Salience of the word as a unit in the perception of language*

CHARLES E. OSGOOD and RUMJAHN HOOSAIN

University of Illinois, Urbana-Champaign, Illinois 61801

Seven interlocking experiments are reported in which both guessing and recognition thresholds for words are compared with those for other linguistic units both smaller than (nonword morphemes and trigrams) and larger than (nominal compounds, ordinary noun phrases, and nonsense compounds) the word. Thresholds were consistently lower for words than for morphemes or trigrams (matched or even much higher in visual usage frequency) and lower for word-like nominal compounds (e.g., *stumbling block*) than for ordinary noun phrases (*copper block*) or nonsense compounds (*sympathy block*). Prior exposure (through two correct recognitions) to ordinary noun phrases, nonsense compounds, and the constituent single words of nominal compounds significantly facilitated subsequent recognition of the single-word constituents, but prior exposure to nominal compounds had no effect whatsoever on subsequent recognition of their single-word constituents. These results as a whole are interpreted as supporting the following conclusions: (1) that the word has special salience in the perception of language; (2) that the reason for this salience is the unique meaningfulness of the word (or the word-like nominal compound) as a whole; and (3) that the mechanism for this salience is the convergence of feedback from central mediational processes with feed-forward from peripheral sensory processes upon the integration of word-form percepts.

The first author (Osgood, 1963, and elsewhere) has argued that the word has special status as a psycholinguistic unit, being simultaneously the largest characteristic unit at the meaningless sensory integration level and the smallest characteristic unit at the meaningful representational level. For language *perception*, the word meets the requirements for sensory integration as a closure mechanism—redundancy of the same sensory “whole” in the Gestalt sense, relatively high frequency of occurrence, and temporal duration well within the limits of cell-assembly reverberation (cf. Hebb, 1949). For language *understanding*, the word is the smallest meaningful unit that accepts almost unlimited insertions at its boundaries (cf. Greenberg’s essay on the linguistic definition of the word, 1957) and hence can combine with other words to form an infinite variety of sentences.

Many experiments with the tachistoscope have testified empirically to an inverse relation between frequency of usage of words and their duration thresholds for recognition (e.g., Howes & Solomon, 1951; Solomon & Postman, 1952; King-Ellison & Jenkins, 1954; Postman & Rosenzweig, 1957). This relation holds when the frequencies of usage are for word *forms*, usually as determined from Thorndike-Lorge tables, independent of their multiple senses or *meanings*. At the level of sensory integration, the word is thus an empty or meaningless form, capable of multiple semantic codings depending upon linguistic context.

There are other identifiable units in the surface forms of printed language, both larger and smaller than the word. Larger units include sentences, clauses, and—closest to the word—noun phrases (like *married man*) that are written as separate words but which

function syntactically as single units. Units smaller than the word include letters, frequent clusters of letters that are not morphemes (e.g., *ple*) and—closest to the word—nonword morphemes such as *pre*.

In linguistic analysis, it is the morpheme, not the word, that is usually considered to be the minimal form having a meaning. If morphemes also function as salient units in the perception of language, then one would expect language users to recognize the morphemic components of words (e.g., *de* + *ceit* + *ful*) independently of and prior to recognition of the words they compose (here, *deceitful*). This is because such morphemes are typically components in many words (e.g., *decrease*, *conceit*, and *harmful* as well as *deceitful*), and therefore their overall frequencies of occurrence must be higher than those of any particular words they appear in. If thresholds are simply a matter of frequency of prior exposure, then it should follow that nonword morphemes or combinations of them would have lower thresholds than either (a) the words they compose or (b) single- or multiple-morpheme words matched in length and general shape but lower in usage frequency.

Two-word noun phrases consist of two separated “wholes” in the Gestalt sense, and yet they function as single units syntactically, just as do words. However, two types of noun phrases can be distinguished: Nominal compounds (like *stumbling block*) have a unitary meaning that is quite different from any simple fusion of the meanings of its single-word constituents (*stumbling* and *block*); this can be shown simply by comparing *he is the stumbling block in this organization* with *stumbling over the block he went sprawling*. Ordinary noun phrases (like *rusty block*) can be characterized as having meanings that represent fusions of the meanings of their compounds (cf. Weinreich, 1966, on “linking”). Semantically speaking, *stumbling block* is not *stumbling* + *block* but *rusty block* is *rusty* + *block*, and this is indicated by the fact that we can make insertions in the

*Requests for reprints should be addressed to Charles E. Osgood, Institute of Communications Research, University of Illinois, Urbana, Illinois 61801.

Method

Materials

All of the morphemes and words used, along with their estimated frequencies of usage, are presented in Table 1. These materials are mainly drawn from a larger set used by J. Robert Smith in a pilot study.¹ In this pilot, only "guessing" thresholds were determined, there was a gross imbalance of words over nonwords in the total materials, and the S population was rather haphazard. Nevertheless, results obtained by Smith were essentially confirmed in the present, more tightly controlled, experiment.

List I: Monosyllabic nonword morphemes (8 items). This list comprises reasonably familiar four-letter morphemes (e.g., *para-*, *-voke*, *-ment*). They may function in either word-initial or word-final positions. For the word-initial morphemes, the frequency estimates given are those of the most frequently used words we could find including them. The frequency estimates given for word-final morphemes are either (a) the summed frequencies of two words belonging to List II used in the pilot study, each including the morphemes in question, or (b) the frequency of the highest frequency word we could find in which the morpheme appears, whichever is larger. These estimates are *minimal*, of course, since they are not summations of the usage frequencies of all words including the morpheme.

List II: Words containing List I morphemes (8 items). For each nonword morpheme on List I, we selected for threshold comparison a word including that morpheme. The frequency of usage per million words for this list—and the other lists of actual words in this experiment—were taken directly from either Thorndike-Lorge (1944) or from West (1966), whichever yielded the larger estimate. Comparison of List I with List II provides a test of the first prediction.

List III: Monosyllabic words matched with List I for length (8 items). These are reasonably familiar four-letter words in English; they are also roughly matched with List I items for shape (e.g., *mani/mane*, *sist/seat*, etc.). Their frequencies of usage vary considerably, but all are used *less* frequently than the minimal estimates for their paired nonword morphemes. This comparison provides one test of the second prediction.

List IV: Multisyllabic nonword morpheme combinations (8 items). Fusions of two- and three-syllable morphemes were used to create multisyllabic nonwords of seven-letter length (e.g., *fam/ness*, *roof/lism*, *plan/lial*). Usage frequencies were estimated by taking the *average* of the frequencies of the two morphemes comprising each form. Listed frequencies were used if the included morphemes were also words (*roof*, *plan*) and frequencies of the most familiar words containing the morphemes were used if they were nonwords.²

List V: Multisyllabic seven-letter words (8 items). Here, words of relatively low frequency of usage were deliberately selected (*article* being an exception). Again, an attempt was made to roughly equate word shapes with those of List IV. Comparison of List IV with List V provides another test of the second hypothesis.

In all cases but one, the minimal estimates of the frequencies of usage of the nonwords are higher than those of the actual words with which they are compared experimentally; the exception is *mani-*, for which the highest frequency word we could find happens to be the one with which it is compared here, *manifest*.

Procedures³

Thirty-two Ss were used in this experiment. They were all

latter without changing the meanings of the original constituents (as in *rusty copper block*) but not in the former (as in **stumbling old block*), although external insertions at their boundaries are acceptable (as in *old stumbling block*)—again, just like single words. If nominal compounds function like ordinary noun phrases (and not like single words) at the level of language perception, then there should be no differences in their thresholds, and those for both types of compounds should be higher than for single words matched in length and usage frequency.

In this paper, we shall present evidence contrary to both expectations about units just smaller than the word and just larger than the word: (1) that tachistoscopic thresholds for recognizing nonword morphemes or combinations of morphemes are *not* lower than those for words in which they are components or single- or multiple-morpheme words with which they are matched, despite the much greater frequencies of usage of the nonword morphemes; (2) that tachistoscopic thresholds for ordinary compounds are *not* as low as those for nominal compounds which function semantically as single words. We shall also present evidence that the words in nominal compounds lose their semantic identity. Data from a number of minor experiments designed as controls for various factors will also be reported. In Discussion, we will argue that the demonstrated salience of the word in language perception is due to two properties deriving from a three-level behavior theory: first, that words and nominal compounds, but neither nonword morphemes nor ordinary two-word compounds, maximize the criterion of *redundancy as Gestalt-like "wholes"* in perceptual experience; second, that the perception of words and nominal compounds, but not nonword morphemes, is facilitated by *distinctive feedback from the representational (meaning) level*. We will come to the very general conclusion that meaning is a primary determinant of perceptual salience.

EXPERIMENT I

The preceding considerations lead to the following predictions, which are tested in Experiment I: (1) Nonword morphemes will have higher thresholds than words in which they are components, even though the morphemic components have higher frequencies of usage. Since this comparison is confounded with the stimulus-length variable—words necessarily containing more letters than included morphemes—a second prediction is tested. (2) Nonword morphemes, or nonword combinations of morphemes, will have higher thresholds than single or multiple morpheme words matched for length in letters and general visual shape, even though the nonword combinations have higher average frequencies of usage than the word combinations.

students, of both sexes, in introductory psychology at the University of Illinois, and they were required to serve as part of their course credit. All Ss had normal eyesight without glasses.

In order to eliminate possible effects of item order, each S was given the 40 experimental items in a different random order—with the constraint that the order of lists, from which items were drawn in rotation, was constant. This constraint was designed to minimize the likelihood of items of the same type appearing in succession (e.g., nonword/nonword, four-letter/four-letter, etc.).

Since exact instructions are important in this type of experiment, they are reproduced as follows: "We are going to see how quickly you can identify some words and some pronounceable non-words in this tachistoscope. You will see each item on a card directly in front of you for a very brief period of time, a flash. Approximately half of the items will be words and half non-words which are easily pronounceable. Possibly you may be able to guess what the item is immediately. If so, say it out-loud. After two correct identifications, I will go on to the next item. Remember, only about half of the items are words, but all are easily pronounceable. If you think you have a good idea of what the item is, say it. If you do not recognize an item on the first try, just say 'again,' and the item will be repeated for a slightly longer period of time. If you still do not recognize it, say 'again' once more. The item will be repeated for as long as necessary for you to recognize it. You are not in competition with anyone, and your responses will not even be treated individually. You should be aware that some items will take longer to recognize than others, but try to use as few repetitions as possible. In other words, an 'educated guess' is perfectly all right. There will be 40 items and a complete run will take about 50 minutes. If you would like to rest between items before going on at any time during the testing, just tell me."

After these instructions were given, Ss were given four practice items, both for familiarization with the types of materials and the procedures and to obtain a rough estimate of each S's general threshold level. These practice items were *bene*, *poly*, *going*, and *comseen*. The luminance of the stimulus field during exposure was 25 mL.

Since some Ss have generally higher thresholds than others, using a constant starting duration and uniform increments presents problems. If the increment is small enough to yield meaningful results for low-threshold Ss, then many tedious repetitions are required for high-threshold Ss before their threshold is even approached. Conversely, using larger increments would sacrifice the possibility of detecting finer discriminations among items for low-threshold Ss. To overcome this problem, we started at a 10-msec exposure, but used different increments—usually 1 msec but varying up to 5 msec—for different Ss, depending on their thresholds on the preliminary trials with the four practice items. The method of ascending limits was used, continuing until Ss gave two consecutive correct responses. All responses were recorded, together with their threshold durations.

Two "thresholds" were obtained for each item for each S: a *guessing* threshold and a *recognition* threshold. The former was included on the assumption that it would favor nonword morphemes or combinations, since even if an actual word of the same general shape were guessed it would still count as the "threshold" for that item. The exposure duration at which the first guess was made—whether correct or not, whether word or nonword—was recorded as the guessing threshold for an item. The duration time for the first of two successive correct identifications was recorded as the recognition threshold for an item.

Results

With regard to the first prediction (that words will be more salient than their included morphemes), the mean

Table 1
Thresholds for Words and Morphemes

	1	2	3	4	
I. Monosyllabic Nonword Morphemes					
ment	130	14	19	mant	(3)
mani	17	15	20	male	(3)
tide	76	13	17	tide	(5)
para	29	13	19	pars	(6)
voke	18	14	18	voice	(3)
sist	129	14	18	mist	(3)
cuse	76	16	20	curse	(4)
sult	407	12	19	suit	(7)
Grand Mean		14	19		
II. Words Containing the Morphemes					
mental	41	13	15	menial	(4)
manifest	17	12	14	—	(2)
fidelity	6	15	16	finally	
paradox	3	14	17	pardon	(2)
provoke	15	13	17	provide	(6)
resist	33	18	18	—	
accuse	28	16	17	—	
insult	21	13	14	—	
Grand Mean		14	16		
III. Monosyllabic Words					
mend	49	14	18	hand	(3)
mane	10	14	19	mand	(3)
pile	50	13	16	name	(5)
pave	14	13	16	pills	(3)
poke	12	14	16	prove	(3)
seat	124	13	20	pose	(3)
fuse	7	14	14	nest	(4)
salt	100	13	14	—	
Grand Mean		14	17		
IV. Multisyllabic Nonword Morpheme Combinations					
famness	135	25	45	famous	(9)
roofism	63	24	35	—	
planial	194	20	23	—	
framive	179	13	44	famine	(2)
artover	606	18	21	actover	(3)
grobion*	—	25	35	question	(4)
lossate	86	22	48	—	
henette	36	22	36	benette	(2)
Grand Mean		21	36		
V. Multisyllabic Words					
dashing	3	15	16	—	
sublime	8	13	16	—	
plaster	16	13	14	plastic	(3)
frantic	12	14	16	frontier	(3)
article	153	13	14	artistic	(3)
problem*	178	13	11	—	
lobster	7	14	15	luster	(3)
harmony	23	13	13	—	
Grand Mean		14	15		

**Grobion* and *problem* are left out in computing grand mean thresholds.

Note—Estimated frequencies of usage per one million words (1), mean duration exposures for first guesses in milliseconds (2), mean duration exposures for correct recognitions (3), and most frequent misperceptions with frequencies in parentheses (4), for monosyllabic nonword morphemes (I), words containing those morphemes (II), monosyllabic words matched in length (III), multisyllabic nonword morpheme combinations (IV), and matched multisyllabic words (V).

guessing thresholds for List I (nonword morphemes) vs the paired items on List II (words containing these

morphemes) showed three of eight item pairs in the predicted direction, one tie and four in the opposed direction (see Column 2, Table 1), and the mean *recognition* thresholds showed seven pairs, as predicted, and one tie (significant at the .008 level by the binomial expansion test). Using a sign test for correlated samples (each S as his own control),⁴ a similar picture emerged. Differences for guessing thresholds were still not significant (14/32 Ss in the predicted direction, 18/32 opposed), but differences in recognition thresholds were significantly as predicted (27/32 with only three Ss contrary— $p < .001$).

