentence (Taljaard & Bosch, 1993). Other studies of languages belonging to the same language family have also identified the frequent use of concords (Allwood et al., 2010; Allwood & Hendrikse, 2003; Pretorius & Bosch, 2003). The overall high proportion of structure formatives that, in isolation, do not have a direct translation in English confirm that core vocabulary needs to be determined from primary data for each specific language, and cannot easily be translated, especially if the two languages in question differ in linguistic structure. Kilgarriff and colleagues (2014) compared frequency lists (comprising between around 7,500 and 9,000 words) between nine different languages, and also found that words that are directly translatable between language pairs constituted less than 50% of the lists for all possible language pairs. Language pairs where the two languages were more closely related in linguistic typology tended to have a higher percentage of words that were directly translatable.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Some similarities to previous studies could be discerned in the proportional distribution of some of the content formative types. For example, the proportion of noun, verb, adjectival, and adverbial roots increased substantially when considering the total core list versus only the top 100 formatives. A similar pattern was seen for nouns, verbs, and adjectives in the study by Boenisch and Soto (2015), where the proportions of these parts of speech increased when comparing the top 100 to the top 300 most frequently used words sampled from conversations of first language English-speaking children.

The findings of the current study underline that decisions about the unit of analysis used in core vocabulary studies need to be made deliberately, taking the linguistic structure of the language under investigation into consideration. Relying on the orthographic space only to determine a Zulu core vocabulary did not result in a core vocabulary with similar characteristics as previously established core vocabularies. Because Zulu is a synthetic, agglutinative language with a conjunctive orthography, smaller units of meaning needed to be identified within orthographic words in order to arrive at a small set of reusable vocabulary items.

Earlier studies in English have not always been explicit as to whether only lexemes are counted or whether grammatical variations such as declensions and inflections are counted separately. Diverging approaches were used in later studies. Boenisch and Soto (2015), for examples, seemed to have counted lexemes, while Trembath et al. (2007) counted grammatical variations as separate words. English is an analytic language with relatively few bound inflectional morphemes, and therefore these two different approaches arguably do not affect results dramatically, as illustrated by similar findings in these two studies despite different methods of counting. Studies in Korean, also a synthetic agglutinative language, have found quite different results depending on whether only lexemes were counted (e.g., Lee et al., 2005) or

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

whether morphological variations were considered (e.g., Shin & Hill, 2016). Lee et al. and Shin and Hill argue for the consideration of morphological variations of words, in order to reflect the complex morphology of Korean on AAC systems. The results of the current study further support the position that studies that aim to identify a generative core vocabulary for synthetic agglutinative languages should take morphology into consideration, because the frequency counts of formatives rather than orthographic words resulted in a core vocabulary that is comparative in number and coverage to previously identified core vocabularies. At the same time, system developers may find information about most frequently used orthographic words in combination with most frequently used formatives helpful, since this information could assist in developing prediction functions for SGDs. By developing algorithms based on word frequency and grammar rules, frequently occurring words could be suggested as a rate enhancement technique once a specific formative has been selected on the system. A similar observation was made by Clendon, Sturm, and Cali (2013) who found that frequently used words often occurred in the same combinations in the texts of beginning writers (e.g., “I like” and “I am”).

Implications for Practice

Incorporation of the core vocabulary into Zulu AAC systems. This study is the first to identify a Zulu core vocabulary list with the intention of proposing it as one resource to select vocabulary for inclusion on aided AAC systems for children from a Zulu language background who have not yet developed conventional literacy skills. The inclusion of the identified 213 core formatives would give access to a system that allows for the generation of novel utterances, using a variety of grammatical structures to convey different meanings. This list of formatives could therefore assist interventionists, parents and teachers in designing Zulu aided AAC systems for children. The question as to whether core vocabulary lists based on speech samples

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

from participants without disabilities are appropriate for children in need of AAC has been posed repeatedly in the literature and by AAC team members (Beukelman, McGinnis, & Morrow, 1991; van Tilborg & Deckers, 2016). While the vocabulary on any AAC system requires customization to reflect the needs, preferences, and personality of the person using the system, it can be argued that a core list based on samples of speakers without disabilities is a useful resource to consult. Various studies in a variety of languages have found highly similar core vocabularies based on the speech of children with and without disabilities (Boenisch, 2014; Deckers et al., 2017; Robillard et al., 2014; see also van Tilborg & Deckers, 2016, for a narrative review), suggesting that children with and without disabilities use the same words frequently in their everyday conversations. This would suggest comparable communication needs and opportunities in order to meet demands of various communication and interaction environments (e.g., the preschool environment). Furthermore, core lists can ensure that items are added to the system that will foster language development, even if a particular construction is not yet within the expressive repertoire of the child. Various authors have cautioned that a noun-heavy AAC system can stunt expressive language development (Binger, 2008; Fey, 2008; Sutton et al., 2002).

