A WORD-FINDING ALGORITHM WITH A DYNAMIC LEXICAL-SEMANTIC MEMORY FOR PATIENTS WITH ANOMIA USING A SPEECH PROSTHESIS

Kenneth Mark Colby, Daniel Christinaz, Santiago Graham, Roger C. Parkison
The Neuropsychiatric Institute/UCLA
760 Westwood Plaza
Los Angeles, California 90024

ABSTRACT

Word-finding problems (anomia) are common in brain-damaged patients suffering from various types of aphasia. An algorithm is described which finds words for patients using a portable microprocessor-based speech prosthesis. The data-structures consist of a lexical-semantic memory which becomes reorganized over time depending on usage. The algorithm finds words based on partial information about them which is input by the user.

I WORD RETRIEVAL PROBLEMS

We are developing an "intelligent" speech prosthesis (ISP) for people with speech impairments [1]. An ISP consists of a small, portable computer programmed to serve a number of functions and interfaced with a speech synthesizer. ISPs can be operated by pressing keys on a keyboard or by eye-movements using a specially-designed pair of spectacles.

How words are stored, organized, and retrieved from human lexical memories constitutes a lively area of research in current psychology, computational linguistics, neurology, aphasiology and other cognitive sciences [2], [3], [4], [5], [6], [7], [8]. Words in a lexical memory can be associated to other words by means of several relations - synonymy, antonymy, part-whole, spatio-temporal contiguity, etc. [3]. It can also be assumed that the process of word-finding for production begins with semantic concepts to which word-signs are connected. Once the word-representation for the concept is found, it is subjected to phonological rules if it is to be spoken, and to graphological rules, if it is to be written. In the final stage of output, articulatory rules governing muscle movements for speech are utilized or rules for hand and finger movements for writing are applied.

Impairment in word expression can be due to failures at any stage of this process from concept to utterance. Our interest here is in those instances in which the speaker has the word, or part of the word, or some information about the word, in consciousness but cannot produce the target word. Our efforts have been directed towards writing and field-testing a computer program which can find words in a lexical-semantic memory, given some working information about the target word.

It has long been known that some aphasic patients, who cannot produce a word, can indicate how many syllables the word contains by finger tapping or squeezing the examiner's hand [9]. Both Barton [10] and Goodglass et. al. [11] reported that aphasics know some generic properties of the target word such as its first letter, last letter, etc. Our own experience with patients having anomic problems have confirmed and extended these observations.

The word-expression disorders in which we are interested are commonly divided into two groups which are (weakly) correlated with the locations of brain lesions and accompanying signs and symptoms. The first group consists of patients with lesions in the anterior portion of dominant cerebral hemisphere. These patients have many concrete and picturable words in consciousness, and they perform well, although slowly, on naming tasks [12], [13]. The naming disruption is part of a more generalized disturbance which includes laborious articulation and often a disruption of productive syntax.

The second group of disorders in this classification scheme involves lesions in the posterior region of the dominant hemisphere. Although these patients often fail to provide the correct name of an object, their substituted response can be related to the target word as described above in the studies of Rinnert and Whitaker [6]. Since the substitute word is so often systematically related to the target word, we thought it might be usable as a clue or pointer in a lexical search for the target word.

II A WORD-FINDING ALGORITHM

The first step involves constructing the data base, a lexicon of words stored on linked lists at various levels. The highest level is a "Topic Area" (TA), representing the idea being talked about. Currently we use 15 topic areas. The TA consists of list of words, each connected in a large lexicon of words, falling within that topic area. Each word within a particular topic area is also linked to a list of selected words, as in the

----

This research was supported by Grant #MCS78-09900 from the Intelligent Systems Program of the Mathematics and Computer Science Division of the National Science Foundation.

following example:

```
(BODY (ACHE (HEAD HURT PAIN STOMACH) )
      ( ANKLE (FOOT  LEG) )
            .
            .
            .
            .
      (WOUND (BLOOD  CUT  HURT ) ) )
```

The organization of the word lists changes dynamically over time according to their usage, as will be described below.

In attempting to find a word, the program first asks the user to identify the topic-area from a list presented to him on a display. We start with an initial set of word-lists but change them in accordance with the individual user's environment. A topic-area specifies where the user is linguistically in the discourse, not where he is physically or socio-psychologically. He is then asked about certain properties of the target word, the questions appearing on the display as follows:

(1) What is the topic area?
(2) What is the first letter of the word?
(3) What is the last letter of the word?
(4) What letters are in the middle of the word?
(5) What word does this word go with?

Question (5) attempts to obtain a word associated to the target word however idiosyncratic it might be. It may even resemble the target word in sound.

Our starting lexical memory of about 1800 words was based on known high-frequency discourse words, associations from Deese [8], word association norms [14] and word-pairs from Rinnert and Whitaker [6] . Word-associations can be highly idiosyncratic to the individual. For example, if one asks 1,000 college students which word they associate to the word table, 691 says chair but the remainder of the responses are distributed over 32 other words including jewels like big, cards, and tennis [15]. Hence, with each user, we add his particular associations and words that do not appear in our starting lexicon. This data is collected by a spouse, friend, research assistant or speech pathologist in conversations with the user.

After the user has offered clues about the target word, we can represent the responses using a clue pattern of the form:

CLUE = TA + L + L + STRING + GWW

where

TA = Topic area
L = Letter
STRING = One or more letters
GWW = GOESWITH word

Assuming that each slot in the pattern is correct and not null, the program first finds a short list of numbered candidate target words (zero to a maximum of 20 words) meeting the input criteria of the clue pattern. The search will be illustrated by cases.