As for the first test of the second prediction (matched words more salient than the nonword morphemes), comparing the mean guessing threshold across all Ss for List I vs List III items, three out of eight items went in the predicted direction and one went against, and there were four ties. Comparing the recognition thresholds, seven were in the predicted direction and only one in the opposite direction (significant at the .035 level). Using the sign test across Ss, guessing thresholds for 21 of the 32 Ss were in the predicted direction, with 9 against and 2 ties (significant at the .02 level) and recognition thresholds for 26 of 32 Ss were as predicted, with only six opposed (significant at the .001 level).

For the second test of the second prediction (multisyllabic words), mean thresholds for guessing were in the predicted direction for six out of seven items (with one against, just fails to reach significance), and for recognition, all seven item pairs of Lists IV and V were in the predicted direction (with *grobion* and *problem* left out in the analysis, see Note 2), significant at the .008 level by the binomial test. On the sign test for correlated samples, comparative guessing thresholds were as predicted for 30/32 Ss and for all 32 for comparative recognition thresholds, both results being significant at well beyond the .001 level.

Results thus generally support expectations. Words are more salient than nonword morphemes—significantly on all comparisons by the more sensitive sign test for recognition thresholds and for all comparisons except words vs their included morphemes for guessing thresholds. It may be noted that differences between morphemes (or morpheme combinations) and corresponding words are consistently greater for recognition thresholds than for guessing thresholds.

EXPERIMENT II

In Experiment I, a greater proportion of items presented were words than nonwords (a ratio of 3/2). It is possible that Ss were even more biased than would normally be expected to give words as guesses (actually, 58% of 325 guessing errors made to nonwords were words, while only 13% of 383 errors made to words were nonwords, 73% of all first guesses being words). In Experiment II, the proportion of nonwords to words will be greater (by a 2/1 ratio). Also in Experiment I, the frequencies of usage of words were deliberately lower

than the minimal estimated frequencies of the nonwords. In Experiment II, we will use words and nonword morphemes which have as nearly equal frequencies as possible. Furthermore, we will include a third category of items for comparison with the words and nonword morphemes—these being nonmorpheme trigrams which have similar frequencies of occurrence (as parts of words) and similar visual shapes as the words and morphemes with which they are compared. These trigrams (e.g., *etu*) share with nonword morphemes the property of not being sensory “wholes” (*etu* as well as *epi* appear only as parts of words like *perpetual*, *epicycle*, etc.); but they are different from nonword morphemes in that they do not have any meanings per se. It is therefore predicted that nonmorpheme trigrams will have higher thresholds than three-letter nonword morphemes, which in turn will have higher thresholds than three-letter words—all sets being equated for frequency of occurrence.

Method

Materials and Procedure

Ten items of each of the three kinds of materials—trigrams, morphemes, and words—were used. These materials, along with their estimated frequencies of usage,⁵ are given in Table 2. The items in each triadic set (e.g., *ple*, *pre*, and *pen*) had as similar a visual shape and frequency of occurrence as we could devise. Following 6 practice items, all 30 items were presented in different random orders to different Ss, with the constraint that items from the same triadic set were separated by at least two items from other sets. The luminance of the stimulus field was 35 m.L.⁶ Five Ss who did not satisfy a criterion of having a 1-msec increment for repeated presentations without too many tedious repetitions were eliminated from the experiment. This had the effect of reducing variance across Ss and lowering mean threshold levels. Twenty-nine Ss were run. Except for modifications appropriate to the types of material, instructions were the same as those given in Experiment I.

Results

In general, words had lower *recognition* thresholds than nonword morphemes, and these in turn had lower thresholds than nonmorpheme trigrams—as predicted. For *guessing*, however, although words had lower thresholds than morphemes or trigrams, differences between morphemes and trigrams were not significant.

Words vs Morphemes

The test for guessing-threshold item-pair means ($6+2-2=$)⁷ is nearly significant at the $p < .05$ level; the sign test result ($25+4-$) is significant beyond the .001 level. For recognition thresholds, the item-pair mean test is significant at the .02 level ($6+0-4=$), and the sign test with Ss as their own controls is significant at well beyond the .001 level ($27+2-0=$). These results confirm the findings of Experiment I, under conditions where more nonwords than words are being presented and where the usage frequencies are roughly equated.

Words vs Trigrams

For guessing thresholds, the item-pair mean test ($7+/3=$) is significant at the .008 level and the sign test with Ss as their own controls ($20+/6-/-3=$) is significant at beyond the .007 level. For recognition thresholds, the item mean test ($9+/1=$) and the sign test over Ss ($29+$) have no exceptions to the predicted direction and are significant at the .002 and beyond the .001 level. This simply testifies to the much higher salience of three-letter words as compared with equally frequent nonword trigrams.

Morphemes vs Trigrams

For guessing thresholds, neither item-pair mean test ($5+/1-/-4=$) nor the sign test over Ss ($17+/12-/-$) yielded significant differences. For recognition thresholds, the item mean test is not significant ($7+/2-/-1=$) but the sign test with Ss as their own controls ($28+/1-/-$) is significant at well beyond the .001 level. Since the only apparent difference between the two kinds of items is that morphemes have some meaning whereas trigrams do not, this result suggests that meaning has more influence on recognizing than on inducing guessing.

That the increased proportion of nonwords (trigrams and morphemes) to actual words in this experiment as compared with Experiment I did influence guessing errors is clearly evident from the fact that here only 35% of the 437 wrong first guesses were words whereas 73% were words in Experiment I. There is a slight favoring of words above the proportion expected from the ratio in the materials in both experiments (73% vs 60% expected in I and 35% vs 33% expected in II), presumably because words are more available as production units. However, despite the two-to-one preponderance of nonword guessing in the present experiment, three-letter words have significantly lower thresholds than either nonword morphemes or trigrams.

EXPERIMENT III

Having compared words with subword units in Experiments I and II—manipulating item lengths, item usage frequencies, and proportions of words to nonwords in the presented materials—in Experiment III, we turn to a comparison of words with supraword units, specifically, nominal compounds which are closest to single-word nominals in linguistic function. Although linguists (cf. Lees, 1960) may describe the origins of such compounds in transformational terms (e.g., *peanut butter* from *butter made from peanuts*), it seems intuitively clear that contemporary speaker-hearers do not perform such transformations, but rather deal with such items as single semantic units (i.e., *peanut butter* is as unitary semantically as *marmalade*).

The fact that such nominal compounds are written as two separate words suggests that they might function

Table 2
Trigrams, Morphemes, and Words

	1	2	3	4
I. Trigrams				
ple	197	14	22	pie (12)
pes	4	14	17	pas (7)
tai	150	13	16	tal (9)
etu	11	14	18	stu (4)
ima	25	15	19	* (2)
ske	36	14	16	she (6)
eam	293	15	20	cam (5)
des	135	14	20	dis (7)
fas	14	13	16	fan (3)
neu	2	14	14	man (3)
Grand Mean		14	18	
II. Nonword Morphemes				
pre	216	13	15	pro (2)
geo	9	14	15	quo (2)
tri	167	13	14	—
epi	17	13	16	api (3)
iso	24	14	17	isc (3)
sym	33	14	17	—
com	301	15	16	con (2)
dia	75	13	20	dis (17)
fam	16	12	15	fan (5)
neo	3	15	19	men (2)
Grand Mean		14	16	meo (2)
III. Words				
pen	161	13	15	pan (11)
gun	5	13	15	man (2)
				pun (2)
tab	164	13	14	—
ego	10	14	16	—
jam	16	13	14	—
six	17	13	14	mix (4)
can	216	13	13	—
die	29	12	17	dis (11)
few	20	13	14	fen (2)
nil	5	12	14	mil (2)
Grand Mean		13	15	all (2)

*ime (2), ina (2), imm (2), ism (2)

Note—Estimated frequencies per one million words (1), mean duration exposures for guesses in milliseconds (2), mean duration exposures for correct recognitions (3), and most frequent misperceptions with frequencies in parentheses (4), for trigrams (I), nonword morphemes (II), and words (III).

perceptually as two independent forms rather than as a single form, in which case their thresholds should be little higher than those of their included words. However, since, as meaningful wholes, the nominal compounds have much lower frequencies of usage, they should have higher thresholds than the words that compose them—which is the prediction made. Since the meanings of component words in *nominal compounds* are quite different from when in isolation, we also predict that, whereas prior exposure to component words will have a marked facilitative effect on subsequent recognition of their compounds, prior exposure to compounds will have little facilitative effect upon subsequent recognition of their components.

Table 3
Words and Nominal Compounds

	I. Nominal Compounds						II. Words						4
	1	2a	2b	3a	3b	4*	1	2a	2b	3a	3b		
looking glass	1	23	16	27	17	looking clean (3)	looking glass	—	14	17	18	20	talking class (3)
post card	1	19	15	20	17	—	post card	AA**	14	18	21	19	class (4)
town hall	1	17	13	22	17	—	town hall	AA	15	15	21	21	past guard (13)
stock market	2	20	14	21	16	stock yard (2)	stock market	A†	17	15	19	17	guard (2)
attorney general	7	23	16	25	17	—	attorney general	AA	16	16	18	17	—
bowling green	1	21	15	26	19	looking glass (2)	bowling green	AA	14	16	23	19	ball (9)
coast guard	1	20	16	24	17	—	coast guard	AA	15	15	19	18	stack (4)
real estate	4	20	16	23	17	—	real estate	AA	17	17	18	20	—
motion picture	5	20	15	22	16	—	motion picture	23	16	16	20	17	stationary (2)
trap door	2	16	16	19	17	—	trap door	AA	14	15	16	17	quarrel (2)
lieutenant commander	1	22	17	28	21	lieutenant colonel (5)	lieutenant commander	1	15	13	16	17	looking queen (2)
stumbling block	1	21	15	25	17	—	stumbling block	AA	14	17	17	17	queen (3)
peanut butter	—	19	14	20	15	—	peanut butter	A	16	20	17	24	closet, green (2)
								AA	16	15	18	16	—
								AA	14	16	15	17	—
								44	16	17	21	22	—
								A	16	16	20	20	notion (2)
								AA	14	15	15	16	pleasure (2)
								42	15	15	19	17	trip (5)
								AA	15	17	17	19	dear (4)
								33	14	13	16	15	limousine (3)
								32	15	16	18	17	—
								—	15	16	18	18	stimulus (2)
								A	14	15	16	19	black (12)
								7	14	16	16	18	penant, peasant (2)
								AA	16	15	19	19	bullet, butcher, butler (2)
Grand Mean		20	15	23	17				15	16	18	18	

*Although guessing threshold times for the compounds include those where Ss reported seeing only one of the constituents of the compounds, such "partial" perceptions are not given in this column.

**Over 100 per million, †over 50 per million

Note—Estimated frequencies per one million words (1), mean duration exposures for guesses in milliseconds for first presentation (2a) and for second presentation (2b), mean duration exposures for correct recognitions for first presentation (3a) and for second presentation (3b), and most frequent misperceptions (for both presentations combined), with frequencies in parentheses (4), for nominal compounds (I) and words (II).

Method

Materials and Procedures

Thirteen nominal compounds that are normally written as separate words, together with the 26 single words making them up, constituted the materials for this experiment. They are listed in Table 3 along with their estimated frequencies of usage, which were obtained from Thorndike-Lorge lists (except for *peanut butter*, which had no separate frequency listing). The same general procedures for obtaining guessing and then recognition thresholds were followed as in Experiments I and II. Luminance of the stimulus field was 25 mL. One group of 15 Ss received a randomized order of the compounds first and then a randomized order of the component words; another group of 15 Ss received the same materials, but in the reverse order. In both cases, Ss were told that they would be shown some single words after the compounds (or the reverse), but not that there would be any relation between them—although, of course, they soon "caught on" to the fact that there was a relation. Thus, in Table 3, "first presentation" for each type of material refers to the data of one group of Ss and "second presentation" to data of the other group. Again, instructions were otherwise similar to those in Experiment I.

Results

We may first compare compounds and single words on their *first presentation* (to the two different groups of 15 Ss): For guessing thresholds (Columns 2a in Table 3), differences in mean threshold durations for the 26 comparisons (i.e., *looking glass* vs *looking* and vs *glass*, etc., for all 13 compounds) were significantly in favor of words, as predicted, at well beyond the .001 level (26+); differences in mean duration for recognition thresholds (Columns 3a in Table 3) were significant, as predicted, at beyond the .001 level (23+/2−/1=). However, on *second presentation* of the compounds (after their single included words) vs the single words (after their including compounds), there was no significant difference for guessing thresholds (this is seen by comparing Columns 2b in Table 3, 6+/14−/6=), and recognition thresholds for compounds were actually significantly lower than those for single words (comparing Columns 3b in Table 3, 5+/14−/7=, $p = .03$). It is thus

clear that words are more salient perceptually than their compounds when presented without prior exposure to related forms, but also that this difference is reversed (significantly so for recognition thresholds) when each type of material follows prior exposure to related forms. That the reversal is not significant for guessing is consistent with the general observation that Ss tended to withhold first guesses for nominal compounds more than for single words.

The triangle points in Fig. 1 show the mean durations for *recognition* thresholds (across the 13 means for compounds and the 26 means for single words) for first and second presentations. It can be seen that the reason for the reversal on the second presentations is a marked facilitation of the perception of nominal compounds following presentation of the single-word components, coupled with a complete absence of facilitation of the perception of single words following presentation of compounds of which they were part. Comparing statistically the first presentation means with the second presentation means for both words and compounds (that is, Columns 2a with Columns 2b and Columns 3a with Columns 3b in Table 3), we find that, whereas the reductions are significant for compounds on both measures [for guessing ($12+/1=$) at beyond the .003 level and for recognition ($13+$) at beyond the .002 level], there is no reduction for single words. In fact, for guessing thresholds, second presentation means are actually longer ($5+/14-7=$, significant at the .03 level), and there is no significant difference for recognition threshold means ($10+/11-5=$).

EXPERIMENT IV

The comparison between nominal compounds and constituent words in Experiment III involved differences beyond simply being double or single "wholes" graphologically: the single words were necessarily shorter than the compounds and also had much higher frequencies of usage as linguistic units. This can be seen by comparing Columns 1 in Table 3; in fact, whereas the frequencies of the nominal compounds were all less than 10 per million, in most cases the constituent words had frequencies in the A and AA categories (more than 50 and 100, respectively, per million). Apparently, nominal compounds tend to be composed from very high-frequency single words.