Any core vocabulary list needs to be supplemented by relevant fringe vocabulary, typically determined by informants or environmental inventories (Boenisch & Soto, 2015; Fallon et al., 2001; Sturm & Clendon, 2004). The 213 formatives identified in this study should therefore not be regarded as a finite set that would cover all the communication needs of a child in need of AAC. Vocabulary items that are reflective of a child's context, interests, group identity, and personality should be included on the AAC system, as should words that allow the child to participate and express himself or herself appropriately and creatively in a variety of

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

contexts. Vocabulary learning is related to broadening of conceptual understanding, and is a predictor of later school success (Biemiller, 2003; Saville-Troike, 1984). As far as possible, AAC systems should support the acquisition and use of a broad vocabulary.

Linguistic unit on Zulu AAC systems. Although many other core vocabulary studies have identified a word- rather than morpheme-based core vocabulary, the results of this study suggest that a word-based core vocabulary in Zulu lacks the generative potential that is typically regarded as a benefit of core vocabulary. Clinically, decisions about the linguistic unit to be included on aided systems for children who have not yet developed conventional literacy skills should be made deliberately, keeping both the learning demands and the linguistic potential of the system as well as the needs and abilities of the child requiring the system and his or her partners in mind.

It could be argued that an AAC system that includes formatives rather than words would pose high metalinguistic demands and may require phonemic skills, because the person using the system would need to segment an utterance into component formatives (not all of which carry meaning by themselves) and recombine the formatives when selecting the elements on the AAC system. In addition, representing structure formatives using graphic symbols may be difficult because these concepts are unlikely to elicit a specific visual image. However, even English systems that include core words would require the representation of structure words such as *to* and *at*—words that are only meaningful within the context of other words and therefore cannot be taught in isolation (Boenisch & Soto, 2015).

If access to a measure of generative language is desired, an AAC system must give access to elements that allow for grammar and morphology. Intervention studies with children and adolescents from an English-language background who used AAC have found that partner

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

strategies such as modeling, recasting, and prompting had facilitative effects on the participants' use of morphemes and correct syntax (Binger, Maguire-Marshall, & Kent-Walsh, 2011; Soto & Clarke, 2017, 2018). Similar strategies may be useful to facilitate the use of a formative-based Zulu AAC system. When formatives do carry a semantic meaning, it may be helpful to explain this meaning. In a study on Blissymbol learning, Schlosser and Lloyd (1993) found that teaching initial semantic elements assisted children without disabilities to discern the meaning of compound symbols made up of these elements. Scaffolds on SGDs, such as prediction functions, and selective availability of the second and subsequent formative based on the initial formative selection may also be considered. Moreover, further analysis of the semantic and grammatical context of formative use in the current and additional future speech samples could be of assistance in this regard.

Limitations and Future Directions

This study had a number of limitations that must be considered. First, the ambient noise on the playground and in the classroom affected the reliability of the transcription negatively, and this could have influenced the results of the study. Similar challenges were described in the core vocabulary study by Robillard et al. (2014). Second, only six pre-schoolers (all falling within the age range of 5–6 years) participated in this study, which limits the generalizability of the results. Although three different sites were used, they were relatively homogenous (rural preschool settings), and were situated in the same geographical region. Variations in language use between urban and rural Zulu speakers have been noted (de Kadt, 2005; Slabbert & Finlayson, 2000). Cross-linguistic contact in urban areas often leads to an increase in the use of loan words and code switching. The decision to focus on children from a rural context was deliberate because these children were less likely to be exposed to multiple languages in their

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

homes and communities; however, the core vocabulary may be somewhat different to that found in more urban settings. Future studies could obtain speech samples from larger groups of participants from diverse geographies to address these limitations.

