Click on Bake to Get Cookies: Guiding Word-Finding with Semantic Associations

Sonya Nikolova[1], Marilyn Tremaine[2] and Perry Cook[1]

[1]Princeton University
Princeton, NJ 08540 U.S.A.
{nikolova,prc}@cs.princeton.edu

[2]Rutgers University
Piscataway, NJ 08854 U.S.A.
tremaine@caip.rutgers.edu

ABSTRACT

It is challenging to navigate a dictionary consisting of thousands of entries in order to select appropriate words for building communication. This is particularly true for people with lexical access disorders like those present in aphasia. We make vocabulary navigation and word-finding easier by building a vocabulary network where links between words reflect human judgments of semantic relatedness. We report the results from a user study with people with aphasia that evaluated how our system (called ViVA) performs compared to a widely used vocabulary access system in which words are organized hierarchically into common categories and subcategories. The results indicate that word retrieval is significantly better with ViVA, but finding the first word to start a communication is still problematic and requires further investigation.

Categories and Subject Descriptors

K.4.2 [**Computers and Society**]: Social Issues—Assistive technologies for persons with disabilities; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces

General Terms

Human Factors, Design, Experimentation

Keywords

Assistive Communication, Aphasia, Adaptive User Interfaces, Semantic Networks, Visual Vocabularies

1. INTRODUCTION

Aphasia refers to a family of acquired communication disorders that impact an individual's language abilities. It affects close to one million people in the United States alone [16] and is usually acquired as the result of a stroke, brain tumor, or other brain injuries. Depending on the area or degree of damage to the brain, the resulting impairments to the ability to understand and produce language vary. For example, some people may speak fluently, but have impaired auditory comprehension while others may have impaired speech, but good reading comprehension. Substantial variations in the type and level of impairment severity can be observed in an individual as well as across different people. Anomia, the inability to access and retrieve words from the brain, is the most common impairment [7]. Even though rehabilitation can improve people's abilities, a significant number of people with aphasia are left with life-long chronic impairments among which anomia persists [10]. Given the importance of language communication in all aspects of daily life, it is not surprising that most individuals with aphasia experience a reduction in their ability to participate in everyday activities with the result that social isolation and depression are relatively common [12].

People with aphasia cope with their inability to communicate using different low-tech strategies such as drawing pictures, writing notes, pointing, mimicking and gesturing. There has also been commercial and research effort to build technological tools to help this user population communicate and thus, regain some of their independence. High-tech assistive communication tools have a number of advantages: 1) they are good for repetitive tasks such as practicing spelling, naming and pronunciation; 2) they provide consistent and unemotional feedback; 3) they are evolving to be mobile and unobtrusive, and thus, can address common stigma issues and assist users in different contexts outside of their home.

In order to be effective for functional communication, an assistive tool needs to provide a rich and expressive vocabulary. To meet this requirement, vocabularies offer extensive collections of words which are cumbersome to navigate. We address this problem by enhancing a basic vocabulary hierarchy with additional associations between words that provide shortcuts in different categories and on different levels. Previously, we showed that by adding semantic associations and associations based on prior word usage, we can significantly shorten the browsing paths between related words [18].We also demonstrated that people without communication impairments find words faster with the enhanced vocabulary and consider interacting with it more satisfactory [19]. In this paper, we demonstrate that a vocabulary hierarchy augmented with associations that reflect human judgments of semantic relatedness enables people with aphasia to find words in the context of a sentence faster and that the associations guide their search in an intuitive way. We present the results from a study comparing word-finding using a commercial hierarchical vocabulary access system and one using our enhanced vocabulary. We discuss the paths people take to reach the target words, the associations they expected to lead them to the target words, users' feedback on interacting with the two organizations and the challenge of searching for the first word to initiate a phrase.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ASSETS'10, October 25–27, 2010, Orlando, Florida, U.S.A.

Copyright 2010 ACM 978-1-60558-881-0/10/10…$10.00.

2. RELATED RESEARCH

It is challenging to enable a user to intuitively and quickly find words and thus, assist him/her in building functional communication. Many people with aphasia cannot always type the desired word in a search box, but instead have to browse through an extensive vocabulary of pictorial representations of concepts until they find the one that expresses their current communication needs. Minimizing the complexity of navigating such a vocabulary and supporting efficient word-finding is essential to the usability of assistive communication tools. In investigating how to create shortcuts between concepts and enable the user to build communication faster, we start with a basic vocabulary organization described below. We then augment the core vocabulary with word associations based on theories that explain how the human mind organizes words.

2.1 The Difficulties of Navigating an Assistive Vocabulary

Previous work on designing assistive communication tools for people with aphasia has shown that vocabularies consisting of thousands of words are challenging to navigate [2, 3]. Most existing assistive visual vocabularies have a lexical organization scheme based on a simple list of words, a list of categories of words or a hierarchy of categories and subcategories.