As a preliminary to Experiment IV, we compared five 12-letter words (the average length of the compounds in Experiment III), namely, *manufacturer*, *longitudinal*, *civilization*, *questionable*, and *acquaintance*, with five 6-letter words (the average length of the constituents), namely, *mainly*, *extort*, *remedy*, *endear*, and *repair*, respectively, matching pairs in frequency of usage. There were no consistent differences in either guessing or recognition thresholds, so apparently length per se (at least at this level) is not a significant variable. In Experiment IV, we control both item length and

Fig. 1. Mean recognition thresholds for nominal compounds and constituent words, and nonsense compounds and constituents as functions of order of presentation.

frequency of usage in a comparison of nominal compounds and single words—but, in this case of course, the single words cannot be constituents of the compounds. If, in Experiment III, the reason for the significantly lower thresholds of single words as compared with nominal compounds, when both were in first order of presentation, was simply the higher frequencies of usage of the former, then there should be no differences between them in Experiment IV.

Method

Materials and Procedures

Omitting *peanut butter* (for which no usage frequency could be determined), the remaining 12 nominal compounds used in Experiment III were matched in letter length (counting the space in the compounds as a letter) with 12 single words of approximately equal frequency of usage.⁸ These materials, along with their frequencies of usage (Columns 1), are given in Table 4. Twenty-one Ss were used, with the compounds and single words being randomized independently for each S. Luminance of the stimulus field was 35 mL, and the usual instructions and procedures for determining guessing and recognition thresholds were followed.

Results

Although the means of the guessing and recognition thresholds, as shown in Table 4 slightly favored words over nominal compounds, neither binomial tests over item-pairs nor sign tests over Ss reached significance. Assigning the + direction to words-lower-thresholds-than-compounds (since this was our original expectation), the item-pair tests are $6+/2-4=$ for guessing and $7+/4-1=$ for recognizing, while the sign tests are $12+/8-1=$ for guessing and $12+/9-$ for

Table 4
Nominal Compounds vs Single Words Matched in Length and Frequency of Usage

I. Nominal Compounds					II. Single Words				
1	2	3	4		1	2	3	4	
looking glass	1	19	22	—	misunderstand	4	17	19	misunderstood (3)
post card	1	17	18	—	pregnancy	1	16	17	—
town hall	1	16	18	—	selective	1	16	18	selection (2)
stock market	2	16	20	—	encyclopedia	4	16	17	—
attorney general	7	18	19	—	transcontinental	3	17	20	—
bowling green	1	18	21	—	approximation	1	16	17	experimentation (3)
coast guard	1	16	19	—	pessimistic	1	16	18	—
real estate	4	15	17	—	bureaucracy	1	17	18	—
motion picture	5	16	17	—	generalization	1	21	23	sterilization (3)
trap door	2	17	17	—	extremist	2	17	19	extremely (2)
lieutenant	1	21	23	lieutenant (2)	characteristically	1	20	21	characterization (4)
commander				colonel	acknowledgement	3	17	19	—
stumbling block	1	20	21	—					
Grand Mean		17	19				17	19	

Note—Estimated frequencies of usage per one million words (1), mean duration exposures for guesses (2), mean duration exposures for correct recognitions (3), both in milliseconds, and most frequent misperceptions with frequencies in parentheses (4), for nominal compounds (I) and single words (II).

recognizing—none of which reach significance at the .05 level. Of the total 252 possible comparisons (for 21 Ss and 12 item pairs), words had lower guessing thresholds in 104 and higher in 100, with 48 ties, and words had lower recognition thresholds in 106 and higher in 98, with 48 ties.

The results of this experiment clearly indicate that the differences favoring single words over nominal compounds in Experiment III are attributable to the extreme differences in frequency of usage of the items as wholes. Nominal compounds thus appear to function perceptually like single words of relatively low frequency. But this poses something of a paradox: if the single-word constituents have such high frequencies, why are they not recognized as readily within compounds, thus—given the semantic constraints within such compounds—lowering the thresholds for the compounds? It appears, to the contrary, that the high-frequency words somehow “lose their identity” within compounds. In this connection it should be noted that, in Experiment III, if an S guessed either of the component words it was recorded as the guessing threshold for that item (this occurred about 40% of the time), yet guessing thresholds equally and significantly favored constituent words. In the last two experiments, we will demonstrate that this “loss of identity” is largely a semantic matter. But first we must eliminate some simple-minded alternative possibilities.

EXPERIMENT V

Although the results of Experiment IV show that the differences between constituent words and compounds on their initial (first in order) presentation in Experiment III were probably due to frequency of usage, they in no way account for the marked difference in facilitation on their second presentation. It will be recalled that, in Experiment III, there was a significant

lowering of thresholds for nominal compounds following exposure to their constituents, whereas prior exposure to the compounds had no effect upon the recognition thresholds for the constituent single words. This phenomenon is presumably related to the “paradox” just noted—that high-frequency words seem to lose their identity within nominal compounds. It is a well-documented fact that reduction of number of alternatives—for example, by informing Ss as to the size and nature of the stimulus class—lowers tachistoscopic thresholds. The question raised by Experiment III thus becomes: why is prior exposure to nominal compounds “less informative” with respect to single-word recognition than the reverse?

But first we must eliminate another possible explanation, which might be termed a “reverse ceiling effect”—that, given the very high usage frequencies of the single words in Experiment III, their ordinary (first presentation) thresholds were already so low that prior exposure to them in compounds could have little added effect. In Experiment V, therefore, we simply compare the effect of prior presentation of the 26 single words used in Experiment III upon thresholds of these same words with the null effect of prior presentation of the same words within nominal compounds obtained in Experiment III. We predict that, in this case, there will be a marked and significant facilitation.

Method

Materials and Procedures

The 26 words making up the compounds in Experiment III were presented to 10 Ss in random orders, varying for each S. Then a random sample of 15 of these 26 words—also in random orders varying with each S—was presented.⁹ The retest words were: *butter, block, estate, green, town, glass, general, market, peanut, real, picture, trap, door, post* and *motion*. As had been the case in Experiment III, the Ss were not told that items were being repeated, and no “break” in the presentations occurred.

However, as would be expected, after only a few items all Ss became aware that items were being repeated (as determined from questioning at the end of the experiment). Luminance of the stimulus field was 25 mL, and the standard threshold determination procedures described earlier were followed.

Results

There was, as predicted, a clear and significant facilitation effect. The mean guessing thresholds for the 15 test items were 16 msec for first presentation and 15 msec for second; mean recognition thresholds were 19 msec for first and 16 msec for second presentation. The by-item binomial tests were significant for both guessing ($12+/3-$, $p < .02$) and recognition ($13+/2-$, $p = .004$). Since the usage frequencies of the items appearing by chance in each of the 15 retest order-positions were highly variable—but each S was still his own control for each item in each position—in Table 5 we present the numbers of Ss for whom retest items in each order position had lower recognition thresholds (top row), higher thresholds (middle row), or equal thresholds (bottom row) with respect to values for the same items on initial presentation. Except for the first and the last positions, all comparisons are in the expected direction (lower thresholds for repeat presentations).

The reason for reversal on the first items repeated is clear from Ss' terminal reports—not expecting repetitions, they delayed reporting until they were “dead sure” it was a repetition of a same word; we have no idea as to why the 15th-order items required longer durations. Since there are no order effects beyond the first two positions, we can assume that Ss “caught on” to the fact of repetition—and reduced the set of alternatives accordingly—at least by the third retest item. These data provide a clear basis for rejecting the “reverse ceiling effect” hypothesis: repetition of the same single words reduces their thresholds significantly, an effect which did not occur in Experiment III when the same words appeared initially as constituents of nominal compounds.

EXPERIMENT VI

We are left with the “paradox” posed by the results of Experiment III—why do high-frequency single words seem to “lose their identity” when presented in nominal compounds? We assume that this is not a “perceptual” matter per se (that is, not failure of sensory integration of the word forms), but rather a “semantic” matter (that is, failure of the meaningful feedback from the representational system for words in compounds to correspond to that normally associated with, and hence facilitating, the single-word perceptions). There are, however, other “simple-minded” explanations which must first be eliminated.

One possibility is that, when Ss shifted from compounds to single words in Experiment III, the

Table 5
Number of Ss Having Lower (Top), Higher (Middle), or Equal (Bottom) Recognition Thresholds on Retest as Compared to Initial Presentation as a Function of Order of Presentation in Retest

Recognition Thresholds	Order of Presentation														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Lower (+)	2	5	8	8	4	8	7	5	7	6	5	9	5	4	2
Higher (–)	6	3	0	2	3	1	1	2	2	1	4	0	3	3	5
Equal (=)	2	2	2	0	3	1	2	3	1	3	1	1	2	3	3

prerecognition “blurs” in the tachistoscope were obviously much shorter in length than the compounds and therefore they “dismissed from their minds” the prior materials. Another possibility is *habituation*. Because the nominal compounds in Experiment III had much higher thresholds than their constituent words, Ss who were presented compounds first may have become habituated to relatively long durations before responding (that is, expecting more repetitions before recognition), and by the time they were presented with the single words such habituation had raised their threshold levels (i.e., Ss would still “wait” for the accustomed numbers of exposures). Conversely, Ss who were shown single words first may have become accustomed to shorter durations and fewer repetitions, and this might have helped to lower times for the subsequently presented compounds.

In Experiment VI, we repeat the design of Experiment III, but with a different type of material—unrelated noun pairs or “nonsense compounds” (e.g., *motion glass*) rather than semantically constrained nominal compounds (*looking glass*). We predict that prior exposure to the unrelated pairs of nouns will facilitate subsequent perception of the single words to approximately the same degree found in Experiment V (simple repetition).

Method

Materials and Procedures

Ten nonsense compounds and their 20 constituent words were used. The second constituent word of each compound was identical with that of one of the nominal compounds of Experiment III (e.g., *actor general* corresponding to *attorney general*, see Table 6), and the corresponding first constituent words were matched in frequency of usage (between *actor* and *attorney*, see Table 7). The same procedures as in Experiment III were used; 15 Ss were shown the single words first and then the 10 nonsense compounds made up of the words, and 13 Ss were shown the same materials in reverse order. The luminance of the stimulus field was 25 mL.

Results

As in Experiment III, the first presentation times for the compounds were significantly higher than those for the constituent words (comparing *actor general* with *actor* and with *general*, etc., for all 10 nonsense compounds, we have $13+/2-/5=$, $p < .004$, for guessing,

Table 6
Nonsense Compounds and Their Constituent Words Compared

	I. Nonsense Compounds						II. Words						
	1a	1b	2a	2b	3		1a	1b	2a	2b	3		
actor	15	16	24	20	motor	(2)	actor	17	15	19	17	motor	(2)
general					general		general	15	13	19	15	—	
motion	15	18	35	25	motion	(5)	motion	17	14	22	16	notion	(4)
glass					glace		glass	15	15	16	17	time	(2)
shade	21	18	28	25	—		shade	15	15	18	17	—	
market							market	18	16	21	19	mother	(2)
tree	15	15	23	20	tree	(7)	tree	15	14	19	15	time	(5)
hall					ball		hall	13	14	17	17	ball	(4)
sympathy	17	15	27	21	stumbling	(5)	sympathy	16	15	18	18	—	
block					black		block	14	14	17	17	black	(12)
medicine	18	16	28	21	—		medicine	16	15	19	14	exercise	(2)
picture							picture	14	15	17	15	pleasure	(2)
plant	16	14	20	16	—		plant	14	14	15	15	—	
card							card	15	14	16	15	—	
cheese	16	18	23	23	—		cheese	16	14	19	17	—	
guard							guard	16	15	17	16	—	
lamp	16	16	22	19	—		lamp	14	14	15	20	—	
door							door	15	15	18	16	deer	(2)
post	19	15	34	21	part	(2)	post	15	14	20	21	past	(6)
estate					estate		estate	18	15	21	20	—	
Grand Mean	17	16	26	21				15	15	18	17		

Note—Mean duration exposures for guesses in milliseconds for first presentation (1a) and for second presentation (1b), mean duration exposures for correct recognitions for first presentation (2a) and for second presentation (2b), and most frequent misperceptions (for both presentations combined), with frequencies in parentheses (3), for nonsense compounds (I) and included words (II).

and 20+, $p < .001$, for recognition, see paired Columns 1a and 2a in Table 6). On second presentation, however, when compounds (after seeing the constituent words) are compared with single words (after seeing the compounds), thresholds were still significantly lower for single words (comparing paired Columns 1b and 2b for compounds vs words in Table 6, $16+/4=$, $p < .001$ for guessing, and $18+/1-1=$, $p < .001$ for recognition). Thus, unlike Experiment III, where we found no facilitation for single words presented after compounds, we do have such a facilitation with nonsense compounds, as shown by the circle points in Fig. 1. Comparing first and second presentation times for the single words, facilitation for guessing thresholds was significant at better than the .006 level ($12+/2-/6=$), and for recognition, at better than the .01 level ($13+/3-/4=$).

There was no significant facilitation of guessing thresholds for these nonsense compounds on their second presentation (comparing Columns 1a and 1b in Table 6, we have $5+/3-/2=$), whereas there was marked facilitation of the meaningful compounds on their second presentation in Experiment III (see Columns 2a and 2b in Table 3). This was largely because guessing thresholds for the first presentations of nonsense compounds were already very low (with Ss apparently soon realizing that they were nonsense and hence "anything is possible"). For recognition thresholds, however, there was significant facilitation (comparing Columns 2a and 2b in Table 6, we have $9+/1=$, $p = .002$),

just as in Experiment III. Noteworthy is the fact that recognition thresholds for *meaningful* compounds (Experiment III) were markedly lower than for *nonsense* compounds (Experiment VI) on both first and second presentations, as can be seen in Fig. 1.

EXPERIMENT VII

Experiment VI shows that the finding in Experiment III—prior presentation of nominal compounds not facilitating perception of their constituent words—cannot be explained in terms of either habituation or Ss "dismissing from their mind" the compounds when they first saw the much shorter length "blurs" subsequently. Still viable, however, is our hypothesis that single words "lose their identity" when embedded in meaningful nominal compounds (e.g., that the ordinary meaning of *block* is lost in *stumbling block*).

