A Visual Recipe Book for Persons with Language Impairments

Kimberly Tee[†]**, Karyn Moffatt**[†]**, Leah Findlater**[†]**, Eve MacGregor**[‡]**,**
Joanna McGrenere[†]**, Barbara Purves**[±]**, Sidney S. Fels**[‡]

Department of Computer Science[†]

Department of Electrical and Computer Engineering[‡]

School of Audiology and Speech Sciences[±]

University of British Columbia, Vancouver, BC, Canada

aphasia-ubc@cs.ubc.ca

Abstract

Cooking is a daily activity for many people. However, traditional text recipes are often prohibitively difficult to follow for people with language disorders, such as aphasia. We have developed a multi-modal application that leverages the retained ability of aphasic individuals to recognize image-based representations of objects, providing a presentation format that can be more easily followed than a traditional text recipe. Through a systematic approach to developing a visual language for cooking, and the subsequent case study evaluation of a prototype developed according to this language, we show that a combination of visual instructions and navigational structure can help individuals with relatively large language deficits to cook more independently.

Categories & Subjects Descriptors: K.4.2 Computers and Society: Social Issues—Assistive technologies for persons with disabilities.

Keywords: Assistive technology, aphasia, multi-modal interfaces, heuristics.

INTRODUCTION

Whether it is preparing a gourmet meal, or simply reheating pre-packaged soup on the stove-top, cooking is a daily activity for many people. Meal preparation is not only a key element of independent living; it can also be an important feature of social identity. However, for the 1.1 million North Americans with aphasia [1], an acquired language disorder with relative sparing of other cognitive abilities, text recipes are often prohibitively difficult. Aphasia is usually acquired as a result of stroke, brain tumor, or other brain injury, and results in impairment of the production and comprehension of speech and written language.

Although language rehabilitation can be beneficial for individuals with aphasia, a significant number are left with

a life-long chronic disability that influences a wide range of activities and prevents full re-engagement in life. The long-term impact of aphasia varies across individuals. However, given the importance of communication in virtually all aspects of everyday life, it is not surprising that most, if not all, individuals experience changes in how they participate in everyday activities, with social isolation and depression relatively common.

One challenge in designing technology for people with aphasia is that individual language abilities can vary greatly due to relative impairments across the four language modalities (speaking, reading, writing, and auditory comprehension) [7]. The Visually Enhanced Recipe Application (VERA) supports these varying abilities through full-text and sound options that allow the primarily visual cooking instructions to be customized to support each individual's strengths. To translate from an original text recipe to its highly visual VERA counterpart we systematically apply a semantic model of cooking instructions and a set of heuristics related to the visual display of such instructions.

This work is part of a larger-scale research project, the Aphasia Project, which is exploring how technology can be used to support people with aphasia in their daily lives. When the Aphasia Project was initially conceived, we conducted informal interviews with aphasic individuals to identify possible ways technology could enhance their independence and social re-engagement. The ability to use a recipe independently was one application area identified (see [12] for more detail).

This paper documents the development of VERA. First, we designed a visual language for communicating cooking instructions to systematically map the instructions from text to a primarily visual representation. We then developed a prototype based on this visual language, with input from eleven non-aphasic participants. Finally, we performed a study with four aphasic participants to assess the effectiveness of the resulting design for individuals of varying language abilities. This evaluation suggested that the combination of visual instructions and navigational structure imposed by VERA helped those with relatively large language deficits to cook more independently.

RELATED WORK

The majority of research on aphasia and technology has focused on the design of devices to assist aphasic individuals in communicating. These augmentative and alternative communication (AAC) devices most often take the form of symbol-based dictionaries, leveraging the retained ability of aphasic individuals to recognize image-based representations of objects. The success of such technology has been limited, perhaps because of its emphasis on individuals with severe aphasia, rather than on those with some retained communicative abilities [8]. Other AAC devices, such as TalksBac [20], target the ability of some higher-functioning aphasic individuals to recognize familiar words and short sentences. In a longitudinal study, TalksBac was shown to be helpful for some individuals, although its reliance on caregivers to maintain the system was problematic.

More relevant to the design of VERA is technology to support aphasic individuals in higher-level daily activities, though much less work has been done in this area. One example is an e-mail system designed to assist cognitively disabled people, including people with aphasia, in composing e-mail messages [17]. The system provides different interfaces with varying levels of support and complexity; though no interface was found overall to be superior, differences in individual preference suggest the need for interface customization for cognitively disabled users. A second example, ESI Planner, incorporates triplets of images, sound, and text to support people with aphasia to independently manage appointments on a PDA [13]. Testing with nine aphasic individuals in an exploratory evaluation showed that participants could complete more tasks correctly using ESI Planner than with an equivalent text-only planner. This suggests that the image and sound support assisted aphasic individuals in comprehending the information presented, a finding we have incorporated into VERA. In a final example, an ethnographic study investigating how a native PDA could be incorporated in one aphasic individual's communication strategies led to the creation of a new file facility application [4].

Limited research has been done to support recipe preparation for aphasic individuals. The C-VIC system [16] uses a subject-verb-object syntax to combine graphic building blocks (pictures and animations) into cooking instructions. For example, to set the oven to 350°F, an image of a chef, an animation of a hand turning an oven knob, and the number *350°* are shown horizontally on the screen. Each step in the recipe is represented on a separate page, and text descriptions of each graphic can be accessed. An evaluation of the system with one aphasic individual used recipes for commercial box mixes in which many ingredients had already been pre-processed and mixed. Though the participant was generally able to prepare the recipes independently, the researchers attributed failures to a lack of global overview and the appearance of the instructions to *suggest* rather than *command*. In contrast to the C-VIC system, VERA uses a heuristics-based approach to displaying visual cooking instructions, allowing for the presentation of more complex recipes than C-VIC. We do not represent a subject in our visual instructions, so these instructions will more likely be interpreted as commands. VERA also provides high-quality color images, optional audio and text support, and overview support.

