

# The abstraction of form in semantic categories

DAVID C. RUBIN, ELLEN R. STOLTZFUS, and KIRSTEN L. WALL  
*Duke University, Durham, North Carolina*

Undergraduates were asked to generate a name for a hypothetical new exemplar of a category. They produced names that had the same numbers of syllables, the same endings, and the same types of word stems as existing exemplars of that category. In addition, novel exemplars, each consisting of a nonsense syllable root and a prototypical ending, were accurately assigned to categories. The data demonstrate the abstraction and use of surface properties of words.

What types of organization do people use when they learn, categorize, or remember? For simple visual stimuli, such as patterns of dots varying randomly around a geometric pattern (Posner & Keele, 1968), people make use of whatever regularity exists. In contrast, for linguistic material, in which meaning can be more easily separated from form, psychologists have provided considerable evidence that people abstract and use meaning to organize information, often to the exclusion of form. The use of meaning as organization for linguistic stimuli, however, may be a result of the regularities available in the stimuli rather than a reflection of human cognition *per se*. We will argue on behalf of this claim, by demonstrating the abstraction and use of surface organization in linguistic material.

One reason that the abstraction of form from linguistic material has been minimized historically is that the surface form appears too transient to be useful. In early studies, Sachs (1967) and Wanner (1974) dramatically demonstrated the rapid loss of the surface forms of text, with many later studies supporting and extending the finding to a wide range of conditions (for recent reviews, see Alba & Hasher, 1983; Brewer & Nakamura, 1984; Gernsbacher, 1985; Wallace & Rubin, 1989). For instance, Gernsbacher (1985) cites over 40 references supporting the rapid fading of surface form. Although many studies have shown that surface form is retained in memory (e.g., Brown & McNeill, 1966; Hunt & Elliott, 1980; Hunt & Mitchell, 1982; Koriat & Lieblich, 1974; Nelson, 1981; Rubin, 1974, 1977), these studies do not provide clear evidence that form is used as the basis for the abstraction of structure. For example, evidence that the sound pattern is retained in the tip-of-the-tongue phenomenon could be taken as evidence for abstraction based on surface form, but it could also be explained by assuming partial recall of the sound of a word.

The general experimental strategy used to show that meaning is the basis upon which people organize their lin-

guistic experience has been to select materials for which meaning was the only clear form of organization present, such as stories written in prose. The results of such experiments were attributed to the nature of human cognition, but they could, perhaps, be attributed to the sampling of stimuli (Petrinovich, 1979, 1989). That is, for materials in which meaning is the most useful form of organization, people might organize using meaning, whereas for materials in which surface form is the most useful form of organization, people might organize using surface form. It followed that our demonstration required materials for which surface form was the most useful organization. For this purpose, we chose a set of semantic categories that contained names with highly organized surface forms. Although similarities among the concepts or objects referred to by the names in these categories are based on meaning, the similarities among the names are based mostly on sound. For example, the names of most nonprescription analgesics have two or three syllables, begin with roots that are not English words, and end with the letters *-in* or *-ol*. In contrast, the names of most laundry detergents are single-syllable meaningful words with no common ending.

To make the strongest demonstration that people use forms of organization other than semantics, we chose (1) a domain that is highly organized by a structure other than semantics; (2) a domain that is learned incidentally, to show that such learning is not just the result of an extreme effort by bright, cooperative, undergraduate subjects; and (3) a task that depends on the knowledge of surface form acting as an "organised mass" (Bartlett, 1932, p. 213), rather than depending on a comparison with a single past experience. In particular, subjects were asked to generate a name for a new exemplar of a category (Carroll, 1985; Rubin & Kontis, 1983; Wallace, 1989).

