AAC to Support Conversation in Persons with Moderate Alzheimer’s Disease

MELANIE FRIED-OKEN1, CHARITY ROWLAND1, DARLENE DANIELS1, MAYLING DIXON1, BRET FULLER2, CAROLYN MILLS1, GLORY NOETHE1, JEON SMALL3, KEVIN STILL1, and BARRY OKEN1

1Oregon Health & Science University, Oregon, USA
2Portland VA Medical Center, Portland, Oregon, USA
3University of California at Berkeley, California, USA

Abstract

Even though we know that external memory aids support communication in Alzheimer’s disease, the components of the communication aids for individuals with Alzheimer’s disease have not been studied systematically. The goal of these two pilot experiments was to examine differences in conversational performance of adults with Alzheimer’s disease related to the presence and absence of an aid, the type of symbol embedded in the aid, and the presence or absence of voice output. In Experiment 1, 30 adults with moderate-to-severe Alzheimer’s disease participated in 10-min conversations with and without personalized AAC boards. There was no effect of AAC, regardless of symbol type, and a deleterious effect of voice output. In Experiment 2, modified spaced-retrieval training preceded conversations, standardized prompts were presented, and semantically-based dependent variables were examined. For the 11 participants in the second experiment, there was a significant effect of AAC, showing that the presence of AAC was associated with greater use of targeted words during personal conversations. We discuss new information about the contribution of AAC for persons with Alzheimer’s disease, and demonstrate how the applied research process evolves over the course of a long-term commitment to a scientific investigation.

Keywords

AAC; Alzheimer’s disease; Dementia; Aided conversation

Introduction

Alzheimer’s disease (AD) is a neurologically degenerative condition characterized by a progressive decline in memory, attention, problem solving, and language (Hebert, Beckett, Scherr, & Evans, 2001; McKhann et al., 2011). The prevalence of Alzheimer’s disease among older adults has become a public health issue worldwide (Thies & Bleiler, 2012). Current reports indicate that 6–10% of all individuals over the age of 65 and 33–40% of individuals by the age of 90 years will present with AD. It is estimated that, by 2050, this number will increase to 13.2 million in the United States (Hebert, Scherr, Bienias, Bennett,
& Evans, 2003). At that time, every state will experience a substantial increase in one or more measures of burden: number of people affected, proportion of the population of all ages affected, and older age distribution of the people with AD (Hebert, Scherr, Bienias, Bennett, & Evans, 2004). A person with Alzheimer’s disease will live an average of 8 years and as many as 20 years or more from the onset of the disease (Alzheimer’s Association, 2002). During years of care giving, families and friends watch their loved ones become forgetful, isolated, and confused.

Language impairment, related to changes to the memory system, is common (Bayles & Tomoeda, 1995), and clinical research has been devoted to describing the expressive language of individuals with AD (Bayles & Tomoeda, 2007). Loss of semantic memory, or access to the words needed to formulate a specific message, is an early symptom that can be especially troublesome during conversation, when a person needs to communicate information or decisions regarding health care, preferences, or satisfaction. Many interventions have been proposed to help a person with AD gain access to lost words. Therapeutic approaches can be divided into two camps: internal memory strategies and external memory strategies (Bourgeois & Hickey, 2009). External memory strategies may include augmentative and alternative communication (AAC) devices, where lexical retrieval or language production may be supported by low- to high-tech devices.

Experimental examination of AAC for persons with Alzheimer’s disease surfaced in the 1990s (Fried-Oken, Rau, & Oken, 2000). For the adult with mild-to-moderate AD, Bourgeois (1992b; 1993; 1994b) pioneered a number of important investigations, proving that external memory aids do, indeed, improve communication. Additional researchers extended the studies to individuals with severe Alzheimer’s disease and demonstrated that, to an extent, external aids support conversation as well (McPherson et al., 2001). External memory aids include notebooks; wallets; communication boards; calendars; and displays that provide photos, drawings, symbols and words for contextualized, relevant interaction (Murphy & Boa, 2012). Bourgeois posited that pairing external aids with familiar and spared skills, such as turning pages and reading aloud, maximizes a person’s opportunity for success because the spared skills rely on automatic memory processes and the stimuli are relevant to a patient’s everyday life. As an individual’s cognitive or sensory deficits increase, the characteristics of the stimuli in the external aids must be adjusted (Bourgeois, 1992a). The use of such memory aids to stimulate conversation about familiar topics is similar to the use of AAC devices to relay messages to conversational partners. In fact, a multimedia communication tool, CIRCA, has been designed that shows one way that AAC can support conversation and social relationships in Alzheimer’s disease (Astell et al., 2008; Astell et al., 2010). CIRCA is a hypermedia-based system where a touch screen database of images, music, and text prompt and support conversation between a person with dementia, his or her care providers, and family. Relying on relatively strong long-term autobiographical memories, a person with dementia and the conversation partner scroll through scenes on the computer for a joint reminiscence activity. Although individuals with AD may have no difficulty speaking and thus have no need for AAC as an expressive communication mode, it is possible that AAC could have a role in supporting semantic memory (Bourgeois & Hickey, 2007). AAC might serve as a cognitive access aid or semantic prime for individuals with AD (Fried-Oken et al., 2000), and personalized AAC symbols may stimulate autobiographical memory, leading to improved conversational behavior. The research reported here investigated these possibilities.

Salient Features of AAC Devices

The Bourgeois (1992b; 1993; 1994a; 1994b) studies hold out the promise that AAC may have an influence on conversation for persons with moderate-to-severe AD. As yet, however, we have little information that would suggest which AAC devices or components
would be most valuable for this population. The symbol system incorporated into the device and the presence or absence of voice output (the spoken word) are two components of any AAC device that might have strong semantic effects on conversation.