But there is another form of "compounding"—ordinary noun phrases, in which constituent words retain their individual meanings (e.g., *copper block* as distinguished from *stumbling block*). An indication that such noun phrases are intuitively felt to be only loosely bound together semantically is that they do not appear as lexical entries in dictionaries (thus we find *stumbling block* but not *copper block*). In this respect, they are like the nonsense compounds, with two semantically distinguishable units; however, they are different from nonsense compounds, and similar to

Table 7
Thresholds for Constituent Words of Nominal Compounds, Ordinary Noun Phrases,
and Nonsense Compounds Used in Experiment VII

I. Words of Nominal Compounds				II. Words of Noun Phrases				III. Words of Nonsense Compounds			
	1	2	3		1	2	3		1	2	3
A* stock	AA	14	16	C street	AA	13	15	B shade	AA	13	14
market	AA	17	23	market	AA	13	21	market	AA	15	20
attorney	23	16	18	combat	18	16	21	actor	28	14	22
general	AA	16	16	general	AA	13	16	general	AA	11	15
trap	42	15	20	steel	A	12	15	lamp	A	12	14
door	AA	15	16	door	A	13	12	door	AA	13	15
B town	AA	15	19	A labor	AA	14	14	C tree	AA	12	14
hall	AA	12	22	hall	AA	12	15	hall	AA	11	13
real	AA	13	20	city	AA	13	13	post	AA	12	14
estate	44	15	19	estate	44	20	23	estate	44	14	21
stumbling	25	13	18	copper	46	13	14	sympathy	43	12	14
block	A	13	16	block	A	13	16	block	A	14	16
C coast	A	17	20	B library	A	12	13	A cheese	45	14	16
guard	AA	14	22	guard	AA	13	14	guard	AA	14	16
motion	A	16	21	fancy	A	12	16	medicine	46	14	15
picture	AA	14	18	picture	AA	12	14	picture	AA	12	14
looking	-÷	12	17	protection	44	14	17	motion	A	14	23
glass	AA	13	16	glass	AA	12	15	glass	AA	13	14
Grand Mean		14	19			13	16			13	16

*The three groups of Ss were each presented items from Group A, B, or C, respectively, of each of the three types of material.

†No available frequency count. Note—Estimated frequencies per one million words (1), mean duration exposures for first guesses in milliseconds (2), and mean duration exposures for correct recognitions (3), for constituent words of nominal compounds (I), ordinary noun phrases (II), and nonsense compounds (III).

nominal compounds, in that the two constituent words together create a definite and congruent meaning (compare *copper block* with *sympathy block*, for example).

In the present experiment, the three kinds of compounds are compared, and it is expected that the ordinary noun phrases will function more like the nonsense compounds, as far as *thresholds* are concerned. Furthermore, insofar as constituent words retain their normal meanings in compounds, their prior presentation as compounds should facilitate perceiving them subsequently as single words. On the other hand, ordinary noun phrases should be more like nominal compounds as far as retention and recall are concerned, by virtue of their greater meaningfulness.

Method

Materials and Procedures

Nine of the nominal compounds from Experiment III and nine of the nonsense compounds from Experiment VI were used together with nine ordinary noun phrases. The constituent words are shown in Table 7. Each triad of compounds had identical second constituents (e.g., *stumbling block*, *copper block*, and *sympathy block*), with the first constituents having matched frequencies (see Columns 1, Table 7). The nine compounds of each kind were divided into three groups at random, and every S was presented three of each kind of compound together with their constituent words (Groups A, B, or C in Table 7).

Seven Ss for each of three combinations (Groups A, B, or C) of nine compounds were first shown the compounds in random order, without being told that their constituent words were to follow. The luminance level of the stimulus field was 25 mL.

However, instead of using the usual method of ascending limits, only two exposure durations were used. Each compound was presented first at a "short" time (usually 30 msec but up to 50 msec), which was estimated to be around the threshold level (after five practice items of *romantic woman*, *steam engine*, *bookmark class*, *electrode head*, and *civil service*). If the S made the correct guess for the entire compound, it was presented a second time at an identical exposure duration. Otherwise, it was repeated twice at an exposure duration double the "short" time (to ensure that the S could see the compound). Thus, each compound was seen twice.

Ss were told that the purpose was to see whether it took the "short" time or the "long" time before they could see each presented compound, explained as belonging to either one of two categories: meaningful or meaningless (the ordinary noun phrases therefore being classified together with nominal compounds, both being distinguished from the nonsense compounds). The short-time/long-time design was used instead of the usual method of ascending limits to avoid possible habituation to the relatively more repetitions and longer durations always needed for compounds (even though Experiment VI had indicated that this effect was not significant). After presenting the nine compounds (three of each kind in random orders), Ss were shown the 18 constituent words in random order, with the constraint that no two words belonging to the same compound were to be presented consecutively. Here, the usual method of ascending limits was used. Finally, Ss were asked to try to recall all the compounds presented in the first part of the experiment.

Results

Across all Ss, the proportions of each of the three kinds of compounds that were recognized correctly at the "short" time was computed.¹⁰ It was 15.9% for nonsense compounds, 23.8% for noun phrases, and

46.0% for nominal compounds [$F(2,60) = 6.6$, $p < .005$]. In other words, nonsense compounds were most difficult to perceive and nominal compounds easiest, with ordinary noun phrases intermediate in this respect but more like nonsense compounds. These results for recognition with "short" vs "long" exposures are consistent with a comparison of the recognition thresholds for nominal compounds in Experiment III and nonsense compounds in Experiment VI (see Fig. 1)¹¹: on first presentations of both types of material, mean thresholds for nominal compounds were lower (23 msec, Table 3, Column 3a) than for nonsense compounds (26 msec, Table 6, Column 2a); comparing the 10 items with common second constituents, e.g., *looking glass* with *motion glass*, and counting lower thresholds for nominal compounds as +, the distribution of results is $7+2-1=$ (which is significant at only the $p < .10$ level, however).

When we look at the proportions of each of the three types of compounds that were correctly recalled at the end of the experiment, a very different picture emerges: 12.7% for nonsense compounds, 30.2% for ordinary noun phrases, and 33.3% for nominal compounds [$F(2,60) = 3.6$, $p < .05$]. The fact that ordinary noun phrases are *recalled* just about as well as nominal compounds (both in contrast to nonsense compounds) suggests that it is the meaningfulness of the former that contributes to ease of coding in memory and hence better recall (it is also possible, of course, that the nonsense compounds are less "imageable").

Turning to the main purpose of this experiment, the mean thresholds for both guessing and recognition were computed for each single word presented, across the seven Ss in each group (see Columns 2 and 3 in Table 7). Note that, although the first constituent words of compounds (e.g., *stock*, *street*, and *shade* with *market*) are not constant across Groups A, B, and C, they were matched closely on frequency of usage (see Columns 1 in Table 7); on the other hand, the second constituents, e.g., *market*, are common across Groups A (given *stock market* first), B (given *shade market* first), and C (given *street market* first). Comparing the constituent words of noun phrases with those of nonsense compounds, we find no differences in mean thresholds ($6+7-5=$) for guessing or ($6+8-4=$) for recognition. Comparing nonsense compounds with nominal compounds similarly, however, the former have significantly lower guessing thresholds at the .002 level ($14+2-2=$) and lower recognition thresholds at the .006 level ($14+3-1=$). And finally, comparing ordinary noun phrases with nominal compounds, we find that the former also have significantly lower guessing thresholds at the .01 level ($11+2-5=$) and recognition thresholds at beyond the .004 level ($14+1-3=$).

These results thus clearly indicate that, although ordinary noun phrases are like nominal compounds in that they constitute meaningful wholes (and are recalled about equally well at the end of the experimental

sessions), they are like nonsense compounds in that their word constituents retain their identity in the compounding process (and equally facilitate subsequent perception of their single-word constituents in contrast to nominal compounds).

DISCUSSION

In introducing his chapter on *Words as Visual Patterns*, Ulric Neisser (1967, p. 105) had this to say: "Few topics in psychology have generated so much heat as the recognition of words . . . Yet despite its liveliness, an author who approaches this subject has some reason to fear that his readers may find it tiresome or even painful . . . there has been no "closure"—no generally acceptable statement of the facts has appeared." We hope to take some small steps in this direction—by pointing up the insufficiency of any single-factor views and by stressing interactions *between* perceptual and meaningful levels of cognitive organization.

Insufficiency of Single-Factor Theories

The combination of two variables—emphasis upon peripheral vs central mechanisms and upon perceptual vs response locus of effects—yields four "types" of single-determinant theories: (1) *peripheral response bias theory*, in which the inverse frequency/threshold relation is attributed simply to the probability of emitting alternative responses, this probability being a function of usage frequency (cf. the "spew hypothesis" of Underwood & Schulz, 1960); (2) *peripheral perceptual bias theory*, in which the same relation is attributed to variations in the probability (strength) of alternative percepts, due to frequency of sensory usage; (3) *central response bias theory*, in which the frequency/threshold relation is attributed to response probability, but the "responses" are mediating processes of some kind (implicit verbalizations, hypotheses, meanings); (4) *central perceptual bias theory*, in which the same relation is attributed to variation in the probability of percepts, but this probability is itself determined by feedback from mediation processes (implicit verbalizations, etc.).

Peripheral Response Bias

The strongest statement of the "spew hypothesis" in this context appeared in a paper by Goldiamond and Hawkins (1958). Their experiment, in which *psuedorecognition* was measured (formless blurs being flashes tachistoscopically), yielded the familiar threshold/frequency function for nonsense words given differential amounts of pretest practice. They interpreted this result as "... challenging a perceptual interpretation of the word-frequency-recognition relationship, where similar procedures are utilized [p. 462]." The implication was that the same hypothesis

would explain threshold phenomena when actual words were flashed. Even on commonsense grounds, this extreme position is untenable. As Neisser observes (1967, p. 166), with vocabularies numbering in the many thousands of words, the probabilities of guessing specific words would be infinitesimal and the process not fast enough to explain ordinary reading. Furthermore, the fact that misperceptions (false guesses) typically have the same general shape as the stimuli being flashes (see examples in our Tables 1 through 6) cannot be handled by a sheer "spew hypothesis."

More experimentally, Neisser (1954) demonstrated that preexposure to words like SEEN did not lower thresholds for their homonyms (same overt responses), like SCENE. Using words that are ambiguous as to noun or verb form-class, but are dominantly used one way or the other (e.g., *ship, land, box* being high noun and *look, can, move* being high verb), Roydes and Osgood (1972) found a significant interaction between induced "set" for noun vs verb and the dominant grammatical coding of the words; note that in this case the words are identical both as visual stimuli and as vocal responses. Zajonc and Nieuwenhuys (1964) took a different tack, testing the Goldiamond and Hawkins implication that real-recognition and pseudorecognition thresholds would display identical functions of prior-exposure frequency: in a direct comparison of real- (stimuli present) with pseudo- (stimuli absent) recognition thresholds, the functions proved to be clearly different.

The *reductio ad absurdum* of peripheral response bias theory in its pure form is that speakers of a language blind from birth should perform before a tachistoscope like speakers of that language with normal vision. A more reasonable version of this position is "fragment theory." Just as preexposure to a subset of the vocabulary reduces the number of alternatives for guessing, so should perception of fragments of words (e.g., initial consonant clusters like ST- - - -) serve to reduce uncertainty and thus lower measured thresholds. However, the nature of the fragments and the reduction in uncertainty they occasion are seldom specified (a rough count of the number of words in the *American College Dictionary* beginning with ST- yields an N of 750). Broadbent and Gregory (1971) used small, fixed vocabularies of three-letter words which the Ss were thoroughly familiar with, and varied the probabilities of the whole words or their included letters independently; they concluded that "the major influence on correct perception is clearly . . . the probability of a complete word rather than that of the letters from which it is composed [p. 13]."

In our Experiment I, the morphemes in the nonword combinations of List IV (e.g., *planial*) were deliberately selected to have much higher average usage frequencies than the multisyllabic words of List V (e.g., *plaster*) and hence should serve as tailor-made "fragments" for guessing (e.g., of *planning, cordial*, etc.). Yet the guessing thresholds—where the first response *whether*

word or nonword determined the recorded threshold value for the item—were significantly higher for the nonwords of List IV than for the words of List V. If "fragment theory" is valid, one wonders why the Ss didn't identify the morphemes separately at durations appropriate to their high frequencies and then simply reel them off (nonword combinations were easily pronounceable). All this is not to deny the validity of "reduction of alternatives" as a factor in tachistoscopic threshold determination (cf. Eriksen, 1956, 1958, 1960, in other connections), but rather to question both its sufficiency and the locus of its effect.

Peripheral Perceptual Bias

We use "perceptual bias" to refer to experience-based modifications on the sensory side of the nervous system (necessarily beyond the projection systems, i.e., Area 17 in the case of vision) that reflect the past frequencies of redundant input events. Hebb's notion of "cell-assemblies," which were assumed to be developed via experience in the postprojection "association areas," was an early proposal of this sort. It is necessary here to distinguish between the notions of *icon* and *percept*: The former (cf. Neisser, 1967, p. 20) is presumably the perseveration for a second or so of a stimulus-produced pattern of activity at the terminus of the projection system; it has been reported, e.g., by Crovitz and Davies (1962), that appropriate eye movements occur during this period, as if the stimulus were still present. The latter is a more central correlate of such projection-level activity (and this correlate can also be activated centrally, in which case it is called an *image*); a percept is also assumed to reverberate over some indefinite (but longer) period, to be susceptible to facilitation via central feedback, and to display certain Gestalt-like properties (e.g., closure). Controversies over the nature of percepts—e.g., template models, feature models, parallel processing models (cf. Neisser, 1967, Chap. 3)—continue but need not concern us here.

Regardless of one's theory of the matter, there is evidence for purely perceptual bias as defined above. One of the most convincing demonstrations was given by Postman and Rosenzweig (1956): Differential pretraining (0, 1, 2, 5, 15 trials) was given on three-letter nonword syllables in either visual or auditory modalities, and subsequent threshold estimations in *either the same or transfer* modalities were made. Although thresholds always fell as a function of amount of prior exercise (presumably a response-bias effect), they were significantly lower when tested in the *same* modalities as practiced (visual-visual, auditory-auditory) than in different modalities (auditory-visual, visual-auditory). In other words, above and beyond a common effect of response bias, there is additional facilitation when testing is within the same modality—which we attribute to peripheral perceptual bias. Convergent evidence is offered by Mewhort (1966), who found that increasing

the spacing between letters decreased the advantage of fourth-order approximations to English (e.g., VERNALIT) over zero-order approximations (e.g., YRULPZOC) in tachistoscopic perception; since the redundancy factor (response bias) remains constant, this again supports a perceptual bias interpretation.