Third, participants were only recorded during preschool activities. Although some authors have argued for a focus on specific activities in core vocabulary studies (e.g., Boenisch & Soto, 2015), sampling conversations in other contexts (e.g., the home) could lead to a more robust core vocabulary that is more representative of words used across contexts. Recording time was limited to 1–3 days per participant. This further limited the heterogeneity of the speech samples obtained. For example, the verb roots *bhala* (write) and *crayona* (to colour with crayons) appeared in the top 213 formatives, which may have been due to the curriculum of the classroom at this stage.

Fourth, this study did not analyze input the participants received from the teacher and other children. The vocabulary references in the Grade R curriculum were also not taken into consideration. The influence of these aspects on vocabulary use may be important to determine in future studies.

Fifth, participant reactivity remains a risk in observational studies. Some precautions were taken to counteract this, such as discarding the first 20 min of each participant's recordings. Also, although a research assistant was on site to assist with any challenges, she was not within sight of the children. All utterances that referred to the recording equipment and the process of recording were omitted from the analysis.

Sixth, no established guidelines exist regarding the criteria for classifying a vocabulary item as part of the core. While the criteria used to identify core vocabulary in this study are similar to those used by others (Boenisch & Soto, 2015; Robillard et al., 2014; Trembath et al.,

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

2007), the commonality score of ≥3 and the frequency count of ≥0.05% as criteria for the inclusion of formatives into core vocabulary must be regarded as somewhat arbitrary. Alternative measures such as grouped frequency counts (cf. Shin & Hill, 2016) may represent more defensible criteria for the determination of a core vocabulary. The decision to remain with the criteria that had been set in most previous studies was made primarily to enable comparisons to the parameters of the core vocabulary identified to those of previously established core vocabularies. An additional analysis of grouped frequency counts may contribute to a clearer data-based distinction between core and fringe vocabulary.

Last, because Zulu AAC systems designed by practitioners to date are based either on translations of English AAC approaches or on Zulu orthographic words, various design-related questions still need to be answered to guide the construction of a core-fringe Zulu AAC system integrating the morpheme-based core vocabulary established in this study. For example, visual representations that would be suited to represent the core vocabulary would need to be chosen or created. Participatory design methodologies may be useful to integrate expertise from various stakeholders and professionals in a number of design cycles and iterations to create one or more prototypes.

Conclusion

Selecting appropriate vocabulary for young children who use AAC is crucially important in order to maximize their learning and participation in a variety of contexts with different communication partners. Core vocabulary lists have been established in various languages and have been used as resources by AAC team members and AAC system developers to assist in this process. The core vocabulary list established for Zulu preschoolers in this study represents the first Zulu core vocabulary resource. On an AAC system, these core formatives could form a

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

grammatical framework that could then be supplemented by personalized fringe vocabulary. In in this way, AAC team members can design AAC systems that allow for a measure of generativity in production. Future studies are needed to establish the learning demands, appropriate teaching methods, as well as effectiveness in enhancing communication competency of such systems.

References

Allwood, J., Hammarström, H., Hendrikse, A., Ngcobo, M. N., Pretorius, L., & van der Merwe, M. (2010). Work on spoken (multimodal) language corpora in South Africa. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, … D. Tapias (Eds.), *Proceedings of the Seventh Conference on International Language Resources and Evaluation* (LREC'10) (pp. 885-889). Malta: European Language Resources Association (ELRA).

Allwood, J., & Hendrikse, A. P. (2003). Spoken language corpora for the nine official African languages of South Africa. S*outhern African Linguistics and Applied Language Studies, 21*, 189–201. doi:10.2989/16073610309486343

Ayres, K., & Ledford, J. R. (2014). Dependent measures and measurement systems. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology* (pp. 124–153). New York, NY: Routledge.

Baker, B. R., & Chang, S.-K. (2006). A Mandarin language system in augmentative and alternative communication (AAC). *International Journal of Computer Processing of Languages, 19,* 225–237. doi:10.1142/S0219427906001438

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Banajee, M., Dicarlo, C., & Stricklin, S. B. (2003). Core vocabulary determination for toddlers core vocabulary determination for toddlers. *Augmentative and Alternative Communication, 19*, 67–73. doi:10.1080/0743461031000112034

Beukelman, D. R., McGinnis, J., & Morrow, D. (1991). Vocabulary selection in augmentative and alternative communication. *Augmentative and Alternative Communication, 7*, 171–185. doi:10.1080/07434619112331275883

Beukelman, D. R., & Mirenda, P. (2013). *Augmentative and alternative communication: Supporting children and adults with complex communication needs.* Baltimore, MD: Paul H. Brookes.

