Author's Copy. 20

Reprinted from
The British Journal of Mathematical and Statistical Psychology,
Vol. 21, Part 1, p. 21, May 1968

A RETEST OF THE RESPONSE-BIAS EXPLANATION OF THE WORD-FREQUENCY EFFECT

By John Morton

Medical Research Council Applied Psychology Research Unit, Cambridge

This paper investigates whether the relation between the frequency of occurrence of a word and its sensory threshold can properly be ascribed to response bias. Seven models for the recognition process are tested against data of Brown & Rubenstein (1961). The models are of two kinds: probabilistic single-threshold models and information-processing models. It is concluded that the data can only be accounted for within the former class of models by assuming a differential effect of the stimulus as well as response bias. For the class of information-processing models, the data require the assumptions that there is equal sensitivity for words of all frequencies and a lower criterion for more common words. The relation between the two acceptable models makes it apparent that the notions of 'stimulus effect' and 'response effect' are by no means as clear as has previously been thought.

1. Introduction

When individual words are presented acoustically in noise or exposed visually for a brief period, the probability of an observer identifying them correctly increases with the frequency of occurrence of the stimulus word in the language. It is at present a matter of dispute as to whether this result is due to stimulus effects or to response bias (Broadbent, 1967).

Ultimately it will be shown in the present paper that the rigid dichotomy between stimulus and response factors is, in part, a semantic artifact, for it is possible to devise information-processing models to which neither of these terms can be singly and unambiguously applied. The dichotomy, as it is usually employed, may be restated in the following way.

1. It is possible that high-frequency words are more distinctive than words of lower frequency either by virtue of their 'pure' stimulus properties or because the recognition system is organized in such a way as to enable common stimuli to be processed more readily than less common stimuli.

2. The response-bias explanation of the word-frequency effect would deny the above possibilities, attributing the results entirely to the observer's tendency to use high-frequency words more often. Thus in the absence of any stimulus information, or if such information were minimal and insufficient for a clear 'percept', the observer would respond according to his predispositions, that is 'guess', and would respond preferentially with high-frequency words.

Brown & Rubenstein (1961) claim that when response bias is controlled the word-frequency effect is eliminated. To eliminate the possibility of response bias affecting the data they introduce the notion of frequency 'intervals', interpreting the response-bias explanation as implying 'that groups of words which differ in word-frequency are random samples from the same population of acoustical or visual parameters'. This assumption, termed the Principle of Acoustical Equivalence, has the force that no information as to the frequency interval of the word can be gained from the stimulus and any word is likely to be confused on the basis of its stimulus properties with a word from any of the frequency intervals.

They divided up the 6500 monosyllabic content words into 13 intervals of 500 words each by their frequency of occurrence. Six subjects were then presented with 1300 words (100 from each interval in random order) at two speech-to-noise ratios (S/N). Three numbers were determined for each frequency interval i:

c_i the number of correct responses to stimuli in the ith interval;

s_i the number of responses whose frequency class corresponded to that of the stimulus (this would include both correct and incorrect responses);

r_i the total number of responses in the ith interval to all stimuli.

TABLE 1. RAW DATA FROM BROWN & RUBENSTEIN

(Maximum values of c_i and s_i are 600. Maximum value of r_i is 7800.)

	S/N=0 db			S/N=10 db		
i	c_i	s_i	r_i	c_i	s_i	r_i
1	21	34	77	325	332	367
2	23	28	81	313	318	361
3	55	68	135	390	394	448
4	29	45	161	344	355	437
5	29	39	159	357	371	450
6	47	62	227	415	424	534
7	55	76	255	414	419	522
8	78	114	409	445	460	617
9	105	145	608	428	447	618
10	101	155	674	427	442	657
11	125	183	924	473	479	742
12	129	213	1301	415	433	739
13	242	388	2428	487	527	883
0			361			425

The values of these three variables at the two S/N are given in Table 1. Brown & Rubenstein derive equations to test the response-bias explanation; these equations do not, however, follow from the premises and are discussed briefly in the Appendix. In addition, they base some conclusions on the linearity of the relation between c_i and s_i. Since the latter is included in the former, such a procedure seems invalid: for the higher S/N it virtually involves plotting c_i against itself. In the present paper a different datum is used: e_i, the number of error responses which are in the correct frequency interval. This equals $s_i - c_i$. The values of e_i at the higher S/N are too small, and are too irregular with respect to i to have value in a graph. The relation between e_i and c_i for the lower S/N is depicted in Fig. 1. The data are plotted in terms of proportions, and the variables will be treated in this way in the body of the paper.

