Independent Word Discovery for People with Aphasia

Mmachi G. Obiorah[1], Anne Marie Piper[2], Michael Horn[1]
[1]TIDAL Lab, [2]Inclusive Technology Lab
Northwestern University, Evanston, IL
obiorahm@u.northwestern.edu, {ampiper, michael-horn}@northwestern.edu

ABSTRACT
Augmentative and Alternative Communication (AAC) devices help people with speech-language impairments due to aphasia express themselves more independently and fluently. However, current AACs are still limited in that they must be pre-programmed with common words or phrases prior to use. When people with aphasia encounter unexpected and unplanned situations or contexts, they may have difficulty using their AAC device to generate the desired speech. We present the design of an AAC application that uses image recognition technology to give people with aphasia just in time access to words. Our design automatically carries out an extensive retrieval of words associated with a captured image. The relevance, accuracy and usefulness of each word can then be verified independently by the user through an intensive retrieval of images associated with each word.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

1. INTRODUCTION
Word finding difficulty is a symptom of aphasia in which an individual is unable to generate spoken language associated with a an idea, object, or action. AAC devices (e.g., Lingraphica[1], Dynavox[2]) attempt to address word finding difficulties by providing a bank of words, often text coupled with images, and organizing those words in a recognizable order. The person with aphasia is then trained to use the device to retrieve words when he/she needs them. Many AAC devices allow the user to expand and customize the device's vocabulary, although this tailoring takes time and speech-language pathologists often perform device customization. AAC devices rely heavily on the presence of other people to explore new words and extend or reorganize the device's vocabulary. In addition, AACs are usually programmed prior to use and therefore may be less useful in novel situations.

Generally, people with aphasia are able to recognize symbols and images in spite of having either expressive or receptive speech-language impairments. This ability to recognize images and symbols is often applied in the design of AAC devices to support the user in sentence composition using captioned icons. Yet, a significant limitation is that both the text and the image sets are static and do not adapt to new stimuli in the user's environment. In our work, we explore a design that supports users of AACs in independently associating texts, images, and sounds in novel situations. We hypothesize that our design will empower users by reducing their reliance on external parties for novel word captioning and support users in exploring their environment.

2. RELATED WORK
While there are many different forms of aphasia and each individual presents slightly differently, word finding is a common challenge across this population. As a result, both commercial products and academic researchers have invested considerable effort in addressing word finding challenges through the use of AACs. Specifically, the ability for an individual to use an AAC effectively depends on their ability to search for, identify, and select the desired word or phrase in time for communication. To help achieve just-in-time speech through an AAC, many existing systems (e.g.Proloquo2go[3], Lingraphica[1]) organize words into hierarchical representations, and words are accessed by navigating through the hierarchy. Hierarchical representations, though useful, can become too deep, complex, and confusing for many users [2].

As an alternative approach, Nikolova et al [2] explored the use of semantic relatedness, or the frequency of occurrence of words together and how much one word brings another to mind. They found that structuring available word choices based on semantic relatedness out performed a hierarchical organization for sentence composition except when finding the initial word in the sentence. Similarly, Epps et al. [1] used location tracking and specialized algorithms to retrieve words from the Internet and supply them to the AAC. Their method complemented the base vocabulary of the AAC and was adaptive based on the location of the user.

Our work, like Epps et al [1], addresses just-in-time word retrieval. However, we do this through image recognition, which supports extensive retrieval of words and allows the user to independently determine the accuracy or relevance of the retrieved words by associating each word with descriptive images.

CONF '22, Jan 1 - Dec 31 2022, Authorberg.

unique doi string will go here

[1] https://www.aphasia.com/
[2] https://www.tobiidynavox.com/
[3] http://www.assistiveware.com/product/proloquo2go