A few research projects have looked at electronic cookbooks in varying contexts without a specific focus on language-impaired populations [2, 3, 9]. CounterActive, for example, projects a display onto the countertop, and uses a multimedia environment (instructions, pictures, movies, and music) to teach people how to cook [9]. The user interacts directly with the countertop, rather than using a mouse or keyboard. The interactive immersive experience it provides is designed for enhancing engagement, rather than facilitating comprehension. It also requires considerably more infrastructure than VERA.

A number of cookbooks developed for beginners, children, or mentally disabled persons use a more visual presentation format than traditional cookbooks. We found many of these useful in developing our initial set of heuristics for visual instructions [e.g., 5, 15, 19]. Some of these books use images to supplement textual instructions, while others present primarily visual instructions, with minimal supporting text. One cookbook, developed specifically for children with language impairment, uses a combination of images and symbols, with text used only for labeling measurements and ingredients [19]. For example, arrows show direction and action, and the addition symbol represents "and".

VISUAL LANGUAGE DEVELOPMENT

Since individuals with aphasia tend to retain their ability to recognize images [18], we hypothesized that a multi-modal format would be easier to understand than a traditional text recipe. As a result, we focused on using a combination of images, symbols, animations, and keywords to visually represent the steps in a recipe, while providing redundant audio and text instructions for each step. The first stage in developing VERA was to translate from a traditional text recipe to a primarily visual, yet multi-modal presentation format. We did this in two steps: (1) creation of a semantic model of cooking instructions to extract the underlying structure and meaning of recipes; (2) development of a set of heuristics for producing visual language phrases based on the semantic model.

Semantic Model

The semantic model of cooking instructions, shown in Figure 1, is based on a survey of a range of recipe types from several cookbooks. The model represents a single cooking instruction but could be extended to model an entire recipe. A semantic breakdown is given for each of the five basic components commonly found in a cooking instruction: measurement, tool, action, ingredient, and
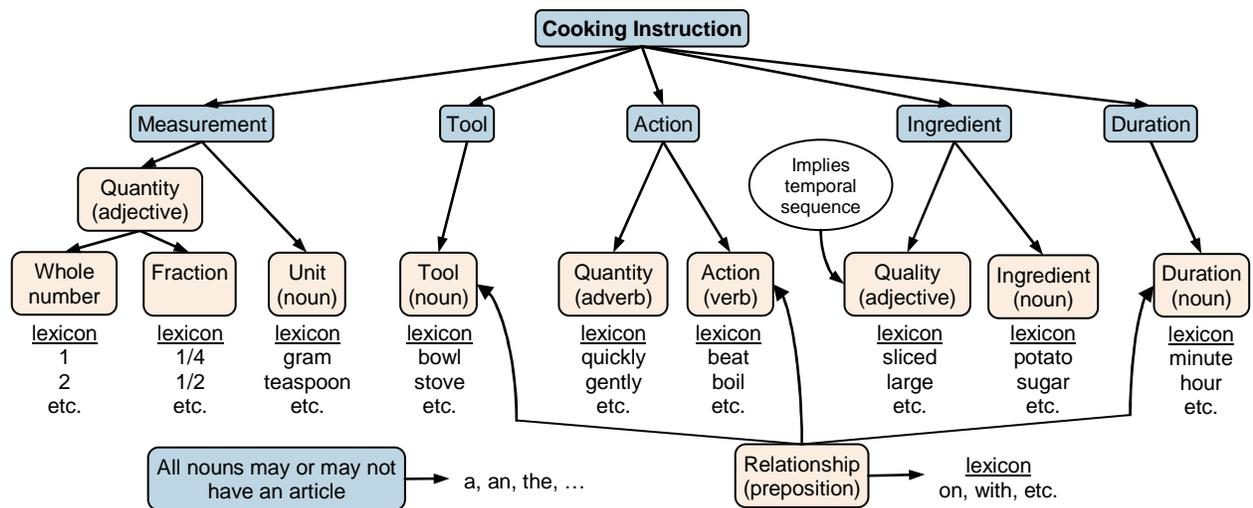
Figure 1. Semantic model for cooking instructions.

duration. The items listed in each component's lexicon could be further organized semantically as a taxonomy; for example, the actions *beat* and *stir* could be subclasses of the action *mix*. Further suggestions of cooking taxonomies can be found in [6] and [11].

The semantic model presented here is flexible enough that it can be applied both to narrative-style text recipes, and to more succinct, imperative-style recipes. We expect it will also prove useful in our future work to develop a compiler that can semi-automatically translate textual cooking instructions into visual cooking instructions.

Heuristics

A set of heuristics forms the basis of our visual language and dictates the layout and content of visual elements that represent components of the semantic model. We used feedback from an informal focus group with four domain experts (two human-computer interaction experts, a psychologist, and a speech-language pathologist) to decide how to visually represent elements of the semantic model and to draft an initial set of heuristics. The set of heuristics was then further refined through informal evaluations with non-aphasic individuals.

