CHAPTER

26

Clinical Utility of a Semantic Categorization Task

Susan T. Jackson
Connie A. Tompkins

At a short course offered by Bartlett, Duffy, and Metter at the 1986 American Speech and Hearing Association (ASHA) convention, one of the take-home messages was that language therapy for aphasic persons is, in general, efficacious. However, it was recommended that clinical aphasiologists attempt to answer more specific questions, such as: For whom will a particular therapy work and why? For whom will a particular therapy not work and why not?

As an example of attempting to specify for whom a particular therapy might not work and why not, we have chosen picture-naming ability as our behavior of interest. We have chosen this behavior for two reasons: (1) Virtually all aphasic persons have some degree of word retrieval difficulty and are likely to receive therapy for this problem, and (2) given that different stages in the word retrieval process have been postulated, word retrieval failures are likely to reflect various etiologies.

Several stages that are common to models of the word retrieval process are exemplified in a model proposed by Howard and Orchard-Lisle (1984). These stages are recognition, semantic representation, the phonological output lexicon, and the motor articulatory program. According to the model, possible reasons for word retrieval failure include a failure of initiation, a disordered phonological output lexicon, raised thresholds in the output lexicon, disconnection of the semantic system from the phonological output system, and deficient verbal semantics.

For our chunk of the word retrieval picture, we decided to hone in on the semantic system itself. One conceptualization of semantic categories postulates organization in semantic circles, with the most typical category representatives close to the center and less typical exemplars fanning out in the periphery. Related semantic categories overlap at their boundaries, with the result that membership in a category can be either relatively clear-cut or more fuzzy as this overlap occurs.

There is evidence suggesting that for some aphasic persons, the semantic system itself is disorganized and that this disorganization is manifested as a difficulty with boundaries between semantic categories (Goodglass and Baker, 1976; Grober, Perecman, Keller, and Brown, 1980; Whitehouse, Caramazza, and Zurif, 1978). For example, Grober and colleagues examined the performance on a semantic categorization task of aphasic subjects with brain damage predominantly anterior to or posterior to the fissure of Rolando. Their stimuli were typical and atypical members of a target category (e.g., fruits), members of a related category (e.g., vegetables), and members of an unrelated category (e.g., clothing). The subject's task was to decide whether each of these items (either in pictorial or written form) belonged to a designated target category. Grober and colleagues defined typical members of the target category and members of the unrelated category as clear-cut instances of membership or nonmembership in the target category. Errors on these items are labelled "clear-cut errors" herein. Atypical members of the target category and members of the related category were considered less clear-cut instances that are distributed at the boundary between semantic categories. Errors on these items are labelled "fuzzy boundary errors" herein. The difference between the aphasic subjects arose at category boundaries; posteriorly damaged subjects made significantly more errors than anteriorly damaged subjects on atypical members of the target categories and members of the related categories.

There are, however, several problems with the Grober and colleagues' (1980) study. First, there was no validation of the relatedness of the categories designated as "related" and "unrelated." Additionally, the typicality of members of the related and unrelated categories was not reported; thus the greater number of errors that were made on members of the related categories could have been due to the related items being more *atypical* than the unrelated items. Finally, there were only two control subjects tested.

More recently, Koemeda-Lutz, Cohen, and Meier (1987) found no differences between aphasic subgroups on various semantic classification tasks. Thus the effects of aphasia on semantic organization are unclear.

The purpose of the present study was to determine whether a semantic categorization task could be used to distinguish a group of aphasic subjects presumed to be semantically disorganized from a group presumed to be semantically intact. Such a task would provide useful information about the presumed etiology of word retrieval difficulties that could be tied to treatment planning.

METHOD

SUBJECTS

Experimental subjects included eight aphasic persons and eight age- and education-matched controls. Aphasic subjects with word-finding difficulty were identified by speech-language pathologists in the Pittsburgh area. All subjects had hearing levels in the better ear of at least 25 dB HL at 500 Hz and 30 dB HL at 1000 Hz and 2000 Hz, as measured by pure-tone air conduction audiometry. Yes/no reliability for six sentence-length questions was 100-percent correct for all subjects. All were native speakers of English. Aphasic and control subjects are described in Table 26-1. Nonparametric analyses showed that there was no significant difference in mean age between the aphasic and control subjects (Mann-Whitney $U = 29$, $p = .75$) or in mean years of education between these two groups ($U = 31$, $p = .88$).