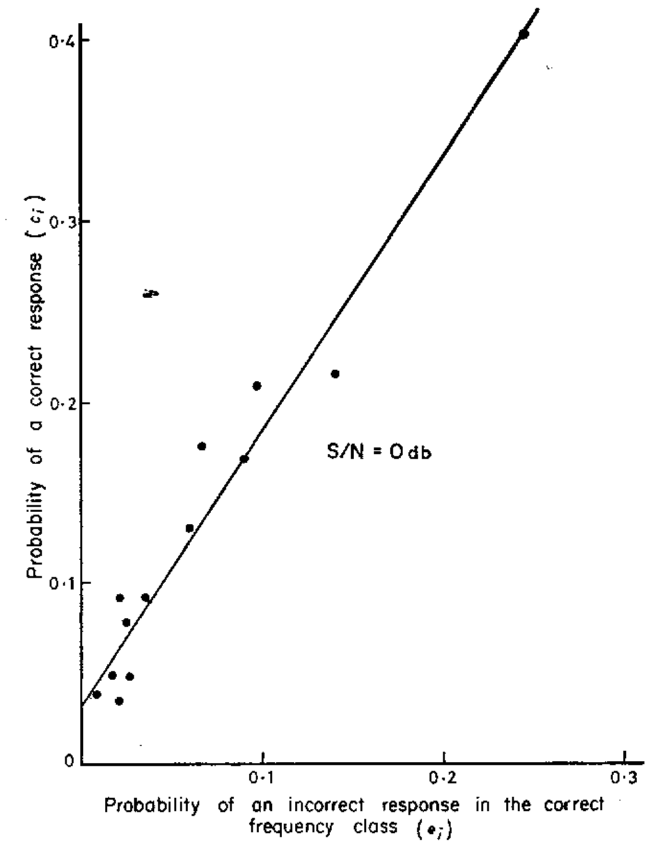

FIGURE 1. Performance in the recognition of visually presented words of different frequencies of occurrence. (Data from Brown & Rubenstein, 1961.)

The line drawn is the best-fitting straight line:

$$c_i = 0{\cdot}033 + 1{\cdot}526\, e_i;$$

and the best-fitting quadratic is:

$$c_i = 0{\cdot}024 + 1{\cdot}755\, e_i - 1{\cdot}356\, e_i^2.$$

A further datum used below is: p_i, the total number or proportion of error responses in the ith interval to *all stimuli*. This equals $r_i - c_i$ in terms of numbers and $r_i - c_i/n$ in terms of proportions. A plot of c_i against p_i is given in Fig. 2 for both S/N values.

Brown & Rubenstein's data will be used to test a number of models for word recognition. Some of these models have been explicitly or implicitly used by other authors (see Broadbent, 1967, for references). Theory E, however, has not previously been examined in any detail.

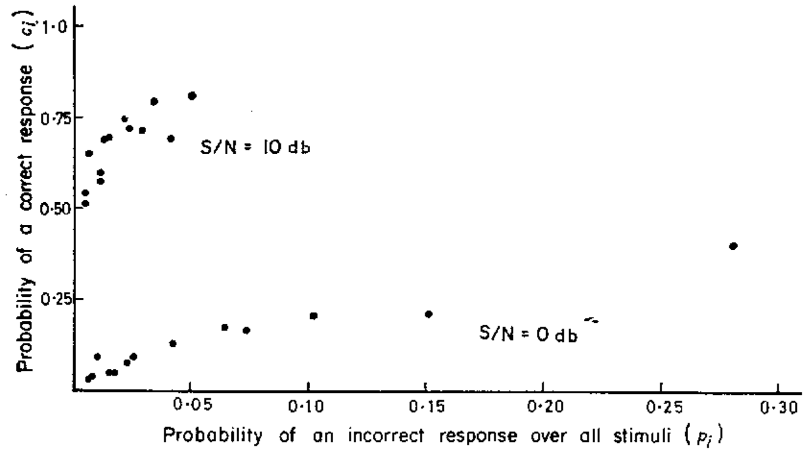

FIGURE 2. Recognition performance with words of different frequencies of occurrence. (Data from Brown & Rubenstein, 1961.)