The major components of the semantic model are represented as follows, with examples shown in Figure 2:

- *Measurements*: Image of a standard set of measuring tools, with the desired measure highlighted by a star.
- *Tools* and *ingredients*: Image (either photograph or drawing) of the tool or ingredient.
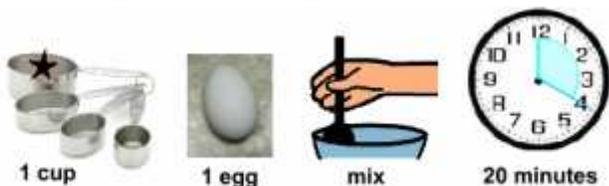

Figure 2: Examples of visual representations for measurement (1 cup), ingredient (an egg), action (mix), & duration (20 min).

- *Actions*: Image or animation to demonstrate the action, or a symbolic representation of the action.
- *Duration*: Animation of a clock face whose hands sweep out the amount of time that needs to pass.

In general, we represent instructions by showing state transitions; the syntax used is *[start-state]—action—end-state*. The start-state may be omitted if the end-state of the last instruction is similar enough that it can be used instead without causing ambiguity or loss of clarity. For example, the instruction *mix the ingredients with a spoon*, can be shown using an implicit start state as: an animation of a hand mixing, an image of a bowl with mixed ingredients, and a transition arrow pointing from the animation to the bowl. The *add* instruction is an exception: a curved arrow is used to denote the addition, with no state transition arrow. Images of these and other instructions are shown with the description of the final prototype—see Figure 3 (next page).

Many reasonable alternatives sometimes arose in developing the heuristics. In these cases, we compared the alternatives informally with seven non-aphasic participants. The participants ranged in age from 16 to 23, and most were undergraduate students. All had previously cooked, both with and without recipes, although three cooked less than once a week. Each alternative visual representation was presented individually, and participants used a Likert scale to rate how effectively a representation conveyed the meaning of a paired textual instruction. In addition, participants specified their preference among the different alternatives. There was little discrepancy between the most effective and most preferred representations. Though the background of the participants limits the results of this evaluation, it provided a basis with which to refine the set of heuristics. Further refinement may be required once an evaluation is performed with a more diverse group of participants. The set of heuristics is as follows:

H1. Measurements and ingredients: Show an ingredient requiring a volume measurement (e.g., 1/2 cup butter) in a separate image from its measurement for clarity, but

grouped together by a box drawn around both. Show an ingredient requiring a count measurement (e.g., two apples) in the quantity desired. To aid in identifying ingredients shown in a can or bottle, display an image of the contents grouped with the image of the can or bottle. Branding may be employed to facilitate identification, but be aware that this may need to change for different geographical locations.

H2. Actions on multiple ingredients: To represent performing an action on multiple ingredients, use a '+' symbol between each ingredient/measurement pair to clearly indicate that the action is to be performed on all the ingredients, not just some.

H3. Visibility of state: The image of any container such as a pan, pot, or bowl should reflect the current state of its contents.

H4. Transitions in state: Display transitions in state in a left-to-right or top-to-bottom fashion.

H5. Increasing/decreasing temperature: Use an image of a knob with a static arrow pointing from the current temperature to the desired temperature to indicate the desired setting.

H6. Speed: Group the action with a scale indicating the speed at which the action should be performed. Though culturally sensitive, the endpoints of the scale could be a hare, representing *fast*, and a tortoise, representing *slow*.

H7. Duration: When an action needs to be performed for a particular duration, group the visual representation of the action in a box with an animated clock displaying that duration.

VERA PROTOTYPE

A prototype of VERA, shown in Figure 3, was implemented in Macromedia Flash MX based on the heuristics described above. In implementing the prototype, we used the

following key aphasia-friendly design principles (adapted from [14]): (1) use a lot of white space and keep text and image sequences simple and concise; (2) use short phrases and sentences, avoiding long words; (3) use a large font size (14 to 18 points); and (4) use sound clips that read the text aloud.

Initial Prototype

Two recipes were implemented: a spaghetti sauce recipe, and a chocolate chip cookie recipe. The initial prototype did not include keywords or text support. The recipes consisted of a series of screens, each of which displayed at most two steps. Navigational arrows allowed users to step forward or backward between screens, and sound clips accompanied each screen. Animations played when a screen first loaded, and could be replayed by clicking on the animation with the mouse. Full text instructions could be shown or hidden by toggling a checkbox in the bottom-left corner of each screen. Finally, audio instructions for each step could be played by clicking on the corresponding speaker icon.

We chose the spaghetti recipe and the cookie recipe to test the effectiveness of our heuristics for different types of recipes. Though both recipes called for the same number of ingredients, the spaghetti sauce was a stove-top recipe requiring many count measurements, while the cookies required baking and used more precise volume measurements. For example, the spaghetti recipe required setting inexact temperatures (e.g., medium heat), whereas the cookie recipe required setting exact temperatures (e.g., 375°F). Each recipe was implemented with 13 major steps.