While there is no consensus on what the best way to organize an assistive vocabulary is, speech-language pathologists suggest that fewer categories are less confusing and easier to navigate [3]. However, populating a category with a sufficient number of words introduces the problem of excessive scrolling. The VocaSpace, featured in an augmentative and alternative communication software product called Proloquo2Go [1], takes this category-centered approach. The words in VocaSpace are organized in functional categories, e.g. greetings and questions, and common word categories such as colors, places, and clothes. Unfortunately, no published evaluation on its effectiveness is available. On the other hand, commercial tools such as TouchSpeak [23] and Lingraphica [13] offer categories of words that are organized in a hierarchy. For example, the *dinner* category is a child of the *meals* category, which is in turn a child of the *food* category.

Each organization has some disadvantages. A hierarchy can help a user build, over time, a mental model of a vocabulary but it often leads to deep and non-intuitive searches. Well-populated common word categories such as a collection of food items could be intuitive as an organization but result in excessive scrolling. More abstract categories, on the other hand, may introduce a sense of disorganization especially since speech-language pathologists suggest that the majority of people with aphasia have difficulties abstracting a word to its superordinate [17].

2.2 Lingraphica's Vocabulary Hierarchy

We base our organization on Lingraphica's vocabulary. Lingraphica is an assistive tool built specifically for people with aphasia. It provides a well-structured vocabulary hierarchy that has been carefully designed and proven useful (via its application in a commercial assistive device widely used across the US). Lingraphica's vocabulary also consists of a set of common words used in daily communication which has evolved over a number of years incorporating feedback from users, their caregivers and speech-language pathologists who prescribe the device and use it to treat patients.

Lingraphica's vocabulary organizes words according to shared contexts that are common in daily life. If you need to find *milk*, for example, you select *kitchen*, then the *refrigerator* category, then you find *dairy*. The icon for *milk* is in *diary*. This organization is not necessarily straightforward for all users. Some people associate the word *milk* with the *drinks* category while others may prefer a *location* association and keep *milk* in the *refrigerator*. This problem is partially addressed by allowing the user to customize the vocabulary categories.

Not unlike other hierarchical organizations, Lingraphica's hierarchy tends to be deep and requires multiple clicks to find even simple words. A particularly bad example is the path to the word *sleep*: *dictionary → actions → daily routines → inhale → sleep*. Even for straightforward paths, the task can easily become cumbersome because a number of hierarchy branches have to be traversed when composing a simple phrase.

2.3 Speaker's Mental Lexicon

To address the problem of cumbersome vocabulary navigation, we appeal to the psychological literature on speakers' "mental lexicon," where words are stored and organized in ways that allow efficient access and retrieval. Every speaker has experienced the inconvenience of temporarily impaired semantic connections (the so-called tip-of-the-tongue (TOT) phenomenon). This inability to retrieve a specific word needed to express a given concept can be due to a variety of causes such as fatigue or interference from a word that is morphologically or phonologically similar to the target word.

Experimental evidence — including evidence from TOT states induced in the laboratory — suggests that words are organized in a speaker's mental lexicon by various similarity relations, in particular, phonological and semantic similarity. For example, subjects in word association experiments overwhelmingly respond with husband to the stimulus wife [14]. Semantic priming [21], a robust and powerful tool for the experimental investigation of cognitive processes, relies on the semantic relatedness of the prime and an experimental target: responses to the target are faster when it is related to the prime as in the classic case doctor-nurse. Spreading network activation models [6] assume that presenting a prime stimulus word activates the corresponding representation in lexical memory and that this activation spreads to other related nodes, thus facilitating the retrieval of related target words.

In the context of our vocabulary, we apply these theories and methods by augmenting a basic hierarchical vocabulary with a collection of semantic associations based on the measure of evocation, i.e. how much one word brings to mind another word. By introducing these meaningful links between words, we hope to make navigation and word-finding more effective and intuitive.

2.4 Semantic Therapy

Other researchers have already recognized the potential of semantic associations in helping people with aphasia rebuild some of the impaired links in their mental lexicon. As part of a treatment, Webster et al., for example, employed a training strategy in which the participant was asked to generate words that are associated with a target verb [24]. This strategy, termed *Semantic Feature Analysis* (SFA), is an approach mostly applied to treating the retrieval of nouns and verbs and has shown great potential [5, 8]. SFA is used to guide the patient in identifying important semantic features of the target word. This exercise helps activate the semantic network that surrounds the target word and consequently aids in its retrieval [5]. In the process of identifying features of the target item, nontargeted and semantically related words may also benefit from the treatment, because they share features that are being accessed or retrained. We build onto these positive findings and create semantic networks that the user can navigate in search of a specific word. These networks model the organization of words in a person's mental lexicon and can compensate for some impaired connections, successfully guiding the user to their desired word.

3. DESIGN OF ViVA

In this section we elaborate on the design of the **Vi**sual **V**ocabulary for **A**phasia, ViVA. ViVA relies on word frequency usage, user preferences and semantic word associations to create a rich vocabulary network that enables effective word-finding. For the experiment described in this paper, we used a subset of the vocabulary and its functionality, but more details on the full version can be found in [18]. We first describe the core set of words that comprise ViVA and then explain what semantic associations were used to build the vocabulary network.

3.1 Core Vocabulary

We selected ViVA's initial vocabulary set to be a collection of commonly used words as well as ones relevant to our target population, people who have aphasia. This was achieved by mining words from two sources: the "core" WordNet (a large-scale semantic database [9, 14]) consisting of frequent and salient words and Lingraphica's visual vocabulary [13].