We chose the generation of a novel name as our task for several reasons. First, we wanted to show that the responses generated to this task could not be accounted for by the retrieval of the name of a single existing exemplar, but rather depended on the use of information abstracted from more than one exemplar. Second, subjects could create names based on semantics for novel concepts such as a particular building, computer files, a person who changes a movie marquee, the procedure of

---

We wish to thank Kirsten Neilsen and Cheryl Pou for help in testing the subjects, Wanda T. Wallace for her discussions throughout the project and Lynn Hasher and our reviewers for comments on the manuscript. Support for this research was provided by NSF Grant BNS-8410124. Reprints are available from David C. Rubin, Psychology Department, Duke University, Durham, NC 27706.

raising  $x$  at least  $y$  inches, cookies described by their recipe, and other cases in which there is little regularity in the form of the existing names to guide the form of the novel response (Carroll, 1985). Third, although we know of no other work showing the purposeful use of surface information to generate new words, we saw hints of the ability to use surface features in other work with linguistic stimuli. For example, errors reported in work on slips of the tongue (Dell & Reich, 1981; Fromkin, 1971) and on the accidental creation of word blends (Bolinger, 1961) often show preservation of surface properties rather than semantic properties.

### EXPERIMENT 1 Defining the Surface Properties of Four Semantic Categories

Experiment 1 was conducted to obtain quantitative normative data on subjects' knowledge of the names of existing exemplars in four semantic domains.

#### Method

A single group of 18 undergraduate volunteers filled out booklets consisting of a cover sheet and four pages. Each of the four pages consisted of the instructions, "In the next 30 seconds, please list all the" (1) "radioactive chemical elements that you can," (2) "different kinds of pastas that you can. In other words, the specific kind of noodle you might order in a fancy restaurant," (3) "brand names of laundry detergents that you can," or (4) "non-prescription pain killers you could buy at a drug store." The instructions were followed by 23 blank lines. The order of the pages was randomized across subjects.

#### Results

The subjects' responses were edited only to correct for obvious spelling errors. Spelling errors that changed the number of syllables were not corrected, because this would change the value on the number-of-syllables measure. Thus, "aspirin" and "asprin" were kept as separate entries. The subjects recalled an average of 4.2 items per category. The items listed and their frequency of listing, if it is greater than one, follow: *radioactive elements*—uranium 13, plutonium 10, radium 5, radon 5, carbon-14 3, californium 2, strontium 2, americium, americium, argon, berkelium, carbon, curium, einsteinium, estinium, hydrogen, lawrencium, lithium, lithuanium,

mendelevium, neptunium, polonium, rarium, rubidium, thallium, titanium, uranium-234, xenonium; *pastas*—spaghetti 15, lasagna 9, fettucini 8, rotini 8, ziti 6, rigatoni 5, linguini 4, ravioli 4, macaroni 3, manicotti 2, angel hair, cannelloni, linovini, mastacholli, mostaciolli, noodle shells, pasta primavera, pastina, pastorini, ritulli, rotelli, tortolini, virmicelli; *laundry detergents*—Tide 16, Cheer 9, Fab 7, All 6, Shout 5, Clorox 4, Fab One Shot 3, Snuggle 3, Wisk 3, Biz 2, Downey 2, Fresh Start 2, Oxydol 2, Surf 2, Bold, Bounce, Dial, Era, Exxon, Fab Shot, Fantastic, Fluffy, Gain, Ivory, Proctor and Gamble, Quick, Snowy, Tempra, Woolite, Zest; *analgesics*—Tylenol 16, Advil 10, Bayer 10, Anacin 7, aspirin 7, Bufferin 5, Excedrin 4, Motrin 4, Ibuprofen 3, Midol 3, Nuprin 3, acetaminophren, aspirin, beer, cortisone, cough syrup, Doan's pills, Exlax, Iuboprin, Medipren, Pamprin, Panadol, Pepto Bismo, St. Joseph's aspirin, vitamins.

Table 1 presents a summary of three properties that distinguish among the four categories: number of syllables, ending, and type of root. In English, the beginnings of words, as well as the endings, show considerable regularity; they function in poetics as alliteration and are frequently recalled in the tip-of-the-tongue phenomenon (Brown & McNeill, 1966; Rubin, 1974). In contrast, the beginnings of the four categories used here show little regularity, and so, little was expected in the generated responses. Thus, although only endings are reported here, the beginnings would probably be important in other semantic categories.