Informal Evaluations

We piloted our design with non-aphasic subjects to discover and repair major design flaws before we did our experiments with aphasic subjects.
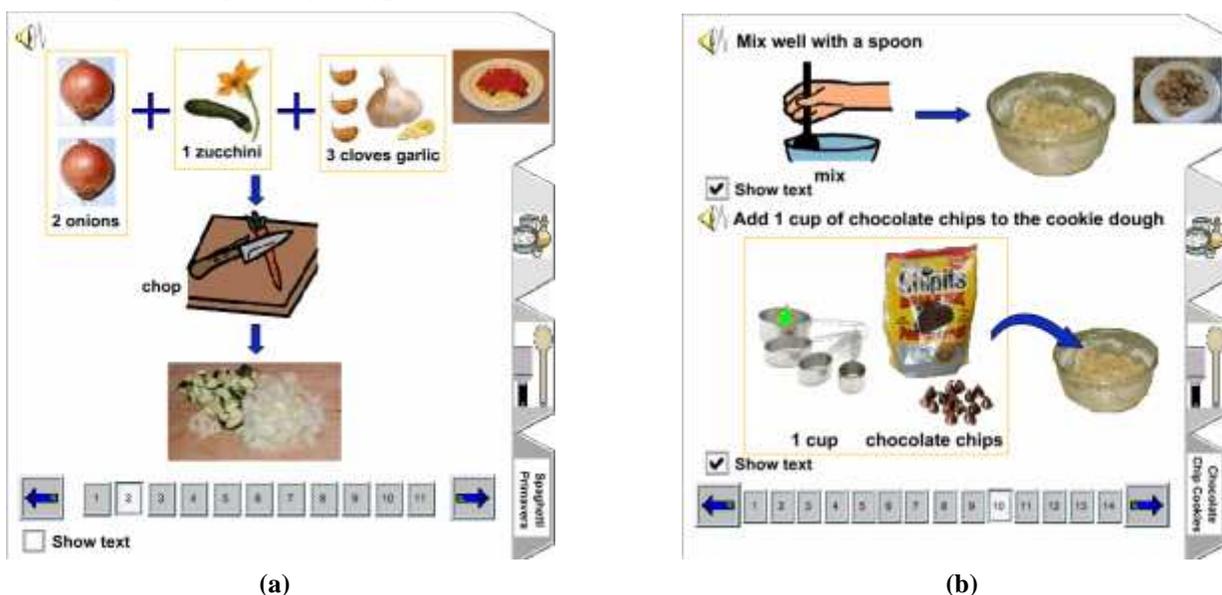

(a)   (b)

Figure 3. Screenshots of the VERA prototype: (a) spaghetti sauce recipe, and (b) cookie recipe displaying full text.

The first round of evaluation, with two undergraduate students experienced in using recipes, focused on how effectively the visual cooking instructions represented their textual counterparts. Participants were asked to interpret the visual cooking instructions, and to rate how confident they were in their interpretation. As we were specifically interested in the effectiveness of the heuristics, participants were not allowed to use the sound functionality. Both the spoon and cup measurements, and the temperature and speed representations were problematic, so we reviewed heuristics *H5* and *H6*; the refined versions of these are presented on the previous page. Overview information was added to the prototype by including tabbed screens to show ingredients, tools, a text-only version of the recipe, and the multimodal recipe itself (see tabs in Figure 3). In addition, two new heuristics were added to accommodate support for customization and navigation:

H8.   Customization   of   modes: In a multi-modal presentation format, allow users to customize the modes displayed.

H9. Navigational context: Use a progress indicator to display the current step with respect to the whole recipe.

In our prototype, the progress bar also allows direct access to other screens, giving users navigational freedom.

A third new heuristic addresses ambiguity concerns by adding keyword support. Many aphasic individuals are capable of reading single words and short phrases [7], so we chose to add minimal keyword support to help remove any remaining uncertainty for those who could use it, without causing a significant visual distraction for those who could not.

H10. Keyword support: Include keywords with each measurement, ingredient and action element to further clarify instructions.

Two additional undergraduate students, who cooked more than twice a week, informally evaluated these changes. This was similar to the first evaluation, but included tasks based on the new overview tabs, and assessed the correspondence between the audio and visual instructions. Minor design flaws found during this evaluation led us to change the images on the overview tabs to exactly mirror those in the recipe screens in terms of size and the inclusion of keywords, which caused some overview sections to be split into multiple screens. No other problems were identified in the recipe portion of the prototype. Neither user reported any difficulties with the audio instructions, which were generally found to be consistent with the visual instructions.

Final Prototype
The final prototype, shown in Figure 3, incorporates all of the changes described above. At this point, the option to have the full text instructions appear above the visual counterpart and beside the sound clip icon was added. Figure 3(a) shows a screenshot of the spaghetti recipe, with the recipe tab open to an instruction to chop some

vegetables and add them to the pan; Figure 3(b) shows the cookie recipe with full text displayed open to a screen with two instructions: a mix instruction, and an add instruction.

EVALUATION METHODOLOGY
We conducted an evaluation to compare how independently participants could cook with the VERA prototype versus a traditional text recipe. In evaluating the final prototype, we also implicitly evaluated the semantic model and heuristics. Although the evaluation followed a structured experimental design so that we could have identified statistically significant trends had they arisen, we were motivated by the qualitative observations case study analyses could provide.

Conditions
The two factors included in the evaluation were *type of recipe* and *presentation format.* Type of recipe was either cookies or spaghetti sauce, the two recipes we had implemented in the design phase of this research. Presentation format was either the VERA prototype or a traditional text-only recipe. To create text-only recipes we used the full-text instructions from the prototype, which had already been adapted to be aphasia-friendly. To minimize performance differences due to readability of the text-only recipe, we reformatted it on a word processor, which is something a caregiver could easily have done. Ingredients and utensils were listed on one page, while instructions were given on a second page. The font size was increased to 14 point, and the instructions were double-spaced.

Participants
Four aphasic individuals, P1 to P4 (one female), participated in the study. The participants ranged in age from 29 to 73. The limited size and variability of our subject pool prevented us from selecting people with equivalent cooking experience. P1 and P2, by self-report, had limited cooking experience and no particular interest in cooking, but were interested in participating in the study for the social interaction. In contrast, P3 and P4 were specifically interested in cooking as they had cooked extensively in the past. All participants were selected to be in a stable condition, at least one year post onset. Although computer experience was not required, all had used computers before acquiring aphasia. P1, P2, and P4 had physical   disabilities   which   prevented   them   from autonomously completing some of the recipe steps.

