

Core Vocabulary Determination for Toddlers

MEHER BANAJEE*, CYNTHIA DICARLO and SARINTHA BURAS STRICKLIN

*School of Allied Health Professions, Louisiana State University, Medical Center, 1100 Florida Avenue,
New Orleans, LA 70119-2799, USA*

The aim of this study was to develop a core vocabulary list for toddlers. Naturally occurring (*i.e.*, unprompted) vocabulary was collected for 50 toddlers, aged from 24 to 36 months, enrolled in five different preschools, during two different activities (play within interest centres and snack time). Results revealed that all 50 children used nine common words across both routines, and that the list contained pronouns, verbs, prepositions and demonstratives. Words representing different pragmatic functions (*e.g.*, requesting, affirming, negating) were also included. Nouns were absent from the list. These data are consistent with similar studies into the core vocabularies of adults, adolescents, and preschoolers.

Keywords: Core; Fringe; Vocabulary; Toddlers

INTRODUCTION

Increasingly, Augmentative and Alternative Communication (AAC) devices are being used with toddlers (children between the ages of 24 months and 36 months) who exhibit expressive communication delays. Several factors have contributed to this trend in the USA, including (a) full implementation of Part C of the Individuals with Disabilities Education Act (IDEA, 1997)¹ which includes policies, procedures, and funding for assistive technology for children birth to 3 years of age with special needs; (b) recent advances in technology that have made AAC devices easier to use, more accessible, and lower in cost; and (c) wide ranging acceptance of recommendations from AAC researchers and practitioners (*e.g.*, Kangas & Lloyd, 1988) to begin implementing AAC strategies with infants (0–24 months) and toddlers (24–36 months) with communication delays before they attain certain prerequisite cognitive skills.

The increased use of AAC with young children creates several challenges for the field, in particular, the identification of suitable vocabularies when devising age-appropriate AAC systems. Some older children and adults may be able to generate their own messages by using the alphabet to spell, for example; however, pre-literate toddlers are unable to generate their own unique messages using letter-by-letter spelling. For these toddlers, significant adults typically select and program vocabularies on AAC devices

using an appropriate representation system (*e.g.*, pictures, icons, or photographs).

According to previous research (Beukelman, McGinnis, & Morrow, 1991; Blackstone, 1988; Morrow, Beukelman, & Mirenda, 1989), there are three main approaches to selecting vocabulary for children: developmental, environmental, and functional. A developmental approach involves the use of developmental vocabulary lists (Fristoe & Lloyd, 1980; Holland, 1975; Lahey & Bloom, 1977; Reichle, Williams, & Ryan, 1981), that are comprised of words chosen from developmental language inventories that have been developed on the basis of language acquisition principles. Knowledge of the development of different word forms (*e.g.*, nouns, verbs) and the number of words that children typically use at a certain age or developmental level is used to determine vocabulary for AAC systems. An environmental approach (Beukelman & Garrett, 1988; Blau, 1983; Carlson, 1984; Fried-Oken & More, 1992; Karlan & Lloyd, 1983; Mirenda, 1985) follows an ecological inventory process, in which words appropriate for specific communication environments (*i.e.*, fringe vocabulary) are identified and programmed on AAC devices. According to Yorkston, Dowden, Honsinger, Marriner, and Smith (1988) fringe vocabulary is specific to each communication environment (*e.g.*, marker, paper, and crayon for an art activity; cookie, drink, and spoon for a snack activity). The third approach, functional communication, interfaces with the pragmatic aspect of language. Vocabularies are

*Corresponding author. Tel.: (504) 942-8200. Fax: (504) 942-8253. E-mail: mbanaj@lsuhsc.edu

chosen based on expressed communication functions such as requesting, commenting, greeting, and protesting.

Identification of core vocabularies for toddlers involves aspects of all three approaches to vocabulary selection. A core vocabulary consists of words common to the vocabularies of peers who are similar in age (Yorkston et al., 1988). Vocabulary lists are based on the language inventories of typically developing toddlers and include the number of words and different word forms that children between the ages of 24 and 36 months typically use. Core vocabularies are small in size and do not change across environments or between individuals. Common words used across all communication environments comprise core vocabulary lists, which include structure words (e.g., want, more) that provide a framework for functional language use.