Case (1). Suppose the target word were steak. The clue pattern might be:

CLUE = FOOD + S + K + A + meat.

The word meat is looked up on the FOOD list and then a search is made on the list of words linked with the word meat which begin with S and end with K and have an A in between. If steak were on the list of words linked to meat, it would be displayed visually and auditorily as:

1. STEAK

When he sees or hears the target the user enters its number (in this case "1.") and the word is then inserted in the utterance of the ISP. (Sometimes, on seeing or hearing the target, the user can utter the word himself.) In this first illustration of the program's operations, we have assumed full and correct entries have been made in the clue pattern and that steak was on the list of words which [GOESWITH] meat. But suppose steak is not on the meat list of FOOD.

Case (2). If steak were not found on the meat list of FOOD, the "meat" part of the clue is ignored and all words under FOOD beginning with (S), ending in (K), and with an (A) in between are retrieved. The word steak might appear in this group and if so, the program will add it to the meat list after the user signifies this is his target word. Thus the lexical memory becomes automatically reorganized over time.

Case (3). If still no acceptable word is retrieved, the "FOOD" part of the clue is ignored and a search is made on all topic-area lists for (S + K + A) words. The word steak might appear in the ANIMAL topic-area associated with the word cow. After the user indicates this is the target word, steak is added to the meat list under FOOD. With repeated usage, steak becomes promoted to the top of the meat list.

Case (4). If the target-word is still not retrieved by an exhaustive search of all topic-areas, the word does not exist in the lexicon and the program ends. One might consider varying the constraints of (S + K + A) clues and searching further but in our experience this is rarely productive. Time is not a problem for the program since an exhaustive search requires only a few seconds. But a large number of candidate words are retrieved when most of the clues are ignored. And it is too time-consuming for the user to search through the list of candidates looking for the desired word.

Case (5). It might be that the user cannot answer completely the questions required by the clue pattern or the entries may be in error. For example, he may not know the first and last letters of steak, but he does know the topic-area is FOOD and the [GOESWITH] word is meat. If steak is on the meat list, the search will succeed. If not, there is no point to displaying all the words under FOOD because the candidate-set is too large. In our experience thus far we have found that at least 2-3 pieces of information are necessary to retrieve a target word.

Further experience may indicate that some users can benefit from variations in the clue pattern. For example, with some patients, we have found it expedient to ask them first for the topic-area and the [GOESWITH] word. If this fails, the program then asks the letter questions. With other patients questions regarding word-size, number of syllables, and "what do you use it for?" may be helpful. We are currently field testing the program with a variety of anomic patients using keyboard or ocular control to study the value of utilizing different clue patterns.* In the meantime, others in the field of communication disorders may wish to utilize and improve on the word-finding algorithm reported here.

Such a program can also be used by speech pathologists as a therapeutic aid. The patient can be given linguistic tasks to practice on using the ISP at home. Repeated exercise and home-practice may facilitate the patient's own word-finding functions.

III SUMMARY

We have described a computer program for an intelligent speech prosthesis which can find words in a lexical-semantic memory given some information about the target words. The program is currently being tested with a variety of anomic patients who have word-finding difficulties. A future report will describe the results of this field-testing.

REFERENCES

[1] Colby, K.M., Christinaz, D., Graham, S. "A Computer-Driven Personal, Portable and Intelligent Speech Prosthesis." Computers and Biomedical Research 11 (1978) 337-343.

[2] Caramazza, A., Berndt, R.S. "Semantic and Syntactic Processes in Aphasia: A Review of the Literature." Psychological Bulletin 85 (1978) 898-918.

[3] Miller, G.A., Johnson-Laird, P.N. Language and Perception. Cambridge: Harvard University Press, 1976.

[4] Goodglass, H., Baker, E. "Semantic Field, Naming and Auditory Comprehension in Aphasia." Brain and Language 3 (1976) 359-374.

[5] Zurif, E., Caramazza, A., Myerson, R., Galvin, J. "Semantic Feature Representation for Normal and Aphasic Language. Brain and Language 1 (1974) 167-187.

[6] Rinnert, C., Whitaker, H.A. "Semantic Confusions by Aphasic Patients." Cortex, 9 (1973) 56-81.

[7] Geschwind, N. "The Varieties of Naming Errors." Cortex 3 (1967) 97-112.

[8] Deese, J. The Structure of Associations in Language and Thought. Baltimore: Johns Hopkins Press, 1965.

[9] Lichtheim, L. "On Aphasia." Brain 7 (1885) 433-484.

[10] Barton, M. "Recall of Generic Properties of Words in Aphasic Patients." Cortex 7 (1971) 73-82.

[11] Goodglass, H., Kaplan, E., Weintraub, S., Ackerman, N. "The "Tip-of-the-Tongue" Phenomenon in Aphasia." Cortex 12 (1976) 145-153.

[12] Wepman, J.M., Bock, R.D., Jones, L.V., Van Pelt, D. "Psycholinguistic Study of the Concept of Anomia." Journal of Speech and Hearing Disorders 21 (1956) 468-477.

[13] Marshall, J.C., Newcome, F. "Syntactic and Semantic Errors in Paralexia." Neuropsychologia 4 (1966) 169-176.

[14] Postman, L., Keppel, G. (Eds.). Norms of Word Association. New York: Academic Press, 1970.

[15] Palermo, D.S., Jenkins, J. Word Association Norms. Minneapolis: University of Minnesota Press, 1964.

----

*We are grateful to Carol Karp and patients of the Speech Pathology Division of Northridge Hospital Foundation, Northridge, California (Pam Schiffmacher, Director) for their collaboration.