Proc. 5th int. Congr. phon. Sci., Münster 1964, pp. 306–311
(S. Karger, Basel/New York 1965).

From the Department of Phonetics, University College London

The Dependence of Stress Judgments on Vowel Formant Structure

By D. B. Fry, London

The experiments reported in this paper attempted to explore the part played by vowel quality in stress judgments obtained from English listeners. Versions of the word-pairs *object*, *contract*, *subject* and *digest* were synthesized in which there was systematic variation of the frequency of the first and second formants in the first syllable of *object*, *contract* and *digest* and the second syllable of *object* and *subject*. Variations in vowel duration ratio were introduced in the same stimuli in order to provide a means of estimating the weight to be assigned to the changes in formant structure. The specification of the test stimuli is contained in the information shown in Table I. The fundamental frequency of the periodic sounds was kept constant at 120 cps. throughout. The overall intensity of syllables was regulated so that the maximum intensity in the two syllables of a test word was equal and a constant difference of 6 db. between formant 1 and formant 2 was maintained throughout. In Table I, f(1), f(2) and f(3) refer to frequencies used in the first syllable of *object*, *contract* and *digest*, f(4), f(5) and f(6) to those in the second syllable of *object* and *subject;* f(1) and f(4) are the formant arrangements likely to give the lowest number of 'noun' judgments. Duration ratios are labelled d(1), d(2) and d(3); d(1) indicates the duration ratio likely to give the lowest number of 'noun' judgments.

The stimuli were made into a listening test in which each stimulus occurred once. They were in random order and were preceded by five practice items. Stimuli succeeded each other at intervals of 2 sec with a gap of 10 sec after each set of 10 items. No carrier sentence or number anouncement was used. Stress judgments were

Table I

Experimental Duration Ratios (V1/V2)

	d(1)	d(2)	d(3)
object	0.45	0.66	1.33
contract	0.2	0.53	0.8
digest	0.63	1.0	1.3
subject	0.25	0.6	1.0

Experimental F1/F2 Values (cps.)

		First Syllable				Second Syllable		
		F1	F2	F1	F2	F1	F2	F3
object	f(1)	570	1380			570	1980	
	f(2)	600	1260			570	1980	
	f(3)	600	1020			570	1980	
contract	f(1)	570	1380			720	1800	2520
	f(2)	600	1260			720	1800	2520
	f(3)	600	1020			720	1800	2520
digest	f(1)	480	2040			570	1980	
	f(2)	840	1440	➢ 570	1800	570	1980	
	f(3)	840	1440	➢ 480	2040	570	1980	
object	f(4)	600	1020			570	1980	
	f(5)	600	1020			540	2160	
	f(6)	600	1020			400	2280	
subject	f(4)	720	1320			570	1980	
	f(5)	720	1320			540	2160	
	f(6)	720	1320			400	2280	

obtained from one hundred subjects who were all young speakers of Southern English, nearly all brought up in the south or in the Midlands.

In presenting the results of these experiments, a recurring difficulty is the problem of comparing the scale of variation in the two dimensions. In the present case, it can be said that the range of variation in F1/F2 was as wide as it could be without introducing a marked unnaturalness in the vowels which provided the end-point of each range; that is to say that in the first syllable of *object* and *contract*, for example, the F1/F2 values at the ends of the range gave as clear an [o] and as clear an [ə] as could be obtained. It has been shown in an earlier paper that context must be expected to play a considerable part in any judgments based on vowel quality. In previous experiments, it has been found consistently that this con-

text effect works in the direction of contrast, that is to say that the perceptual difference between vowels is increased by their juxtaposition (*Fry et al.*, 1962). If we assume that the same tendency holds good with regard to the stimuli in the present experiments, this would mean that an [o], f(3), would sound more back and more open when it followed the corresponding f(2) or f(1), and an [ə], f(1), would sound more central when it followed the corresponding f(2) or f(3). The effect would therefore be to force the judgments at one end of the scale further in the direction of 'noun' and at the other end, further in the direction of 'verb', thereby increasing the apparent effect of the variation in F1/F2. That some influence of this kind was at work is suggested by the rather anomalous results obtained with the middle F1/F2 value in some word-pairs, where in one case f(2) produced slightly more 'nouns' than f(3) and in two other cases the same number of 'nouns' as f(1).

The relative effect of the formant changes and the duration changes can be judged from the range of percentage 'noun' judgments given in Table II. In every case the increase in 'noun' judgments produced by the change in duration ratio is greatly in excess of that produced by the change in formant frequencies. There is no

Table II

	Percentage 'nouns'	
	f(1)–(3)	d(1)–(3)
object	45–62	32–72
contract	35–52	12–72
digest	33–56	14–76
	f(4)–(6)	d(1)–(3)
object	60–65	39–84
subject	57–63	31–82

doubt that in the conditions of this experiment, the weight of the duration cue is very considerably greater than that of the formant structure cue. This fact is expressed in Figure 1 where the 'noun' scores for all word-pairs are pooled and the effect of the three steps of duration change and three steps of formant change are abstracted. The difficulty of comparing or equating the two scales has already been pointed out; here the steps of change are set off arbitrarily on the horizontal scale as though they were equal, with the duration