Both core and fringe vocabularies are important for communication purposes; however, children appear to use core vocabulary more frequently than fringe vocabulary (Beukelman, Jones, & Rowan, 1989). In a study of the frequency of word usage by six preschoolers, Beukelman et al. (1989) analyzed language samples for common or core words. They found that the 25 most frequently occurring words accounted for 45.1% of the sample collected. Fifty of the most frequently occurring words represented 50% of the sample, and 85% of the sample included 250 of the most frequently used words. Some examples of these frequently occurring words included want, eat, and go—verbs, demonstratives, prepositions, and adverbs. Nouns were not among the common or core words most frequently used by preschoolers within the study sample.

Despite evidence that nouns are not among core vocabulary used by preschoolers, Adamson, Romski, Deffenbach, and Sevcik (1992) reported that clinicians typically select nouns representing foods and objects as first symbols when designing AAC systems. According to Adamson et al. (1974) clinicians reported that nouns are chosen because they are considered to be easiest to teach and assess and are of considerable functional use to the communicator. In addition, the clinicians often omitted other words (e.g., want, more, help) that regulate interaction from augmentative communication systems and are harder to teach and represent on communication systems. When Adamson et al. (1974) added these action words (in addition to the nouns) to communication boards used by young males with moderate to severe intellectual disabilities, the frequency with which they used these boards increased from 2 to 41%. The Adamson et al. (1974) study is one of

several recent studies that have demonstrated that combining core and fringe vocabulary words increases the frequency of AAC use (e.g., Beukelman et al., 1991; Yorkston, Dowden, Honsinger, Marriner, & Smith, 1989).

Researchers have attempted to identify lists of words that could be included in a core vocabulary for a variety of people who use AAC, including adults (Balandin & Iacono, 1998); adolescents (Adamson et al., 1992); and preschoolers (Beukelman et al., 1989; Fried-Oken & More, 1992). Of these studies, Beukelman et al. (1989) is the most relevant to the identification of a core vocabulary for use by toddlers, given its focus on the vocabularies of preschool children. Beukelman and his colleagues audio-recorded and transcribed the spoken communication samples of six nondisabled preschool children (3 years 10 months to 4 years 9 months) in three different classrooms. Three of the participants were male and three were female. Teachers nominated these children for participation in the study because they were 'active verbal participants in the preschool program' (p. 244).

Conversational samples collected by inconspicuously audio-recording the target children in the classroom across six sessions were analyzed for vocabulary commonality. A commonality score of 6 indicated that all six participants produced the targeted word, whereas a score of 1 indicated that only one participant produced the word. Twenty-five words were identified as the most frequently occurring words (i.e., words that obtained a commonality score of 6). These words were mainly verbs, prepositions, pronouns, demonstratives, and articles. They also represented different semantic functions, including affirmation, negation, nomination (or labeling), and interrogation. Pragmatic functions represented included recurrence, termination, requesting actions, and establishing and maintaining joint attention. No nouns were noted in this list of 25 words.

Published studies that identify core vocabularies for toddlers could not be found in AAC or related literature. Accordingly, the Beukelman et al. (1989) study served as a foundation for the present study, whose purpose was to begin the process of identifying a core vocabulary for toddlers by collecting language samples (during play activities and functional classroom routines) from speaking toddlers and analyzing these samples for common words. For the purposes of this study, a core vocabulary list was defined as a list of words used by toddlers across all activities during both play within interest centers and snack time activities. The specific research questions for the study were: (a) Does the vocabulary used by

toddlers differ across different activities? (b) Are common words used by toddlers across different activities? (c) What are the common words used by toddlers across different activities? and (d) What kind of syntactic, and pragmatic and semantic functions do these common words serve?

METHOD

Participants

Fifty toddlers between the ages of 24 and 36 months served as participants in the study; 34 were girls and 16 were boys. The participants were recruited from five daycare centres/nursery schools in different socioeconomic areas (urban and suburban regions) within a large metropolitan area. In addition to meeting the criteria of age and enrolment in the selected child care centres, parent consent was obtained for each of the children who participated.