A language assessment for each participant was done using the Western Aphasia Battery [10], a standardized battery widely used to assess language impairments in aphasic individuals. Due to an unexpected change in P3's health after the cooking sessions and before the scheduled language assessment, we were unable to perform a standardized assessment with him; thus, comments on his language ability come from qualitative observations made by the researchers. The severity of the other three participants' impairments in terms of reading, auditory

Table 1. Language assessment results for 3 participants; ratings refer to the degree of impairment.

Participant	P1	P2	P4
Speech	moderate	moderate	moderate
Auditory comprehension	moderate	mild	moderate
Reading	moderate	moderate	moderate
Writing	severe	moderate	severe

comprehension, speech, and writing abilities based on the assessment results are shown in Table 1.

Experimental Design

Given the small number and variability of our participants, we chose a within-subjects design. Each participant (except P3, see above) completed three sessions, none of which lasted more than 90 minutes. The first two sessions were cooking sessions, where participants cooked with one of the recipes paired with a presentation format. The order of presentation for type of recipe and presentation format was counterbalanced. Video recordings were made during both cooking sessions to capture the participants' interactions with the system and the researcher. Two researchers were present for each cooking session: one to conduct the session and one to observe. The third session consisted of a language assessment conducted by a certified speech-language pathologist. The three sessions took place on different days.

Apparatus and Procedure

The cooking sessions were conducted in the participant's home or in the home of the researcher. Although for consistency we would have preferred to use a single location, we needed to be sensitive to the needs of our participants. Some of our participants were not willing to have researchers in their home (P2, P3), while others, for reasons of mobility and comfort, required that we come to them (P1, P4). There were strengths and weakness to each of the locations. In the participant's home, the participant gained the advantage of using a familiar kitchen and the comfort of familiar surroundings; however, it was much harder to minimize distractions and interruptions. Neither of the participants who used their own kitchen lived alone, so family members sometimes interrupted the experiment.

All ingredients and cooking tools required for the recipe were provided and laid out for the participant before the session began. Some additional ingredients not used in the recipe were included as decoys. At the beginning of a cooking session with the prototype, the researcher explained and demonstrated how the program worked. The prototype ran in Internet Explorer 6.0 on a Tablet PC during the evaluation, and could be placed on a table or counter-top. This procedure was piloted with one non-aphasic 72-year-old female before finalizing the details of the study. In the study, participants were given time to explore the program until they felt ready to begin. This training period took no longer than six minutes for any of the participants.

To ensure that each participant had a positive experience, we measured performance based on the number of interventions required to complete the recipe, rather than allowing the participant to make errors and potentially spoil the outcome of the recipe. The researcher provided assistance when asked, offered assistance when the participant seemed lost, and intervened when the participant was about to make an irrecoverable error.

Video Analysis

To identify interventions and record qualitative observations, we analyzed the videotaped sessions. Interventions were classed as high-level interventions (HIs) or low-level interventions (LIs). The following were classed as HIs because without the researcher's help, the recipe would not likely have been completed successfully:

- The participant was unable to identify the correct ingredient, measurement, tool, or action; the researcher showed or demonstrated it to the participant.

- The participant attempted to make an incorrect measurement, action, or ingredient choice; the researcher questioned and/or corrected the participant.

- The participant skipped a step; the researcher prompted him or her to go back to the previous step.

The following interventions were considered LIs, because either they were not critical to the outcome of the recipe, or given enough time, the participant would have most likely succeeded with the step:

- The participant performed an action for too long; the researcher then prompted the participant to stop.

- The participant struggled with reading an instruction or word. For the text-only recipe, the researcher folded the paper to make only the current instruction visible, or read it aloud. For the prototype, the researcher showed how to play the sound clip or view the full text.

- The participant seemed confused or unsure of what to do; the researcher then prompted the participant to check the previous, current, or next instruction.

Physical assistance was noted, but was not counted as an intervention. The focus of this study was to evaluate the visual cooking language and prototype design, not each participant's physical capabilities. Food processors, electric can openers, and other gadgets could provide physical assistance to those who need it.

At the end of each cooking session, the researcher conducted a short, semi-structured interview to obtain information about the participant's perception of how well the session went. The interview at the end of the first session also included questions about the participant's background, and the interview at the end of the second session included questions about the participant's preferences on the two presentation formats.

Dependent Measures

The quantitative measures of performance were:

- *Steps requiring LIs only*: number of steps that required at least one low-level intervention, but did not require any high-level interventions.

- *Steps requiring HIs*: number of steps that required at least one high-level intervention, and 0 or more low-level interventions.

The qualitative measures self-reported in the semi-structured interviews were:

- *Ease of use*: the format (prototype vs. text-only) that was easiest to follow.

- *Comfort level*: the format that was most comfortable to use.

- *Preference*: the format preferred overall.

- *Future use*: the format that would be preferred if the participant were to cook using recipes in the future.

RESULTS

Given the large individual differences inherent in this population of users and the small number of participants used in this study, we did not expect to identify strong trends in behavior, which is reflected in the results. Instead, we present general findings followed by the analysis as four individual case studies.

General Findings

The overall intervention counts, shown in Table 2, do not show a clear trend. P1 and P4 required more HIs for the text recipe, P3 required the same number of HIs for each condition, and P2 required more HIs when he used VERA.

For further analysis, each step in the recipe was classified according to the semantic model, and each element was considered as a possible cause for an HI. For example, an HI occurring during the step *add one tablespoon of sugar to bowl* could be caused by one of the following elements: ingredient (e.g., salt instead of sugar), measurement (e.g., 1 tsp instead of 1 tbsp), or action/tool (e.g., adding sugar to the wrong bowl). Note that we have classified HIs according to four categories, while there were five primary

Table 2. Number of steps with interventions: (LI Only) steps with only low-level interventions; and (HI) steps with at least one high-level intervention. Each recipe had 13 steps in total.