2. Single-Threshold Theories

The first group of models examined are single-threshold models. According to these, the subject guesses in the absence of a percept.

Theory A. Either a word is recognized or it is not. If it is not, then the observer guesses according to his dispositions which are independent of the stimulus. By the response-bias theory, the probability of a word being recognized, x, increases with S/N but is independent of the interval i. If the stimulus is not recognized, then the observer guesses, the proportion of guesses in interval i being g_i. Thus with n intervals and M words in each interval, given a stimulus in the ith interval, it follows that:

$$s_i = x + (1-x)g_i, \tag{1}$$

$$c_i = x + \frac{1}{M}(1-x)g_i, \tag{2}$$

$$e_i = \left(1 - \frac{1}{M}\right)(1-x)g_i. \tag{3}$$

Therefore, eliminating g_i from eqns. (2) and (3):

$$c_i = \frac{1}{M-1}e_i + x. \tag{4}$$

Also, given stimuli equiprobable in i,

$$r_i = (x/n) + (1-x)g_i \tag{5}$$

and

$$p_i = \left(1 - \frac{1}{nM}\right)(1-x)g_i. \tag{6}$$

Therefore,

$$c_i = [n/(nM-1)]p_i + x. \tag{7}$$

Eqn. (4) should correspond to Fig. 1; however, the slope of the data line is too steep (since in the equation $M = 500$ and $n = 13$). Likewise, the slopes predicted for the relations depicted in Fig. 2 from eqn. (7) are too steep, even assuming linearity. The simple theory is therefore incorrect.

Theory B. If the stimulus contains information relating to its own class then the proportion of responses in the correct class increases. This affects the second term on the right-hand side of eqns. (1), (2) and (3), but leaves eqn. (4) and the slope of eqn. (7) unchanged. The modification of Theory A must therefore also be abandoned.

Theory C. It is possible, and intuitively obvious, that on a number of occasions, while insufficient stimulus information is available to enable the observer to recognize the stimulus, there will, however, be enough to enable him to make a choice from a restricted set of alternatives. Thus if the initial phoneme /t/ is heard clearly, only words beginning with 't' will be considered by the observer. By the theory of acoustical equivalence, however, this restricted set of alternatives will be equally distributed among the separate response intervals. The forms of eqns. (1) and (5) will therefore remain identical. Eqn. (2) will become:

$$c_i = x + \frac{1}{m}(1-x)g_i, \tag{2 a}$$

where m is the number of words in the reduced set in each interval. Thus eqn. (4) becomes:

$$c_i = \frac{1}{m-1}e_i + x, \tag{4 a}$$

and eqn. (7) becomes:

$$c_i = [n/(nm-1)]p_i + x. \tag{7 a}$$

In eqns. (4 *a*) and (7 *a*), m has a lower limit of unity, since it refers to the number of possible words *in the correct class* in the original definition of c_i; the occurrence of m in the last two equations is determined entirely by this definition. Clearly, in the limit the strict definition of the principle of acoustical equivalence breaks down, for it is possible that only two words are possible responses, one of which will be the correct one. The modified principle would state that the other word is on average equally likely to come from any *other* interval. This limiting modification does not, however, affect the equations.

Given that $m \geqslant 1$, the slope of the function defined by eqn. (7 *a*) has a maximum of $13/12 = 1{\cdot}083$. If the (unlikely) assumption is made that the data in Fig. 2 are indeed linear, the equations of the best-fitting straight lines are:

$$c_i = 5{\cdot}09\, p_i + 0{\cdot}56 \quad (\text{for S/N} = 10 \text{ db}),$$

$$c_i = 1{\cdot}29\, p_i + 0{\cdot}05 \quad (\text{for S/N} = 0 \text{ db}).$$

Both these lines have slopes in excess of the permitted maximum and model C is thus untenable on the basis of Brown & Rubenstein's data.