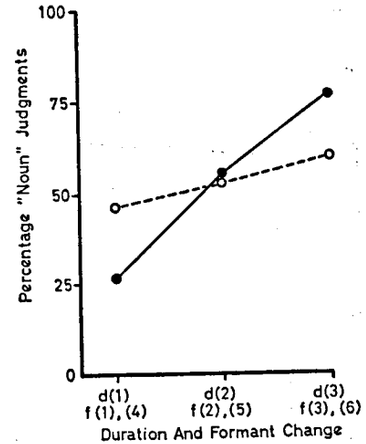

Fig. 1. The effect of changes in duration ratio and in F1/F2 plot on the percentage of "noun" judgments, scores for all word-pairs and all subjects pooled. The continuous curve refers to duration ratio and the dotted curve to formant structure.

and formant values giving the greatest number of 'verb' judgments nearest the origin. In the case of the formant values, f(1) and (4), f(2) and (5), and f(3) and (6) are pooled.

It seemed clear from preliminary experiments that there was a certain asymmetricality with respect to the effect of formant change in the first and the second syllables of the words; vowel change in the first syllable had a greater effect on the judgments than change in the second syllable. This fact is shown in Figure 2 where the effects of f(1)–(3) and of f(4)–(6) are plotted separately. This difference might be explained simply by the fact that in the particular words employed in the test the formant change in the second syllables was smaller than that in the first, judged either phonetically or by the crude test of the difference between F1 and F2 in cycles per second. The change from [o] to [ə] would probably strike many people as being perceptually greater than the change from [e] to [i]; in the stimuli used in the test, the frequency difference between F1 and F2 changed in the first progression from 420 to 810 cps. and in the second from 1410 to 1880 cps., a much smaller proportional change. When listening to the synthetic stimuli, however, it is difficult to resist the impression that there is something in the situation which renders vowel change in the second syllable really less effective than change in the first syllable. It is not easy to see how this can be check-

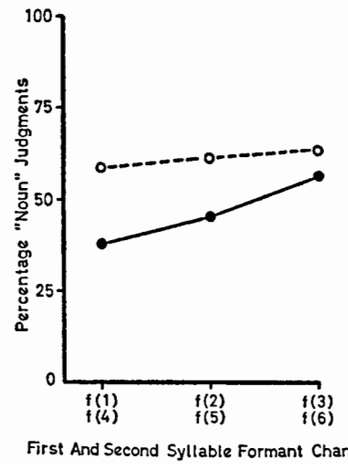

Fig. 2. The effect of changes in formant structure in the first and second syllable. The continuous curve refers to change in the first syllable, and the dotted curve to change in the second.

ed further by experiment except perhaps by synthesizing a series in which the vowel quality in the second syllable of such word-pairs as *subject* and *object* varied from [e] to [ə]. The latter pronunciation has a certain regional flavour but it might be possible to find out whether the effect could be enhanced in this way.

The Relative Weight of Formant Structure and Duration Ratio

The difficulty of arriving at any quantitative estimate of the relative weights of stress cues is aggravated in the case of formant changes by the fundamental problem of formulating a reasonable method of quantizing distances in the F1/F2 space. Any arrangement of scales chosen for this purpose is necessarily somewhat arbitrary and, in the absence of much fundamental work on the perceptual scales relevant to speech, is not likely to be very closely related to impressions of quality. In the present experiments, an attempt to arrive at some notion of the relations between the duration and the formant cues has been made by using the method outlined in a previous paper (see *Fry*, 1958) of plotting the logit values for percentages resulting from changes in duration ratio and in F1/F2 plot. In this way it is possible to find out approximately what change in duration ratio has an effect on the stress judgments equal

to that produced by the total change of formants used experimentally in any word-pair. These computations show that the swing in listeners' judgments due to the whole range of formant shift in the first syllable of *object* and of *digest* could be effected by a change in duration ratio of about 0.4 in the same words; for the first syllable of *contract*, the equivalent duration ratio change is about 0.16 and for the second syllable of both *object* and *subject*, about 0.1. If these ratios are compared with the ranges shown in Table I, it will be seen that they constitute a small proportion of the total ranges used experimentally. Values given previously as a basis for comparing the effects of duration and intensity in the same word-pairs were as follows: a difference of 20 db. between the two syllables of the stimulus was about equivalent to a difference in duration ratio of 0.4 for *object*, 0.16 for *digest*, 0.35 for *contract* and 0.6 for *subject*. A difference of 20 db. is rather large, and also it must be remembered that the stimuli in the two experiments were not identical, but a comparison of the two sets of figures suggests at least that the formant structure cue for stress may in fact be less effective than the intensity cue. A firmer conclusion on this point should, however, await the results of a direct experimental comparison of the two kinds of cue in action.

References

Fry, D. B.: Experiments in the perception of stress. Language and Speech *1:* 126–152 (1958).

Fry, D. B.; Abramson, A. S.; Eimas, P. D. and *Liberman, A. M.:* The identification and discrimination of synthetic vowels. Language and Speech *5:* 171–189 (1962).

Author's address: Prof. D. B. Fry. Department of Phonetics, University College London, *London* (England).