Memory aids must contain some level of symbolization, but there has been no systematic examination of the effect of different symbol types (or language representation types) on conversation in AD (Fried-Oken et al., 2000). Bourgeois (1994a) and Bourgeois, Burgio, Schulz, Beach, and Palmer (1997) determined that varying the size of the pages, the size of the print font, and the number of pages, helps maintain the accuracy and automaticity of a memory book without specific instruction. Neustadt (2001), in describing functional interventions that she initiated for her husband with AD, states that different text fonts increased the probability of reading success. Abrahams and Camp (1993) used orthography to successfully teach name recall with spaced retrieval to individuals with AD. Wilson, Dagenais, and Rubin (2001) found that persons with moderate AD could be redirected for topic maintenance when presented with 2-dimensional (2D) pictures relating to their activities. And Leseth, Beesley, and Collier (1995) used 3-dimensional (3D) items from personal events or activities (such as candy wrappers and sewing scraps) in an attractive box to reduce agitation in persons with AD. There are no investigations about whether a specific symbol set (either print, pictures, or small objects) influences AAC use in conversation by persons with AD. voice output is another variable that has not been examined previously as an AAC feature for persons with AD.

**Purpose**

The overall goal of this research was to investigate the effect of AAC on the conversations of individuals with moderate AD, examining the influence of the type of symbols incorporated into the device and the presence of voice output. As we collected data on the first 30 participants, we observed that voice output seemed to have an unexpected negative effect on conversational success (Fried-Oken et al., 2009). It was at this point that we elected to analyze the data; results from these first 30 participants constitute Pilot Experiment 1. These results suggested major changes in our study protocol, which were implemented for the final 11 participants, which constitute Pilot Experiment 2. The methods and results for each experiment are reported in the next section. The discussion section presents the combined results of the two efforts and justifies the need for further work in this area, considering that most AAC studies have small subject groups. Taken together, they provide a glimpse into the realities of conducting applied research over time with a diverse subject group that demonstrates complex cognitive-communication impairments.

**PILOT EXPERIMENT 1**

The goals of Pilot Experiment 1 were (a) to determine whether AAC devices individualized for participants with moderate AD would improve conversational performance, (b) to examine the difference in conversational performance related to three types of symbols incorporated into the AAC device, and (c) to examine the difference in conversational performance related to the presence or absence of voice output in the AAC device. We predicted that use of an AAC device with 2D or 3D symbols would improve conversational performance as compared to no AAC device or an AAC device with print-only symbols. We had no hypothesis regarding the effect of voice output on conversational success.

Independent variables were AAC support, symbol type, and voice output. The two levels of AAC support were present versus absent; the three levels of symbol type were print alone (print), 2D symbols with printed labels (2D + print), and 3D symbols with printed labels (3D + print); and the two levels of voice output were present versus absent. AAC support was varied within participants, while symbol type and voice output were varied between
Method

Participants—Pilot Experiment 1 participants were the first 30 adults with AD who met the inclusion/exclusion criteria for the larger research effort. All exhibited moderate AD, based on NINCDS-ADRDA criteria (McKhann, Drachman, & Folstein, 1984). They were recruited from the Layton Aging & Alzheimer’s Disease Center at Oregon Health & Science University (OHSU), one of the 30 Alzheimer’s disease centers funded by the U.S. National Institute on Aging. Inclusion criteria were: diagnosis of probable or possible AD by a board certified neurologist according to DSM-IV criteria (American Psychiatric Association, 1994); Clinical Dementia Rating (CDR) = 1 or 2 (Hughes, Berg, Danziger, Coben, & Martin, 1982); Mini Mental Status Examination (MMSE) = 5–18 within 6 months of enrollment in the study (Folstein, Folstein, & McHugh, 1975); visual acuity better than 20/50 O.U. (as performed in the Layton Center); hearing loss < 40dB (as performed in the Layton Center); and English as the primary language. Exclusion criteria were: history of other neurologic or psychiatric illness (no CVA, reported alcohol abuse, traumatic brain damage, and reported recent significant psychological or speech/language disorder).

A total of 23 females and 7 males, with a mean age of 74 years (range = 50–94 years) participated. All participants identified their race as White, except for one whose race was identified as African American. The mean MMSE score was 12 (range = 5–18, out of a possible 30), placing them well within the moderate AD stage. The Functional Linguistic Communication Inventory (FLCI) (Bayles & Tomoeda, 1994) was administered to document degree of language impairment. The mean FLCI score was 61 (range = 27–85, out of a possible 88).

Procedures

Consenting, Testing, and Assignment to Condition: All sessions were conducted in participants’ residences (private family homes or residential care facilities). A total of six visits were conducted with each participant. During the first one or two sessions, the consenting process was completed, with administration of the FLCI and re-administration of the MMSE and/or CDR, if scores were more than 6 months old. Participants were randomly assigned to one of the six symbol-type/voice output conditions. Table I shows the total number of conversations conducted and the number of participants for each of the six conditions (print alone, 2D symbols + print, 3-D symbols + print; all with and without voice output). The unequal cell sizes resulted from the fact that the study was terminated earlier than intended (as explained below), while the randomization scheme was applied to the larger participant sample that we had targeted initially.