Theory D. A combination of models B and C still yields eqn. (4 *a*) and a form of eqn. (7 *a*) with the same slope. This theory cannot be accepted either. It would appear, then, that threshold models which deny the stimulus any role in the word-frequency effect cannot account for the data.

Theory E. By denying the common assumption of the above group of theories, the primary data can be accounted for quite satisfactorily. It will be accepted, then, that the proportion of correct responses resulting from the stimulus alone, x_i, is a function of the interval i as well as of S/N, initially assuming that the stimulus contains no information as to its interval and that the stimulus is either recognized or guessed at. The following equations are then obtained:

$$c_i = x_i + \frac{1}{M}(1 - x_i) g_i, \tag{8}$$

$$e_i = \left(1 - \frac{1}{M}\right)(1 - x_i) g_i, \tag{9}$$

$$p_i = \left(1 - \frac{1}{nM}\right)(1 - X) g_i, \tag{10}$$

where $X = \sum_{i=1}^{n} (x_i/n)$. From these equations the following can be derived:

$$c_i = \frac{1}{M+1} e_i + x_i, \tag{11}$$

$$c_i = p_i \frac{n(1 - x_i)}{nM(1 - X) - (1 - x_i)}. \tag{12}$$

Since x_i is now a function of i, eqn. (8) can now represent Fig. 1 without the assumption that M is small, and the slope limitations on the form of Fig. 2 no longer apply. Equally it is not necessary for either function to be linear. There are then no gross objections to this model.

If x_i is eliminated from the modified eqns. (2) and (3), then

$$c_i = 1 - e_i \frac{M - g}{g(M - 1)}. \tag{13}$$

Thus

$$g_i = \frac{e_i M}{e_i + (1 - c_i)(M - 1)}. \tag{14}$$

If M is large, we may simplify and rearrange this equation to obtain:

$$g_i = e_i/(1 - c_i). \tag{14 a}$$

Values of g_i were calculated from this equation and are tabulated in Table 2 (*a*). If the model were accurate, the sum of the values of g_i plus g_0, the probability of producing a response outside the classification system (i.e. blanks, partial responses and polysyllabic words) should be unity. This value is, in fact, exceeded at both S/N. This is not a consequence of the approximation, since the effect of reducing the value of M in eqn. (11) is to increase the calculated values of g for given values of c_i and e_i. The values of g being too high could result from the values of either c_i or e_i being too high compared with the values of the other. One possible explanation of this is that words do carry some information concerning their frequency of occurrence which would have the effect of increasing e_i relative to c_i. A second anomaly in Table 2 (*a*) is that values of g for high frequency words ($i = 10$ to 13) are proportionally lower at the high S/N. These two points will be discussed below.

TABLE 2. VALUES FOR g_i CALCULATED FROM (*a*) EQN. (14 *a*), (*b*) EQN. (15)

	(*a*) $g_i = e_i/(1 - c_i)$		(*b*) $g_i = p_i/(1 - c_i)$	
i	S/N = 0	S/N = 10	S/N = 0	S/N = 10
1	0·023	0·026	0·008	0·016
2	0·009	0·017	0·009	0·019
3	0·024	0·019	0·012	0·023
4	0·028	0·043	0·020	0·036
5	0·018	0·058	0·019	0·036
6	0·027	0·049	0·027	0·046
7	0·038	0·027	0·030	0·042
8	0·069	0·097	0·049	0·067
9	0·081	0·111	0·074	0·074
10	0·108	0·087	0·084	0·089
11	0·122	0·047	0·118	0·104
12	0·178	0·097	0·173	0·126
13	0·408	0·354	0·323	0·154
Sum	1·133	1·032	0·946	0·834
Others	0·054	0·166	0·054	0·166
Sum	1·187	1·198	1·000	1·000

If M is large, then $c_i \doteqdot x_i$, and $p_i \doteqdot (1-X)g_i$ from eqns. (8) and (10). The value of X can then be estimated from the average number correct and we may write:

$$g_i = p_i/(1-C), \tag{15}$$

where $C = \sum c_i/n$. These estimates of g are given in Table 2 (*b*). The form of eqn. (15) ensures that these values sum to unity. As with Table 2 (*a*), the values of g_i for higher values of i are consistently lower at the higher S/N.