For each property in each category, a distribution of values from responses across all subjects was produced. Table 1 shows the modal value for each property in each category and the percentage of responses that had that value. The full distribution of these properties is given later in Tables 2, 3, and 4 where they are compared with the distributions obtained from generated novel instances.

Endings were defined as a letter or cluster of letters that ended at least one multisyllabic word. Single-syllable words were defined as having no ending. The set of endings was chosen by looking for the longest cluster that ended the largest number of words across the four categories. For instance, *-ium* was used as an ending because it ended almost as many words as did *-m* and *-um* but many more than *-nium*. Each ending category contained only endings that were the same both in spelling and in sound

Table 1  
Surface Properties of Four Semantic Categories

Category	Exemplars	Property					
		Syllables		Endings		Root	
		Modal Value	%	Modal Value	%	Modal Value	%
Radioactive elements	uranium plutonium	4	54	<i>-ium</i>	84	proper noun	60
Pastas	spaghetti lasagna	3	53	<i>-i</i>	86	no English meaning	100
Laundry detergents	Tide Cheer	1	75	none	75	English word	83
Analgesics	aspirin Tylenol	3	53	<i>-in</i>	40	no English meaning	62

pattern. Thus the *-i* of *linguini* and the *-y* of *Ivory* were separate endings, as would have been the *-y* of *way* and the *-y* of *Ivory*.

Roots, which were taken to be the first syllable or syllables of a word, were classified as belonging to one of the following four categories: English words with meanings outside the category (e.g., *Cheer*, *Tide*), misspelled English words (e.g., *Downey*, *Wisk*), proper nouns (e.g., *californium*, *americanium*), or nonwords that have no meaning outside the category (e.g., *fettucini*, *Tylenol*). Roots were counted as proper nouns even if their original ending was altered to allow a new ending to be added (e.g., *lawrencium*).

It is evident from examination of Table 1 that the four categories chosen differ with respect to the number of syllables, endings, and type of roots. Because the subjects, and not the experimenters, determined which items were included (Rubin & Olson, 1980), the data from Experiment 1 show the information subjects have about the surface form of the four categories. The question to be addressed next is whether or not individual subjects could abstract and use this information.

## EXPERIMENT 2

### The Generation of Novel Category Exemplars

Experiment 2 was conducted to determine whether the names that subjects generate for novel exemplars have distributions of surface properties similar to the distributions obtained from the subjects who recalled existing exemplars in Experiment 1.

#### Method

**Subjects.** Thirty-five undergraduates who were enrolled in an introductory psychology course received credit toward a course requirement. None of the subjects had participated in Experiment 1.

**Procedure.** The experiment was conducted in groups. The first page following the cover page of a booklet contained the following instructions.

You are interviewing with a top marketing firm. In order to test your aptitude for the business, you have been given a short test. Following the conventions commonly used, please make up a reasonable one-word name for each of the four items requested below. These names should be appropriate: an ideal name should fit in so well with those existing products in that category that it is indistinguishable from real names, perhaps to the point that someone new to the country would not know the name was not real. Please don't be silly.

- a new kind of pasta noodle \_\_\_\_\_.
- a new non-prescription pain killer, to be sold in drug stores \_\_\_\_\_.
- a newly discovered radioactive chemical element \_\_\_\_\_.
- a brand name for a new laundry detergent \_\_\_\_\_.