Vocabulary Selection: During the first two sessions, we queried participants, their familiar caregivers, and other direct-care staff regarding autobiographical topics that participants had enjoyed discussing in the recent past but now had difficulty discussing. To assist with topic selection, we presented a list of approximately 100 typical events (e.g., traveling, grandchildren, a famous person) that was developed by Svoboda (2002). We guided each participant to select one topic that he or she was comfortable discussing in detail. Some of the chosen topics included: a job as a store detective, gardening, the family farm, and summers at the lake. Once the topic was selected, we asked for suggestions regarding the vocabulary needed to converse about it; ultimately, 16 content words or 2-word phrases most needed to discuss the topic were generated with each participant. The topic and the associated 16 words were used for all conversations conducted with the participant.
**AAC Device, Symbol Types, and Voice Output:** The AAC device was a Flexiboard™, chosen because it is a touch-sensitive membrane board that is physically appealing to elderly participants (it is made of natural wood and titanium). It can be programmed with digitized speech output; includes software to develop vocabulary overlays; is light-weight, portable, large enough to display 16 tangible symbols + printed words; and user-friendly. Once the symbol type/voice output, topic, and vocabulary were determined, we created an overlay for the AAC device that incorporated the designated symbols for the 16 selected vocabulary items and recorded spoken labels using the Flexiloader™ software for those assigned to voice output conditions. The overlay presented the 16 items in one of three symbol types: print, 2D + print, or 3D + print (Figures 1-3). Print symbols were 1- to 2-word, 24-point printed labels. 2D + print symbols were approximately 2 in × 2 in (5.08 cm × 5.08 cm) colored photographs with 1–2 words, using 24-point printed labels (e.g., *Blue Lake, grandfather*). 3D + print symbols were small items representing the selected vocabulary, with 1–2 words using 24-point printed labels (e.g., a Hershey’s kiss + the word *chocolate*).

Family members approved all symbols and often provided photographs or other materials for the 2D and 3D symbols. For those assigned to voice output-present conditions, digitized speech presented the 1–2 word printed labels. The Flexiboard uses a Microsoft™ Office Sound Recorder³ (PCM 22.050 kHz, 8 bit, Mono; sound playback/recording is Intel™ Integrated audio⁴) to digitize and store spoken phrases. The voice output, programmed into the Flexiboard using the Flexiloader software, was played over two speakers placed next to the Flexiboard.

**Conversations:** Each participant engaged in 10 conversations (5 with and 5 without the AAC device) conducted over a period of 5–8 weeks. Sessions generally occurred once a week. During each visit, two conversations were held with the participant, one with and one without the AAC device. All conversations for a given participant were conducted by the same research assistant (RA). The order of control and experimental conditions were systematically alternated from session to session and counterbalanced across participants to control for order effects. A 10-min rest was provided between the two conversations held during each session. Conversations involved a predictable structure, with a greeting, introduction to the topic, introduction to the AAC device (if present), posing of questions and comments to prompt conversation about the selected topic, and closing grammar. For conversations with AAC support, the RAs began the conversation by drawing the participants’ attention to the device and labeling the symbol in each of the four corners of the board, activating the voice output if it was present. This strategy was employed to encourage participants to attend to the board, since many people with moderate AD may not process pictures or words merely placed in front of them. There was no specific training of the 16 symbols. In all conditions, RAs provided at least 5 sec for the participant to respond to each conversational prompt; if no response was forthcoming, the RAs supported the conversation using a *downshifting* strategy, which has proven effective with people who use AAC (Light, Beesley, & Collier, 1988). Each downshift provided a little more information about the response that was being solicited. For instance, if the initial probe was a general query about grandchildren (designed to elicit a comment about grandson, Matthew, whose picture was on the board), the first downshift might include the information that one of the grandchildren lives next door, and the next might include the information that this is a

---

¹The Flexiboard software was manufactured by Handitek AB of Sweden, and distributed by ZYGO-USA at www.zygo-usa.com of Fremont, California, USA.
²The Flexiloader software was manufactured by Handitek AB of Sweden, and distributed by ZYGO-USA at www.zygo-usa.com of Fremont, California, USA.
³The Microsoft Office Sound Recorder is a product of Microsoft Corporation, Redmond, WA USA.
⁴The Intel Integrated Audio is a product of the Intel Corporation Santa Clara, CA USA.
grandson who loves baseball. Each conversation was videotaped by a second RA; videotaping was terminated after 10 min.

Dependent Variables—Conversations were coded using a social communication framework that draws heavily on the work of Clark and Brennan (1991); Clark (1996; 1999); and Clark and Fox Tree (2002). The participant’s utterance, defined in relationship to the conversational turn, was the unit of analysis. An utterance was defined as a proposition completed, abandoned, or interrupted within the bounds of a conversational turn. The utterance was coded according to the signal track that it reflected. Signal track may be either main (analogous to Clark’s primary track), or collateral (analogous to Clark’s secondary track), or a combination of the two. Main-track utterances (e.g., “We always had fun at the lake in the summer.”) involve attempts to ground the main proposition of the conversational exchange. Collateral-track utterances involve attempts to manage the conversation itself, as opposed to its propositional content. Explanatory collaterals (e.g., “I can’t remember what I was trying to say.”) advance the conversation by attempting to clarify, re-phrase, or relay how the speaker is coping with the conversation; they advance the conversation by managing it for both the speaker and listener. Flag collaterals (e.g., a false start such as, “I said, he said, I say, I…”) serve as signals that the speaker is having difficulty with the conversation, but do not provide any insight into what’s wrong with the conversation. In addition to the signal-track coding, two other variables were coded. First, one-word utterances spoken by the participant were tallied because they imply a paucity of speech and minimal response to conversational prompts. Second, physical references to the AAC device were coded to quantify use of the Flexiboard, such as touching a symbol on the board or pointing to it (but excluding passively resting the hand on it). We also tracked the total number of utterances. Thus, five dependent variables were examined: number of utterances, percent of flag collateral (percent of all utterances that included flag collateral), percent of explanatory collateral (percent of all utterances that included explanatory collateral), percent of one-word utterances, and number of references to AAC device (only possible in conditions where the AAC device was present).