3. Information-Processing Models

The above models have as their basis the probability of certain events taking place. If instead we regard word recognition in terms of information processing we obtain different equations, use different parameters, and discover different implications for the notions of stimulus effect and response effect.

The basic unit to be considered is one which accumulates evidence from the stimulus. When more than a certain amount of evidence is available, the word is recognized. On the assumption that all the units are inherently noisy, we can treat them in terms of signal-detection theory (Green & Swets, 1966). The two relevant parameters are then the amount of stimulus information available, d', and the thresholds or criteria of the recognition units, β.

Theory F. A model which was purely stimulus based, in the sense described in the introduction, would say that the value of d' was a function of the frequency of occurrence of the word as well as of S/N. 'Active' theories of speech perception (Halle & Stevens, 1962; Liberman, Cooper, Harris & MacNeilage, 1963) would have to predict such a result. It is uncertain as to whether these theories would have the criteria dependent on frequency.

Theory G. An alternative model is one in which d' remains constant for all words but the value of β changes with frequency. In such a model, less stimulus evidence is required for a high-frequency response.

By analysing the data under discussion in terms of response strengths (Luce, 1959), a form of analysis roughly equivalent to signal-detection theory but considerably easier to handle mathematically, it will be shown that model F cannot handle the data, whereas model G can.

A matrix can be constructed showing, for a stimulus of a particular frequency interval, the relative response strengths for the correct response, for incorrect responses of all intervals, and for other responses. This matrix is given in Table 3. V_i is the relative response strength for a word in the absence of a stimulus and is related to the criteria; α is a multiplicative factor representing the stimulus strength which is related to d'. The value of α represents a contribution to the response strength of the correct response compared with the *average* of other words in the interval. This model also assumes the acoustic equivalence of intervals. The probability of a particular response is then given by the response strengths for that response divided by the sum of all response strengths for that stimulus. Given a stimulus in interval i, the proportion of correct responses, c_i, is given by:

$$c_i = \frac{\alpha V_i}{(\alpha-1)V_i + (MZ+1)}, \tag{16}$$

where $Z = \sum V_j$ $(j = 1, \ldots, i, \ldots, n)$, which we may more simply write as

$$c_i = \alpha V_i / T_i. \tag{16 a}$$

The proportion of incorrect responses in the correct frequency interval, e_i, is given by:

$$e_i = [(M-1)V_i]/T_i. \tag{17}$$

From these two equations the following relation is obtained:

$$c_i = e_i[\alpha/(M-1)]. \tag{18}$$

Table 3. Response Strengths

Stimulus interval	Correct	Incorrect responses by interval					Other responses	Total
		1	...	i	...	n		
1	$\alpha_1 V_1$	$(M-1)V_1$	...	MV_i	...	MV_n	1	T_1
i	$\alpha_i V_i$	MV_1	...	$(M-1)V_i$	...	MV_n	1	T_i
n	$\alpha_n V_n$	MV_1	...	MV_i	...	$(M-1)V_n$	1	T_n

Fig. 1 shows this prediction to be largely upheld. The possible negative curvature and the non-zero intercept could be interpreted simply as due to the subjects recognizing words which they had never before seen and so could not produce as incorrect responses; they would effectively be nonsense syllables. Any recognition model would have to treat such words as exceptional cases. It might be noted that such an explanation would serve to account for the non-unit sum of g_i in Table 2 (*a*).

If model F were correct, Fig. 1 would be positively accelerating since α would increase as i (and so c_i and e_i) increased. This prediction is emphatically rejected by the data.

4. A Comparison of Models E and G

The conclusion that α is constant across i may seem to be in conflict with the result from model E where it was seen that x_i had to vary with i. However, while α can only be interpreted in terms of the amount of information available from the stimulus, x_i merely specifies the proportion of correct responses which result from the stimulus. From eqns. (8), (9), (16) and (17) the relations among the parameters in the two models are given by:

$$x_i = \frac{(\alpha-1)V_i}{(\alpha-1)V_i+(MZ+1)}, \tag{19}$$

$$g_i = MV_i/(MZ+1). \tag{20}$$

Thus x_i represents a combination of both stimulus and response strength factors and g_i is proportional to V_i, being specifically the probability of making a response in the ith interval in the absence of a stimulus, i.e. when $\alpha=1$.