All of the participants were screened using the Ages and Stages Questionnaires (ASQ), a parent-completed child-monitoring system (Bricker & Squires, 1999). The ASQ indicated that participants were functioning at age-appropriate developmental levels, used a variety of two to three word utterances, spontaneously initiated interaction, maintained interaction by taking turns, terminated interaction appropriately, and consistently followed simple one-step directives and some two-step directives without gestures.

Setting

Participants were enrolled in nursery schools and day care programs located within inner city and suburban areas. All programs shared common features: (a) the classroom schedule included at least one free play and one snack time activity during the day; (b) care and education were provided by at least one teacher and one teacher assistant; (c) classroom environments were organized by interest centers (e.g., blocks, dramatic play, art); (d) the classroom schedule provided for both small group and large group activities; and (e) some activities were led by an adult (e.g., snack time) whereas others were child-directed (e.g., center time).

Although materials across each of the classrooms differed, each of the classrooms had some common materials. As an example, the block centers contained a variety of different blocks (e.g., LegosTM, cardboard blocks) and building materials (e.g., pop beads). Materials in the dramatic play centers included dress up materials (e.g., sunglasses, beads, shoes) and cooking

utensils (e.g., a stove, pots, and pans). The art centers included materials such as paper, crayons and markers, whereas the manipulative area contained different cause-and-effect toys. Each of the classrooms also included areas for reading books.

Across all of the classrooms, snack time activities took place at designated tables. After the children had washed their hands, they were asked to sit at the snack table where they were served by a teacher or assistant. They were given a choice of juice or milk to drink, but they were not given a choice of snack items; however, the children could request more snack or drink. When finished, the children were required to clean their area and place trash items in a garbage can. During snack time, the children were restricted to the snack table; however, during free play activities they were free to move from interest center to interest center. Teachers and teaching assistants interacted with the children during both snack time and play within interest centers.

Apparatus

Three voice-activated tape recorders with lapel microphones (Radio ShackTM Optimus CTR-115) were used to record the language samples. Voice-activated tape recorders helped to record words spoken by the target toddler only. Adult and peer speech was too distant for the recorder to be activated. The toddlers wore the tape recorders at the waist in a small bag. A lapel microphone was plugged into the tape recorder and clipped to the collar of the toddler. High quality microphones were used, in order to compensate for difficulties in understanding tape-recorded toddler speech; this provided a clear recording of the speech used by the toddlers.

Procedures

Data collection

Data were collected using the procedure outlined in Beukelman et al. (1989). This procedure involved audiotaping interactions among the target children, the classroom staff, and other classroom children during two different categories of activities on three separate days. One category of activities included child-directed play across five different classroom interest centers (e.g., blocks, dramatic play). The second category involved an adult-directed activity, snack time. Each activity lasted for approximately 20 min. During free play, children were allowed to play freely within any of the interest centers. Audiotapes were reviewed for the first 150 utterances

within interest centers and snack time activities across all three days. These 150 utterances included the first 25 words used by each child across the two activities on each of the 3 days.

Data were collected after the children had become accustomed to wearing the microphones and tape recorders. After the first 2–3 days, most children (and some of their peers not included in the study) asked to wear the tape recorders and would talk about them with adults in the centers. It took an average of 2 weeks across the five daycare centers/nursery schools for the children to wear the apparatus and resume their typical play behaviors without distractions from the recorders. Data were not collected during this phase of the study.

Data analysis

The language samples recorded during both categories of activities on all 3 days from all 50 children were analyzed. Three students enrolled in a communication disorders Master's level program were trained by the first author to develop a written, verbatim transcription of all of the language samples. During the transcription process, audiotapes were stopped after each utterance and a verbatim transcription was completed of the utterance. Unintelligible utterances were omitted from the transcription. If intelligibility problems were identified during any point in the day, the entire day's recording was omitted and was not used for transcription.

Analyses were conducted to examine commonality among the words across activities and children (as outlined by Beukelman et al., 1989). Each new word was given a score of 1. If the same child used a particular word in both activities, the word was given a score of 2. A word used in both

activities on all 3 days was given a score of 6. In addition, words with the same commonality score were ranked according to frequency of use, which was defined as the percentage of the number of times each word was used in the language sample. Using the method outlined by Miller (1989), type-token ratios (number of different words divided by the total number of words for each activity) were calculated for all 50 children per activity for each day. Average type-token ratio scores were also reported for all 50 children per activity across all 3 days. These ratios were compared with type-token ratios of 3 year old children as reported by Miller (1989).