Beukelman, D., Jones, R., & Rowan, M. (1989). Frequency of word usage by non-disabled peers in integrated preschool classrooms. *Augmentative and Alternative Communication*, *5*, 243–248.

Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology, 24*, 323–335. doi:10.1080/02702710390227297

Binger, C. (2008). Grammatical morpheme intervention issues for students who use AAC. *Perspectives on Augmentative and Alternative Communication, 17*, 62–68. doi:10.1044/aac17.2.62

Binger, C., Maguire-Marshall, M., & Kent-Walsh, J. (2011). Using aided AAC models, recasts, and contrastive targets to teach grammatical morphemes to children who use AAC. *Journal of Speech Language and Hearing Research*, *54*, 160. doi:10.1044/1092-4388(2010/09-0163)

Boenisch, J. (2014). Kernvokabular im Kindes- und Jugendalter: Vergleichsstudie zum Sprachgebrauch von Schülerinnen und Schülern mit und ohne geistige Behinderung und

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Konsequenzen für die UK [Core vocabulary in childhood and adolescence: Comparative study in language use of students with and without cognitive disabilities and implications for AAC]. *UK und Forschung 3*, 4-23.

Boenisch, J., & Soto, G. (2015). The oral core vocabulary of typically developing English-speaking school-aged children: Implications for AAC practice. *Augmentative and Alternative Communication, 31*, 77–84. doi:10.3109/07434618.2014.1001521

Clendon, S. A., Sturm, J. M., & Cali, K. S. (2013). Vocabulary use across genres: Implications for students with complex communication needs. *Language, Speech, and Hearing Services in Schools, 44*, 61–72. doi:10.1044/0161-1461(2012/10-0112)b

Dada, S., Murphy, Y., & Tönsing, K. (2017). Augmentative and alternative communication practices: A descriptive study of the perceptions of South African speech-language therapists. *Augmentative and Alternative Communication, 33*, 189–200. doi:10.1080/07434618.2017.1375979

Dark, L., & Balandin, S. (2007). Prediction and selection of vocabulary for two leisure activities. *Augmentative and Alternative Communication, 23*, 288–99. doi:10.1080/07434610601152140

de Kadt, E. (2005). English, language shift and identities: A comparison between "Zulu-dominant" and "multicultural" students on a South African university campus. *Southern African Linguistics and Applied Language Studies, 23*, 19–37. doi:10.2989/16073610509486372

Deckers, S. R. J. M., Van Zaalen, Y., Van Balkom, H. & Verhoeven, L. (2017) Core vocabulary of young children with Down syndrome. *Augmentative and Alternative Communication, 33*, 77-86. doi:10.1080/07434618.2017.1293730

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Department of Basic Education. (2018). *CAPS for foundation phase*. Retrieved from https://www.education.gov.za/Curriculum/CurriculumAssessmentPolicyStatements(CAPS)/CAPSFoundation/tabid/571/Default.aspx

Doke, C. M. (1939). *Textbook of Zulu grammar* (3rd ed.). London, United Kingdom: Longmans Green.

Fallon, K. A., Light, J. C., & Kramer Paige, T. (2001). Enhancing vocabulary selection for preschoolers who require augmentative and alternative communication (AAC). *American Journal of Speech-Language Pathology, 10*, 81–94. doi:10.1044/1058-0360(2001/010)

Fey, M. (2008). Thoughts on grammar intervention in AAC. *Perspectives on Augmentative and Alternative Communication, 17*, 43–49. doi:10.1044/aac17.2.43

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.

Kilgarriff, A., Charalabopoulou, F., Gavrilidou, M., Johannessen, J. B., Khalil, S., Johansson Kokkinakis, S., … Volodina, E. (2014). Corpus-based vocabulary lists for language learners for nine languages. *Language Resources and Evaluation, 48*, 121–163. doi:10.1007/s10579-013-9251-2

Kosch, I. M. (2006). *Topics in morphology in the African language context*. Pretoria, South Africa: UNISA.