A series of papers by Haber and Hershenson raise the question of whether perceptual bias is to be interpreted in terms of projection-level mechanisms (in the *icon*) or in terms of experience-based modifications (in the *percept*)—or both. The basic technique was to present a given verbal stimulus repeatedly at a constant flash duration, well below full-word recognition, with Ss instructed to report after each exposure what they could perceive *in terms of letters* (“seeing”) rather than what they thought the word might be (“guessing”); the intertrial intervals were “never less than 8 sec” (Haber & Hershenson, 1965). The basic result was evidence for what has been termed “microgenesis”—that is, gradual increase in the subjective clarity of the stimuli.

The results of a series of experiments can be summarized succinctly as follows: With flash duration constant at several levels (at Ss' minimum threshold and -5, +5, +10, and +15 msec around it), stimulus clarity increased with repetitions according to a negatively accelerated function (except for duration = -5), even if the early flashes appeared to be blanks (Haber & Hershenson, 1965). When English and pronounceable Turkish seven-letter and three-syllable words were compared (Hershenson & Haber, 1965), microgenesis was obtained in both cases, although clarity was consistently higher for English words at all repetition frequencies—the latter effect probably being due to “...cognitive feedback to the perceptual system [p.45].” To eliminate the obvious possibility of response bias, Ss were either shown the stimulus word for 5 sec prior to the test repetitions (prior knowledge) or not (no prior knowledge), and frequent vs rare English words were compared under these conditions (Haber, 1965); prior knowledge completely eliminated differences in subjective clarity for frequent vs rare words, but the microgenesis effect appeared both with and without prior knowledge. Hershenson (1969a) required Ss to maintain fixation at the centers of seven-letter displays (both words and nonwords), followed the same repetition and reporting procedures, and also varied presence or absence of prior knowledge; the typical growth functions were obtained, but clarity was also a function of closeness to the fixation point [on the fourth (middle) letter]. Notable was the finding that the closeness-to-fixation-point function (a) was least apparent for real English words (with no prior knowledge), and (b) was markedly reduced with prior knowledge (for all stimuli). Shifting the fixation point (from Letter 2, to Letter 4, to Letter 6) produced appropriate shifts in the locus of highest perceptibility (Hershenson, 1969b).

Certain of these findings, particularly the relation of

perceptibility to fixation point, clearly indicate a role of projection-level mechanisms—which is not surprising, of course—but this is obviously not a sufficient explanation of microgenesis. For one thing, there is the long interexposure intervals (about 10 sec) over which it occurs; for another, there are the differences in clarity in favor of English over Turkish words at all levels of repetition; for yet another, there is the fact that the effect of retinal location is markedly reduced with real English words or by prior knowledge. On the other hand, the fact that microgenesis is primarily an effect upon *percepts*, rather than due to response bias, seems guaranteed by its persistence under conditions of complete and immediate prior knowledge, as well as by the subjective experience of increasing clarity reported by Ss. Haber (1965, p. 286) suggests that there is a similarity between *ontogenesis* of percepts (Hebb's development of cell assemblies) and *microgenesis* of percepts (a kind of recency effect upon the tuning up of cell assemblies); but to conclude thusly would require demonstration that *random* repetitions of the verbal stimuli (rather than successive repetitions of each stimulus) also produce microgenesis. As far as we are aware, this remains to be shown.

In any case, it is clear that such peripheral perceptual bias is also insufficient as an explanation of the frequency/threshold relation.¹² When prior knowledge of the stimuli is *not* given, frequent and rare verbal stimuli show threshold differences throughout levels of repetition. The results of the present experiments are also relevant: In Experiment I, the higher frequency nonword morphemes and morpheme combinations had higher thresholds than the words with which they were compared; in Experiment II, three-letter words had significantly lower thresholds than either three-letter morphemes or trigrams equated for visual usage frequency; in Experiment III, nominal compounds like *peanut butter* (linguistically functioning like single words) had lower thresholds than nonsense compounds like *motion glass* in Experiment VI on their first presentations—yet their component words showed no threshold differences on *their* first presentations. None of these findings is explicable in terms of pure perceptual bias. But, again, insufficiency is not evidence against validity of this mechanism as a contributory factor.

Central Response Bias

Here, the “responses” are central mediating processes rather than peripheral motoric processes (overt or subvocal), but the basic determinant of the frequency/threshold function is still variations in the numbers and probabilities of responses in hierarchies of alternatives (i.e., degrees of uncertainty in the information theoretic sense). Neisser (1967, p. 119) states that the trouble with his homonym experiment (1954) was that “...no one who speaks of ‘response

bias' really has such a restricted definition in mind"—where the definition in question was *overt verbal report*. But he also might have added that no one who espouses response bias theory has been explicit about the *nature* of these central mediating processes, usually referring vaguely to "inner speech."

Most of the evidence that has been interpreted as supporting central response bias indicates the "central" character of the effect clearly enough but is ambiguous with respect to its locus—whether upon responses or upon percepts. For example, Postman and Conger's (1954) demonstration that meaningful three-letter words like MAN yield the typical frequency/threshold function whereas meaningless trigrams like MAS do not is ambiguous as to response or perceptual locus. Similarly, although a series of studies by Haber and his associates (cf. Haber, 1966) clearly indicate that "sets" for perceiving particular attributes of stimulus patterns depend upon central coding of the displays, they do not indicate whether the effect of the coding is upon responding alternatives or upon perceiving alternatives. The most direct attempt has been to compare the effects of giving "sets" just prior to and just after stimulus exposure—on the assumption that the former can affect perception as well as central response (memory, meaning, etc.) but the latter can affect only central processes (see Haber, 1966, for a review of this work). However, this assumption depends crucially upon the duration of the icon, and no one apparently has both varied the interval between exposure and postset and filled it with visual masking noise. Perhaps the strongest case for unambiguous response bias is Haber's (1965) demonstration that giving Ss immediate prior knowledge of the verbal stimulus completely eliminates differences in perceptual clarity for frequent vs rare English words; if S is provided with the precise response alternative (and N alternatives equals 1), then the frequency/threshold relation does disappear.

However, it is again obvious that central response bias is an insufficient mechanism. The many demonstrations of perceptual learning without verbal responses (or even specifiable responses of any kind) being involved are relevant here (see E. J. Gibson, 1953, for a review). The evidence presented above for pure perceptual effects—Postman and Rosenzweig's same vs transfer modalities/threshold results, Mewhort's visual spacing effects, the Haber-Hershenson demonstrations of microgenesis despite giving prior knowledge—is equally evidence *against* the sufficiency of any pure response-bias interpretation.

Central Perceptual Bias

Put simply, the difference between central-response and central-percept positions is this: in response-bias theory, mediating processes directly influence what one *says*; in perceptual-bias theory, mediating processes influence what one *sees* and only indirectly thereby

what one *says*. Again, most of the evidence purporting to demonstrate perceptual bias is ambiguous as to the locus and nature of the effect—including much of the work under the aegis of the New Look School (really, a very Old Look School, going back at least to Külpe's (1904) studies on the effects of attention on perception). Eriksen, in particular (1958, 1960, 1962), has made critical analyses in terms of the potential response biases in many of these types of studies.

Although the Neisser (1954) study with *homonyms* is unambiguous with respect to the insufficiency of the *peripheral* response-bias position, it is ambiguous as far as the central response-bias position is concerned. This is because the fragments in "set" words (SE— in SEEN) are different from those in their homonyms (SC— in SCENE), as Eriksen and Browne (1956) have argued in interpreting related studies. The Roydes and Osgood (1972) experiment with *homographs* which can function as either noun or verb (e.g., MAN dominantly coded noun; CAN dominantly coded verb; etc.) does seem to provide unambiguous evidence for central perceptual bias, since the word fragments are identical for the noun "set" (MAN as a noun) and the verb "set" (MAN as a verb) conditions; yet MAN had significantly lower thresholds under the noun "set" and CAN under the verb "set" (i.e., an interaction between "set" and dominant form-class coding).

The experiments reported here also provide unambiguous evidence for perceptual bias (with potential response bias controlled): in Experiment III, prior exposure to constituent words significantly lowered thresholds for their nominal compounds (e.g., COAST GUARD), but prior exposure to nominal compounds had no effect upon the thresholds of their constituent words; Experiment V demonstrated that prior exposure to the single words used in Experiment III *did* significantly lower thresholds for samples of the same words (the usual reduction-of-alternatives effect); and Experiment VII directly compared perceptual facilitation effects of prior exposure to nominal (COAST GUARD), ordinary (LIBRARY GUARD), and nonsense (CHEESE GUARD) combinations—the subsequent thresholds for the constituents of nominal compounds being significantly higher than for those of the other two types (with no differences between the latter). Since reduction of single-word response alternatives should have been equivalent for the three types of two-word combinations, we conclude that some central process must have influenced the perception of words in nominal compounds differently from words in ordinary and nonsense compounds. But central perceptual bias theory is also insufficient. It will not, for example, account for the results of Goldiamond and Hawkins (1958) when "blurs" replace verbal stimuli; nor will it account for the microgenesis effects obtained by Haber, Hershenson et al.

Fig. 2. A three-level model of information processing (with feedbacks).

An Unbiased Bias Theory

In the history of psychology, the term "perceptual" has been applied to phenomena that can be adequately handled with projection-level dynamics (like contour formation and brightness contrast), to phenomena that require the postulation of postprojectional integrative mechanisms (like closure), and to phenomena that rather obviously depend upon even more central mechanisms (e.g., "perceiving" emotion in a face or meaning in a TAT card). Although, as Hebb (1968, p. 468) points out, both anatomical and functional differences make the distinction between *sensation* (projection level) and more central processes relatively easy to maintain, distinction between what we shall call *perception* (integration level) and *meaning* (representation level) is not easily made anatomically and has rarely been made explicit functionally. Yet incorporation of the total data on word (and other) perception requires both a distinction between the latter levels and a specification of the modes of interaction between them.

Osgood (1957a,b, 1963) has argued that, on functional grounds, at least three levels of neural processing must be postulated, applying to both input (S) and output (R) sides of the behavioral equation. Figure 2 provides a schematic representation of such a

model. Most peripherally, the *projection systems* transfer information from receptor surfaces inward to sensory cortex and from motor cortex outward to effector surfaces. Since projection systems are essentially isomorphic and unmodifiable by experience, surface stimulations can be viewed as being in one-to-one relation with more central patterns of *sensory signals* (symbolized by $\dot{s}-\dot{s}-\dot{s}$ in Fig. 2), or *icons*, and surface respondings similarly with patterns of *motor signals* ($\dot{r}-\dot{r}-\dot{r}$), or *motons* (for lack of any established term). It is assumed that these signals at the cortical termini of sensory and motor projection systems have connections with neurons in still more central *integration systems* (probably the sensory and motor "association" areas). The relations between projection-level signals and integration-level events are *not* assumed to be isomorphic, but the interrelations among integration-level events *are* assumed to be modifiable by experience.

Without implying any particular neurology of the matter (however, see Hebb, 1949, and elsewhere for one possibility), the following functional principle for organization within integration systems can be suggested: *the greater the frequency with which sensory signals ($\dot{s}-\dot{s}-\dot{s}$) or motor signals ($\dot{r}-\dot{r}-\dot{r}$) have cooccurred in input or output experience, the greater will be the*

tendency for their postprojectional ($\bar{s}\bar{s}\bar{s}$) or preprojectional ($\bar{r}\bar{r}\bar{r}$) correlates to activate each other centrally. This principle could be viewed as a behavioristic attempt to account for the phenomena of perceptual organization (long the stronghold of Gestalt psychology), on the one hand, and of motor organization (as emphasized by Lashley, 1951, in his Hixon Symposium address), on the other. We refer to sensory integrations ($\bar{s}\bar{s}\bar{s}$) as *percepts* and to motor integrations ($\bar{r}\bar{r}\bar{r}$) as *programs* (following Lashley's, 1951, notion of "motor programming"). What this principle says, in effect, is that repeated redundancies in either sensory inputs or motor outputs will come to be mirrored in evocative integrations (closures, higher frequency) or predictive integrations (tunings, lower frequency) in the nervous system. Since this principle operates on both sides of the behavioral equation—producing both perceptual skills (e.g., word-form percepts) and motor skills (e.g., syllabic programs)—it can fairly be called an "unbiased" peripheral bias.

We refer to the most central mediating system (see Fig. 2) as the *representational level*, and the critical concept here is *meaning*. The total meaning of a sign (r_M/s_M) is conceived to be the near-simultaneous activation of a distinctive subset of mediator components [$m_1 \dots m_5 \dots m_{12} \dots m_i \dots m_n$]¹³; these componential subsets can also be construed as "bundles of semantic features" with no change in the functional properties of the theory. A principle relating meanings to both percepts ($M \leftrightarrow \bar{s}\bar{s}\bar{s}$) and programs ($M \leftrightarrow \bar{r}\bar{r}\bar{r}$) is as follows: *the greater the frequency with which representational mediation processes (meanings) have cooccurred with either sensory integrations (percepts) or motor integrations (programs) in experience, the greater will be the tendency for the former (r_M/s_M) to activate the latter ($\bar{s}\bar{s}\bar{s}$ and $\bar{r}\bar{r}\bar{r}$), and vice versa*. As suggested by the use of the symbol r_M/s_M (or simply M), the representational or symbolic level is indifferent with respect to the usual sensory/motor distinction, the same mediating process being functionally a "response" to input from the sensory integration system and functionally a "stimulus" to output in the motor integration system. It thus provides an "unbiased" central bias, tuning up both response programs (directly) and stimulus percepts (via feedback) equivalently. We may now consider some of the implications of this model for word perception.

Reciprocal Facilitation Effects

It will be noted that the statement above ends with "...and vice versa." This reflects the fact that representational mediators ($r_M = M = s_M$) can function as either subsequent or antecedent elements in associative relations, and this implies four types of facilitative effects:

(1) $\bar{s}\bar{s}\bar{s} \rightarrow M$. This "feed-forward" is the encoding of

percepts into meanings.¹⁴ In early language development, perceptual signs which have already acquired meanings are typically present while verbal signs are used by adults, and the word percepts are thereby cooccurrent with appropriate mediator component sets; in mature language performance, word perception under ordinary reading and listening conditions must be followed by appropriate meanings in milliseconds (probably well under 200 msec, given the within-cortex distances involved). *In this theory, percepts of words are meaningless until encoded into semantic feature patterns*. That the word is a meaningless unit at the integration level is suggested by the fact that the inverse frequency/threshold relation holds when frequency is summed over all form classes and senses of words, as in the Thorndike Lorge tables; since, as Zipf (1945) has shown, the shortest and most frequent words also have the largest numbers of different meanings, the frequency/threshold effect should tend to flatten out if total frequencies were divided by numbers of different meanings.