#### Results

The subjects used their knowledge of the surface properties of the four categories, as is shown, in qualitative terms, by their responses: *radioactive elements*—saturanium 2, byronium, chemx, delirium, diplonium, dukinium, elanine, farogenium, fissionate, gammagon, glowium, inflectium, kryptonium, lithium, lodranium, m-x, neuric, nucleum, nukinium, oops, quotron, radactium, radche, radioactanium, raliun, tibilium, trosphorus, uphorium,

uradion, veron, xenibium, xeno-x6, xyrgon, zanonioium; *pastas*—pastaroni 2, aldentini, catolini, linguista, maglini, masagnetti, mega-oodle, minestrageggi, mussolini, noodleoni, padodill, pastille, priatini, rastini, raziti, reguccini, rigatalini, rodelli, rodollos, ronacelli, rotolini, spaghettironi, sparoli, spiroghetti, spiro-gyro, sporini, squareoni, spaghetti, stralimi, stronguini, tortelucci, twirls, zigatoli, zoodle; *laundry detergents*—sparkle 3, all-clean, blanca, blitz, brightex, brite, bubbles, clean, clean n sweet, cleen, clenzeit, easywash, erabil, finally, fresh, kleen, le fleur, lift, line fresh, novax, phoron, powerclean, rainbow, resolve, schwab, sheen, snazzy, splash, stainaway, sunshine, vibrant, wave, zeld; *analgesics*—pain-away 2, acheaway, actrol, aftrex, ahh, anadospin, anapin, axadrin, banvel, codinium, curacil, entrex, kainer, matril, narcotan, ouch-away, ouchbegone, pacurine, painaway, pain-free, painprin, phodeine, pomatril, repadryl, sedomyn, shaknrin, sisprin, soothine, sport, supain, total cure, tranquil, valdol, zepatin. One response, the generated pasta "spaghetti," was removed from the analysis, because the subject who generated it probably would have recognized his or her response as the name of an existing pasta.

Tables 2, 3, and 4 display the distributions of responses for each value shown in Table 1. Although all items were included in the listings of responses just given, multiword responses were excluded. Only single-word responses (including hyphenated words) were scored in the quantitative analyses presented in the tables, both because the instructions specified single-word responses and because the number-of-syllable, root, and ending measures have different meanings for single- and multiword responses.

A visual inspection of Tables 2, 3, and 4 shows that in almost every case, the generated values from each category are closer to the recalled values from that category than they are to the recalled values from other categories. For instance, only the radioactive elements are characterized by having a large percentage of recalled and generated responses that are four syllables (Table 2), end in *-ium* (Table 3), and have proper nouns for roots (Table 4). In addition, where semantic characteristics of the generated exemplars are a regularity of the existing category names, such as in the laundry-detergent category, these semantic characteristics are also followed.

There is no ideal statistic to summarize the degree of agreement between the data from Experiments 1 and 2.

Table 2  
The Distribution of Responses (in Percent) as a Function of the Number of Syllables

Category	Experiment	Number of Syllables					
		1	2	3	4	5	6
Elements	recall	0	12	23	54	9	2
	generate	6	15	21	48	6	3
Pastas	recall	0	8	53	38	1	0
	generate	3	6	38	44	9	0
Detergents	recall	75	20	5	0	0	0
	generate	39	45	16	0	0	0
Analgesics	recall	1	41	53	5	0	1
	generate	6	48	39	6	0	0

**Table 3**  
The Distribution of Responses (in Percent) as a Function of Their Endings

Category	Experiment	Endings								
		-ium	-on	-ol	-in	-i	-a	-y	other	none
Elements	recall	84	12	0	0	0	0	0	4	0
	generate	61	15	0	0	0	0	0	18	6
Pastas	recall	0	0	0	0	86	14	0	0	0
	generate	0	0	0	0	81	3	0	12	3
Detergents	recall	0	1	3	0	0	1	7	12	75
	generate	0	3	0	0	0	3	6	48	39
Analgesics	recall	0	0	24	40	0	0	0	34	1
	generate	3	0	6	23	0	0	0	61	6

**Table 4**  
The Distribution of Responses (in Percent) as a Function of the Type of Root

Category	Experiment	Type of Root			
		Word	Misspelling	Proper Noun	Nonword
Elements	recall	2	3	60	35
	generate	21	3	15	61
Pastas	recall	0	0	0	100
	generate	28	0	3	69
Detergents	recall	83	8	0	9
	generate	74	16	0	10
Analgesics	recall	30	7	0	62
	generate	32	10	0	58

A chi-squared statistic cannot be applied, because there would be many cells with expected frequencies of zero and collapsing these cells would remove much of the effect. A correlation is appropriate, but because the number of degrees of freedom is two less than the number of categories, not two less than the number of subjects, the small number of categories (six for the number of syllables, nine for the endings, and four for the type of roots) makes this test less powerful than would be ideal. Nonetheless, we use matrices of correlation coefficients to summarize the results.