Data Set and Reliability Procedures—The total data set was comprised of 300 conversations. Of these, we chose to analyze the last four for each participant (for a total of 120 conversations), since RAs were more familiar with participants and their topics during the later conversations. The first 5-min segment of each 10-min conversation was discarded to permit familiarization between participants and conversational partners (which had to be re-established at each visit because participants did not remember the RAs); the remaining 5 min of the conversation were coded; this was determined to be sufficient, based on the research of Bourgeois (1992b) with a similar population.

The coding scheme was administered using the Observer 5.0 software developed by NOLDUS. The coding configuration was loaded into this software package and the videotaped conversations were coded as they were viewed through the program. Two conversations per participant, one AAC-supported and one unsupported, were systematically selected for reliability analyses, totaling 22% of the data. Conversations were coded by three RAs who neither conducted nor videotaped the conversations. RA1 served as the standard for the other two RAs; thus, reliability was evaluated for RA1/RA2 and for RA1/RA3 pairs. RA2 and RA3 coded only AAC-supported or unsupported conditions for any given participant, to avoid bias related to perception or expectation of contrasting performance with and without the AAC device. Inter-observer agreement for these 120 conversations...
Results

General Characteristics of Participant Conversations—Overall, the behavior of individual participants was remarkably stable across conversations and conditions, although some characteristics varied widely between participants. The number of utterances per 5-min conversation ranged widely between participants, with the least talkative participant averaging 32 utterances and the most talkative averaging 75 utterances across four conversations. Some conversations involved primarily one-word utterances (uh-huh, yes, no), while other conversations did not include any one-word utterances. At one extreme, a participant averaged 68% of one-word utterances across four conversations, while at the other extreme, a participant averaged only 10% of one-word utterances. Overall, while flag collaterals comprised only 16% of utterances and explanatory collateral comprised only 7% of utterances, some conversations were composed of over 40% of each type of collateral. Finally, references to the AAC device, which averaged only three per conversation, ranged from 0–24. Ten of the 30 participants never referenced the device, while one participant averaged 16 references across the two AAC-supported conversations that were analyzed.

Main Effects of AAC Support, Symbol Type, and Voice Output—Factorial MANOVAs were calculated with each of the fixed effects as independent variables: (a) AAC support (present or absent), (b) symbol type (print, 2-D + print or 3-D + print), and (c) voice output (present or absent). MMSE and FLCI scores were entered as covariates.

Effect of AAC Support: The effect of AAC support was not significant across the five dependent variables, yielding Wilks’ lambda = .958, \( F(1, 116) = .976, p < .436 \). Table II shows the means and SDs for each dependent variable for the 60 conversations conducted with AAC and the 60 conducted without AAC.

Effect of symbol type: The effect of symbol type was not significant across the five dependent variables, yielding Wilks’ lambda = .788, \( F(2, 55) = 1.291, p < .245 \). Table III shows the means and SDs for each dependent variable for the AAC-supported conversations conducted with each of the three symbol types (print, 2D + print, 3D + print).

Effect of Voice Output: The effect of voice output was significant across the five dependent variables, yielding Wilks’ lambda = .756, \( F(1, 56) = 3.363, p < .010 \). Univariate tests showed that there were significantly fewer total utterances and significantly more one-word utterances when AAC devices included voice output: total utterances, \( F(1, 56) = 7.604, p < .008 \); one-word utterances, \( F(1, 56) = 8.679, p < .005 \). Table IV shows means and SDs for each dependent variable for the AAC-supported conversations conducted in each of the two voice-output conditions (present versus absent).

Research Modifications—The results of Pilot Experiment 1 suggested a number of modifications to the research protocol that might produce more promising results because our initial hypotheses regarding the facilitative effect of AAC and of specific symbol types were not supported. The one statistically significant effect, related to the presence of voice output, was not anticipated and is discussed in detail in Fried-Oken et al. (2009). Voice output, a feature embedded in the AAC devices used by half of the participants, appeared to distract participants and depress performance. We, therefore, chose not to include voice output in the adapted protocol for Pilot Experiment 2.
Pilot Experiment 1 had demonstrated that placing a customized AAC device in front of an individual with AD does not necessarily lead to its use in conversation. Participants did not produce significantly different numbers of utterances or one-word responses, nor did they produce significantly different percentages of collateral track utterances to manage their conversations. Furthermore, in conversations supported by AAC, the type of symbol used did not significantly affect the conversational variables measured. Clearly, for this population of users, the presence of an AAC device without training does not improve naturalistic conversations about specific topics. We suspected that providing an AAC training exercise might address the attentional difficulties of participants. We decided to add a training component and began an active literature review to determine the best intervention to include, given our population, goals, and methods. Our search led us to the Academy of Neurologic Communication Disorders and Sciences (ANCDS, 2011). We determined that the work of Brush and Camp (1998) in the area of spaced retrieval (SR) was appropriate. The goal of SR, according to Sohlberg and Mateer (2001), is to alleviate specific problems in participation associated with a memory impairment rather than to restore memory processes. SR training has been used previously to teach individuals to use external memory aids. Bourgeois et al. (2003) compared the effectiveness of SR and a modified cueing hierarchy to teach persons with mild-to-moderate AD to use an external aid, and found that significantly more goals were attained and maintained at 1 week and 4 months post-training using SR. Camp, Foss, O’Hanlon, and Stevens (1996) and Brush and Camp (1998) also demonstrated some success with SR training for learning external memory aids in persons with mild-to-moderate AD. We posited that SR training would increase the likelihood that persons with AD would attend to the customized AAC device during personal conversations.