Interrater reliability

Reliability was calculated on 20% of all word lists across both activities (free play and snack time). The first author conducted reliability checks across language samples collected from one center activity and one snack time activity of at least 10 of the children. Reliability scores were obtained by dividing the number of agreements between each student and the first author by the total number of agreements and disagreements multiplied by 100. Mean reliability for sample transcription, across all students, was 91% (range = 86–95%). Mean reliability for the first student was 89% (range = 86–92%), for the second student it was 91% (range = 89–93%), and for the third student it was 93% (range = 91–95%).

RESULTS

Table 1 shows the list of words that achieved a commonality score of 6 (nine words), 5, and 4.

TABLE 1 Words with commonality scores of 6, 5, and 4 and their frequency of use

Commonality Score					
6		5		4	
Words	Frequency	Words	Frequency	Words	Frequency
I	9.5	mine	5.8	a	4.6
no	8.5	the	5.2	go	4.2
yes/yeah	7.6	is	4.9	what	3.1
want	5.0	on	2.8	some	2.3
It	4.9	in	2.7	help	2.1
that	4.9	here	2.7	All done/ finished	1.0
my	3.8	out	2.4		
you	3.2	off	2.3		
more	2.6				

Note: Frequency is presented as a percentage.

The frequency of use of these words was converted into a percentage score by dividing the total number of words and multiplying by 100. As is evident from Table 1, eight common words were used by most of the toddlers across most of the settings, and six common words were used by some of the children across some of the settings.

Table 2 presents type-token ratios for each activity on each day, as well as average scores of each activity across all 3 days. These type-token ratios were compared with those developed for this age group (Miller, 1989) and were found to be age appropriate. Ratios obtained during snack time activity were lower than those obtained during free play activities because of the limited number of choices provided during this adult-led activity.

The data were analyzed for syntax, semantic, and pragmatic functions using the procedures developed by Miller (1989) for analyzing free-speech samples. The core vocabulary was found to serve different syntactic, semantic, and pragmatic functions. Core vocabulary words contained demonstratives (that), verbs (want), pronouns (my), prepositions (on), and articles (the). No nouns were found in this list. Semantic functions included use of agents (I), objects (you), labeling objects (that) and actions (go), possession (my), affirmation (yes), negation (no), location (in), interrogation (what), quantity (some), and termination (finished). Pragmatic functions expressed included initiating interaction by attracting attention (you), maintaining joint attention (this), indicating recurrence (more), and terminating interaction (finished).

DISCUSSION

Vocabulary selection is a difficult process when designing age-appropriate AAC systems for young children who do not speak. This is especially true for children who are still preliterate and, therefore, are unable to express their needs and wants using traditional orthography (i.e., either selection of whole words or individual letters to spell words). The literature review indicated that some core words are used across

different activities of older children (Beukelman et al., 1989), adolescents (Adamson et al., 1992), and adults (Balandin & Iacono 1998), but information was not available for toddlers. In the present study, we examined vocabulary words used by 50 toddlers, in an attempt to redress this gap in the literature.

The results of this study revealed that nine common words were used across child-directed free play and adult-directed activities within nursery school and day programs. A further analysis of the language sample revealed the use of words to express different parts of syntactic, semantic, and pragmatic functions. A lack of nouns was noted in the common words used across different activities. This finding seems logical because activities in a typical classroom contain different materials and toys. Furthermore, this finding reflects those obtained by Beukelman et al. (1988), whose vocabulary lists similarly contained very few if any nouns. The addition of words from other syntax classes (e.g., verbs, demonstratives, and pronouns) helped to increase frequency of use of the communication systems (e.g., Beukelman et al., 1991; Yorkston et al., 1989).

In the present study, the types of words in the core vocabulary appear to be similar in syntax, semantic, and pragmatic functions to those identified by previous investigators of core vocabulary for preschoolers (Fried-Oken & More, 1992), adolescents (Adamson et al., 1994), and adults (Balandin & Iacono, 1998). The nine core words identified by this research project were all included in the 25 most frequently used words identified by Beukelman et al. (1989). The similarities to past research help strengthen the premise that a common core vocabulary can be applied across activities and environments.