Lingraphica represents each concept with an icon that combines text, a pictorial representation of the concept, and speech output of the text. After finding the intersection of WordNet's most frequent and salient words and Lingraphica's vocabulary, we used the pictorial representations for coarse disambiguation. For each word, we selected WordNet's meaning that corresponded to Lingraphica's visual representation. For example, Lingraphica has images for two meanings of *glass*: a) drinking glass, and b) a brittle transparent solid. Thus, both senses were considered and matched to the appropriate icon. The resulting vocabulary consisted of approximately 1300 words that we organized according to Lingraphica's hierarchy. In the following section, we explain how this set was reduced to fit the purposes of our experiment.

3.2 Evocation

To augment the basic vocabulary hierarchy with meaningful links between words, we concentrated on the measure of evocation, i.e., how much does one word bring to mind another word. Evocation is particularly useful for adding cross-part-of-speech links that allow for connections among entities (expressed by nouns) and their attributes (encoded by adjectives). Similarly, events (referred to by verbs) can be linked to the entities with which they are characteristically associated. For example, the intuitive connections among *traffic*, *congested*, and *stop* can be clearly conveyed using evocation. By introducing such links based on evocation in the vocabulary, we attempt to improve vocabulary navigation, speed up phrase composition and increase users' satisfaction in interacting with the vocabulary.

We collected scores for strength of evocation for 100,000 word pairs through a large-scale online experiment that asked untrained annotators to provide ratings [18]. The data, which correlated well with ratings gathered from trained annotators, was then used to augment the core vocabulary hierarchy with links reflecting human judgments of semantic relatedness.

4. EXPERIMENTAL TASK DESIGN

Our first evaluation of ViVA simulated vocabulary usage and sentence construction using sentences collected from blogs of elderly people. The experiment found that, compared to the paths available in Lingraphica, ViVA shortened approximately 52% of the browsing paths between words in a sentence [18]. We define a browsing path to be the number of words retrieved between $word_n$ and $word_{n+1}$ in a sentence. To investigate whether people would take advantage of these shorter paths, we conducted a preliminary experiment evaluating ViVA with participants without language impairments. We asked the participants to find as quickly as possible the missing words in a number of phrases, using two vocabularies. The first vocabulary that we call **LG** has Lingraphica's hierarchical organization and the second one, **ViVA**, inherits the same hierarchy but is augmented with semantic associations which translated into related words that participants could see once they had clicked on an icon.

The results of the experiment showed that participants took significantly less time to find words with ViVA, taking shorter paths guided by the provided semantic associations. All participants agreed that having related words helped them find words faster and most thought that finding words in ViVA was less confusing than searching in LG [19]. Using ViVA, people tended to search for words via related-word links instead of trying to locate what category the word should belong to. Next, we describe how we adapted the experimental task used in the study with non-aphasics to evaluate ViVA with people with aphasia.

4.1 Stimulus Construction

We followed three high-level guidelines in redesigning the task for aphasic participants. The first guideline is avoiding fatigue – people with aphasia tend to get tired fast, one hour is therefore a reasonable time frame for an experiment. This meant that we had to reduce the number of phrases and target words. The second guideline is eliminating any factors that may intimidate the participants, e.g., making the participants aware that we were measuring speed or error rate. Thus, instead of asking people to find the words as fast as possible, we encouraged them to take their time searching and did not specify the maximum number of target words for the experiment. The third guideline is making the experience rewarding by ensuring that participants can successfully complete most of the tasks. To address this consideration, we simplified the phrases and provided additional context so that it was easier for people to guess the missing words.

We simplified the phrases by using only words that are frequently used according to the British National Corpus [22]. To be able to compose a simple sentence or a phrase, we paired nouns with verbs and nouns with adjectives. Because presenting only a phrase to the participants provided very limited context, we added an additional step to the protocol to emphasize the context of the target words and to stimulate the part of the mental lexicon where they could be found. We created a set of scenarios consisting of an image and a phrase related to it. The images were selected from Google Images [11] such that they evoked the target words. This selection was accomplished by typing the verb/noun or noun/adjective pairs in the search box and selecting the best image from the first page of search results. Results were filtered based on image quality and how well they represented the desired terms. For example, the image judged best for *bake* and *cookies* is the one in the upper half of the stimulus shown on Figure 1. We then constructed sentences that related to the image.

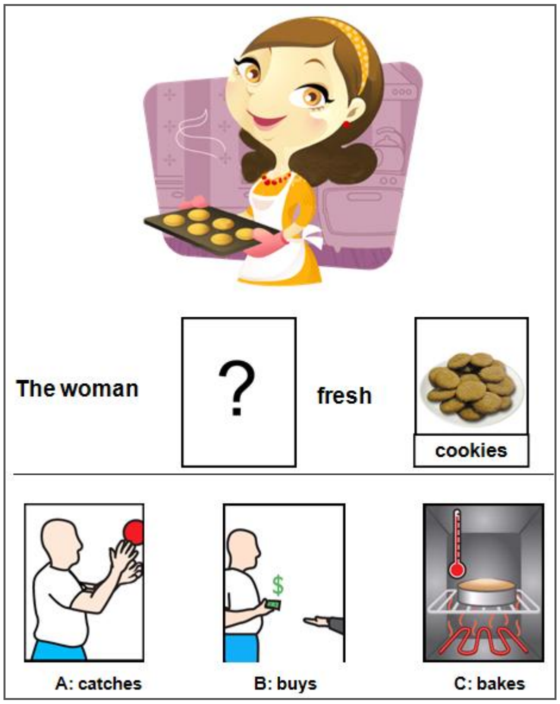

Figure 1. Example stimulus: based on the image and sentence context, the participant selects the missing word.