TABLE 26-1. DESCRIPTIVE CHARACTERISTICS OF APHASIC AND CONTROL SUBJECTS

	Aphasic	*Control*
Age (years)	50–76 ($\bar{X}$ = 62.8)	51–75 ($\bar{X}$ = 62)
Education (years)	8–20 ($\bar{X}$ = 13.5)	10–16 ($\bar{X}$ = 13.2)
Gender	Males = 7; Females = 1	Males = 6; Females = 2
Months post-onset	6–33	N/A
Boston Naming Test (max = 60)	5–46	53–60
BDAE comprehension percentile	38–97.5	92.5–97.5
Peabody Picture Vocabulary Test percentile	6–99+	26–74

A group of 12 different non-neurologically-damaged adults validated some of the experimental stimuli.

VALIDATION PROCEDURE

Validation of experimental stimuli preceded the experiment proper. The 12 subjects involved rated the relatedness of numerous semantic categories on a seven-point scale from 0 to 6. Six categories that were highly related to one other category *and* completely unrelated to one other category were designated as the target categories for the experiment.

STIMULI

The stimuli were fashioned after Grober and colleagues (1980) with the following changes: (1) the relatedness of the categories was validated; (2) members of each category (target, related, and unrelated) were either highly typical, moderately typical, or atypical according to published norms (Battig and Montague, 1969); and (3) the stimuli were presented in a combined auditory-visual mode to maximize input potential.

Eighteen stimuli were associated with each target category — six members each of the target, related, and unrelated categories. Table 26-2 shows an example of target, related, and unrelated categories and their members.

TABLE 26-2. AN EXAMPLE OF TARGET, RELATED, AND UNRELATED CATEGORIES AND THEIR MEMBERS

Target (clothing)	*Related (kind of cloth)*	*Unrelated (alcoholic beverage)*
Shirt	Wool	Beer
Socks	Cotton	Whiskey
Shorts	Terrycloth	Ale
Bra	Flannel	Martini
Pajamas	Organdy	Bloody Mary
Earmuffs	Suede	Drambuie

PROCEDURE

The subjects were presented with a card on which the target category name was printed. One at a time, the 18 stimuli were presented in printed form on separate cards, while the examiner read each word aloud. The subject's task was to indicate (verbally or by pointing) whether or not each of the 18 stimulus words belonged to the target category.

Responses were scored as accurate or inaccurate. Like Grober and colleagues (1980), "clear-cut errors" were defined as errors on typical members of the target category and members of the unrelated category. "Fuzzy boundary errors" were defined as errors on atypical members of the target category and members of the related category.

RESULTS

Nonparametric analysis indicated that aphasic subjects made significantly more fuzzy boundary errors than clear-cut errors ($t = 0, p = .01$). The same was true for the control subjects ($t = 0, p = .01$). This result was not due to a word frequency effect, as there was no significant difference in the mean frequency of fuzzy boundary and clear-cut items on the experimental task ($F [1, 86] = 2.01, p = .16$).

Visual inspection of the data showed the range of fuzzy boundary errors for the aphasic group to be quite large. One subset of four aphasic subjects had errors within the range of the control group; they are designated as "high-level" aphasic subjects. A second subset of aphasic subjects had errors that fell outside the control subjects' range; they are designated as "low-level" aphasic subjects. Nonparametric analyses revealed that there was an overall significant difference among the number of fuzzy boundary errors made by control, high-level aphasic, and low-level aphasic subjects (Mann-Whitney U = 9.05, p = .01). Further analyses revealed that there was no significant difference in the number of fuzzy boundary errors made by the high-level aphasic and control subjects (*U* = 13.5; p = .67), that low-level aphasic subjects made significantly more fuzzy boundary errors than high-level aphasic subjects (Mann-Whitney U = 0, p = .02), and that low-level aphasic subjects made significantly more fuzzy boundary errors than control subjects (*U* = 0, p = .01).

The finding that the low-level aphasic subjects made more fuzzy boundary errors than the high-level aphasic subjects was not due to a word frequency effect. The median frequencies of fuzzy boundary items that were missed by each group did not differ significantly (Mann-Whitney U = 5, p = .72). This differential performance of the high- and low-level aphasic subjects also cannot be explained on the basis of a typicality effect. Subjects in both groups made significantly more errors on atypical items than on moderately typical or highly typical items (F [1, 6], p = .04), but there was no interaction between aphasic group and degree of typicality of errors.