ID	Session	Recipe	Platform	LI Only	HI
P1	1	spaghetti	text	3	6
	2	cookie	prototype	2	2
P2	1	cookie	prototype	2	5
	2	spaghetti	text	0	0
P3	1	spaghetti	prototype	2	4
	2	cookie	text	0	4
P4	1	cookie	text	1	4
	2	spaghetti	prototype	1	1

elements in the semantic model. In doing the classification, we noticed that tools are often tightly coupled with actions because they are objects of an action, a relationship not reflected in the model. Due to this tight coupling, we found it difficult to distinguish between a tool and an action as a possible source for an HI. As such, we have collapsed them into a single category: *action/tool*. Table 3 shows these counts compared to the observed HIs.

Overall, we found that a disproportionate number of interventions were in the measurement category. Measurements made up 21% of the instructions, but accounted for 48% of HIs. The cause of this discrepancy is unclear. It could simply be due to some participants' lack of cooking experience, or, for those who cooked at the researcher's home, the use of unfamiliar measuring tools. It is also possible, however, that the measurement and ingredients heuristic (*H1*) is not as effective as we had anticipated. To make the measurements clearer, one option would be to add colour-coding to the images in VERA. For further clarity, the measuring tools in the user's home could be correspondingly coded.

We next turn to the case study analyses.

Participant 1

P1 is a 29-year-old woman, two years post-onset. She has difficulty understanding both written and spoken sentences, and expresses herself primarily through individual words and gestures. Before acquiring aphasia, P1 occasionally prepared simple meals such as sandwiches, but did not use recipes. Since acquiring aphasia, she no longer cooks.

P1 seemed to particularly benefit from the navigational structure imposed by VERA. In the text-only condition, P1 had difficulty keeping track of her place in the recipe, sometimes failing to fully complete a step, and sometimes skipping a step entirely. Of the six steps in the first session requiring HIs, four were related to navigating the recipe. However, with VERA P1 did not have any comparable navigational difficulties.

Overall, P1 was considerably more independent when working with VERA, completing 9 of the 13 steps correctly without intervention (compared to 4 out of 13 with the text recipe), and only requiring HIs for two steps.

During interviews at the end of the first and second sessions, P1 was asked about ease of use and preferences for the two recipe formats. Responses were at times difficult to interpret with confidence, as she had some difficulty understanding the questions. When asked after the first session about the text-only format, she replied that

Table 3. Proportion of high-level interventions, by category. Note: the 6 navigational HI's observed are not included here.

Category	Observed HI's	Possible HI's
Measurement	48%	21%
Action/Tool	24%	52%
Ingredient	21%	29%
Duration	0%	6%

she needed help. However, it was not clear even after further questioning whether she meant help in following instructions or help in carrying them out. After the second session, she replied "no" when asked if she had any trouble using VERA. Although she had only made use of the extra text or sound functionality when prompted by the researcher, she replied "good" when asked whether the extra text was useful; she described the sound clips as "bad", adding "loud" with a gesture to suggest they needed to be louder. She had some difficulty with a question asking her to compare ease of use of the two formats, but eventually answered "same". However, when asked about which she would like to use in the future, she indicated the text-only recipe. It was not possible to determine with confidence why she preferred it, although she said "yes" when her mother eventually suggested that the text only recipe was less complicated.

Participant 2

P2 is a 51-year-old man, two years post-onset. He has relatively good auditory comprehension and expresses himself using short, incomplete sentences.

In the VERA session, measurement errors accounted for three of the five steps requiring HIs, which may have been due to his using an unfamiliar set of measuring spoons and cups, or to a lack of baking experience. In the text-only session, he completed the entire recipe without intervention.

During the interview at the end of the first session, P2 indicated that he found VERA easy to use and did not have any problems with it. He felt the images were sufficiently self-explanatory, and as a result he had not used the sound or extra text. At the end of the second session, P2 expressed satisfaction with the text recipe, feeling that he could read the text with relative ease. When asked about his preferences, he found both recipe formats equally easy to use, he liked both formats equally, and he would use both formats were he to cook in the future. When asked which was more comfortable to use, however, he indicated the text-only recipe.

Participant 3

P3 is a 57-year-old man, one year post-onset. Although he has some difficulty with word retrieval and recognition, he is generally capable of reading short sentences and individual words. He maintains relatively fluent speech and communicates effectively. Subjectively, as language testing could not be done, P3 has the least language impairment of the four participants. P3 cooked regularly before acquiring aphasia, and currently cooks from scratch two to three times a week.

Although P3 needed the same number of HIs (four) in both conditions, he made slightly fewer errors overall with the text-only recipe. In both conditions, his HIs involved mixed up ingredients, measurements, or heat settings, which may have been a due to a lack of familiarity with western cuisine. Prior to the study, P3 had never cooked spaghetti sauce or baked cookies.

In his second cooking session, P3 forgot his glasses and initially expressed concern that he might be unable to read the text-only recipe. However, when given the opportunity to postpone the session to a later date, he preferred to continue without his glasses. The researcher read aloud the first instruction for him when asked to, but he required no help reading subsequent instructions. Note for consistency that P3 did not have his glasses for the first session either, but had not been concerned.

At the end of the first session, P3 reported that he did not have any trouble using VERA. He found the extra text helpful in confirming the meaning of images, and felt the sound would be helpful if it were louder. At the end of the second session, P3 reported that he had no problems using the text-only recipe, except that he found reading difficult because he had forgotten his glasses. When asked about his preferences, P3 stated that he found it easier to cook with VERA, as the image and keyword text combinations facilitated understanding. He liked both formats equally, but felt more comfortable cooking with VERA because when reading was difficult, it was still possible to recognize objects from the images. For long-term use in the future, he predicted that he would also prefer the prototype.