The resulting collection of 30 images and the corresponding sentences was evaluated by a linguist and two computer scientists, all with experience in designing assistive technology for people with aphasia. They were asked to view each image, read the sentence following it and based on the picture and the context of the sentence, guess the missing word. In cases of ambiguity, they listed alternative words that the image, accompanied by the sentence context, might evoke.

Based on this first round of evaluation, some images were replaced and the sentences were adapted so the scenarios were less ambiguous. The improved scenarios were then presented to two people with aphasia. They were first shown the image; then the first author read the sentence out loud; finally we asked them to guess the missing word. Having people produce what we had designated as the correct word turned out to be time-consuming and confusing. Even the slightest ambiguity slowed down the participants significantly and they hesitated to finalize their guess. Thus, we felt it was necessary to constrain the task further by introducing multiple choice answers.

Each scenario was made to fit on an A4 sheet of paper (see Figure 1 for an example); we showed the image first, followed by a short sentence. One of the target words in the sentence was represented by a question mark icon and the other by the corresponding Lingraphica icon. The user was given three choices. One of the choices was the correct answer, one could be used in the sentence, but not in the context implied by the image, and one choice was obviously wrong.

4.2 Pilot Testing

The final set of tasks and the interface for accessing and browsing the vocabulary (described in the following section) was also evaluated by two staff members at a local support center for people with aphasia where we planned to recruit participants. One staff member is a research speech language pathologist and the other one is a senior computer coach. They suggested an improvement on the interaction with the interface which we describe in section 5.1.4.

Two people with aphasia were recruited to pilot the study. They were given thirty minutes to find the missing words in 15 phrases with each one of the two vocabularies, ViVA and LG. They completed only five sentences for each vocabulary condition so we reduced the number of phrases to seven. Participants got easily discouraged if they could not find the target words in a few clicks so we eliminated all words that were deeper in the hierarchy (more than seven clicks away from the top level of the hierarchy). We also eliminated words that were in categories we judged less intuitive. For example, the word *broken* is reached by navigating through *dictionary* → *other* → *modifiers* → *state* → *broken*.

5. WORD-FINDING WITH ViVA AND LG

The purpose of the study was to compare word-finding in a static vocabulary hierarchy and in a vocabulary hierarchy adapted with links between words reflecting semantic relatedness. Our goal was to investigate whether people take advantage of the shortcuts possible through the provided semantic associations and how evocation guides their search. Based on ViVA's previous evaluations, we formed the following hypotheses:

H1. *The paths participants take to find words with ViVA will be significantly shorter than the paths in LG.* We expected that people will take advantage of the related words provided by ViVA which will guide them more quickly to the desired word.

H2. *Study participants will rate ViVA's organization less confusing than LG.*

H3. *Study participants will find it easier to locate words in ViVA than in LG.*

We hypothesized that the related words available in ViVA will speed up the search and also provide more satisfactory vocabulary navigation for the participants.

5.1 Methodology

5.1.1 Task

Participants had to find a number of words using two different vocabularies. They were presented with fourteen scenarios (similar to the one shown in Figure 1) comprised of an image followed by a sentence related to the image. They were then asked to identify, from a choice of three words, the word that was missing in the sentence. Once they had chosen the correct word, participants had to find in the vocabulary the two target words (*bakes* and *cookies* in Figure 1).

There were seven sentences per vocabulary condition. In the ViVA condition, 60% of the possible pairs were directly associated (the second word was displayed in the first word's related words panel). The rest of the words were associated through a common related word. The scenarios were distributed across the two conditions so that the depth of target words and the optimal path between target words in a sentence was balanced.

5.1.2 Vocabulary Conditions

The first vocabulary which we call **LG** provided a hierarchical organization. LG's vocabulary is a subset of 200 words of Lingraphica's vocabulary (Lingraphica's vocabulary consists of approximately 5,000 words). The subset was chosen such that it provided paths to all target words. The maximum depth of a path to a missing word was seven.

The interface to the vocabulary enabled the user to search for words by clicking on different categories. Figure 2 illustrates the home screen for the LG vocabulary interface. The vocabulary is accessed through the dictionary icon on the home screen. Clicking on the down arrow button leads the user to the subcategories of a specific concept which in turn takes the user deeper into the hierarchy. For example, if the user is trying to complete the phrase *bakes cookies*, to find *bakes*, she would have to traverse the portion of the hierarchy shown in Figure 3 and click on *dessert*, the parent node of *cookies*. Once the user has found *bakes*, she can click on the plus button icon and the choice will be reflected in the lower portion of the screen where the target words, indicated by a grey icon and a questions mark on top, are displayed (see Figure 4). To draw the participants' attention on the target words, the interface displayed a phrase comprised only of the words (and prepositions where applicable) extracted from the original sentence. The arrow button to the left of the phrase enabled the participant to skip phrases.