To produce Table 5, the percentage of responses generated in Experiment 2 in each category of Tables 2, 3, and 4 was predicted from the percentage of responses recalled in those tables from Experiment 1. If the data from the generate experiment from one category were predicted by the data from the recall experiment for that same category and by no other category, then the correlation matrices in Table 5 would have large correlations along their main diagonals and small correlations elsewhere. For the number-of-syllables data from Table 2 and the ending data from Table 3, the largest correlation for every row and for every column falls on the main diagonal. For the type-of-root data from Table 4, this is true for two of the four rows and columns. The correlations in the latter matrix, however, are each based on only four observations. Another way to summarize the data is to note that 9 of the 12 statistically significant correlations in the first three matrices of Table 5 fall on the main diagonal even though there are three times as many off-diagonal as there are diagonal correlations.

To form one measure representing the combination of the three properties, the three sets of percentages of responses from Tables 2, 3, and 4 were combined, making one 19-item-long profile for each generated and re-

called category (i.e., 6 syllables + 9 endings + 4 roots = 19 items). The last matrix of Table 5, which correlates the generated and recalled responses for this composite measure, is a way of looking at how well the three properties used together can distinguish among the categories. The composite measure also has the advantage of having much greater statistical power, because it has more degrees of freedom. For the composite measure, all 4 correlations on the main diagonal were .80 or greater (all  $p$ s <

**Table 5**  
Correlations among the Percent of Generated and Recalled Responses for Three Measures

Generated Category	Recalled Category			
	Elements	Pastas	Detergents	Analgesics
Number of Syllables from Table 2 ( $df = 4$ )				
Elements	.99*	.73	-.32	.16
Pastas	.90*	.94*	-.40	.35
Detergents	-.37	-.19	.73	.43
Analgesics	.07	.42	-.01	.95*
Ending from Table 3 ( $df = 7$ )				
Elements	.97*	-.24	-.11	.18
Pastas	-.18	.98*	-.14	-.19
Detergents	-.21	-.25	.68*	.23
Analgesics	-.14	-.24	.03	.76*
Type of Root from Table 4 ( $df = 2$ )				
Elements	.28	.96*	-.06	.93
Pastas	.06	.92	.12	.99*
Detergents	-.70	-.30	.99*	.17
Analgesics	-.17	.85	.25	1.00*
A Composite Measure of Syllables, Endings, and Roots ( $df = 17$ )				
Elements	.80*	.40	-.09	.31
Pastas	.12	.94*	-.07	.35
Detergents	-.31	-.20	.81*	.32
Analgesics	-.03	.32	.13	.90*

\* $p < .05$ .

.0001) and all 12 other correlations were .40 or less (all  $p$ s > .05). Thus, when the three properties considered in Tables 1-4 are combined, they differentiate among the four categories and show that the subjects did use the properties of existing exemplars in generating novel ones.

### EXPERIMENT 3

#### Assigning Novel Exemplars To Categories

Another way to demonstrate a sensitivity to surface forms is to have subjects judge the category membership of novel exemplars. Although the data obtained in this way are not as rich as the generated responses, they provide a simple quantitative index of people's ability to use surface information.

#### Method

**Subjects.** Forty-eight undergraduates enrolled in an introductory psychology course received credit toward a course requirement. None of the subjects had participated in Experiment 1 or Experiment 2.