Lee, Y. M., Kim, Y. T., & Park, E. E. (2005). A preliminary study for the core and fringe AAC vocabulary used by elementary school students. *Communication Sciences and Disorders, 10*, 134–152.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Lewis, P. M., Simons, G. F., & Fennig, C. D. (Eds.) (2016). *Ethnologue: Languages of the world* (19th ed.). Retrieved from https://www.ethnologue.com/statistics

Light, J. C., & McNaughton, D. (2012). Supporting the communication, language and literacy development of children with complex communication needs: State of the science and future research priorities. *Assistive Technology, 24*, 34–44. doi:10.1080/10400435.2011.648717

Liu, C., & Sloane, Z. (2006). Developing a core vocabulary for a Mandarin Chinese AAC system using word frequency data. *International Journal of Computer Processing of Oriental Languages, 19*, 285–300. doi:10.1142/S0219427906001530

Lund, S. K., & Light, J. (2007). Long-term outcomes for individuals who use augmentative and alternative communication: Part III—contributing factors. *Augmentative and Alternative Communication, 23*, 323–335. doi:10.1080/02656730701189123

Lund, S., Quach, W., Weissling, K., McKelvey, M., & Dietz, A. (2016). Assessment with children who need augmentative and alternative communication (AAC): Clinical decisions of AAC specialists. *Language, Speech, and Hearing Services in Schools, 48*, 56–68. doi:10.1044/2016_LSHSS-15-0086

Madrid, P., & Antonio, P. (2016, August). *Creating a non-English core vocabulary: Challenges faced while developing the Spanish Proloquo2Go vocabulary*. Paper Presented at the 17th Biennial Conference of the International Society for Augmentative and Alternative Communication, Toronto, Canada.

Marvin, C. A., Beukelman, D., & Bilyeu, D. (1994). Vocabulary-use patterns in preschool children: Effects of context and time sampling. *Augmentative and Alternative Communication, 10*, 224–236.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Miller, J., & Iglesias, A. (2012). *Systematic analysis of language transcripts (SALT).* Middleton, WI: SALT software LCC.

Morrow, D., Mirenda, P., Beukelman, D., & Yorkston, K. M. (1993). Vocabulary selection for augmentative communication systems: A comparison of three techniques. *American Journal of Speech-Language Pathology, 2*, 19–30.

Nyembezi, C. L. S. (1982). *Uhlelo lwesiZulu* [Zulu program] (4th ed.). Pietermaritzburg, South Africa: Shuter & Shooter.

Owens, R. E. (2016). *Language development: An introduction* (9th ed.). Boston, MA: Pearson Education.

Paulos, G., & Msimang, C. T. (1998). *A linguistic analysis of Zulu* (1st ed.). Cape Town, South Africa: Via Afrika.

Pretorius, L., & Bosch, S. E. (2003). Computational aids for Zulu natural language processing. *Southern African Linguistics and Applied Language Studies, 21*, 267–282. doi:10.2989/16073610309486348

Prinsloo, D. J., & De Schryver, G. (2001). Corpus applications for the African languages, with special reference to research, teaching, learning and software. *Southern African Linguistics and Applied Language Studies, 19*, 111–131. doi:10.2989/16073610109486279

Robillard, M., Mayer-Crittenden, C., Minor-Corriveau, M., & Bélanger, R. (2014). Monolingual and bilingual children with and without primary language impairment: Core vocabulary comparison. *Augmentative and Alternative Communication, 30*, 267–278. doi:10.3109/07434618.2014.921240

Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly, 18*, 199-219. doi:10.2307/3586690

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Schlosser, R. W., & Lloyd, L. L. (1993). Effects of initial element teaching in a storytelling context on Blissymbol acquisition and generalization. *Journal of Speech and Hearing Research, 36,* 979-995. doi:10.1044/jshr.3605.979

Shin, S., & Hill, K. (2016). Korean word frequency and commonality study for augmentative and alternative communication. *International Journal of Language and Communication Disorders, 51,* 415–429. doi:10.1111/1460-6984.12218

Slabbert, S., & Finlayson, R. (2000). "l'm a cleva!": The linguistic makeup of identity in a South African urban environment. *International Journal of the Sociology of Language, 144*, 119–135. doi:10.1515/ijsl.2000.144.119

Smith, M. M. (2015). Language development of individuals who require aided communication: reflections on state of the science and future research directions. *Augmentative and Alternative Communication, 31*, 215–233. doi:10.3109/07434618.2015.1062553