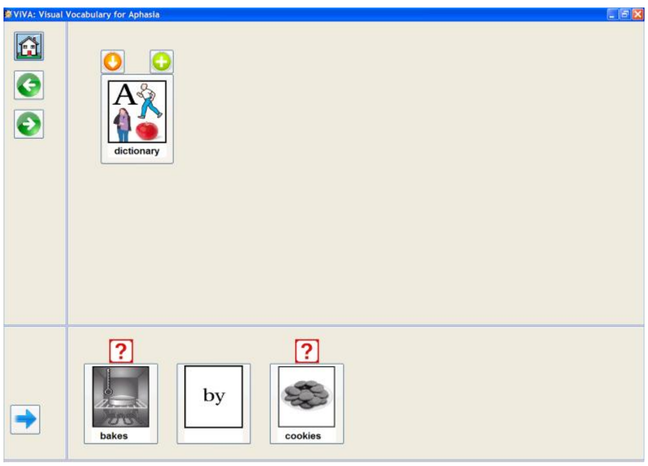

Figure 2. Home screen of the LG vocabulary interface with the phrase to be completed displayed at the bottom.

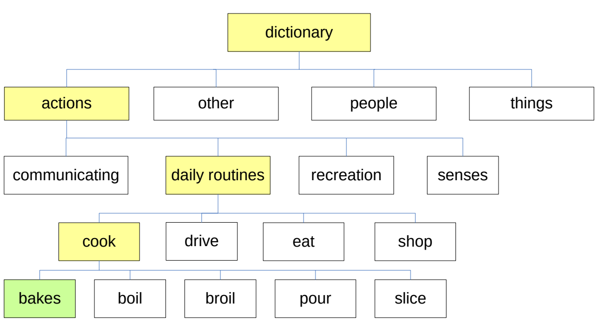

Figure 3. Hierarchical search leading to *bakes*.

ViVA, the second vocabulary, implemented the LG hierarchy but also provided links between words based on the evocation data. No additional words were added to the core 200 words. Moderate to strong evocation was considered sufficient for creating a link between two words. The displayed associations were limited to a maximum of five words.

The interface for browsing ViVA had the same layout and functionality as the LG interface with one exception. When browsing the vocabulary, words related to the concept the user clicked on were displayed in a related-words panel in the upper part of the navigation screen. Figure 5 shows what the user sees after clicking on *cook* and finding *bakes*; *kitchen* and *tasty* are related to *cook*. In addition, the target word *cookies* can be reached using the association *tasty* → *cookies*.

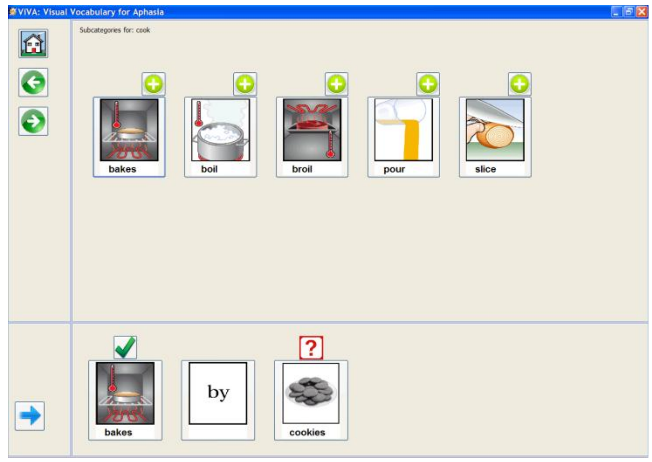

Figure 4. *Bakes* is checked as found after the user clicks on the plus button above the bakes icon.

Figure 5. ViVA provides words related to the one the user has clicked on, e.g., *food* and *kitchen* are related to *cook*.

5.1.3 Procedure

The experiment was designed to last approximately 60 minutes and the time was divided into two 30-minute slots. Participants were first introduced to the task using an aphasia friendly consent form describing the study's components. When all questions were answered and their signature was collected, they were given a brief tutorial on how to use the first interface which was randomly assigned (ViVA or LG). The scenarios were printed on A4 sheets of paper and bound in a booklet. The investigator would show the scenario drawing attention to the image first, then read out loud the sentence, omitting the missing word, and let the participant make a choice. Once the participant had chosen the correct answer, their attention was redirected to the computer screen and the vocabulary interface where they had to find both of the target words. Participants were encouraged to try the first scenario as an example before completing the remaining six. They were allowed to skip words or whole phrases if they felt like they had tried hard enough and the task was becoming frustrating. After completing the first set of phrases or after the first half hour had passed, participants were asked questions about their experience with the vocabulary they had interacted with and given five minutes rest. They were then the presented with the second interface and the differences were highlighted. After completing an example with the second vocabulary, they completed the last six scenarios. At the end of the study the investigator asked a number of background questions as well as questions related to the participant's experience with the second vocabulary. Participants were compensated $10 which they could keep or donate to the center where the study was run.