We also wondered whether, in the naturalistic conversations, the RAs were unknowingly providing additional support in the AAC absent condition to compensate for the lack of external support and joint reference for the conversations. If so, this might mask the AAC effect. We posited that results would be different if the RAs used a standard pre-determined set of prompts for all conversations to maintain uniform support across conditions.

Finally, we questioned whether the dependent variables that we examined in Pilot Experiment 1 were not particularly sensitive to the manipulation of AAC support. Based on research cited previously on communication in AD, we thought that semantic variables would more closely align with our premise that AAC serves as a semantic prime. To address these concerns, we made four major changes to the protocol for Pilot Experiment 2: (a) the RAs used a standard protocol to structure the conversation, (b) we eliminated the voice output condition, (c) we added a training component before each experimental conversation, and (d) we changed our dependent variables. These changes are now discussed.

**PILOT EXPERIMENT 2**

The goal of Pilot Experiment 2 was to determine whether implementing the previously noted changes to the research protocol would result in a clearly facilitative effect of AAC. For this experiment, the independent variable was AAC support. The three levels of this variable were: (a) control (conversations conducted without an AAC device), (b) primed control (conversations conducted without an AAC device that were preceded by a SR priming exercise), and (c) primed AAC (conversations conducted with an AAC device that were preceded by a SR priming exercise). (We did not include a simple AAC condition because Pilot Experiment 1 had shown that providing AAC without training was not effective.) The level of AAC support was varied within participants. The dependent variables were related to the use of targeted topical vocabulary and the number of physical references to the AAC device. We also examined the number of partner utterances and percent of partner questions...
to gauge consistency of partner behavior, using the standardized system of conversational prompts that was implemented. We had determined in Pilot Experiment 1 that we did not need to hold five conversations in each condition for reliable results. We decided to hold two conversations per condition in Pilot Experiment 2.

**Methods**

**Participants**—Participants were the next 11 adults who met the criteria established for Pilot Experiment 1. The 11 participants included 8 females and 3 males with a mean age of 73 (range = 60–85). All participants identified their race as White, except for one whose race was identified as Asian. The mean MMSE score was 16 (range = 14–18). The mean FLCI score was 73 (range = 61–84).

**Procedures**—Procedures were identical to those in Pilot Experiment 1, except for the modifications described in the upcoming section.

**Conversations:** During each of six sessions, one conversation was held with the participant under one of the three AAC support conditions (control, primed control, and primed AAC). We conducted two conversations per condition. Sessions occurred 2–3 times a week depending on the availability of each participant. The order of conditions was systematically alternated from session to session and counterbalanced across participants to control for order or practice effects. In primed control and primed AAC conditions, the conversations were preceded by an SR priming exercise using the 16 vocabulary items that had been chosen for the participant’s topic. The AAC device was presented and then removed before the primed control condition. For the SR task, the RA pointed to each symbol on the board and spoke its label. The participant repeated this behavior one symbol at a time until all of the symbols had been pointed to and spoken aloud by the participants. To standardize the RA’s behavior across conditions, all conversations involved 10 standard probes targeting the same 10 vocabulary items randomly selected from the 16 on the AAC device. The 10 probes prepared for one of the participants are included in the Appendix as an example.

**Dependent Variables**—we used the SALT, Semantic Analysis of Language Transcripts, software (Miller, 2000) to transcribe and code the conversations. The speech of both participant and partner was coded, along with any physical references to the AAC device by the participant. The utterance, defined in relationship to the conversational turn, was the unit of analysis. The SALT Phonological Unit method was used to define the bounds of an utterance, based on evidence of thought completion in conjunction with rising/falling segmentation and the presence of a pause.

Utterances were coded for the topicality of content words (i.e., nouns, verbs, adjectives, and adverbs). A corpus consisting of all content words used in any of the six conversations was developed for each participant, devoid of information as to which words were used in which conversations or conditions. From this corpus, words were identified as being either (a) targeted (identical to or synonymous with vocabulary on the AAC device), (b) highly related to those symbols, or (c) unrelated. Once the entire corpus of content words had been rated for topicality, the SALT software was used to calculate the number of times in each conversation a participant used words identified as targeted or highly related.

Three variables related to topicality were calculated: number of targeted words used, percent of targeted words used (out of total words), and percent of related words (targeted plus highly related words out of total words). Number of references to AAC device was also coded, as in Pilot Experiment 1. Two conversation-level dependent variables related to RA behavior were used to assess the consistency of RA behavior: number of utterances...
produced by the partner and percent of questions produced by the partner (out of total number utterances), which are automatically tallied by SALT. Thus, six dependent variables were examined: (a) number of targeted words used, (b) percent of targeted words used, (c) percent of related words, (d) number of references to AAC device, (e) number of partner utterances, and (f) percent of partner questions. The first three variables were used to evaluate the main effect of AAC support, the fourth variable quantified the degree to which participants used the AAC device when it was present, and the remaining two variables were used to describe the consistency of partner behaviors using the new standardized conversational protocol.

**Data Set and Reliability Procedures**—All six conversations were coded and analyzed for each participant, for a total of 66 conversations. The three topicality measures did not require observer judgment or reliability evaluation. The 16 targeted words were simply entered into SALT for each participant and the software identified how many times those words were used. Relatedness ratings were based on ratings that were applied to every conversation, (that is, if the word *auto* was rated as highly related to the symbol for *CAR* that was included on the AAC device, *auto* was automatically counted by the software as a highly related word in all conversations). For the other three dependent variables, reliability was assessed between two RAs for one AAC-supported and one AAC-unsupported conversation per participant. Inter-observer agreement for these conversations (number agreements/[number agreements + disagreements]) averaged 99% for number of partner utterances, 91% for references to AAC device and 97% for percent of questions produced by the partner.