Snodgrass, M. R., Stoner, J. B., & Angell, M. E. (2017). Teaching conceptually referenced core vocabulary for initial augmentative and alternative communication. *Augmentative and Alternative Communication, 29,* 322-333. doi:10.3109/07434618.2013.848932

Soto, G., & Clarke, M. (2017). Effects of a conversation-based intervention on the linguistic skills of children with motor speech disorders who use augmentative and alternative communication. *Journal of Speech, Language, and Hearing Research, 60,* 1980–1998. doi:10.1044/2016_JSLHR-L-15-0246

Soto, G., & Clarke, M. T. (2018). Conversation-based intervention for adolescents using augmentative and alternative communication. *Augmentative and Alternative Communication, 34*, 180–193. doi:10.1080/07434618.2018.1490926

Spiegler, S., van der Spuy, A., & Flach, P. A. (2010a). Ukwabelana–An open-source

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

morphological Zulu corpus. *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010),* 1020–1028. doi:10.1109/ANZIIS.1995.705747

Spiegler, S., van der Spuy, A., & Flach, P. A. (2010b). *Additional material for the Ukwabelana Zulu corpus.* Retrieved from http://www.cs.bris.ac.uk/Research/MachineLearning/Morphology/Resources

Statistics South Africa. (2012). *Census in brief*. Pretoria, South Africa: Author.

Sturm, J. M., & Clendon, S. A. (2004). Augmentative and alternative communication, language, and literacy: Fostering the relationship. *Topics in Language Disorders, 24*, 76–91. doi:10.1097/00011363-200401000-00008

Sutton, A., Soto, G., & Blockberger, S. (2002). Grammatical issues in graphic symbol communication. *Augmentative and Alternative Communication, 18,* 192–204. doi:10.1080/07434610212331281271

Suzman, S. M. (1990). *Language acquisition in Zulu* (Doctoral dissertation, University of the Witswatersrand, Johannesburg, South Africa). Retrieved from http://nrfnexus.nrf.ac.za/handle/20.500.11892/106780

Taljaard, P. C., & Bosch, S. E. (1993). *Handbook of isiZulu* (2nd ed.). Pretoria, South Africa: Van Schaik.

Thistle, J. J., & Wilkinson, K. M. (2013). Working memory demands of aided augmentative and alternative communication for individuals with developmental disabilities. *Augmentative and Alternative Communication, 29*, 235–245. doi:10.3109/07434618.2013.815800

Thistle, J. J., & Wilkinson, K. M. (2015). Building evidence-based practice in AAC display design for young children: Current practices and future directions. *Augmentative and Alternative Communication, 31*, 124–136. doi:10.3109/07434618.2015.1035798

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284. https://doi.org/https://doi.org/10.1080/07434618.2019.1692902

Trembath, D., Balandin, S., & Togher, L. (2007). Vocabulary selection for Australian children who use augmentative and alternative communication. *Journal of Intellectual and Developmental Disability, 32*, 291–301. doi:10.1080/13668250701689298

van Tilborg, A., & Deckers, S. R. J. (2016). Vocabulary selection in AAC: Application of core vocabulary in atypical populations. *Perspectives on Augmentative and Alternative Communication, 1*, 125–138. doi:10.1044/persp1.SIG12.125

Wilkinson, K. M., & Hennig, S. (2007). The state of research and practice in augmentative and alternative communication for children with developmental/intellectual disabilities. *Mental Retardation and Developmental Disabilities Research Reviews, 69*, 58–69. doi:10.1002/mrdd

Witkowski, D., & Baker, B. (2012). Addressing the content vocabulary with core: Theory and practice for nonliterate or emerging literate students. *Perspectives on Augmentative and Alternative Communication, 21*, 74–81. doi:10.1044/aac21.3.74

Yorkston, K. M., Dowden, P. A., Honsinger, M. J., Marriner, N., & Smith, K. (1988). A comparison of standard and user vocabulary lists. *Augmentative and Alternative Communication, 4*, 189–210. doi:10.1080/07434618812331274807

Endnotes

[1] Grade R (Reception Year) refers to a non-compulsory year of schooling prior to the first compulsory year of schooling (Grade 1) that children are required to enter into in the year they turn 7.

Postprint version of:
Mngomezulu, J., Tönsing, K. M., Dada, S., & Bokaba, B. (2019). Determining a Zulu core vocabulary for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 35*(4), 274–284.
https://doi.org/https://doi.org/10.1080/07434618.2019.1692902