5.1.4 Apparatus

We ran the experiment on two laptop computers connected to external monitors. We omit their detailed specifications which are irrelevant to the results' analysis because we were not interested in task completion speed or accuracy in interacting with the interface (e.g., target acquisition accuracy) across subjects.

The computer coach at the center where participants were recruited suggested that to make interaction with the interface easier for all participants (especially those with motor impairments that prevented them from using a mouse), we should let the participants point at the interface. To have the participants feel in control of the interface and the task (and due to the lack of a touch screen), we simulated a touch screen experience. The investigator navigated the mouse tracking the participant's pointing and clicked when the participant applied pressure to the screen. The simulation worked well in that a few participants did not realize that the investigator was controlling the interaction. Those that eventually did, continued applying pressure to the screen since that translated into a mouse click.

5.1.5 Participants

We recruited 20 participants from a local support center for people with aphasia. Two of them helped us pilot the study and two did not finish it so their data was excluded from the analysis. We also excluded the data for two additional participants, randomly chosen, to counterbalance the order of presentation of the vocabularies. Participants met two selection criteria: 1) impaired speaking abilities and 2) good comprehension abilities. Participants were informally evaluated by staff at the center and were all medium- to high-functioning. In addition, we recorded their self-reported communication abilities. On average, participants found it easy to understand what is said in a conversation (mean of 3.8 on the scale of 1, very difficult, to 5, very easy) and somewhat difficult to express their thoughts, wishes and needs (mean of 2.7). Most participants had aphasia due to a stroke, between 3 and 11 years post onset. Two participants have had aphasia less than a year, but their performance did not differ significantly from the rest. The age range for participants was between 40 and 89 with the majority of people in the 50-59 range.

5.1.6 Qualitative and Quantitative Measures

We logged each interaction with the interface automatically – which word was selected (from the main dictionary or from the related words) and which button (down arrow, add, home, back, forward, or next phrase) was clicked on. This was done to track what paths people took to find each target word and whether they took advantage of the associations in ViVA.

During the study, the investigators also kept notes of what words were skipped or particularly difficult to find and recorded participants' comments during task completion. At the end of the experiment, participants were asked to fill out a questionnaire collecting their demographic information and feedback on their experience with the two vocabularies.

5.1.7 Design

A within-subjects design was used. Thus, each participant used both vocabularies to search for words. Order of presentation was counterbalanced and participants were randomly assigned to conditions. The independent variable was the type of vocabulary used to retrieve words and the dependent variable was length of path taken to reach the target words.

6. RESULTS

6.1 Quantitative Analysis

A 2x2 (vocabulary x presentation order) repeated measures ANOVA on path length to the target words revealed that there was no significant main or interaction effects of order of presentation of the vocabularies and thus no evidence of a learning effect. This lack of order impact contradicts the results from the study with non-aphasic participants. We suspect that this is due to the limited number of words that the aphasic participants had to search for (14 as opposed to more than 80 per vocabulary condition in the previous study). We might have observed a learning effect if our aphasic participants had more time to explore the vocabularies.

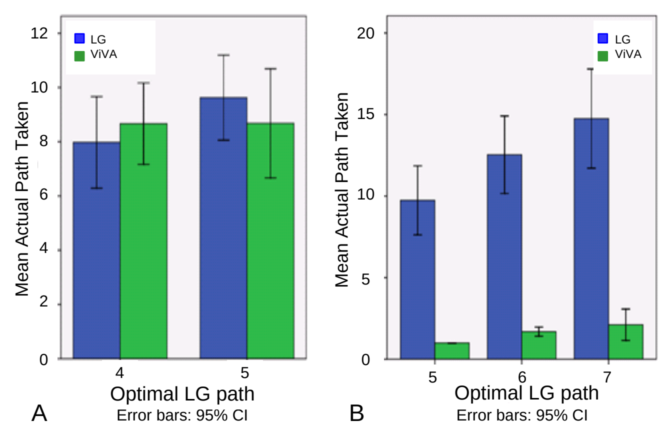

Figure 6A & 6B. Comparison of the optimal (LG) path length to a target word, measured in clicks, and the actual path length resulting from using ViVA and LG to find the word. Graph A shows the average path length to the first word participants searched for; graph B shows the average path length to the second word.

As seen in Figure 6B, participants took advantage of the provided associations and once they found one of the words in a phrase, fewer clicks were required to find the second word and complete the sentence. The average path users took to connects concepts in ViVA was significantly shorter than in LG ($F(1, 13) = 60.58$, $p < 0.01$), supporting the first hypothesis.

However, this did not hold for finding the first word in a sentence (Figure 6A), highlighting the difficulty of locating which category words belong to. Even though Lingraphica's hierarchy mostly consists of common categories such as food and clothes, some categories such as modifiers were difficult to decipher. Thus, people may have a hard time finding words that are hidden under unusual categories, for example *sleep* which is under: *dictionary* → *actions* → *daily routines* → *inhale* → *sleep*. ViVA offers the advantage of finding a word without having to know which category it belongs to. For example, many people did not know *milk* was under: *house* → *kitchen* → *refrigerator* → *dairy*, but could still find it through related words such as *coffee* and *tea*.