**Procedure.** A single page contained the following instructions.

You are interviewing with a top marketing firm. In order to test your aptitude for the business, you have been given a short test. Several new products handled by this firm are in the process of being named and released to the public. Present employees with this firm have selected tentative names. The names were selected so that the public would be able to determine the category of product from the name alone. You are instructed to match the four new names to the most appropriate category by drawing a line between the name and the category heading. Each name and category is to be used only once.

Below the instructions, on the right-hand side of the page, were four exemplars. Each of these exemplars was made up of one of four single-syllable nonsense syllables that did not appear as the root of any of the exemplars that subjects listed in Experiment 1 (*bax*, *jek*, *naz*, and *vub*) and one of four endings (*-ini* for pastas, *-inium* for radioactive elements, *-irin* for analgesics, and no ending for laundry detergents). Because the roots were all CVC nonsense syllables, the type of root was not a cue in this experiment, leaving only the number of syllables and endings as cues to category membership. The endings are all one syllable longer than those listed in Table 3, because they were chosen from the results of Experiment 1 to provide both a prototypical ending and a prototypical number of syllables for each category. For each booklet, nonsense-syllable roots and endings were combined by sampling randomly without replacement, thereby producing four exemplars. Sample sets of four exemplars included *baxini*, *jekinium*, *nazirin*, and *vub*; and *bax*, *jekirin*, *nazini*, and *vubinium*. On the left side of the page were the four category headings with which the exemplars were to be matched.

Two sets of the four factorial combinations of roots and endings were made for a total of 48 single-page booklets. The order of the exemplars and category titles were random, except that an order was not used if its correct answer was four parallel lines drawn straight across the page. We avoided such cases, because we were concerned that if subjects were faced with four parallel straight lines, they would assume they had done something wrong and would become confused.

#### Results

The overall chance level probability for guessing which category went with the name that contained the proto-

typical surface features of that category is .25. The observed probability is .92 [ $t(47) = 22.16$ ,  $p < .0001$ ]. Of the 48 subjects, 41 were correct on all four of their responses. With only ending and syllable information, subjects can accurately judge category membership.

It is important to note that although the novel exemplars used were easily categorized, they were not the best category exemplars. One reason is that, in order to control for semantic factors that appear regularly in at least one of the categories (laundry detergent), we used nonsense syllables that gave no root information. Although excluding root information meant that not all of the names would be modal exemplars, it allowed us to minimize semantic cues to category membership.

Another reason why the exemplars used in Experiment 3 were certainly not the best exemplars possible is that we identified only three properties to study in this paper, ignoring other potentially important surface properties. Such properties are not always easily identified (Koriat & Lieblich, 1974), or easily applied to many categories as was required for our analyses, so we therefore chose only three easily defined and applied properties for our four word categories. It is likely, however, that other properties, such as the stress pattern of a word, are important for these and other categories, making our results an underestimate of the power of surface properties. For example, it is likely that there is more to the sound of good names for analgesics than the set of all three syllable words with nonword roots that end with *-in*.

### DISCUSSION

In Experiment 2, novel exemplars were generated that had distributions of surface features similar to those of the recalled exemplars of Experiment 1. In Experiment 3, novel exemplars were assigned with high accuracy to categories on the basis of surface form. Our subjects have thus demonstrated that surface forms, as well as semantics, can be abstracted and used when the task and the stimuli encourage it. Nearly every individual demonstrated this sensitivity to regularities in surface form—a sensitivity that is too often overshadowed in the cognitive literature by the emphasis placed on meaning.