Participant 4

P4 is a 73-year-old man, two and a half years post onset. He has moderate speech, reading, and auditory comprehension deficits. Before acquiring aphasia, P4 cooked frequently, making a variety of foods using fresh ingredients and occasionally using a recipe, but he rarely cooks now.

In his first cooking session, P4 made cookies using the text-only recipe. Measurement problems accounted for two of the four HIs. This may have been due to P4's measuring spoons, which were unlabeled. During the interview after the session, P4 initially reported having no trouble with the text-only recipe but later added that he found the text a bit hard to read and the instructions sometimes unclear. However, he did not need any help from the researcher to read instructions.

In his second cooking session, P4 made spaghetti using VERA. Because he was in a wheelchair and could not view the screen when it was flat on the table in tablet form, he used the prototype with the computer in laptop form; that is, the laptop was open with the keyboard between him and the screen. Although the screen was less stable in this form, he seemed to have little trouble using the pen to interact with the prototype. Only one step required an HI; P4 found the visual representation of *simmer* unclear.

During the interview at the end of the second session, P4 responded that the visual instructions were not hard to understand. He did not find the extra text useful, and he was unable to hear the sound. When asked about his preferences, he found both recipe formats equally easy to

follow. He also felt equally comfortable using both. He initially appeared to indicate that he liked the text-only recipe format better, although he tried unsuccessfully to qualify that preference. Further exploration during the third session with the speech-language pathologist revealed that he liked VERA's format better, but he enjoyed making the cookie recipe (which was in the text-only condition) more because he was physically able to complete more steps on his own than he could with the spaghetti recipe.

DISCUSSION

While the prototype did not help all of our participants in cooking more independently, the data suggests that the most severely impaired participants (P1 and P4) did best with the prototype, while the more mildly impaired participants (P2 and P3) either did best with the text or did roughly equally well on both. However, this result is complicated by the fact that the P1 and P4 happened to use the prototype second, whereas P2 and P3 used it first. All participants required the same or fewer interventions in their second session, suggesting a possible learning effect. While some of the improvement between sessions may be due to presentation order, we feel that for P1 and P4 some of the improvement was due to the support provided by the prototype.

That one of the more mildly impaired individuals (P2) did worse using VERA is a concern and suggests the need for further work to improve usability. The following sections discuss these and other issues that have arisen from our work.

Design Issues

Several usability errors were observed during the sessions. Although these were not counted as interventions, we discuss them here. Three of the four participants attempted to click a button with their finger, rather than the pen, suggesting that touch screen technology (as used in stylus-based devices) may be more intuitive than the electromagnetic technology used in the pen-based Tablet PC. Two participants double-clicked the sound icon, causing the sound clip to be played twice, which was particularly problematic because the sound streams would overlap, making the sound impossible to understand. One participant double-clicked the navigation arrow twice, inadvertently skipping steps. Three participants had trouble using the pen to click on buttons, requiring more than one try to activate the button on eight different occasions. Two of the participants could not find the extra text once it had been turned on; the researcher had to point it out to them.

A few general problems with the heuristics were noted. One common problem was that in several places, a *stir* instruction occurred on the screen following an *add* instruction. Participants would often stir automatically after the *add*. When they advanced to the next step, they would be confused by the *stir* instruction, and some participants would even go back and stir a second time. To remedy this, we have added another heuristic:

H11. Dependent action: Show closely related actions on the same screen, space permitting.

While we were expecting the sound and full text support to play a secondary role to the visual instructions, we were surprised by how little participants made use of this functionality. It is possible that participants were not yet familiar enough with the prototype to effectively use the customization mechanisms for sound and full text. For the full text support, however, it is also possible that the keywords in the visual instructions provided adequate textual support, so participants did not need the full text instructions. We also suspect that the lack of sound use was primarily due to usability issues, such as the volume being too low. The poor auditory quality was a limitation of the particular Tablet PC used, which could easily be remedied with the use of external speakers. In addition, our observations of participants interacting with the sound clips suggested that the granularity of the sound clips may be too large; that is, it may be more useful for participants to play a single word or phrase aloud than to hear the entire instruction read at once. In a future iteration of VERA we would like to provide user control over this granularity. That there were so many remaining usability problems with the sound functionality highlights a key limitation in our iterative design approach: by using non-aphasic undergraduate students for usability testing, we suspect that key usability issues were missed, especially as nonoptimal sound quality may be particularly problematic for individuals with auditory comprehension deficits and/or mild hearing impairment. This issue could have been partially mitigated by using age-matched non-aphasic individuals as they would have been more similar to our target audience at least in terms of hearing ability.

Methodological Issues

The results were almost evenly split when participants were asked about their recipe format preferences. No one found the text-only recipe easier than VERA, but only two individuals found VERA easier. In general, P1 preferred the text-only, P3 and P4 VERA, and P2 rated the two formats equally. These findings, however, require some qualification in view of the issues associated with interviewing individuals with aphasia. Impairments in comprehension and production of spoken language pose challenges both in conveying and in interpreting nuances of meaning that could potentially influence interpretation of results. This was particularly evident in the interviews for P1 and P4.

We were surprised that none of the participants reported having any major difficulties with the text or the prototype, despite their actual performance. We suspect that due to our intervention strategy participants may have been less aware of their performance than they would otherwise have been. That is, because no participant was allowed to make an irreversible error, all may have perceived their performance

to be satisfactory. As a result, participants may not have shown a strong preference for either of the formats.