Clinical Implications

The results from the present study indicate the need to include words that enable young children to meet a variety of syntactic, semantic, and pragmatic functions on their communication devices. Some words that meet these needs might be difficult to graphically represent, which may result in their being omitted from the initial overlays developed for communication systems. Use of words that are difficult to represent graphically may be taught to young children by modeling the use of the words within activities. In addition, consistently pairing the picture or symbol (e.g., the Picture Communication Symbol for 'want') with the word programmed on the device should help to teach a child to use the same symbol to request objects.

TABLE 2 Average type-token ratios across participants per activity for each day

Activity	Day 1	Day 2	Day 3	Day 4
Snack time	0.41	0.42	0.41	0.4133
Free time	0.44	0.43	0.44	0.4433

Words that were less frequently used by the toddlers were also identified (Table 1). The words from these lists included an extended core group of words to draw from for vocabulary selection for communication overlays to be used on voice output communication devices. Thus, if a toddler was able to use more than nine words, the word lists with commonality scores of 5 and 4 (Table 1) were used. These words also were found within the 50 most frequently used words as identified by Beukelman et al. (1989).

Research Implications

Although the results of the present study appear to be promising, they should be interpreted with caution because of certain limitations. First, the size of the sample was small (i.e., 50 toddlers); second, the sample used was a convenient one (i.e., language samples were collected from daycare or nursery centers with which the authors had previous relationships); and third, the sample involved more girls than boys, and the participants were predominantly Caucasian, which means that the core vocabulary of the sample may not be representative of the core vocabulary used by children of different ethnic, cultural, or socioeconomic backgrounds.

In addition, because the vocabulary was collected across activities in daycare/nursery school settings, it may not be representative of core vocabularies used by children across different environments (e.g., home, playgrounds, and grocery stores). Marvin, Beukelman, Brockhaus, and Kast (1994), found that children use different topics in the preschool setting than at home, and argued that this probably resulted from being exposed to different toys, people, and routines (e.g., circle time in school versus bath time at home). However, some overlap of vocabulary across the two environments would be expected as a result of similarities between routines (e.g., meal times or toy play). Routines in homes (e.g., dressing, bathing) may include different materials and interactions that could create the need to use different vocabulary words than those used during routines in daycare centers. Children who play on playground equipment that requires them to use gross motor movements and activities may need to use different words while interacting with their peers and other adults than they would during indoor activities, such as those utilized in this study. Accordingly, just as there has been the case for topics (Marvin et al., 1994), there is a need to investigate vocabularies across many types of environments, in order to ensure the validity of a given core vocabulary.

In addition, words identified as core vocabulary for toddlers who are not disabled may or may not

be appropriate for use by toddlers with expressive communication delays. The core vocabulary list identified in the present study needs to be used with children who rely on AAC because they either experience communication delays or are unable to use speech, in order to determine how useful the vocabulary is for them across different activities.

Another potential limitation of the present study is that only the first 25 words expressed by each child per day per activity were used in the analyses. These did, however, combine into a corpus of 150 words in total for use in the analyses. Typically, the middle 25 words are included in a language sample (Miller, 1989); however, this procedure was not used because some language samples did not have sufficient content. Some children, for example, produced only approximately 25 words within the 20 min activity. Although type-token ratios (calculated for all 50 children per activity for each day) were found to be age appropriate when compared to those reported by Miller (1989), further investigation using larger vocabulary samples may be warranted.

Further research is needed to investigate the effectiveness of integrating core vocabulary words with fringe vocabulary words on communication devices. Researchers (Fristoe & Lloyd, 1988; Holland, 1975; Lahey, 1977) have suggested that core words and fringe vocabulary words be included in the first lexical words selected for language intervention. Additionally, researchers and practitioners have recommended that fringe words appropriate for different activities be used together with core words in order to develop a rounded communication system that could be used across various activities and daily routines (e.g., Beukelman et al., 1991; Yorkston et al., 1989). Systematic investigations with toddlers are needed to determine the utility of AAC devices programmed with core words only, fringe vocabulary words only, and core and fringe words integrated within the system or stored separately in the system in a way that may be easy to retrieve and recall (e.g., in a different area for each child or page). Future studies are also needed to evaluate the utility of the core vocabulary identified in this study on communication devices used by a variety of toddlers across a variety of activities.