**Results**

We evaluated the main effect of AAC support by examining differences observed on the three topicality measures for the three AAC support conditions. We also examined characteristics of partner behavior in order to evaluate the consistency of their conversational support across the three AAC support conditions. Finally, we counted the number of references to the AAC device to describe the degree to which participants appeared to use AAC support.

**Main Effect of AAC Support**—The main effect of AAC support was evaluated by comparing the three conditions of control, primed control, and primed AAC. A linear mixed model was run for each dependent variable separately. The models were adjusted for the effect of multiple observations within subjects. Condition was specified as a fixed effect, and MMSE and FLCI scores were entered as covariates. A restricted maximum-likelihood algorithm estimated the parameters and appropriate standard errors. All estimates were calculated with the MIXED procedure in SPSS v.15. Table V shows means and standard deviations of the three dependent variables for conversations conducted under each of the three conditions. Number of targeted words used (the raw number of words represented on the AAC device that were used) showed a significant difference for condition, yielding, $F(2, 95) = 3.232, p < .044$. Pairwise comparisons (Least Significant Difference) revealed that the differences between the primed AAC and control condition and the primed AAC and primed control condition were significant ($p < .029$ and .032, respectively), while the difference between the two control conditions was not significant. This indicates that, as predicted, participants included in their conversations significantly more of the 16 targeted words when they had the AAC device than when their conversations were not supported by AAC. Even though the standardized conversations targeted the same specific vocabulary under all conditions, and the SR exercise used in the primed control condition involved practicing the 16 targeted vocabulary items, only the presence of the AAC device appeared to improve lexical access. The percent of targeted words used (out of the total number of words used), a
measure of the extent to which the total vocabulary was characterized by use of the targeted words, also showed a main effect of condition, yielding, \( R(2,93) = 3.870, p < .024 \). Pairwise comparisons (LSD) revealed that the differences between the primed AAC and control condition, and the primed AAC and primed control condition, were significant \( (p < .013 \) and .027, respectively), while the difference between the two control conditions was not significant. However, this effect was quite focused and did not extend to the use of highly related words. For the percent of all related words (which included both the targeted words and words highly related to them) the effect of condition was not significant, \( R(2,90) = 2.528, p < .085 \).

**Consistency of Partner Behavior**—To determine whether the conversational protocol was used consistently by RAs across conditions, we monitored the number of utterances and number of questions produced by the RAs. The two conversational partners achieved remarkable consistency across conditions using the standard protocol. The mean number of utterances was virtually identical for the two partners (159 versus 160 per 10-min conversation) for each conversation condition (control, primed control, and primed AAC conversations), as was the mean number of questions asked (41 versus 42 per condition).

**Use of the AAC Device by Participants**—The mean rate of reference to AAC device was 13 per conversation. All participants in Pilot Experiment 2 made some references to the AAC device and only four of the AAC-supported conversations involved no references to the device at all. Clearly, Pilot Experiment 2 participants paid far more attention to the AAC device than did the Pilot Experiment 1 participants, who averaged only three physical references per AAC-supported conversation.

**Discussion**

This work provides important information about AAC treatment for individuals with moderate AD, as well as insight into the research process and experiments with diverse adults with complex cognitive-communication impairments. Pilot Experiment 1 results presented researchers and clinicians with some clear direction and significantly altered our subsequent research questions, methods, and data analysis. Conversational behaviors did not change significantly when a customized AAC device was placed in front of persons with moderate AD. Participants did not produce significantly different numbers of utterances or one-word responses, nor did they produce significantly different percentages of utterances to manage conversations. Furthermore, the type of symbol on AAC devices did not significantly affect the conversational variables measured. Most importantly, our results suggest that merely providing an AAC device to an individual with moderate AD does not necessarily lead to its use in conversation. In other words, AAC without training is not assistive. This adage has been used repeatedly, for many years (R. Creech, personal communication, 1992) but until now has not been supported by evidence for this population. This observation is supported by the low frequency of references to AAC documented in AAC-supported conversations in Pilot Experiment 1. In half of the conversations, participants never referred to the AAC device at all and eight participants never referenced the device in any of their conversations.

McPherson et al. (2001), in an attempt to improve conversational behaviors of adults with severe AD by providing customized external memory aids to five elders, provided two reasons why participants with severe AD did not improve target behaviors. First, they suggested that the elders demonstrate perceptual and attention problems that interfere with the use of an external device for conversation. This certainly must be considered for the participants with moderate AD. Second, they suggested, as we do, that more practice and specific instruction in the use of the external memory aids might support improved
expression. Bourgeois, Dijkstra, Burgio, and Allen-Burge (2001) stated that devices and strategies alone do not ensure communication success for the population of adults with AD. Indeed, this study provides vital empirical evidence that follow-up treatment is warranted to establish functional use when a device, even one that is customized with autobiographical information, is provided for an individual with cognitive-communication impairments.

Pilot Experiment 2 clearly proved that, when participants were provided with AAC priming in the form of spaced-retrieval exercises at the beginning of each experimental session, they used the Flexiboard much more frequently. Pilot Experiment 2 participants referred to the device 4 times more on average than did the Pilot Experiment 1 participants, who received no training on using the device. Training encouraged participants to pay attention to the device; merely placing an AAC device in front of a person does not lead to its use in conversation.

**Impact of AAC on Access to Topical Vocabulary**—A significant difference between the primed AAC condition and the two control conditions was found for two dependent variables (number of targeted words used and percent of targeted words used) but not for percent of all related words. Participants used significantly more targeted words to discuss the designated topic when the AAC device was present than when it was absent. This is clinically relevant for intervention guidelines, as primed words may be measured for treatment outcomes during topic-specific conversations. It implies that successful treatment outcomes for participants with severe anoma also may be characterized by this lexico-semantic variable. A valid measure for AAC treatment outcomes is the number of words produced in spontaneous conversation about a specific topic that appears on AAC tools.