6.2 Subjective Feedback and Expectations

In order to gather feedback on participants' experience with the two vocabularies, we asked three five-point Likert scale questions after they had completed the tasks with each one of ViVA and LG. Friedman's test revealed ($N = 14$, $X^2(1) = 7.00$, $p = .008$), in support of our second hypotheses, that ViVA's organization is less confusing ($M = 2.7$, $SD = 0.9$) than LG's ($M = 3.4$, $SD = 0.9$). Even though both vocabularies rely on the same basic hierarchy, the fact that participants found ViVA less confusing to navigate suggests that having the associations resulted in a more satisfactory experience with the vocabulary.

Our third hypotheses was also supported (based on Friedman's test, $N = 14$, $X^2(1) = 5.33$, $p = .021$), because participants felt that it was easier to find words with ViVA ($M = 2.5$, $SD = 1.0$) than with LG ($M = 3.2$, $SD = 0.8$). Finally, participants agreed that having related words helped them find words faster ($M = 3.9$, $SD = 0.7$). While this claim is only partially supported by the data (all participants found it challenging to navigate to the first they targeted), the positive feedback suggest that based on the associations, ViVA was perceived faster in general.

Participants were not intentionally encouraged to talk out loud while completing the tasks (to avoid overwhelming them with additional tasks and because speaking is challenging for a few of them), but some gave us feedback throughout the experiment. This revealed certain expectations they had about associations between words that we had not provided (e.g., *"the* ***mom*** *could* ***make tea*** *... or the* ***dad"*** while searching for tea while in the family category). These expectations were further revealed in analyzing the paths people took to find certain target words. Some of the associations that were most frequently expected are listed in Table 1 along with a few other examples.

Table 1. Some direct associations were expected but missing in the vocabulary. For example, 5 people clicked on *leisure* expecting to find *fly* and 4 expected the association *food-milk*.

word_1	word_2	#
leisure	fly	5
food	milk	4
family	shop	4
food	shop	4

word_1	word_2
recreation	bake
family	play
backyard	baseball
time	fly

Based on user feedback and the investigators' observations during the experiment, a number of approaches in interacting with the vocabulary emerged. Some participants ***explored the vocabulary*** by clicking on icons without any evident plan of action. Others ***memorized the organization*** while browsing to help them find words faster in the subsequent tasks. A few participants were ***guided by the associations***, and they based their next move solely on what was displayed on the current screen. Some participants ***formed associations*** and expectations in advance and let their intuition about where a word should be found guide them. It would be interesting to explore these approaches and user profiles further in order to understand how to design better vocabularies and provide better user-vocabulary interaction.

7. DISCUSSION AND FUTURE WORK

The study results revealed that participants found words with ViVA significantly faster because they took advantage of the semantic associations used to augment the basic vocabulary hierarchy. However, the results also highlighted the problem of using preset categories to organize the words in the vocabulary. It was difficult to find the first word in a sentence both for ViVA and LG, because they relied on the same basic hierarchy of categories. To facilitate searching for the first word, people could create their own categories or access points to the vocabulary, but this would add additional burden on the user. When asked whether they would prefer organizing the vocabulary in categories themselves, participants' response on a five-point Likert scale was neutral ($M = 3.0$, $SD = 1.3$). Further investigation in what the higher levels of the hierarchy should be is required. One possibility of addressing this problem is to provide custom access points to the vocabulary that reflect the person's profile, but also branch out to more general words in the vocabulary.

Even though the number of missing words as well as the optimal path to find them was consistent across conditions, we thought it was unreasonable to do a pair wise comparison of the average path length participants took to complete each phrase. While the user can find *flower* and *cookies* in the same number of clicks, the path to *cookies* is easier to predict: *things* → *food* → *dessert* → *cookies*. The word *flower*, on the other hand, is reached along the path: *things* → *house* → *backyard* → *flower*. Naturally, different users expect to find different things in different categories depending on factors such as life style and personal experiences. Having the words stored in preset categories provides a stable vocabulary organization which facilitates learning over time, but it also often does not make sense to users who tend to get easily discouraged by unsuccessful attempts to find a word.

There is compelling evidence that while performing a task, people's behavior and their decision on what to do next is strongly influenced by the current context [20]. This notion known as situated action [20] has been of great importance to interface design in shaping the idea that people take advantage of contextual queues when manipulating an interface to accomplish a task. At each stage of the process, users check the environment to make a decision on what action to take next instead of forming a detailed plan in advance. Applying situated action to the problem of vocabulary organization supports our hypothesis that instead of expecting the users to memorize the vocabulary organization in order to be able to navigate it effectively, we should create the right context at each step of the navigation. The reported results suggest that creating this context by mimicking how words are organized in a speaker's mental lexicon has the potential to assist users with aphasia with word-finding. All users took advantage of the provided associations but also a few expected certain associations that were not available to lead them to the target word. Thus, user preferences and vocabulary usage patterns should also be taken into consideration in organizing the vocabulary.