(2) $M \rightarrow \bar{r}\bar{r}\bar{r}$. This "feed-forward" is the decoding of meanings into programs for overt motor expression. In early language development, meanings produced by the perceptual signs of familiar objects (KITTY being seen) are typically present while imitative vocalizations (saying "kitty") are being made, and hence mediator component sets cooccur with appropriate motor programs for the particular language; in adult language users, there must again be near-simultaneous tuning up of motor programs for expression of meaning states—but overt vocalization is often inhibited, of course. This $M \rightarrow \bar{r}\bar{r}\bar{r}$ "feed-forward" is central response bias in this theory.

(3) $\bar{r}\bar{r}\bar{r} \rightarrow M$. This is a nonobvious, potential "feed-backward" from motor programs to the meaning system. In theory, if particular motor programs are set in motion by any means other than the normal "feed-forward" $M \rightarrow \bar{r}\bar{r}\bar{r}$ route (with or without overt expression), then the actor must tend to also experience the meanings appropriate for the programs (but probably less intensely). The other routes may be reflexive actions (the startle reflex, for example), posturings (doing an "Indian War Dance"), and imitations (spread of panic and empathic behaviors generally). One is reminded, of course, of the James-Lange theory of the emotions (one feels afraid because he runs); in recent times, Greenwald (1970) has rejuvenated William James's notion of ideomotor action.

(4) $M \rightarrow \bar{s}\bar{s}\bar{s}$. It is this "feed-backward" from meanings to percepts that constitutes central perceptual bias in the present theory. The impoverished sensory signals provided in ordinary scanning of the environment or by tachistoscopic flashes of words are susceptible to many alternative perceptual integrations. Any additional inputs to the integration system which, via neural convergence, facilitate one percept as against its

competitors should lower the closure-threshold for that particular perceptual integration. One would also predict from this theory that prerecognition guesses should tend to share semantic features with the correct word, particularly when they have the same general visual shape; to borrow an example from Neisser (1967, p. 133), if the impoverished signals of flashed EASTER are more likely to be guessed as SACRED than as PESTER, then semantic subception is supported, and a number of studies suggest this to be the case (Postman, Bruner, & McGinnies, 1948; Eriksen, Azuma, & Hicks, 1959; Johnson, Thomson, & Frincke, 1960).¹⁵ The basic issue, however, is whether such evidence requires a central *perceptual* bias ($M \rightarrow \overline{s-s-s}$) or can be explained in terms of central *motor* bias ($M \rightarrow \overline{r-r-r}$), which would be a form of "response bias." This theory includes both possibilities, and in fact predicts that *both* effects will usually be operative in word perception studies.

Semantic Feedback and Word Thresholds

How, in theory, would $M \rightarrow \overline{s-s-s}$ feedback lower tachistoscopic thresholds? Without prior knowledge or familiarization, Ss commonly report that at a certain point words seem to suddenly "leap into awareness" as wholes (Neisser, 1967, p. 108), but such is not the case for meaningless letter combinations. In theory, at that point in a series of exposures of increasing duration where there is enough information in the icon to activate enough integrative correlates to activate some M, *if this M is correct* (i.e., associated in prior experience with the stimulus word form), then its near-synchronous feedback will summate with the reverberating sensory integration and produce full closure—and the percept is suddenly clear. If the M is incorrect, then the feedback cannot converge with the icon-produced integration and there is no "leap into awareness."

Furthermore, if the integration associated with the incorrect M is similar in visual form to the actual stimulus, then *this* integration is enhanced and blocks recognition. This would be the reason why a reversed letter *in a word* takes longer to perceive than one *in a nonsense item* (Postman, Bruner, & Walk, 1951) and why a "trick" card (a black six of hearts) is *misperceived* as a normal card (black six of spades or red six of hearts) *before* it is correctly perceived (Bruner & Postman, 1949). This interplay between icon-determined and meaning-determined inputs to the sensory integration system is reminiscent of the Gestalt notions of *restraining forces*, peripherally determined, and *cohesive forces*, centrally determined, as they interact in perception (cf. Osgood, 1953, pp. 202-208).

Words vs Morphemes

This theory provides a basis for differences in the perceptual salience of words and morphemes. Speaking linguistically, morphemes are the smallest units of

language having meanings. It is in this way that nonword morphemes like *pre-* differ from equally frequent and visually familiar trigrams like *ple-*; the form *pre-* has semantic features which might be characterized as ⁺Time and ⁺Sequence (i.e., "prior" as in *preconscious* and *prevent*), whereas the form *ple-* does not (compare *plead*, *pleasant*, and *plenary*). But the features associated with nonword morphemes are both very limited in number and have *generalized* association with many words of otherwise diverse meanings (e.g., *confer*, *conduct*, *conclude*, *concert*). Words, on the other hand, have richer semantic coding (*husband*, for example, being simultaneously ⁺Concrete, ⁺Animate, . . . ⁺Sex, ⁺Marital), and these features serve to rather uniquely distinguish the meaning of the given form from most others in the language.

Referring back to Fig. 2, it can be seen that, whereas the feedback from s_M (total meaning) is precise, that from a single mediator component (m_i) is necessarily diffuse, having been *part* of the semantic feedback to many different word forms. This means that, at a near-threshold level of icon information, even if the meaning component(s) appropriate to a morpheme stimulus are activated, the feedback will be facilitative for a sizeable set of visual forms as well as for the morpheme form itself.

The three-level (with feedback) theory sketched above is functionally similar to theories of many who have worked in the field of perception generally (e.g., Bruner, Postman et al; Ames, Cantril, Hastorf, Ittleson, Kirkpatrick et al) as well as some who have worked primarily on word perception. Both Hersenson and Haber (1968) and Neisser (1967), for example, seem to arrive at positions which postulate interaction between peripheral and central (feedback) factors. Hersenson and Haber (1968) conclude their review of the role of the percept and perceptual processes in word recognition by saying: "The percept can best be thought of as a result of integrative and constructive processes . . . it is necessary to conclude that information is extracted from the percept and is encoded into a form useful in memory contact and analytic processing [which] may take part in the subsequent processes of search, matching, and identification which are probably involved in the recognition task [p. 23]." Neisser (1967, pp. 94-97, 114-115) also talks about central feedback into perceptual processing, using the notion of "figural synthesis," and relates it to such phenomena as physiognomic perception (e.g., attributing meanings to facial expressions), hallucination, illusion, and the conviction of Ss that letters missing from or distorted in familiar words in the tachistoscope were really *there* in the flash (that they saw an O in WRD or WURD).

The difference lies in the explicitness with which the nature of central coding and its feedback are specified. The closest Neisser comes to defining "figural synthesis" is to suggest that it has an intimate relation to focal attention (pp. 94-97) and that ". . . Penfield's electrodes

may have touched [its] mechanism [p. 169].” He also states that “genuinely visual effects will appear only if the cognitive unit with which the subject is familiar . . . is the unit actually employed in his visual synthesis.” but the nature of cognitive units remains obscure. Haber (1966) says that “...most visual stimuli are remembered by being encoded into previously learned linguistic units, usually words . . . [which] takes place while the stimulus (or its brief short-term memory) is still present [pp. 345-346].” and Hershenson and Haber (1968) equate what they call the “coded representation” of the percept ambiguously with “...visual image, verbal image, stored list, meaning structure, etc. [p. 23].” It seems obvious that central representations of “words” must be auditory, visual, or motor images, and *images are not a code*; they do not elaborate on “meaning structure.” It is the analysis of *meanings* into componential representational processes that enables us to make distinctive predictions about differences in perceptual salience of words vs morphemes and of nominal compounds (word-like) vs ordinary noun phrases.¹⁶

Convergent Evidence

Garner, Hake, and Eriksen (1956) have made the significant methodological point that, since we can directly observe only what goes into the organism (stimuli) and what comes out (responses), *convergent experimental operations* “... which allow the selection or elimination of alternative hypotheses or concepts which could explain an experimental result [p. 150]” are required. Ideally, convergent operations should be *orthogonal* (that is, utilize independent dimensions of variation) and yet involve tests of the same alternative hypotheses; operations that are *parallel* (that is, make further tests along the same dimension of variation) cannot converge to differentiate among the alternative hypotheses.

Within the studies reported here, Experiments I (single- and multisyllabic nonword morphemes vs matched words) and II (trigrams vs three-letter nonword morphemes vs three-letter words matched for usage frequency and shape) would fall somewhere between these extremes, using different types of material but testing along the same dimension; the same would hold for Experiments III (nominal compounds vs constituent words) and VI (nonsense compounds vs constituent words). On the other hand, the contrast between Experiments I/II (comparative saliences) and Experiments III/VI/VII (differential facilitation effects of prior exposure) does approach the independence ideal. In this final section, we will show that the findings reported here and independent findings of others do converge upon three theoretically significant conclusions: (1) that the humble word does have a special salience in language perception; (2) that meaning plays the crucial role in determining this salience; and

(3) that meaning exerts its effect via feedback into the perceptual process, quite apart from whatever effect it may ordinarily have upon response alternatives as well.

For Salience of the Word

Experiment I demonstrated that—despite loading visual usage frequency in favor of morphemes and despite using guessing thresholds (assumed to favor morphemes) as well as recognition thresholds—words had lower thresholds than morphemes in nearly all comparisons (e.g., *mental* and *mend* more salient than *ment*, and *plaster* more salient than *planial*). With the ratio of nonwords to words in the materials increased from two-fifths to two-thirds and with three-letter stimuli now equated for both usage frequency and shape, Experiment II demonstrated that words (like *pen*) had significantly lower thresholds than either nonword morphemes (like *pre*) or nonmorpheme trigrams (like *ple*); however, it was also found that the morphemes had significantly lower *recognition* thresholds than trigrams (but not lower guessing thresholds), suggesting that meaning has more influence upon recognizing than upon guessing (morphemes being more meaningful than trigrams).

Moving to units larger than the word, Experiment IV demonstrated that nominal compounds (like *post card*), having unique meanings as *wholes*, have just as low guessing and recognition thresholds as single words matched for length and usage frequency (like *pregnancy*). However, when the first-presentation recognition thresholds for nominal compounds like *stumbling block* (Experiment III) were compared with those for nonsense compounds like *sympathy block* (Experiment VI), the former had lower recognition thresholds than the latter¹⁷—*yet the thresholds for their single-word constituents on first presentation showed no difference whatsoever* (means were 18 msec in both cases, see Fig. 1). It thus appears to be the “word-like-ness” of the nonword morphemes as compared with the trigrams (in Experiment II) and of the nominal compounds as compared with the nonsense compounds (in Experiments III vs VI) that is responsible for their relative salience.

There is considerable evidence converging on the conclusion that words have special salience and that “word-like-ness” increases salience. Neisser (1967, 106-107) notes the well-established fact that the “span of apprehension” can encompass only four or five unrelated letters, yet familiar words as long as 20 letters can be readily apprehended within a 100-msec exposure. The greater the approximation to English (i.e., the more word-like), the lower the identification threshold (Miller, Bruner, & Postman, 1954). In this connection, it might be noted that pronounceability per se is not the whole explanation; in our own Experiments I and II, nonword morphemes, combinations of morphemes, and trigrams were all easily pronounceable (e.g., *ment*, *lossate*, *ple*).

Along quite different lines, Novik and Katz (1971) have shown that the scanning speed (search) for single letters is faster with words as the stimuli than with nonwords; furthermore, the slope of increasing search time as a function of size of the displays (number of letters) was much flatter for words than for nonwords.

Particularly relevant are studies by Reicher (1968) and Wheeler (1970). Reicher compared recognition for single letters presented alone with their recognition in either four-letter words or nonwords (quadrigrams of the same letters), using forced-choice between the correct letter and another letter which, in the real-world condition, would also yield a meaningful word (e.g., — — — D vs — — — K for testing D vs WORD vs ORWD). He found within-word letter recognition to be significantly superior to single-letter recognition, but within-quadrigram letter recognition to be inferior to single-letter recognition (although not significantly so). Wheeler modified Reicher's conditions in ways designed to test five alternative hypotheses which might explain Reicher's results *without* assuming any interactive processes between letters in larger language units. None of these alternatives was upheld, yet letters in words continued to be more readily perceived at a high level of significance. He concludes that "... word recognition cannot be analysed into a set of independent letter recognition processes [and] it seems appropriate to stop trying to explain away the phenomenon and, instead, to consider the implications [of intraword facilitation] for models of the human recognition system [p. 78]." Wheeler notes that neither his data nor Reicher's require that the larger units be words, but our results pretty effectively eliminate either trigrams or morphemes as candidates.

The general conclusion from the recognition-threshold results reported above seems to be that the more "word-like" a language unit, the higher will be its perceptual salience under tachistoscopic conditions. But what is the basis for the special salience of the word and, to lesser extent, word-like units? Even the "fragment" version of peripheral response bias seems to be eliminated by the results of our Experiments I and II (words vs higher frequency morphemes and words vs equal-frequency morphemes and trigrams, respectively). Since there is nothing in our version of peripheral perceptual bias—sensory integration as a function of experienced frequencies of cooccurrence of sensory signals—that would distinguish between words and other high-frequency units of language, the results of Experiment III vs Experiment VI (in which nominal compounds have lower thresholds than do nonsense compounds, yet their constituent words display no differences), as well as of Experiments I and II, cannot be accounted for. There remains the possibility that a purely perceptual, Gestalt-like principle of "common fate" is operating; single words and nominal compounds like *stumbling block* have redundant parts in the sense of coming and going together as wholes, whereas

morphemes, trigrams, and nonsense compounds do not. There is also the visual spacing of words in many languages to distinguish them perceptually from nonword morphemes and trigrams (but other languages, like Thai, get along nicely without such spacing). Both perceptual and response versions of central bias also remain viable.

For the Role of Meaning

The set of interlocking experiments reported here offer rather convincing evidence that *meaningfulness of units as wholes* is the critical determinant of perceptual salience. In Experiment II, we found that the more richly and uniquely coded three-letter words were more salient than the sparsely and nonuniquely coded morphemes, which in turn were more salient than the semantically uncoded trigrams. Similarly, for units larger than the word, the nominal compounds (which, like words, are uniquely and richly coded as wholes) in Experiment III have thresholds lower than those for the nonsense compounds (which are essentially meaningless as wholes) in Experiment VI, yet the thresholds for nominal compounds are as low as those for single words matched in length and frequency of usage, as shown in Experiment IV. At a commonsense level, it is obvious that we not only read and listen for meaning, but also scan the nonverbal environment of perceptual signs for their significance as well; it is not therefore surprising that we should "look" for meaning in the tachistoscope.