The discovery that voice output does not facilitate and may even impede conversation for this population is an important one for intervention (Fried-Oken et al., 2009). Remember that the participants were not dysarthric or nonspeakers, so that voice output did not serve as an expressive communication mode. In voice output conditions the spoken label was heard only when the participant had selected a symbol on the AAC device, and then only if he or she pushed the symbol hard enough to activate the voice output (the RAs would provide assistance if participants did not push hard enough). The user had to select a symbol first, before the word was even spoken. Therefore, voice output could not serve as a semantic prime, although it might serve to reinforce the accuracy of the semantic choice after the fact.

The fact that the voice output caused reduced verbal output and appeared to distract Pilot Experiment 1 participants has clear clinical implications. This population of elders may not be familiar with the talking photo frames, speaking computers, voice-output information kiosks, and talking stuffed animals that are available on the market today. For a number of participants, the novelty of the output caused them to stop conversing and repeatedly press the symbols. Others simply ignored the output, perhaps because they couldn’t hear it, or because it was emitted by speakers that were separate from the AAC device and therefore may not have been associated with touching a symbol on the board.

**Limitations**—We acknowledge a number of limitations in this study that suggest additional research directions for AAC and AD. The design of Pilot Experiment 1 was affected by the unequal cell sizes for the six AAC conditions, which was related to the fact that the randomization scheme was applied to a larger sample of participants than was targeted initially (at the time that we decided to suspend the first study, due to a lack of a main effect for AAC, an uneven number of participants had been assigned to the six conditions). It is clear that persons with moderate-to-severe AD may not attend to AAC intervention because of significantly impaired perceptual and attentional skills and the effects of late stages of the disease. Further neuropsychological examination of attention and
perceptual skills is warranted to better understand participant behaviors. Characteristics of participants in the two experiments were different in terms of MMSE (cognitive) and FLCI (language) skills. Both of these scores were higher, on average, for Pilot Experiment 2 participants, which might have contributed to their relative success in the AAC condition.

Since both the procedures and the dependent variables changed from the first to the second study, the only direct comparison possible is the difference in the number of references to the AAC device observed. One might speculate that, had we continued the first study in order to include more participants and equalize cell sizes, a main effect of AAC might eventually have been revealed. However, after 2 years of study involving a relatively large number of participants, this did not appear to be a promising avenue of exploration. The clear results obtained from the second study, with only a small number of participants, confirmed the wisdom of this decision. To a great extent, these limitations are an expected result of a long-term research enterprise with a group of individuals who have severe neurode-generative language and cognitive impairments. There was little evidence on which to base the initial methodology within the AAC research field. Design changes during the 3 year duration of this study should have been expected, and do support the social validity of this research.

**Future Directions**—While evidence supports the value of low-tech AAC for individuals with AD when spaced-retrieval training is conducted before conversations, to date there is no empirical data on the use of high-tech, speech-generating devices for this population. As with the severe, chronic, aphasia population, questions arise for this clinical group as the general population is presented with new technologies and novel storage methods (Fried-Oken, Beukelman, & Hux, 2012). Mobile computing, and devices that produce good quality voice output, such as tablet computers and smart phones, must be examined for this population. As individuals with AD come to the clinical task with experience and skills in using daily technologies, and families bring questions about the feasibility of using these new technologies to AAC specialists, we must provide them with research-based answers. It is possible that individuals with AD who have been exposed to computers in their daily lives will respond more favorably to voice output.

**AAC for other Dementia Syndromes:** The research reported here addresses a population of adults who presented with moderate-to-severe AD. This is the most common dementia syndrome. There are many other dementia syndromes, such as frontotemporal dementia, Lewy Body disease, vascular dementia, and mild cognitive impairment, where people present with a different set of symptoms that require different intervention tools and strategies. Research should be directed toward the effectiveness of AAC intervention for these populations.

**Longitudinal Studies and Staging AAC Interventions:** As with any progressive neurodegenerative condition, staging of interventions becomes a crucial question that must be addressed over time. Fried-Oken (2008) proposed staged interventions for individuals with primary progressive aphasia (PPA) that combine a restorative rehabilitation approach with a communication support approach. We are currently comparing the effectiveness of AAC in conversations for individuals with PPA and moderate AD, and are finding promising results for the PPA cohort (Fried-Oken, Rowland, & Gibbons, 2010). We still need to determine whether early adoption of AAC tools and strategies improves performance over time for individuals with progressive cognitive-communication impairments. For example, should adults with mild cognitive impairment learn to use personally relevant contextualized photographs, even though they can still spell and use a computer keyboard adequately? Should we change the organization of the visual language for adults during different stages of their disease as they lose visual-perceptual, attentional,
and receptive/expressive language skills? Is initial literacy status a factor in system design if we know that users will lose their reading abilities? Is there a predictable set of supports that should fade away or be reinforced during the disease progression? Future research will address many of these issues, as we learn more about the how disease progression interacts with the nature of communication supports over time.

**Acknowledgments**

We acknowledge support from the U.S. Department of Health and Human Services, National Institutes of Health grant R21 HD47754 and P30 AG008017; the U.S. Department of Education, National Institute on Disability and Rehabilitation Research grant H133G040176; and the 2004 Oregon Tax Check-Off Alzheimer’s Research Grant. We thank the participants and their caregivers for inviting us into their homes and permitting us to document their conversations.

**Appendix**

Examples of 10 Initial Questions and Structured Probes used for every Conversation with Participant GS. Topic: Pets

1. Which family dog played soccer? *Brandy*
   Which dog did the neighborhood children ask to play with?
   This dog knew where the soccer lines were, what dog was it?