If eqn. (20) is summed over i we obtain:

$$\sum_i g_i = MZ/(MZ+1). \tag{21}$$

However, it has already been observed that the sum of the obtained values of g is greater than unity. Thus theory G also requires some modification to account completely for the data if the explanation following eqn. (18) is not acceptable. Two possibilities remain to be examined: that there is some information in the stimulus concerning its interval, and that an incompletely recognized stimulus results in a reduced set of alternatives.

For theory E, this would give equations such as:

$$c_i = x_i + \frac{1}{m}(1-x_i)[y_i+(1-y_i)g_i], \tag{22}$$

where m is the size of the reduced set and y_i is the proportion of occasions when a stimulus in the ith interval is not recognized but the correct interval is obtained from the stimulus. It is reasonable to assume that both m and y_i are functions of the S/N. From such equations we can derive values for g_i from the relation:

$$g_i = \frac{me_i - y_i[(1-c_i)(m-1)+e_i]}{(1-y_i)[(1-c_i)(m-1)+e_i]}. \tag{23}$$

As m is reduced, the values of g_i increase, and as y increases g decreases. When Brown & Rubenstein's data are entered in this equation, it turns out that there are no values of y and m such that $\sum g < 1$ for a constant y, which also give positive values for all g_i; this comes about since, if y is increased to reduce the values of g_i, the right-hand negative side of the numerator exceeds the positive part for low values of i. Therefore, within this model, the value of y must be a function of i; that is, an unrecognized high-frequency word gives more information about its frequency interval than does a low-frequency word. This modification does not affect the acceptability of theories C and D since eqn. (7 *a*) is unaffected.

With theory G, interval information would be expressed by a multiplicative factor on the response strengths for all responses in the correct interval. This factor will be called δ. A relationship between c_i and e_i is obtained which is

unchanged by δ but does include the size of the reduced set of alternatives m. This is

$$c_i = e_i[\alpha/(m-1)]. \tag{24}$$

The equivalent of eqn. (20) is the following:

$$\frac{mV}{mZ+1} = \frac{me_i}{(m-1)[\delta(1-e_i)]+me_i}. \tag{25}$$

This equation must produce positive values for V_i. As δ is increased, the value of the function falls, and as m decreases the value rises: this is analogous to the properties of eqn. (23). For any m, δ can be adjusted to fulfil the restrictions on the sum of the left-hand side of the above equation (cf. eqn. (21)) without having to postulate that δ varies with i.

By either of theories E and G, Brown & Rubenstein's data can only be accounted for by assuming that the stimulus contains information about its interval. With theory E, it is in addition necessary to propose that such information is greater with high- than with low-frequency words. Such a proposition could work in practice through partial information and as such would be associated with the size of the reduced set which was considered above. For example, if all that was recognized clearly was the letter 'z', then the set of possible responses would be relatively biased in favour of low-frequency words. On the other hand, to give a hypothetical example, since the necessary statistics are not available: if the letter sequence '–the–' were recognized clearly, the set of alternatives could be biased in favour of high-frequency words.

The large change in the relative sizes of the entries in Table 2 at the higher S/N, which affects g_i in theory E and V_i in theory G, requires a further construct. The observation amounts to saying that as more stimulus information becomes available, proportionately fewer high-frequency words are produced erroneously. This could be accounted for in both the theories by postulating that whereas the amount of interval information given by a stimulus is greater for high- than low-frequency words, the difference becomes less at higher S/N ratios. This would apply to either y_i or δ in the two theories.

An alternative way of explaining this finding would be to assume that the recognition-system, while requiring less positive stimulus information for a high-frequency response to be made, also requires less negative stimulus information for such a response to be rejected. For lower-frequency words the system would be relatively indifferent to negative information, being only concerned to accept positive evidence. Such a proposition would require a far more detailed specification of the recognition system to be tested.

5. Conclusions

It has been shown that while a simple guessing model cannot account for the word-frequency effect, a more complex one can fit the data. The disadvantage of this theory compared with the one derived from signal-detection theory is that there are two monotonically related variables in the theory, x_i and g_i. Model G, on the other hand, predicts the relation between c_i and e_i from its initial postulates and requires for the main part only one variable, V_i. To this extent, then, model G is preferable. The testing of more detailed predictions will have to wait upon the collection of more precise data.

