

Question-Asking Strategy for People with Aphasia to Remember Food Names

Shiho Arima, Shingo Kuroiwa, Yasuo Horiuchi

Graduate School of Advanced Integration Science, Chiba University

Daisuke Furukawa

Kimitsu Chuo Hospital

kuroiwa@faculty.chiba-u.jp

Abstract

We have developed a personal vocabulary assistant system as an Android smartphone application for aphasic people who can recognize objects and understand what they are but have difficulty finding the names of objects. Our system assists these aphasic people so that they can record the names of daily meals. We performed a pre-test using the prototype system with one patient and found two reasons for erroneous answers: pushing wrong button and the difference between the definition of the food/dish of the database and that of the users. Therefore, we proposed a method for dealing with them and implemented the system by improving the database and introducing a probabilistic model. When we performed the experiment with three patients, the system was able to lead the user to the correct name even when there were some erroneous answers.

Keywords

Aphasia, question-asking, Android application, word recall difficulty

Introduction

We have developed a personal vocabulary assistant system as an Android smartphone application for aphasic persons who can recognize objects and their meanings but have difficulty finding the names of objects. Our target patients can often recall the names when given clues such as written characters. The system supports aphasic persons by enabling them to take a photograph of the food that they want to talk about later and record the name of the object. The aphasic persons can name the food during a chat with their family or friends by looking at the photo and the object name displayed on the system.

The computer-aided visual communication (C-VIC) system is one of the famous augmentative and alternative communication (ACC) systems for aphasia (Steele et al., 409-426). Many assistant systems are conceptually based on C-VIC (van de Sandt-Koenderman et al., 459-474). These systems support conversation for specific tasks in situations such as shopping or visiting a doctor by using images and written text that have been prepared by non-aphasic people including speech-language-hearing therapists (ST). In contrast, our system helps aphasic people to prepare the system data that he/she wants to talk about and provides an opportunity to communicate. This paper shows our proposed feature to support recording the object name for aphasic persons.

Proposed system

The system is used as a personal vocabulary assistant system. Our proposed feature supports aphasic people so that they can record the names of daily meals. In general, an ST skillfully asks specific questions for better communication to get words out of aphasic persons. Furukawa et al. (100) studied question strategies to efficiently elicit words from aphasic persons. We implemented a question strategy based on this study. The system generates yes/no questions

and narrows down the candidates in accordance with the answers from the user. After the user has answered some questions, the system shows 5 candidates on the display. The user can record the name by touching one of them if the object name is found in the candidates.

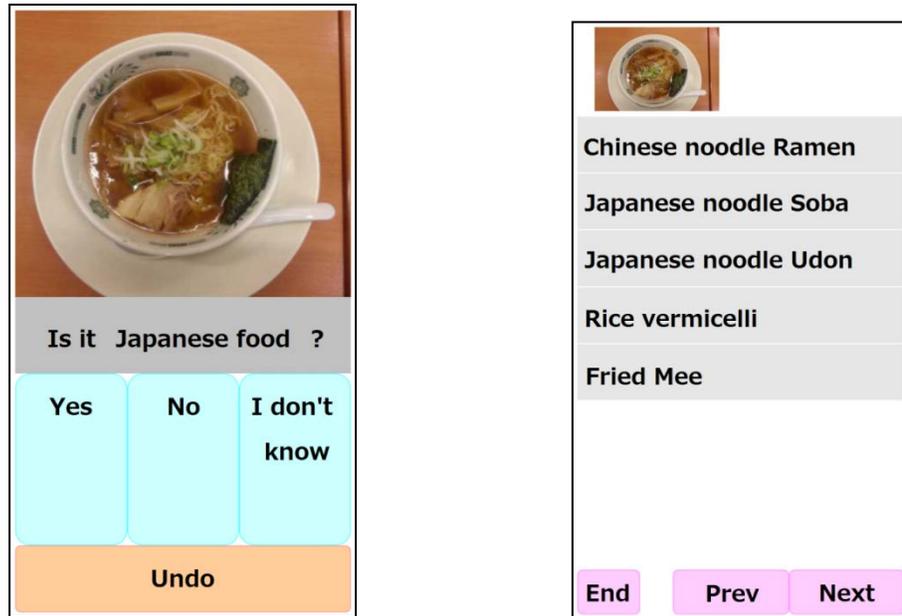


Fig. 1. Question screen (left) and screen of food name candidates (right).