Our claim that subjects generated their novel exemplars in Experiment 2 by abstracting and organizing surface information from existing exemplars would be weakened if it were possible for subjects to generate their responses from a single exemplar and general knowledge of the language (e.g., Aronoff, 1976) rather than from domain-specific information. To show that this is not an alternative, consider how subjects would generate novel instances of analgesics using only the exemplar *aspirin*. The subjects could either keep the beginning and change the ending or they could change the ending and keep the beginning. There is, however, no single rule that would apply to all categories. Endings were kept in three of the four categories studied here, but this need not always be the case. If subjects were asked to generate a name for a new

computer chain, they might keep the common *P.C.* or *Compu* beginning and vary the ending. If they were asked to name a new bran cereal, they might keep the common *bran* morpheme either as a beginning or as an ending. It is clear from examining many analgesics that the ending is shared and not the beginning, but if subjects based their responses on only one exemplar and on a general knowledge of English, they would not know this. Moreover, it is not even clear from a single instance where the beginning ends and the ending begins. Is it *as-pirin*, *aspi-rin*, or even the nonsyllabic break of *aspir-in*? Again, it is impossible to tell from one instance, but it is not difficult to abstract that *as-pirin* is the wrong choice from the set: *aspirin*, *Anacin*, *Bufferin*, *Excedrin*, *Ibuprofen*, *Motrin*, *Nuprin*, and *Pamprin*. The list of generated responses in the Results section of Experiment 2 makes it clear that the subjects prefer to keep the ending and that they break the word appropriately to leave an *-in* or a *-rin*.

Some generated instances do appear to be based primarily on one exemplar, such as *anapin* (from *Anacin*) or *painprin* (from *Pamprin*). Even for these generated instances, the changes made to the single instance follow the regularities found in the domain as a whole. If such changes were based *only* on the single instance, then rules for modifying words in English, in general, rather than in the particular domain, should have been used. It is not at all obvious how to formulate such general rules to produce the domain-specific changes made by the subjects. Thus, even when a subject appears to base a response on a single exemplar, the changes made to produce the novel response require consideration of the domain as a whole.

The clear use of surface information and its consistency across subjects suggests that the forms of the words were actually "meaningful," although not in the traditional semantic sense. The number of syllables, for example, is surface information by almost any definition. The endings, however, are more complex. Webster's unabridged dictionary (1976) lists one of the endings, *-ium*, as a suffix meaning chemical element (but note *aquarium*, *atrium*, *bacterium*, *cranium*, *delirium*, *gymnasium*, *opium*, *stadium*, and *tedium*, among others). Given the same reasoning, in the next edition of Webster's unabridged dictionary, *-i* might be listed as a suffix indicating pastas, and *-in* and *-ol* might each be listed as a suffix indicating non-prescription analgesics. That is, the consistent use of these endings to refer to categories may mean that we are witnessing the formation of new English morphemes, or perhaps in some cases what Bloomfield (1933) terms "root-forming morphemes" (e.g., the *fl-* in *flash*, *flare*, *flame*, or *flicker* denotes a moving light, while the *gl-* in *glow*, *glare*, *gloom*, *gleam*, and *glint* denotes a stationary light). Any surface property used consistently in this way may become "meaningful," in a broader sense, over time. Nonetheless, most psychologists would classify the three factors studied in this paper as surface features. If they would blur the old distinctions and consider them as part of the meanings of words, then our basic claim in this

paper, that people use organization outside of the old meaning of meaning, would be made.

In studying the remarkably accurate and long-lived memory that people have for ballads, popular songs, and counting-out rhymes (Hyman & Rubin, 1990; Kelly & Rubin, 1988; Wallace & Rubin, 1988a, 1988b, 1989, in press), we have found a similar reliance on surface structure for organization in memory. Recall errors are less likely to occur where there are regular surface forms, such as alliteration, rhyme, and rhythm, and the errors that do occur tend to preserve or distort toward the surface forms. In memory for material in these domains, meaning is important, but so are poetics. Thus, meaning is only one of many constraints that together preserve information in memory. The present study supports these findings and suggests that such an abstract appreciation of surface form is not limited to these particular genres of poetic language.