Our final analyses were aimed at exploring whether any of the demographic or behavioral variables would predict performance on the semantic categorization task. We ran univariate nonparametric correlations on all of the variables listed in Table 26-3, and only overall auditory comprehension percentile on the Boston Diagnostic Aphasia Examination (Goodglass and Kaplan, 1983) correlated significantly with number of fuzzy boundary errors (rho = −.76, p = .04).

TABEL 26-3. VARIABLES CORRELATED WITH FUZZY BOUNDARY ERRORS

Variable	*Spearman's Rho*
Age (years)	−.02
Education (years)	.30
Months post-onset	.11
Boston Naming Test	.17
Peabody Picture Vocabulary Test percentile	−.26
BDAE single-word comprehension percentile	−.33
Phrase-length ratio	−.56
Concepts/minute	−.71
BDAE overall comprehension percentile	−.76*

*p < .05.

To determine whether the addition of any these variables would better predict performance on this task, we ran a multiple regression analysis with overall auditory comprehension and concepts per minute entered into the equation. We chose concepts per minute as an additional variable because it had the next highest value of rho. The multiple r of .87 was significant ($p = .03$), and the variable of concepts per minute explained an additional 18 percent of the variance over and above that explained by auditory comprehension alone. The results of this multiple regression analysis are very tentative, as it was based on the data of only eight aphasic subjects.

DISCUSSION

Our findings suggest that there are some aphasic persons whose performance on this task falls within the range of normal control subjects' performance and another group of aphasic persons whose performance falls outside of this range. Grober and colleagues (1980) have referred to these subjects as semantically intact and semantically disorganized, respectively. One possible interpretation of these findings is that the word retrieval difficulties of the semantically intact group is not due to a deficit in the semantic system. It would be necessary to test other stages in the word retrieval process to examine other possible etiologies of the word retrieval deficit of these subjects.

We selected subjects neither on the basis of anatomical lesion site nor on the basis of diagnostic type of aphasia; rather we included subjects only on the basis of the naming behavior itself to allow any differential performance on the task to occur naturally rather than forcedly. We were, however, still very much interested in what factors would predict who might be semantically disorganized or intact. Our analyses showed that overall auditory comprehension percentile was the variable that was most highly correlated with performance on our task. Although correlation does not imply causality, the question arises as to whether the root of "semantic disorganization" lies in the input to the semantic system, rather than in the semantic system itself. The design of this study will not allow us to evauate this possiblity; the data do, however, indicate that there is not a 1 : 1 correspondence between auditory comprehension percentile and number of fuzzy boundary errors.

The finding that concepts per minute improved the prediction of fuzzy boundary errors was unexpected. Concepts per minute is a measure of the efficiency with which content is communicated. The *number* of concepts was not significantly correlated with number of fuzzy boundary errors, indicating that *efficiency* is truly important. A possible interpretation of this finding is that if a person's semantic system is intact, it takes less time to search through the semantic lexicon, and hence (all other things being equal) this person will be quicker at communicating ideas than someone who is semantically impaired.

The main reason for our interest in determining the etiology of word retrieval failure is that we would choose a treatment for this language problem based on its presumed etiology. One of the most common therapies for word retrieval difficulty is cuing, and the majority of the cuing literature concerns the efficacy of phonemic cues (Gainotti, Silveri, Villa, and Miceli, 1986; Kohn and Goodglass, 1985; Love and Webb, 1977). However, we are more interested in the efficacy of semantic cues because they are one of the most effective self-cuing strategies — both in terms of the word retrieval success of the aphasic person (Marshall, 1975, 1976) and in terms of their value for listeners (Tompkins and Marshall, 1982). Our prediction is that an externally provided semantic cue would not be effective for semantically impaired aphasic persons in eliciting the correct name of a picture because accessing a disorganized system would only lead to an impaired response. We cannot predict the success of semantic cues for semantically intact subjects because a breakdown at another stage in the word retrieval process may preclude a picture being named correctly in response to a semantic cue.

The identification of the presumed etiology of disordered language performance is an important first step in the process of refining our investigations of treatment efficacy. If we can understand the basis of an observed symptom, then our treatment decisions and our predictions of treatment efficacy can be guided by data and theory rather than by intuition and speculation.