We now have some evidence for the validity of semantic features (r_m components, in theory), derived from purely linguistic analysis, which is independent of tachistoscopic threshold effects. One experimental technique (Sara Smith, 1971) utilized a "word-finding" task, in which a *target word* (interpersonal verb of a given feature composition) is presented with missing letters, e.g., —X—LO—T, either alone (control), along with *one cue word* (BULLY), along with *two cue words* having *redundant* semantic information with respect to the target word (BULLY and PERSECUTE), or along with two cue words having *nonredundant* information (BULLY and IMPOSE ON); for each such set, the *number* of features on which target and cue words differed in coding was also varied. The results were: (1) the less the number of features differentiating cues and targets, the shorter the time for word-finding; (2) nonredundant (complementary) cues facilitated word-finding more than did redundant cues. Another study (W. R. Smith, 1972) utilized a "false recognition" technique but used the same feature-coded interpersonal verbs as materials; in this case, the probability of falsely "recognizing" verbs not previously shown in the list increased with the number of features shared by old and new words. The above studies were carried out in our own laboratory. Completely independent studies by Delos Wickens and his associates (see particularly Wickens & Clark, 1968; Wickens, 1972) utilized as the

dependent variable release from proactive inhibition in short-term memory occasioned by a shift in materials; whereas shift in terms of many variables (e.g., upper- to lowercase printing, syntactical markers, verb tenses) produced little disinhibition, shifts from positive to negative coding (or the reverse) on the affective meaning features obtained in semantic differential work (Evaluation, Potency, and Activity) produced very marked and significant disinhibition.

Although the numerous studies on the effects of semantic "set" upon recognition thresholds for relevant vs irrelevant words are ambiguous with respect to perceptual vs response central bias, they do constitute convergent evidence for the role of meaning. Particularly convincing, however, is a study by Paivio and O'Neill (1970): Values of words on each of three variables—frequency (or rated familiarity), imagery concreteness, and meaningfulness (*m*)—were manipulated (high vs low values) while the other two were held constant. As would be predicted, measured frequency of usage (or rated familiarity) had the greatest influence on tachistoscopic thresholds; as should have also been expected, imagery concreteness was unrelated to thresholds when both frequency and meaningfulness were controlled—because images are of *perceptual* signs of objects (e.g., the visual form of a seen lion), not of their *verbal* signs (the visual form of the word LION). What was not expected by Paivio and O'Neill, but what is of peculiar significance to us, was the fact that *meaningfulness* (*m*) had a significant effect upon thresholds *when all other variables were controlled*; furthermore, in another experiment, both frequency and meaningfulness were varied orthogonally and a significant interaction between them was found—if frequency was high, then variation in *m* made little difference in thresholds (and conversely), but if frequency was low, then *m* had a marked effect (and conversely). Noble's *m* (meaningfulness) has been shown to be highly correlated with *polarization* (extremeness, or intensity of meaning) in semantic differential ratings (Jenkins, 1960). Thus, the critical point here is that, when the usage frequencies of words are relatively low, the sheer intensity of their mediating *Ms* significantly influences the thresholds for recognizing them. However, the *locus* of this meaningfulness effect—whether upon seeing (central perceptual bias) or upon saying (central response bias)—remains ambiguous.

For Semantic Feedback

Nominal compounds like *stumbling block* represent an emergent composition of semantic features (a kind of idiom, cf. Weinreich, 1966), and they function like single words linguistically, accepting insertions freely at their boundaries but not within (Greenberg, 1957); ordinary noun phrases like *copper block* are momentary conjunctions (linkings, according to Weinreich) of features that are congruent semantically; nonsense

compounds like *sympathy block* constitute either incongruent conjunctions of features or at least surprising ones (like *actor general*).

In Experiments III, VI, and VII, we compared the effects of prior exposure to these three types of noun phrases upon lowering the subsequently tested thresholds of their constituent words. In Experiment III, although prior exposure to the constituent words greatly facilitated subsequent perception of their nominal compounds, prior exposure to the compounds had no effect whatsoever on subsequent perception of their constituent words. (That this was not a "reverse ceiling effect"—constituent words already being near their minimal possible thresholds on their first presentations—was shown in Experiment V, where prior exposure to the same constituent words *did* significantly lower subsequent thresholds for a sample of them.) In Experiment VI, however, prior exposure to *nonsense* compounds also resulted in significant lowering of thresholds for their constituent words (see Fig. 1)—which meant that neither shifting from "long blurs" (compounds) to "short blurs" (single words) nor habituation to the longer recognition times of compounds could be responsible for the failure of *nominal* compounds to yield facilitation.

On the face of it, this is a most puzzling phenomenon: In Experiments III and VI, nominal and nonsense compounds, respectively, are carried through two correct recognitions before their constituent words are tested, yet there is facilitation in the latter (nonsense) case but not in the former (nominal) case. Why does "response bias" (reduction in alternatives) occur for nonsense compounds but not for nominal compounds? [A related paradox appeared in Experiment I, where the higher frequency nonword morpheme combinations of List IV had *higher* thresholds than the matched multisyllabic words of List V, and in Experiment IV, where the thresholds for nominal compounds (like *motion picture*) were *not lower* than those for single words matched in length and usage frequency (like *generalization*), despite the much higher frequencies of their components (*motion* and *picture*)—why are such familiar constituents not recognized individually, thereby lowering thresholds for the wholes?] In Experiment VII, we compared the effects of prior exposure (two correct recognitions) of nominal compounds (*stumbling block*), nonsense compounds (*sympathy block*), and ordinary noun phrases (*copper block*) upon the subsequent thresholds of their constituent single words. Ordinary noun phrases are congruent, meaningful combinations, in which, however—like nonsense compounds but unlike nominal compounds—the individual words retain their normal meanings (*copper block* is semantically equivalent to *copper* + *block*). As predicted, both ordinary noun phrases and nonsense compounds produced significantly more facilitation in perceiving their constituent words than did nominal compounds, and there were no

differences between the two former types.

These interlocking results lead to the following conclusions: (1) Words (and nonword morphemes) "lose their identities" when they appear in larger compounds which have unique meanings as wholes. That what is lost is their *semantic* identities is clear from the fact that the visual forms of the single words (or morphemes), once the compounds are recognized, are identical. (2) What is "reduced" by the prior exposure to compounds is the set of alternative *meanings*, not the set of alternative percepts (sensory) or programs (motor). This is shown by the fact that if the meanings of the constituent words are *not* included in the set (as is the case when nominal compounds are presented first), then prior familiarization with the visual forms of the single words has no facilitative effect upon their subsequently measured thresholds. So it appears that *constancy of meaning* between forms in compounds and forms in isolation is the necessary condition for facilitation of prior exposure on subsequent threshold.

However, there still remains the question of whether meaning is affecting "saying" directly (central response bias) or "seeing" via feedback as well (central perceptual bias). In other words, is it necessary to postulate central tuning of *percepts* *per se*? Other findings in Experiment VII bear on this matter: It will be recalled that to guarantee two prior recognitions of the compounds, they were each flashed once at a "short" (30-50 msec) duration and, if recognized, once again at this duration, but if *not* recognized, flashed twice at a "long" (double the "short") duration. A significantly larger percentage of nominal compounds (46%) were recognized at the "short" exposure than was the case for either nonsense (16%) or ordinary (24%) compounds. At the end of the experiment, Ss were tested for their recall of all compounds originally presented for familiarization; here, ordinary noun phrases were recalled nearly as well (30%) as nominal compounds (33%), and both significantly better than nonsense compounds (18%). Since recall depends upon adequate encoding of the stimuli at the time of exposure, and this was essentially equivalent for nominal and ordinary compounds, we cannot explain their difference in ease of recognition in terms of central response bias. Rather, it appears that it is the distinctive semantic feedback to the percepts of the word-like nominal compounds (*looking glass, real estate, attorney general*) that makes them "leap into awareness."

Hershenson and Haber (1968, p. 11) report that in their situation (repetitions at the same exposure duration) "... Ss are convinced that the stimulus energy is increasing—they say that the stimulus looks *brighter* and *clearer* with each exposure." In the same paper, in discussing the relations between "seeing" (clarity) and "saying" (recognizing), they report a reanalysis of earlier data (Haber & Hershenson, 1965), in which the trial on which recognition first occurred is designated *trial zero*, trials prior being given appropriate negative (-1, -2,

etc.) and trials after recognition positive designations (+1, +2, etc.). They state that "no greater clarity was found immediately after recognition than immediately before it—the rate of increase in clarity was clearly continuous through the point of correct recognition [pp. 22-23]." This would seem to be evidence against any "leap into awareness," but there are at least two reasons to question this evidence: First, the Ss were set for letter clarity judgments (giving their word "guess" after letter-reporting), and they were habituated to the gradual clearing with each exposure; second, we would expect the rate of clarity increase to be rather constant up to just prior to recognition (Trial -1) and then to "leap" at the point of recognition (Trial 0) and continue at a new rate. In other words, a continuous (new) rate from immediately prior through immediately post (-1 to 0 to +1) is actually expected from theory, but this is not the right test to make—rather the -2/-1 rate should be compared with the -1/0 rate.

The theory proposed here makes another prediction about subjective clarity that is relevant: Since heard words and seen words with the same referents are assumed to have near identical meanings (indeed, are probably as close as we come in language to true "synonyms"), one would expect signs in one modality to facilitate perception of near-threshold corresponding signs in the other. Apparently, no one has measured thresholds simultaneously in both auditory (masking noise) and visual (tachistoscopic exposures) modalities *for the same words* and compared them with single-modality threshold functions. However, W. M. Smith (1965a, b) found that near-synchronous, above-threshold presentation of auditory verbal stimuli—either spoken by E or by S himself, repeating E—facilitated tachistoscopic recognition of the "same" stimuli in the visual mode; only the thresholds for matched, not mismatched, three-syllable words yielded this effect. A response-bias explanation seems to be ruled out by the fact that this facilitative effect peters out as the visual exposure *is moved toward the end* of the temporal envelope of auditory stimulation.¹⁸

Evidence for such a cross-modality facilitation effect, but in the reverse direction, has recently been reported by Osgood and McGuigan (1973). A number of years ago, the first author (1957a) described how, while listening to a multivoice male chorus in a Gilbert and Sullivan operetta recording with his wife (who was following the libretto with her finger), what was unintelligible gibberish auditorily became perfectly clear when he glanced down at the printed words. It was not simply that he now *understood* what the chorus was singing—it suddenly *sounded* clear.

In the McGuigan laboratory, this "Gilbert and Sullivan" effect was put to test, using operatic choral materials: With surface electrodes placed on chin, lip, tongue, and right calf for recording EMGs (to check for peripheral motor feedback), Ss listened via earphones to 40-sec selections of choruses (rated either Hi-I or Lo-I

intelligibility, independently) while a blank slide (10 sec), a clearly printed matching *or* nonmatching slide (10 sec), another blank slide (10 sec), and then another matching *or* nonmatching slide (10 sec) were shown in succession. As a measure of *subjective clarity*, Ss pressed a button during periods when "they could clearly hear and understand the singing": as a measure of *objective clarity*, Ss were given a recognition test at the end of each 40-sec passage, this test including short phrases from the auditory materials (accompanied by either matching *or* nonmatching slides), from the nonmatching visual materials, and from other operettas. Ss were instructed to "pay attention to the music."

In essence, the results were as follows: Ss did reliably report the subjective "Gilbert and Sullivan effect," signaling understanding of the Lo-I passages 60% of the time when accompanied by matching slides and only 11% of the time when accompanied by nonmatching slides ($p < .05$). Recognition test scores (objective clarity) were higher while signaling understanding (48%) than while not (22%), this also being significant ($p < .05$). That Ss were paying attention to the auditory material and not the visual is indicated both (a) by the fact that they not only signaled understanding for Hi-I passages 86% of the time with matching slides, but also 84% of the time with nonmatching slides, and (b) by the fact that recognition-test scores for visually presented material on *nonmatching slides* during choral passages was very low (only 10% of the items). Since tongue EMGs were also significantly higher during signaled understanding, the possibility that long-loop motor feedback contributed to the subjective clarity effect remains viable—but central facilitation of auditory clarity via meaningful visual matching is demonstrated in either case.

In ordinary semantic feedback within the same modality (e.g., in tachistoscopic word perception), we have the same (single) icon and the same (appropriate) meaning (or M). In cross-modality semantic feedback (as in the W. M. Smith and Osgood and McGuigan studies above), we have different *icons* (one visual and one auditory) but the same M, and hence the possibility of reciprocal facilitation in both directions. One might expect the same effects from *similar* Ms—synonyms of test words, despite gross differences in visual shape, should lower thresholds when presented in the auditory modality (and vice versa)—but as far as we are aware this remains to be demonstrated.

What about where we have the same (single) icon but *different* Ms? The Roydes-Osgood study (1972), using grammatically ambiguous noun/verb forms, is a case in point; an interaction between form-class "set" and dominant codings of the words was found, but this is confounded somewhat by the fact that many of the homographs had very similar meanings as nouns or verbs (e.g., MAN, VIEW, NAME, MOVE, CALL). Nonverbal perception provides more convergent evidence, however: Ambiguous figures, like Boring's "wife and

monther-in-law," provide one clear case where the icon is identical (albeit ambiguous) yet the *percept* shifts with changes in M: in fact, with multiply ambiguous figures (like Walter Miles's Kinephantoscope) what one "sees" can be readily influenced by feeding the S different words, i.e., meanings (cf. Osgood, 1953, p. 221). Along different lines, Bruner, Postman, and Rodrigues (1951) showed that the apparent hue of an "orange" stimulus (determined by matching it with adjustments of red and yellow sectors on a color-wheel) could be influenced appropriately by varying the shape of the stimulus (e.g., like a tomato vs like a lemon), which, of course, changed its meaning as a perceptual sign. Similarly, Hastorf (1950) showed that the apparent distance of an ambiguous rectangular form could be influenced appropriately by telling Ss it was a calling card, an envelope, and so forth; it is also reported that Ss agree quite well on their estimated distances of meaningful objects of known size like packs of cigarettes, but not on "meaningless" objects of ambiguous size like star shapes (cf. Kilpatrick, 1952).