The nature of the model underlying theory G is such that it is impossible to talk about the results being due either to 'stimulus effects' or 'response effects'. Because of what are traditionally called 'response predispositions', the system requires less stimulus information to produce the same internal event for a high-frequency word than for a low-frequency word. It is not necessary, however, to talk in terms of a two-stage process; the system described in theory G is a single-stage process, as far as the word-frequency effect is concerned. Equally, the discussion as to whether the word-frequency effect has its origins in the frequency of experience or the frequency of emission of the words (Rubenstein & Aborn, 1960) is unnecessary. Within the preferred system, such a distinction becomes irrelevant (Morton, 1964; Morton & Broadbent, 1967). These conclusions agree with Broadbent (1967), who examined theories A, C, F and G, concluding that only theory G could account for his data.

The model underlying theory G can also be applied to the interaction of stimulus and context information in word recognition and the predictions derived from the theory are reasonably upheld (Morton, 1967). One version of this model has been stated axiomatically by Morton (1969) 8.

Appendix

In discussing their Principle of Acoustic Equivalence, Brown & Rubenstein claim 'a response from a particular interval will be given just as often to words from one (different) interval as to words from another (different) interval'. They express this statement as the equation:

$$e_{i/j} = br_i + d \quad (j \neq i),$$

where $e_{i/j}$ is the probability of a response in the ith interval to a stimulus in the jth interval.

In fact, the above equation does not follow from the principle of acoustical equivalence. Consider the case of four word-frequency intervals which gives rise to the matrix in Table 4, where the entries correspond to the number of responses; and w, x, y and z are responses in the correct interval. Brown & Rubenstein's equation implies that $a=h=k$, and $b=e=l$, since $e_{i/j}$ is constant for j, the (incorrect) stimulus interval. The principle of acoustical equivalence, however, only implies that the frequency interval of an unrecognized stimulus has no effect upon the frequency interval of the response. That is, that the *ratios* of responses in two incorrect intervals remain constant; thus $a/b=k/l$. The equations they subsequently derive and their test of their data are thus invalid.

Table 4

		Response interval				
		1	2	3	4	Total
Stimulus interval	1	w	a	b	c	N
	2	d	x	e	f	N
	3	g	h	y	i	N
	4	j	k	l	z	N

Acknowledgements

This article was partly written while the author was on leave at the Center for Research on Language and Language Behavior, Ann Arbor, Michigan, U.S.A. and supported by a grant from the U.S. Office of Education. I am grateful to Dr H. Rubenstein for making his original data available to me. This article has benefited from discussions with Dr D. E. Broadbent.

References

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychol. Rev.* **74,** 1–15.

Brown, C. R. & Rubenstein, H. (1961). Test of response bias explanation of word-frequency effect. *Science N.Y.* **133,** 280–281.

Green, D. M. & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Halle, M. & Stevens, K. N. (1962). Speech recognition: A model and a program for research. *I.R.E. Trans. Inf. Theory* **IT-8,** 155–159.

Liberman, A. M., Cooper, F. S., Harris, K. S. & MacNeilage, P. F. (1963). A motor theory of speech perception. In C. G. M. Fant (ed.), *Proceedings of the Speech Communication Seminar.* Stockholm: Speech Transmission Laboratories, Royal Institute of Technology.

Luce, R. D. (1959). *Individual Choice Behavior.* New York: Wiley.

Morton, J. (1964). A preliminary functional model for language behaviour. *Int. Audiol.* **3,** 216–225.

Morton, J. (1969). On the interaction of information in word recognition. *Psychol. Rev.* **76,** (in press).

Morton, J. (1968). Grammar and computation in language behaviour. In *Progress Report No.* 5, C.R.L.L.B., University of Michigan.

Morton, J. & Broadbent, D. E. (1967). Passive vs. active recognition models or Is your homunculus really necessary? In W. Wathen-Dunn (ed.), *Models for the Perception of Speech and Visual Form.* Cambridge, Mass.: M.I.T. Press.

Rubenstein, H. & Aborn, M. (1960). Psycholinguistics. *Ann. Rev. Psychol.* **11,** 291–322.