Figure 1 (left) shows the question screen. The system displays a photograph the user has taken above the question about it. Below the question there are three buttons for answering: “Yes”, “No”, and “I don’t know”. The user replies to the question by touching one of the buttons. Then, the system shows the next question. The user can go back to a previous question and reply again by using the “undo” button.

After the user replies to about 10 to 30 questions, the system outputs 5 candidates as shown in Figure 1 (right). The system saves the name that the user selects from them. If there is no right name, the user can see the next 5 candidates by using the “next” button. He/she can also see the previous 5 candidates by using the “prev” button. If the candidate is not readily found, the user can finish this application by using the “end” button.

We registered data for 600 different kinds of foods/dishes in the database. As an attribute, we prepared the type of food/dish, food genre, cuisine, and ingredients. The system generates questions based on all these attributes.

Question strategy

We can calculate the conditional probability of a candidate food/dish name when the system observes an answer for a question:

$$P(r_n|q, a_o) = \frac{P(q, a_o|r_n)P(r_n)}{\sum_{r_i \in R} P(q, a_o|r_i)P(r_i)}, \quad (1)$$

where $r_n \in R$ is the food/dish name that is registered on the system, R is the set of names, $q \in Q$ is a question, Q is the set of questions, and $a_o \in A = \{\text{yes, no}\}$ is the content of an observed answer. Here, $P(q, a_o|r_n)$ is the likelihood of observing an a_o for q given r_n .

We should consider that aphasic persons reply with erroneous answers. Actually, we performed a preliminary test with one aphasic person and found two reasons for an erroneous answer. First, the definition of the food/dish of the user was different from that of the database. For example, when the system asked the user “Are there carrots with it?” for Hamburg steak, he replied “Yes” though the database had registered “no carrots with the Hamburg steak”. Second, a user might make a reply that was different from his/her intention. For example, when the system asked “Is it Japanese?” for “Sushi”, the user thought “Sushi is Japanese food” but touched the “No” button.

To consider these erroneous answers, we introduce a_T as a random variable expressing the true intention of the user to q . Then, we can calculate the likelihood:

$$P(q, a_o|r_n) = \sum_{a_T \in A} P(a_o|q, a_T)P(q, a_T|r_n), \quad (2)$$

where $P(a_o|q, a_T)$ is the term in consideration of the case where a_T is different from a_o and $P(q, a_T|r_n)$ is the term in consideration of the case where the definition of the food/dish for a user is different from the definition of the database. Then, we set formula (3) as constraints:

$$\sum_{a_T \in A} P(a_o|q, a_T) = 1 \quad (3)$$

We define $P(a_o|q, a_T)$ as the probability of observing a_o when the intention of the user for q is a_T . For example, $P(no|Is it main dish?, yes)$ refers to the probability that the intention of the user was supposed to be “Yes” for “Is it a main dish?”, but he/she mistakenly touched the “No” button.

In this paper, we assume that the error of the answer does not depend on the content of the question and answer. Then, we define $P(a_o|q, a_T)$ as follows:

$$P(a_o|q, a_T) = \begin{cases} \theta, & a_o = a_T \\ 1 - \theta, & a_o \neq a_T \end{cases} \quad (4)$$

where θ is the arbitrary probability. For simplicity, we wrote θ instead of $P(a_o|q, a_T)$ when $a_o = a_T$ in this paper.

$P(q, a_T|r_n)$ is the probability of the intention of the user for q given r_n . For example, $P("Is there beef in it?", yes|Hamburg steak)$ refers to the probability that the intention of the user is "Yes" when the system shows “Is there beef in it?” for Hamburg steak.

The above-mentioned probability will be set based on a large quantity of observation data. However, we cannot judge which reason causes the erroneous answer from the observed result. Therefore, in this paper, when the definition of the user was different from the database, we assumed it was an erroneous answer caused by another reason. We set θ to 0.84 based on preliminary experiments in this paper.