ACKNOWLEDGMENTS

This research was supported by a grant from the American Speech-Language-Hearing Foundation, and a Grant-in-Aid of Research from Sigma Xi, the Scientific Research Society.

REFERENCES

Battig, W. F., and Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 80,* 1–46.

Dunn, L. M., and Dunn, L. M. (1981). *Peabody Picture Vocabulary Test — Revised.* Circle Pines, MI: American Guidance Service.
Gainotti, G., Silveri, M. C., Villa, G., and Miceli, G. (1986). Anomic with and without lexical comprehension disorders. *Brain and Language, 29,* 18–33.
Goodglass, H., and Baker, E. (1976). Semantic fields, naming, and auditory comprehension in aphasia. *Brain and Language, 3,* 359–374.
Goodglass, H., and Kaplan, E. (1983). *The assessment of aphasia and related disorders.* Philadelphia: Lea & Febiger.
Goodglass, H., Quadfasel, F. A., and Timberlake, W. H. (1964). Phrase length and the type and severity of aphasia. *Cortex, 1,* 133–153.
Grober, E., Perecman, E., Kellar, L., and Brown, J. (1980). Lexical knowledge in anterior and posterior aphasics. *Brain and Language, 10,* 318–330.
Howard, D., and Orchard-Lisle, V. (1984). On the origin of semantic errors in naming: Evidence from the case of a global aphasic. *Cognitive Neuropsychology, 1,* 163–190.
Kaplan, E., Goodglass, H., and Weintraub, S. (1976). *Boston Naming Test, Experimental Edition.* Boston: Veteran's Administration Hospital.
Koemeda-Lutz, M., Cohen, R., and Meier, E. (1987). Organization of and access to semantic memory in aphasia. *Brain and Language, 30,* 321–337.
Kohn, S. E., and Goodglass, H. (1985). Picture-naming in aphasia. *Brain and Language, 24,* 266–283.
Love, R. J., and Webb, W. G. (1977). The efficacy of cueing techniques in Broca's aphasia. *Journal of Speech and Hearing Disorders, 42,* 170–176.
Marshall, R. C. (1975). Word retrieval strategies of aphasic adults in conversational speech. In R. H. Brookshire (Ed.), *Clinical aphasiology* (Vol. 5, pp. 278–284). Minneapolis, MN: BRK.
Marshall, R. C. (1976). Word retrieval behavior of aphasic adults. *Journal of Speech and Hearing Disorders, 41,* 444–451.
Tompkins, C. A., and Marshall, R. C. (1982). Communicative values of self-cues in aphasia. In R. H. Brookshire (Ed.), *Clinical aphasiology* (Vol. 12, pp. 75–82). Minneapolis, MN: BRK.
Whitehouse, P., Caramazza, A., and Zurif, E. (1978). Naming in aphasia: Interacting effects of form and function. *Brain and Language, 6,* 63–74.

DISCUSSION

Q = question; A = answer; C = comments.

Q. You talked about the subjects in terms of their comprehension level. Did you have any data on lesion site?

A. We did not have lesion site data. We chose not to categorize subjects on that basis.

Q. Did you have a measure of fluency?

A. Yes. We used the phrase-length ratio (Goodglass, Quadfasel, and Timberlake, 1964), which was not significantly correlated with performance on our task.

Q. In general, do you think that difficulty in determining fuzzy boundaries or insensitivity to those fuzzy boundaries necessarily means that these subjects have a semantic disorganization?

A. No. There could be a number of reasons why a person was having that sort of difficulty. The visual semantic system may be impaired rather than the verbal semantic system, or there may be a disconnection between the visual and verbal semantic systems.

Q. In my experience, patients who make semantic errors in naming (like "chair" for "table"), are frequently very good at categorizing even atypical members of the category. In your terminology, those patients would be semantically intact. Is it possible, in fact, that they could be semantically organized, but nonetheless semantically impaired? I have in mind something like an underspecified semantic representation, such that they don't have a representation that allows them to distinguish between "chair" and "table," for instance.

A. Yes, or a representation that is inadequate — a partial representation. It's possible.

Q. Where did your concepts per minutes measure come from?

A. Yorkston and Beukelman in the 1977 *Clinical Aphasiology Conference Proceedings.*