What about cases where we have *both different* (but similar) icons *and different* Ms? In word perception, this is the condition for competition and interference (and raising of thresholds): Havens and Foote (1963) have shown that rare words with familiar "shapemates," as Neisser calls them, have higher thresholds than equally rare words without such high-frequency "shapemates"—presumably due to facilitation of the percept of the higher frequency form. In the area of nonverbal perception, Osgood (1957a) has suggested that "regression toward the real object" (Thouless, 1931)—a well-known phenomenon in studies of perceptual constancy—is due to the fact that the meanings of perceptual signs are most strongly associated with the sensory integrations their objects produce *when being attended to* (e.g., that icon of an APPLE object when held at crooked-arm's inspection distance); in the continuous series of possible apparent sizes, that percept most facilitated by the convergence of icon-determined and meaning-determined inputs to the integration system will therefore be some compromise between them. And we are once again back to the Gestalt notion of interaction between "restraining" (iconic) and "cohesive" (meaningful) forces in the determination of what we perceive.

REFERENCES

- Broadbent, D. E., & Gregory, M. H. P. Effects on tachistoscopic independent variation of word probability and of letter probability. *Acta Psychologica*, 1971, 35, 1-14.
- Bruner, J. S., & Postman, L. On the perception of incongruity: A paradigm. *Journal of Personality*, 1949, 18, 206-228.
- Bruner, J. S., Postman, L., & Rodrigues, J. Expectation and the perception of color. *American Journal of Psychology*, 1951, 64, 216-227.

- Crovitz, H. F., & Davies, W. Tendencies to eye movement and perceptual accuracy. *Journal of Experimental Psychology*, 1962, 63, 495-498.
- Doherty, M. E., & Keeley, S. M. On the identification of repeatedly presented, brief visual stimuli. *Psychological Bulletin*, 1972, 78, 142-154.
- Eriksen, C. W. Subception: Fact or artifact? *Psychological Review*, 1956, 63, 74-80.
- Eriksen, C. W. Unconscious processes. In M. R. Jones (Ed.), *Nebraska symposium on motivation*. Lincoln: University of Nebraska Press, 1958.
- Eriksen, C. W. Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, 1960, 67, 279-300.
- Eriksen, C. W. Figments, fantasies, and follies: A search for the subconscious mind. In C. W. Eriksen (Ed.), *Behavior and awareness*. Durham: Duke University Press, 1962.
- Eriksen, C. W., Azuma, H., & Hicks, R. B. Verbal discrimination of pleasant and unpleasant stimuli prior to specific identification. *Journal of Abnormal & Social Psychology*, 1959, 59, 114-119.
- Eriksen, C. W., & Browne, C. T. An experimental and theoretical analysis of perceptual defense. *Journal of Abnormal & Social Psychology*, 1956, 52, 224-230.
- Ferguson, G. A. *Statistical analysis in psychology and education*. New York: McGraw-Hill, 1966.
- Garner, W. R., Hake, H. W., & Eriksen, C. W. Operationalism and the concept of perception. *Psychological Review*, 1956, 63, 149-159.
- Gibson, E. J. Improvement in perceptual judgments as a function of controlled practice or training. *Psychological Bulletin*, 1953, 50, 401-431.
- Goldiamond, I., & Hawkins, W. F. Vexierversuch: The log relationship between word-frequency and recognition obtained in the absence of stimulus words. *Journal of Experimental Psychology*, 1958, 56, 457-463.
- Greenberg, J. H. *Essays in linguistics*. Chicago: University of Chicago Press, 1957.
- Greenwald, A. G. Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review*, 1970, 77, 73-99.
- Haber, R. N. The effect of prior knowledge of the stimulus on word recognition processes. *Journal of Experimental Psychology*, 1965, 69, 282-286.
- Haber, R. N. Nature of the effect of set on perception. *Psychological Review*, 1966, 73, 335-351.
- Haber, R. N., & Hershenson, M. The effect of repeated brief exposures on the growth of a percept. *Journal of Experimental Psychology*, 1965, 69, 40-46.
- Hastorf, A. H. The influence of suggestion on the relationship between stimulus size and perceived distance. *Journal of Psychology*, 1950, 29, 195-217.
- Havens, L. L., & Foote, W. E. The effect of competition on visual duration threshold and its independence of stimulus frequency. *Journal of Experimental Psychology*, 1963, 65, 6-11.
- Hebb, D. O. *The organization of behavior: A neurophysiological theory*. New York: Wiley, 1949.
- Hebb, D. O. Concerning imagery. *Psychological Review*, 1968, 75, 466-477.
- Hershenson, M. Stimulus structure, cognitive structure, and the perception of letter arrays. *Journal of Experimental Psychology*, 1969a, 79, 327-335.
- Hershenson, M. Perception of letter arrays as a function of absolute retinal locus. *Journal of Experimental Psychology*, 1969b, 80, 201-202.
- Hershenson, M., & Haber, R. N. The role of meaning in the perception of briefly exposed words. *Canadian Journal of Psychology*, 1965, 19, 42-46.
- Hershenson, M., & Haber, R. N. The role of the percept and of perceptual processes in word-recognition. Unpublished paper, 1968.
- Howes, D. H., & Solomon, R. L. Visual duration threshold as a function of word probability. *Journal of Experimental Psychology*, 1951, 41, 401-410.
- Jenkins, J. J. Degree of polarization and scores on the principal factors for concepts in the semantic atlas. *American Journal of Psychology*, 1960, 73, 274-279.
- Johnson, R. C., Thomson, C. W., & Frincke, G. Word values, word frequency, and visual duration thresholds. *Psychological Review*, 1960, 67, 332-342.
- Kilpatrick, F. P. (Ed.) *Human behavior from the transactional point of view*. Hanover, N.H.: Institute for Associated Research, 1952.
- King-Ellison, P., & Jenkins, J. J. The durational threshold of visual recognition as a function of word-frequency. *American Journal of Psychology*, 1954, 67, 700-703.
- Kölpe, O. Versuche über Abstraktion. Berlin International Congress of Experimental Psychology, 1904, 56-68.
- Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon Symposium*. New York: Wiley, 1951.
- Lees, R. B. *The grammar of English nominalizations*. Bloomington: University of Indiana Press, 1960.
- McGinnies, E., Comer, P. B., & Lacey, O. L. Visual recognition thresholds as a function of word length and word frequency. *Journal of Experimental Psychology*, 1952, 44, 65-69.
- Mewhort, D. J. K. Sequential redundancy and letter spacing as determinants of tachistoscopic recognition. *Canadian Journal of Psychology*, 1966, 20, 435-444.
- Miller, G. A., Bruner, J. S., & Postman, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology*, 1954, 50, 129-139.
- Neisser, U. An experimental distinction between perceptual process and verbal response. *Journal of Experimental Psychology*, 1954, 47, 399-402.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- Novik, N., & Katz, L. High-speed visual scanning of words and nonwords. *Journal of Experimental Psychology*, 1971, 91, 350-353.
- Osgood, C. E. *Method and theory in experimental psychology*. New York: Oxford, 1953.
- Osgood, C. E. A behavioristic analysis of perception and meaning as cognitive phenomena. In J. S. Bruner (Ed.), *Contemporary approaches to cognition*. Cambridge, Mass: Harvard University Press, 1957a.
- Osgood, C. E. Motivational dynamics of language behavior. In M. R. Jones (Ed.), *Nebraska symposium on motivation*. Lincoln: University of Nebraska Press, 1957b.
- Osgood, C. E. On understanding and creating sentences. *American Psychologist*, 1963, 18, 735-751.
- Osgood, C. E. Interpersonal verbs and interpersonal behavior. In J. Cowan (Ed.), *Studies in language and thought*. Tucson: University of Arizona Press, 1970.
- Osgood, C. E. Exploration in semantic space: A personal diary. *Journal of Social Issues*, 1971, 27, 5-64.
- Osgood, C. E., & McGuigan, F. J. Psychophysiological correlates of meaning: Essences or tracers? In F. J. McGuigan and R. A. Schoonover (Eds.), *The psychophysiology of thinking*. New York: Academic Press, 1973.
- Osgood, C. E., & Sebeok, T. A. (Eds.) *Psycholinguistics: A survey of theory and research problems*. *Journal of Abnormal & Social Psychology (Supplement)*, 1954, 49, Part 2.
- Paivio, A., & O'Neill, B. J. Visual recognition thresholds and dimensions of word meaning. *Perception & Psychophysics*, 1970, 8, 273-275.
- Postman, L., Bruner, J. S., & McGinnies, E. Personal values as selective factors in perception. *Journal of Abnormal & Social Psychology*, 1948, 43, 142-154.
- Postman, L., Bruner, J. S., & Walk, R. D. The perception of error. *British Journal of Psychology*, 1951, 42, 1-10.
- Postman, L., & Conger, B. Verbal Habits and the visual recognition of words. *Science*, 1954, 119, 671-673.

- Postman, L., & Rosenzweig, M. R. Practice and transfer in the visual and auditory recognition of verbal stimuli. *American Journal of Psychology*, 1956, 69, 209-226.
- Postman, L., & Rosenzweig, M. R. Perceptual recognition of words. *Journal of Speech & Hearing Disorders*, 1957, 22, 245-253.
- Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 1970, 81, 275-280.
- Roydes, R. L., & Osgood, C. E. Effects of grammatical form-class set upon perception of grammatically ambiguous English words. *Journal of Psycholinguistic Research*, 1972, 1, 165-174.
- Shannon, C. E., & Weaver, W. *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949.
- Solomon, R. L., & Postman, L. Frequency of usage as a determinant of recognition thresholds for words. *Journal of Experimental Psychology*, 1952, 43, 195-201.
- Smith, S. W. Word-finding in the presence of semantically related cues. Doctoral dissertation, University of Illinois, 1971.
- Smith, W. M. Visual recognition: Facilitation of seeing by saying. *Psychonomic Science*, 1965a, 2, 57-58.
- Smith, W. M. Visual recognition: Facilitation of seeing by hearing. *Psychonomic Science*, 1965b, 2, 157-158.
- Smith, W. R. Influence of semantic feature similarity on false recognition of interpersonal verbs. Doctoral dissertation, University of Illinois, 1972.
- Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words*. New York: Columbia University Press, 1944.
- Thouless, R. H. Phenomenal regression to the real object. *British Journal of Psychology*, 1931, 21, 339-359.
- Underwood, B. J., & Schulz, R. W. *Meaningfulness and verbal learning*. New York: Lippincott, 1960.
- Weinreich, U. Explorations in semantic theory. In T. A. Sebeok (Ed.), *Current trends in linguistics*. The Hague: Mouton, 1966.
- West, M. A. *A general service list of English words*. London: Longmans, Green, 1966.
- Wheeler, D. D. Processes in word recognition. *Cognitive Psychology*, 1970, 1, 59-85.
- Wickens, D. D. Characteristics of word encoding. In A. W. Melton and E. Martin (Eds.), *Coding processes in human memory*. New York: Winston, 1972.
- Wickens, D. D., & Clark, S. Osgood dimensions as an encoding class in short-term memory. *Journal of Experimental Psychology*, 1968, 78, 580-584.
- Zajonc, R. B., & Nieuwenhuys, B. Relationship between word frequency and recognition: Perceptual process or response bias? *Journal of Experimental Psychology*, 1964, 67, 276-285.
- Zipf, G. K. The meaning-frequency relationship of words. *Journal of General Psychology*, 1945, 33, 251-256.
- procedure used was a sign test for two correlated samples (Ferguson, 1966).
5. Frequency estimates were obtained from tables in Underwood and Schulz, 1960.
6. Due to a change in bulb, the luminance for Experiments II and IV (35 mL) was somewhat different from that in all others (25 mL), but this will have no effect on the results, since comparisons across experiments are limited to those having the same luminance level.
7. To save space, we henceforth simply summarize the number of comparisons (either for paired item means or for Ss over all item pairs involved) which are in the direction predicted (+), are opposed (-), or are tied (=), and indicate the significance level (by binomial expansion test or by sign test for correlated samples).
8. Since no single word of near-equal frequency could be found to match the 20-letter *lieutenant commander*, we used an 18-letter word, *characteristically*.
9. The only reason for sampling 15 words rather than retesting all 26 was practical—to fit the experimental task into the usual subject hour.
10. Since we were not using the usual method of ascending limits, perceptual thresholds were much higher than for the same items in Experiments III and VI (e.g., only 15.9% successful recognitions of nonsense compounds with exposure durations of 30 msec or higher as compared with a mean threshold of 26 msec for the same compounds in Experiment VI).
11. This comparison across experiments is possible because identical luminance levels (25 mL), types of Ss, and general procedures were employed.
12. Although a recent paper by Doherty and Keeley (1972) is considered by the authors to be critical of the Hershenson, Haber et al conclusions, there are a number of reasons to discount it: (a) In many cases, nonverbal stimuli (e.g., Landolt rings) were used, and even the verbal materials were small sets of letters (A, T, U); (b) the interexposure intervals were much shorter than the 10 sec or so of Hershenson et al, since Ss responded only after sets of exposures, and possibly were within icon reverberation limits (hence tapping projection-level effects).
13. This is not the place for a detailed discussion of how the meanings of signs and their mediator components are established; the interested reader is referred to Osgood, 1971 (pp. 11-18), for a concise, recent statement.
14. Over the past decade, there has been a reversal in the usage of the terms "encoding" and "decoding"—from "reading out of" and "reading into" a system, respectively (cf. Shannon & Weaver, 1949; Osgood & Sebeok, 1954), to the opposite uses; here, we follow present usage, hence percepts are *encoded* into semantic representations.
15. Arabic would provide an ideal test here, since in this language *radicals* (usually triplets of consonants, e.g., N-Z-R) actually signal semantic classes, members of which are formed by inserting *serviles* (usually vowels, e.g., NaZaRa (he looked), NaZiR (inspector), NaZaRan (viewing), NaZaRi (visionary), etc.).
16. It is interesting that, in his important *Cognitive Psychology* (1967), Neisser specifically discusses *meaning* only in connection with reading (pp. 134-137), and the concept doesn't even appear in the subject index!
17. It is worth noting that the lower *guessing* thresholds for nonsense compounds—due to Ss' catching on to the fact that they were nonsense and hence unpredictable—did not contribute to lowering their *recognition* thresholds.
18. Personal communication from W. M. Smith to F. J. McGuigan (see Osgood & McGuigan, 1973).

NOTES

1. Smith's study was designed in connection with the first author's graduate seminar in psycholinguistics; we wish to express our indebtedness to him for providing us with both materials and preliminary data.

2. The use of *grobion* was an error in the preparation of materials that failed to be caught (*grob-* not being an English morpheme). The "test" against *problem* is not included in the statistical comparisons of Lists IV and V, since it would tend to favor our hypothesis.

3. Unless otherwise specified below, the types of Ss and general procedures for subsequent experiments were the same as given here for Experiment I.

4. This is a more sensitive test, comparing the means of individual Ss for all items of a given type with their means for another type, each S serving as his own control. The statistical

(Received for publication February 19, 1973;
revision received September 27, 1973.)