Using the above-mentioned formulation, the system generates a question with the expectation by the user answering that as much information as possible will be extracted. When the system observes a_o for q , we can calculate the entropy using formula (1):

$$H'(q, a_o) = \sum_{r_i \in R} -P(r_i | q, a_o) \log P(r_i | q, a_o). \quad (5)$$

We can calculate the expectation of the information obtained when the system observed a_o for q :

$$I'(q, a_o) = H - H'(q, a_o), \quad (6)$$

where H is the current entropy. In this way, we can calculate the expectation $I'(q)$ of the information to be provided by the next expression:

$$I'(q) = \sum_{a_o \in A} P(q, a_o) I'(q, a_o). \quad (7)$$

The question displayed by the screen \tilde{q} is chosen using formula (8).

$$\tilde{q} = \operatorname{argmax}_{q \in Q} I'(q). \quad (8)$$

In this paper, the system shows 5 candidates when the cumulative probability of them exceeds 0.5 or the number of questions reaches 31.

Experiment with three patients

We conducted evaluation experiments to confirm whether three aphasic persons could record the correct object name using our proposed system for the photographs for which they could not name. For this experiment, we prepared 30 photographs for each subject and selected the photographs of the food/dish that they have eaten and thought they could name if they saw

the written word. Table 1 shows the characteristics of the three subjects and each experimental condition.

Table 1. Characteristics of three subjects and experimental conditions

Subject	Age	Gender	Symptoms	Photos for
Ken	40s	Male	Sensory	10
Hiro	70s	Male	Apraxia of	3
Hana	50s	Female	Apraxia of	3

Results

Table 2 shows the number of successful trials, average number of questions when patients can record the name, and number of trials with erroneous answers. In successful trials, the system made from 11 to 31 questions to show the correct name. Besides, we confirmed that the system always showed all patients the correct names although there were erroneous answers for each trial.

Table 2. Successful trials, average number of questions, and trials with erroneous answers

Patient	Success trials/All trials	Average number of questions	Trials with erroneous answers
Ken	$9/10$	19.8	4
Hiro	$2/3$	21.3	3
Hana	$3/3$	17.7	2

Discussion

The system successfully provided the right food name in the candidate list 14 times during 16 trials for the aphasic persons to select. In one unsuccessful trial, the system provided the right food name, but Hiro could not select it. In another unsuccessful trial, Ken ended answering the questions. He was not able to continue answering the questions because he thought he had answered all the ingredients of the dish. In this case, the system needed to ask for the ingredient that had not been used to narrow down the candidates. However, it is considered to be unable to match human intuition.

The average number of questions in each trial was 19.7. In seven trials without erroneous answers, the average number of questions was 12.7. If we use the previous version that does not consider erroneous answers in these trials, the average number of questions reduced to 9.0. However, if we use the previous version for all trials, only seven trials were successful during 16 trials. If the number of successful trials increased like this by increasing the number of questions by 3 or 4, it can be concluded that the new system is better.

Conclusions

We have developed a personal vocabulary assistant system as an Android application for aphasic people who can recognize objects and their meanings but have difficulty finding the names of objects. To help these aphasic people remember food names, the system generates yes/no questions and narrows down the candidates in accordance with their answers. Experimental results showed that three subjects with aphasia successfully recalled the food names 14 times during 16 trials using a system that considers erroneous answers. In the first trial, we observed an ST letting out an exclamation of admiration when an aphasic person found the right name and uttered it by himself. Enabling him to do something that he was unable to do

alone had a huge impact not only on him but also on the people around him including us. Now, we are planning to improve the system using image recognition technology, which can reduce the number of questions asked until presenting the right name.

Works cited

- Furukawa, D., Sachiyo, M., Tomoyuki, K., “Appropriate question strategy to an aphasic person (1) - Basic study of the question to narrow down (Shitsugoshosha heno tekisetsu na shitsumon-houhou ni tsuite (1) – shiborikomu shitsumon no kisoteki kenkyu),” Japanese association of Speech-Language-Hearing Therapists 8 (2007): 100. Print. (in Japanese)
- Steele, R. D., Michael, W., Robert, T. W., Maria, K. K., and Gloria, S. C., “Computer-based visual communication in aphasia,” *Neuropsychologia* 27.4 (1989): 409-426. Print.
- van de Sandt-Koenderman, W. M. E., Jiska, W., Sandra, M. W., Hugo, J. D., and Gerard M. R., “High-tech AAC and severe aphasia: Candidacy for TouchSpeak (TS),” *Aphasiology* 21.5 (2007): 459-474. Print.