Note

- 1 Part C of the IDEA provides funding for the provision of developmental services such as special instruction, speech, and occupational and physical therapy to children with disabilities

References

- Adamson, L., Ronski, M., Deffenbach, K. & Sevcik, R. (1992). Symbol vocabulary and the focus of conversations: Augmenting language development for youth with mental retardation. *Journal of Speech and Hearing Research*, 35, 1333–1343.
- Balandin, S. & Iacono, T. (1998). A few well-chosen words. *Augmentative and Alternative Communication*, 14, 147–161.
- Beukelman, D., McGinnis, J. & Morrow, D. (1991). Vocabulary selection in augmentative and alternative communication. *Augmentative and Alternative Communication*, 7, 171–185.
- Beukelman, D., Jones, R. & Rowan, M. (1989). Frequency of word usage by nondisabled peers in integrated preschool classrooms. *Augmentative and Alternative Communication*, 5, 243–248.
- Beukelman, D. & Garret, K. (1988). Augmentative and alternative communication for adults with acquired severe communication disorders. *Augmentative and Alternative Communication*, 4, 104–121.
- Blackstone, S. (1988). Vocabulary selection: Current practices and a glimpse of the future. *Augmentative Communication News*, 1, 1–5.
- Blau, A. (1983). Vocabulary selection in augmentative communication: Where do we begin? In H. Winitz (Ed.), *Treating language disorders: For clinicians by clinicians* (pp. 205–234). Baltimore: University Park Press.
- Bricker, D. & Squires, J., (1999). *Ages & stages questionnaires (ASQ)*. Baltimore: Paul H. Brookes.
- Carlson, F. (1984). *Picsyms categorical dictionary*. Unity, ME: Baggeboda Press.
- 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, 41–54.
- Fristoe, M. & Lloyd, L. (1980). Planning an initial expressive sign lexicon for persons with severe communication impairment. *Journal of Speech and Hearing Disorders*, 45, 170–180.
- Holland, A. (1975). Language therapy for children: Some thoughts on context and content. *Journal of Speech and Hearing Disorders*, 40, 514–523.
- Individuals with Disabilities Education Act Amendments of 1997. Part C, Section 632.
- Karlan, G. & Lloyd, L. (1983). Considerations in the planning of intervention: Selecting a lexicon. *The Association of the Severely Handicapped Journal*, 8, 13–25.
- Kangas, K. & Lloyd, L. (1988). Early cognitive skills as prerequisites to augmentative and alternative communication use: What are we waiting for? *Augmentative and Alternative Communication*, 8, 211–219.
- Lahey, M. & Bloom, L. (1977). Planning a first lexicon: Which words to teach first. *Journal of Speech and Hearing Disorders*, 42, 340–349.
- Marvin, C., Beukelman, D., Brockhaus, J. & Kast, L. (1994). “What are you talking about?” Semantic Analysis of preschool children’s conversational topics in home and preschool settings. *Augmentative and Alternative Communication*, 2, 75–860.
- Miller, J. (1989). *Assessing Language Production in Children*. Baltimore: University Park Press.
- Mirenda, P. (1985). Designing pictorial communication systems for physically able-bodied toddlers with severe handicaps. *Augmentative and Alternative Communication*, 1, 58–64.
- Morrow, D., Beukelman, D. & Mirenda, P. (1989). *Augmentative and alternative communication: Three vocabulary selection techniques*. Paper presented at the 1989 ASHA annual convention, St. Louis, MO: ASHA.
- Reichle, J., Williams, W. & Ryan, S. (1981). Selecting signs for the formulation of an augmentative communicative modality. *Journal of the Associations for Persons with Severe Handicaps*, 6, 69–84.
- Yorkston, K., Honsinger, M., Dowden, P. & Marriner, N. (1989). Vocabulary selection: A case report. *Augmentative and Alternative Communication*, 5, 101–108.
- Yorkston, K., Dowden, P., Honsinger, M., Marriner, N. & Smith, K. (1988). A comparison of standard and user vocabulary lists. *Augmentative and Alternative Communication*, 2, 189–210.