The simplicity of the phrases and the limited vocabulary allowed us to design a fun, manageable task for our participants. This is also a drawback, because to really understand how people navigate the vocabulary and search for words, we need to present them with more challenging tasks. This study was only the beginning of our exploration and we plan to investigate the emerging issues using more complex designs and tasks.

8. CONCLUSION

Through a controlled study, we showed that augmenting a basic vocabulary hierarchy with semantic associations helps people with aphasia navigate to related words more effectively. Using a scenario comprised of an image and a sentence related to what was portrayed in the image, we provided study participants with contextual tasks. We asked them to find the missing words in the given sentences using two types of vocabulary organizations. One vocabulary, LG, was organized hierarchically and the other one, ViVA, augmented the basic hierarchy with human judgments of semantic relatedness. Participants found ViVA less confusing to navigate and agreed that the associations that it provided helped them find words faster.

ViVA provides richer relationships between words based on human judgments of semantic relatedness which often cross parts of speech. This makes it more likely to address successfully different user needs compared to a static hierarchical organization. Although we do not argue that our approach produces the best organization, our study suggests that it can improve vocabulary navigation and word-finding for users with lexical access impairments as those present in aphasia.

9. ACKNOWLEDGMENTS

We would first like to thank our study participants and the Adler Aphasia Center in Maywood, NJ for their help and support. The authors would also like to thank Christiane Fellbaum for her insightful feedback and Xiaojuan Ma for helping run the study. The research has been funded in part by Microsoft Research and the Kimberly and Frank H. Moss'71 Research Innovation Fund of the Princeton School of Engineering and Applied Sciences. We thank Lingraphicare Inc. for giving us access to their vocabulary hierarchy and icons.

10. REFERENCES

1. AssistiveWare. Proloquo2go. http://www.assistiveware.com/.
2. Beukelman, D.R & Mirenda, P. Augmentative and alternative communication: Management of severe communication disorders in children and adults. Brooks Publishing Company, 1998.
3. Boyd-Graber, J., Nikolova, S., Moffatt, K., Kin, K., Lee, J., Mackey, L., Tremaine, M., & Klawe, M. 2006. Participatory design with proxies: developing a desktop-PDA system to support people with aphasia. Proc. CHI '06, 151–160.
4. Boyd-Graber, J., Fellbaum, C., Osherson, D. & Schapire, R. 2006. Adding dense, weighted, connections to WordNet. In Sojka, P., Choi, K.S., Fellbaum, C., Vossen, P., eds.: Proc. Global WordNet Conference'06, 29–35.
5. Boyle, M., & Coelho, C. A. 1995. Application of semantic feature analysis as a treatment for aphasic dysnomia. American Journal of Speech Language Pathology, 4, 94–98.
6. Collins, A.M. & Loftus, E. F. 1975. A spreading-activation theory of semantic processing. Psychological Review, 82: 6, 407–428.
7. Davis, G. A. 2000. Aphasiology: Disorders and clinical practice. Boston: Allyn & Bacon.
8. Edmonds, L. A., Nadeau, S. E. & Kiran, S. 2009. Effect of Verb Network Strengthening Treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia. Aphasiology, 23: 3, 402–424.
9. Fellbaum, C. 1998. WordNet: An Electronic Lexical Database, chapter: A semantic network of English verbs. MIT Press, Cambridge, MA.
10. Goodglass H, Wingfield A. 1997. Anomia: Neuroanatomical and cognitive correlates. Academic Press, San Diego, CA.
11. Google Inc. Google Images. http://images.google.com/
12. Kauhanen, M.L., Korpelainen, J. T., Hiltunen, P., Määttä, R., Mononen, H., Brusin, E., et al. 2000. Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke. Cerebrovascular Diseases 10, 6, 455–461.
13. Lingraphicare Inc. Lingraphica. http://www.aphasia.com/.
14. Miller, G. A. 1990. Nouns in WordNet: A lexical inheritance system. Int. Journal of Lexicography, 3: 4, 245–264.
15. Moss, H & Older, L. 1996. Birkbeck Word Association Norms. Psychology Press.
16. National Aphasia Association. Aphasia FAQ. http://www.aphasia.org. Available Online February 2010.
17. Personal communication, Jan. 20, 2010.
18. Nikolova, S., Boyd-Graber, J., Fellbaum, C. & Cook, P. 2009. Better vocabularies for assistive communication aids: connecting terms using semantic networks and untrained annotators. Proc. ASSETS'09, 171-178.
19. Nikolova, S., Ma, X., Tremaine, M. & Cook, P. 2010. Vocabulary navigation made easier. Proc. IUI'10, 361-364.
20. Suchman, L.A. 1987. Plans and Situated Actions: The Problem of Human-Machine Communication. Cambridge Press, Cambridge, MA.
21. Swinney, D. Lexical access during sentence comprehension: (Re)consideration of context effects. Journal of Verbal Learning and Verbal Behavior, (18):645{659, 1979.
22. The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/
23. TouchSpeak & TypeSpeak. http://www.touchspeak.co.uk/
24. Webster J., Morris J. & Franklin S. 2005. Effects of therapy targeted at verb retrieval and the realisation of the predicate argument structure: A case study. Aphasiology, 19: 8, 748–64.