## REFERENCES

- ALBA, J. W., & HASHER, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*, 203-231.
- ARONOFF, M. (1976). *Word formation in generative grammar*. Cambridge, MA: MIT Press.
- BARTLETT, F. C. (1932). *Remembering: A study in experimental and social psychology*. London: Cambridge University Press.
- BLOOMFIELD, L. (1933). *Language*. New York: Henry Holt.
- BOLINGER, D. L. (1961). Verbal evocation. *Lingua*, *10*, 113-127.
- BREWER, W. F., & NAKAMURA, G. V. (1984). The nature and function of schemas. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 119-160). Hillsdale, NJ: Erlbaum.
- BROWN, R., & MCNEILL, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning & Verbal Behavior*, *5*, 325-337.
- CARROLL, J. M. (1985). *What's in a name? An essay in the psychology of reference*. New York: Freeman.
- DELL, G. S., & REICH, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning & Verbal Behavior*, *20*, 611-629.
- FROMKIN, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, *41*, 27-52.
- GERNSBACHER, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, *17*, 324-363.
- HUNT, R. R., & ELLIOTT, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General*, *109*, 49-74.
- HUNT, R. R., & MITCHELL, D. B. (1982). Independent effects of semantic and nonsemantic distinctiveness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *8*, 81-87.
- HYMAN, I. E., JR., & RUBIN, D. C. (1990). Memorabentia: A naturalistic study of long-term memory. *Memory & Cognition*, *18*, 205-214.
- KELLY, M. H., & RUBIN, D. C. (1988). Natural rhythmic patterns in English verse: Evidence from child counting-out rhymes. *Journal of Memory & Language*, *27*, 718-740.
- KORIAS, A., & LIEBLICH, I. (1974). What does a person in a "TOT" state know that a person in a "don't know" state doesn't know. *Memory & Cognition*, *2*, 647-655.
- NELSON, D. L. (1981). Many are called but few are chosen: The influence of context on the effects of category size. In G. H. Bower (Ed.), *The psychology of human learning and motivation: Advances in research and theory* (Vol. 15, pp. 129-162). New York: Academic Press.
- PETRINOVICH, L. (1979). Probabilistic functionalism: A conception of research method. *American Psychologist*, *34*, 373-390.
- PETRINOVICH, L. (1989). Representative design and the quality of generalization. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Every-*

- day cognition in adult and later life (pp. 11-24). Cambridge: Cambridge University Press.
- POSNER, M. I., & KEELE, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353-363.
- RUBIN, D. C. (1974). Within word structure in the tip-of-the-tongue phenomenon. *Journal of Verbal Learning & Verbal Behavior*, *14*, 392-397.
- RUBIN, D. C. (1977). Very long-term memory for prose and verse. *Journal of Verbal Learning & Verbal Behavior*, *16*, 611-621.
- RUBIN, D. C., & KONTIS, T. C. (1983). A schema for common cents. *Memory & Cognition*, *11*, 335-341.
- RUBIN, D. C., & OLSON, M. J. (1980). Recall of semantic domains. *Memory & Cognition*, *8*, 354-366.
- SACHS, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, *2*, 437-442.
- WALLACE, W. T. (1989). *The generation of a new ballad*. Unpublished manuscript, Duke University, Fuqua School of Business.
- WALLACE, W. T., & RUBIN, D. C. (1988a). Memory of a ballad singer. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues: Vol. 1. Memory in everyday life* (pp. 257-262). New York: Wiley.
- WALLACE, W. T., & RUBIN, D. C. (1988b). "The wreck of the old 97": A real event remembered in song. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 283-310). Cambridge, U.K.: Cambridge University Press.
- WALLACE, W. T., & RUBIN, D. C. (1989). *Poetic organization in memory for verse*. Unpublished manuscript, Duke University, Fuqua School of Business.
- WALLACE, W. T., & RUBIN, D. C. (in press). Characteristics and constraints in ballads and their effects on memory. *Discourse Processes*.
- WANNER, E. (1974). *On remembering, forgetting and understanding sentences: A study of the deep structure hypothesis*. The Hague: Mouton.
- Webster's third new international dictionary of the English language unabridged* (1986). Springfield, MA: Merriam-Webster Inc.

(Manuscript received July 13, 1989;  
revision accepted for publication May 8, 1990.)