One alternate evaluation approach would have been to test single independent instructions instead of a recipe as a whole. With such an approach, we would have been able to allow participants to make mistakes, which possibly would have resulted in greater self-awareness of performance. However, this approach would have had less ecological validity.

An inherent assumption in this work is that independence is of value for our participants. While independence is commonly cited as a major rehabilitative goal, the importance of support and care can be in direct conflict with this goal. We suspect that this may have been the case for P1. As a result, she may have preferred the text-only format because it required more social interaction and collaboration with the researcher. Independence is also a difficult issue for assistive technology as the desire to do things the "normal" way can impede the acceptance of such technology. Thus, the text-only format may also have been appealing due to its appearance of being more "normal".

CONCLUSIONS AND FUTURE WORK

We have derived a visual language for communicating cooking instructions and have used it in a prototype recipe application, VERA. An evaluation with four aphasic individuals suggested that the combination of visual instructions and navigational structure imposed by VERA helped those with relatively large language deficits to cook more independently.

Our results also suggest a possible effect of learning on the number of interventions required to successfully complete a cooking session. A longitudinal study involving three or more cooking sessions should be used to explore this possibility and to determine whether the advantages of VERA improve in such a context. In addition, this would reduce the effect of differing levels of prior cooking experience and provide participants with a better idea of which presentation format they prefer. Currently, only a modest preference for VERA is discernible.

We would like to develop two more-similar recipes for future evaluations. In this evaluation, we chose to use a cooking and a baking recipe to test if the set of heuristics was generalizable; however, the differences hindered a comparison of the text and VERA presentation formats. In addition, we still believe that the multimodal sound and full-text support could be useful for some individuals with aphasia, so we wish to address the usability problems that were encountered with those features.

REFERENCES

1. The Aphasia Institute: What is Aphasia? Available 08/09/2004 at http://www.aphasia.ca/about/whatis.html.

2. Bradbury, J.S., Shell, J.S., Knowles, C.B. Hands on cooking: Towards an attentive kitchen. *Ext. Abstracts CHI 2003*, ACM Press (2003), 996-997.

3. Chen, L., Cheng, S., Birnbaum, L., Hammond, K.J. The Interactive Chef: A task-sensitive assistant. *Proc. IUI 2002*, ACM Press (2002), 234.

4. Davies, R., Marcella, S., McGrenere, J., Purves, B. The ethnographically informed participatory design of a PDA application to support communication. To appear in *Proc. ACM Assets 2004*, ACM Press (2004), 153-160.

5. Elliott, D. *Step-by-Step*. SAGE Publishing, Olds, Alberta, Canada, 1990.

6. The Art of Cooking, Easily Explained for Knowledge Representation Experts. A Semantic Model in LOOM. http://www.reitter-it-media.de//compling/index.html.

7. Goodglass, H., Kaplan, E., Barresi, B. The assessment of aphasia and related disorders (3rd Ed.). Lippincott Williams & Wilkins, Philadelphia, PA, 2001.

8. Hux, K., Manasse, N., Weiss, A., Beukelman, D. Augmentative and alternative communication for persons with aphasia. In *Language Intervention Strategies in Adult Aphasia* (4th Ed.) Lippincott Williams & Wilkins, Philadelphia, PA, 675-689.

9. Ju, W., Hurwitz, R., Judd, T., Lee, B. CounterActive: An interactive cookbook for the kitchen counter. *Ext. Abstracts CHI '01*, ACM Press (2001), 269-270.

10. Kertesz, A. *Western Aphasia Battery*. Grune and Stratton, New York, NY, 1982.

11. Lehrer, A. *Semantic Fields and Lexical Structure*. North-Holland Publishing Company, Amsterdam, 1974.

12. McGrenere, J., R. Davies, L. Findlater, P. Graf, M. Klawe, K. Moffatt, B. Purves, and S. Yang. Insights from the Aphasia Project: Designing Technology For and With People who have Aphasia. *Proc. CUU 2003*, ACM Press (2003), 112-118.

13. Moffatt, K., McGrenere, J., Purves, B., Klawe, M. The participatory design of a sound and image enhanced daily planner for people with aphasia. *Proc. CHI 2004*, ACM Press (2004), 407-414.

14. Queensland University Aphasia Groups: Web Developer's Guidelines. Available 08/09/2004 at http://www.shrs.uq.edu.au/cdaru/aphasiagroups/Web_Development_Guidelines.html.

15. Steed, F. R. *A Special Picture Cookbook*. H & H Enterprises Inc., Lawrence, KS, 1974.

16. Steele, R. D., Weinrich, M., Carlson, G. S. Recipe preparation by a severely impaired aphasic using the C-VIC 2.0 interface. *Proc. RESNA 1989*, RESNA Press, 1989, 218-219.

17. Sutcliffe, A., Fickas, S., Sohlberg, M. M., Ehlhardt, L. A. Investigating the usability of assistive user interfaces. *Interacting with Computers*, 15 (2003), 577-602.

18. Thorburn, L., Newhoff, M., Rubin, S. Ability of subjects with aphasia to visually analyze written language, pantomime, and iconographic symbols. *AJSLP*, 4 (1995), 174-179.

19. Triebel, J. S., Manning, M. C. *I Think I Can Learn to Cook or I Can Cook to Think and Learn*. Academic Therapy Publications, San Rafael, CA, 1976.

20. Waller, A., Dennis, F., Brodie, J., Cairns, A.Y. Evaluating the use of TalksBac, a predictive communication device for nonfluent adults with aphasia. *INT J LANG COMM DIS*, 33(1) (1998), 45-70.