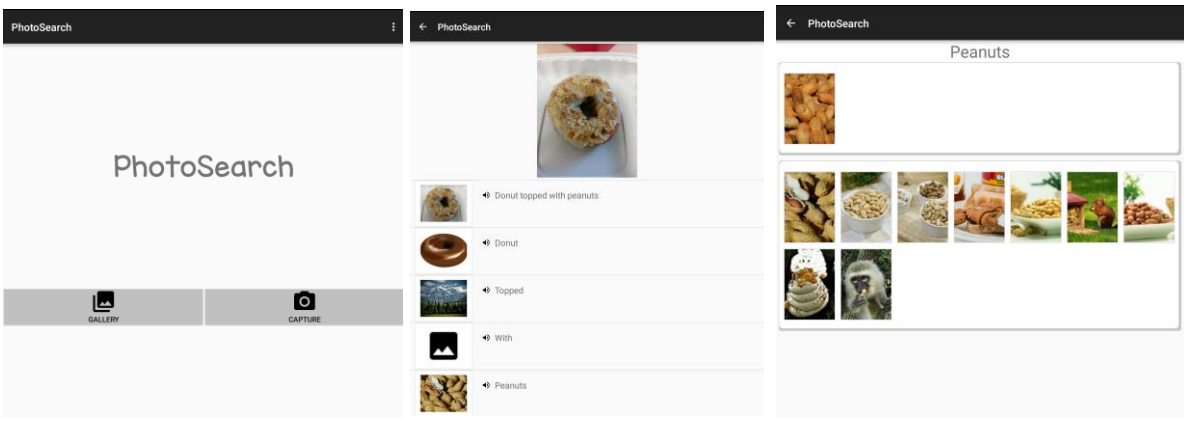

Figure 1: *Left:* The user can either select an existing image from the gallery or take a new photo. *Center:* The captured photo is captioned. In this case, it is captioned as "Donut topped with peanuts". Each word is listed below and associated with a descriptive image. *Right:* When the word "Peanut" is clicked more images that describe the word "Peanut" are retrieved. The images are ordered according to a prediction confidence.

3. DESIGN OF PHOTOSEARCH

To better understand the process of speech-language therapy and word finding challenges, we have observed seven adults with aphasia during a total of 43 hours of therapy. Additionally, we collaborated with four clinicians. Our observations and discussions with clinical experts focused on becoming familiar with aphasia and the process of therapy; understanding how current AACs are used within therapy and how they can empower/disempower people with aphasia; and investigating how AAC's can be better designed to support people with aphasia. We took field notes from our observations and discussions with experts. Preliminary analysis of these data led to the design of a new approach to a word finding application.

4. SYSTEM DESIGN

We created an Android application called PhotoSearch, which captions images automatically. Image captioning for PhotoSearch is powered by the CloudSight[4] API. Users take a photo of their environment and the system returns text-based tags associated with that image. The tags include nouns associated with objects in the image as well as adjectives, such as the color of the object and type of material of the object (e.g., wood, paper). As a result, a single image capture serves as an access point to a variety of words (see Figure 1). Since people who have aphasia vary in their ability to access written and spoken language but may still have the ability to recognize images, the application presents each term to the user along with a separate image for each term (retrieved from the Openclipart[5] project and Pixabay[6]). The application then provides text-to-speech functionality for each word-image description. The descriptive image helps the user to distinguish and identify the meaning of each retrieved word while also serving as a link to more images associated with the word. If the user clicks the descriptive image, the application displays a page with even more descriptive images associated with this initial word (as in Figure 1 where 'Peanut' was clicked). This is intended to help the user to independently establish a level of confidence in the accuracy of the retrieved word as well as iteratively and independently find the desired word. Images are ordered on the display screen according to a prediction algorithm confidence that is based on content based image retrieval.

5. CONCLUSION AND FUTURE WORK

Our ongoing work involves working with speech-language pathologists to use this application within a clinical setting. We are first testing our application with several speech-language pathologists to further inform the design, understand potential usability problems, and identify use cases. In addition, feedback from initial tests will inform our design iterations and our final design, which we plan to test with people with aphasia.

Even though our application tries to help the user establish confidence in the association of a word with an image, word sense disambiguation is a difficult problem to solve. We intend to further explore establishing confidence in the accuracy of the retrieved word by including a social component to the application where a structured request for help from a caregiver or family members can be made by the user of the application.

The loss of independence in communication is one of the most difficult aspects of aphasia. PhotoSearch applies image recognition and captioning to better support people with aphasia in discovering new words independently in diverse contexts.

6. ACKNOWLEDGMENTS

We thank NUCASLL for their support with data collection and for their advice and feedback throughout this study.

7. REFERENCES

[1] C. Demmans Epps, J. Djordjevic, S. Wu, K. Moffat, and R. M. Baecker. Towards providing just-in-time vocabulary support for assistive and augmentative communication. *In Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*, IUI '12, pages 33-36. ACM, 2012.

[2] S.Nikolova, M. Tremaine, and P. R. Cook. Click on bake to get cookies: Guiding word-finding with semantic associations. *In Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '10, pages 155-162. ACM, 2010.

[4] https://cloudsight.ai/
[5] https://openclipart.org/
[6] https://pixabay.com/