2. Who cooked food for Brandy? *Hannah*
   This person never had a dog before so she just cooked its meals?
   The breeder was astonished at the dog, who was it that fed the dog?

3. Where did you live when you had Brandy? *Germany*
   What country did you live in while you had Brandy?
   You worked at a school in this country?

4. What cat would take walks with Brandy? *Sheba*
   You had a cat that would walk with the dog?
   What cat was it that loved Brandy?

5. Which daughter gave you Sheba? *Kathy*
   This daughter had to fly back to US to get the Sheba?
   This daughter was in college so she gave you Sheba, who was it?

6. What dog was left abandoned? *Jeeves*
   You really didn’t care for this dog at first, which dog was it?
   Which dog had some behavior problems?

7. What was Jeeves favorite dog toy? *Stuffed Animals*
   He took very good care of this special toy, what was it?
   Other dogs would tear this apart but not Jeeves, what was it?

8. You had this breed of dog when you were growing up, what was it? *Scottish Terrier*
   This dog is known for being feisty and very loyal?
This dog breed came from a European Island?

9. What is the name of the dog that lives with you now? *(Flash)*
This dog likes to sleep with you in your bed, which dog is it?
This dog is very friendly?

10. Which daughter does Flash belong to? *(Janey)*
This daughter lives here with you what’s her name?
She is your youngest daughter?

References


Bayles, KA.; Tomoeda, CK. Improving the ability of Alzheimer’s patients to communicate. Canyonlands Publishing; Phoenix, AZ: 1995.


Leseth L, Beesley M, Collier B. Utilizing an AAC system to maximize receptive and expressive communication skills of a person with Alzheimer’s disease. ASHA AAC Special Interest Division Newsletter. 1995:7.


McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr. Kwas CH, Phelps CH. The diagnosis of dementia due to Alzheimer’s disease: Recommendations from the national institute on


Figure 1.
Sample communication board with print symbols.
Figure 2.
Sample communication board with 2-D + print symbols.
Figure 3.
Sample communication board with 3-D + print symbols.
Table I
Number of Pilot Experiment 1 Conversations Conducted and Participants (n) for each Condition.

<table>
<thead>
<tr>
<th>Voice output</th>
<th>Symbol type</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Print</td>
<td>2D+ print</td>
<td>3D+ print</td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>50 (n = 5)</td>
<td>50 (n = 5)</td>
<td>20 (n = 2)</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>60 (n = 6)</td>
<td>50 (n = 5)</td>
<td>70 (n = 7)</td>
<td></td>
</tr>
</tbody>
</table>
Table II
Means (SDs) of Dependent Variables for Pilot Experiment 1 Conversations Comparing Presence and Absence of AAC Device (120 Conversations).

<table>
<thead>
<tr>
<th>AAC</th>
<th># Utterances</th>
<th># 1-word utterances</th>
<th>% Flag collateral</th>
<th>% Explanatory collateral</th>
<th>References to AAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>51 (15)</td>
<td>32 (18)</td>
<td>14 (10)</td>
<td>8 (8)</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Absent</td>
<td>49 (13)</td>
<td>30 (16)</td>
<td>17 (11)</td>
<td>7 (6)</td>
<td>NA</td>
</tr>
</tbody>
</table>

*Augment Altern Commun. Author manuscript; available in PMC 2013 December 01.*
### Table III

Means (SDs) of Dependent Variables for each of Three Symbol Types in 60 Pilot Experiment 1 Conversations (with AAC).

<table>
<thead>
<tr>
<th>Symbol type</th>
<th>#Utterances</th>
<th>% 1-word utterances</th>
<th>% Flag collateral</th>
<th>% Explanatory collateral</th>
<th>References to AAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Print</td>
<td>52 (11)</td>
<td>29 (15)</td>
<td>11 (8)</td>
<td>6 (6)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>2D + print</td>
<td>55 (15)</td>
<td>36 (16)</td>
<td>17 (10)</td>
<td>7 (6)</td>
<td>5 (7)</td>
</tr>
<tr>
<td>3D + print</td>
<td>45 (18)</td>
<td>33 (23)</td>
<td>17 (12)</td>
<td>10 (12)</td>
<td>2 (5)</td>
</tr>
</tbody>
</table>
Table IV
Means (SDs) of Dependent Variables for each of Two Voice Output Conditions in 60 Pilot Experiment 1 Conversations (with AAC).

<table>
<thead>
<tr>
<th>Voice output</th>
<th>#Utterances</th>
<th>% 1-word utterances</th>
<th>% Flag collateral</th>
<th>% Explanatory collateral</th>
<th>References to AAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>46 (10)</td>
<td>35 (16)</td>
<td>15 (10)</td>
<td>9 (7)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Absent</td>
<td>54 (16)</td>
<td>30 (19)</td>
<td>14 (11)</td>
<td>7 (9)</td>
<td>3 (6)</td>
</tr>
</tbody>
</table>
### Table V

Means (SDs) for Three Semantic Variables for each of Three Conditions in 66 Pilot Experiment 2 Conversations.

<table>
<thead>
<tr>
<th>Condition</th>
<th># Targeted words</th>
<th>% Targeted words</th>
<th>% Related words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>28 (16)</td>
<td>4 (2)</td>
<td>15 (10)</td>
</tr>
<tr>
<td>Primed control</td>
<td>28 (13)</td>
<td>4 (2)</td>
<td>14 (11)</td>
</tr>
<tr>
<td>Primed AAC</td>
<td>38 (24)</td>
<td>6 (4)</td>
<td>10 (6)</td>
</tr>
</tbody>
</table>