ACOUSTIC CORRELATES OF STRESS

JOHN MORTON and WIKTOR JASSEM

Applied Psychology Research Unit, Cambridge and Acoustic Phonetics Laboratory, Polish Academy of Sciences, Poznan

The use of the term ' stress ' is reviewed, and inconsistencies are pointed out. An experiment is described in which synthetic nonsense syllables of the forms /sisi/, /sɔsɔ/ and /sasa/ were played to subjects who were asked to judge the position of the stress. The parameters of fundamental frequency, intensity and duration were varied systematically.

The results revealed a high level of consistency between the (naive) subjects. Variations in fundamental frequency produced far greater effects than variations in either intensity or duration, a syllable being marked as stressed if it differed from the ' context ' fundamental. A raised fundamental was more efficient than a lowered one. In general the more intense and longer syllables were more likely to be marked as stressed, the exception being when a syllable was reduced in duration by 40%, when some listeners marked it as stressed.

The use and limitations of the results and method are discussed.

The concept of ' stress ' may be considered in several aspects ; its phonetic nature, its structural (phonemic) status, its morphological or syntactical function (or both), its distribution within rhythmic units (often co-extensive with word) and its distribution within intonation pattern units (often co-extensive with sentences or clauses and syntactic parts thereof). It would appear that the most basic question is that of *what stress is*. It seems to be pointless to consider what stress does in a language or where it falls without being reasonably sure what it is that falls or does something. Although stress is such an important concept not only for phonetics but also for general linguistics, the discussion of its nature is often given little room in works dealing with language generally or even with general phonetics and the result leads to some confusion. To take an example or two almost at random, Bloomfield (1957) describes the nature of stress in one short subordinate clause : " Thus, in English, when we combine several simple elements of speech into a word of two or more syllables, we always use a secondary phoneme of *stress* which consists in speaking one of these syllables louder than the other or others . . . " (p. 90). K. L. Pike (1943) avoids the issue completely in the first part of his *Phonetics*, which is a pity because it is still the most instructive critical analysis of general phonetic concepts. Admittedly, Pike, in this volume, is almost exclusively concerned with the ' segmental ' or ' sound quality ' features, but he does not (and probably could not) ignore stress altogether in this book. Yet all he has to say about its nature (again an almost marginal remark) is : " A *stress group* is a sequence of several syllables one of which, the stressed syllable, has a much stronger initiator pressure than do the others ". O. von Essen (1953) devotes several paragraphs to stress (he uses the term " Akzent " rather than the vernacular " Betonung ") without formally defining it. Its nature is vaguely implied in these opening explanations : " Bei gesprochenen Wörtern, die aus mehreren prosodischen Einheiten bestehen, pflegt eine Lautgruppe durch irgendeine Art der *Hervorhebung* über die anderen hinauszuragen und sie damit zu übergreifen, sie ist das beherrschende Glied, die übrigen sind die beherrschten Glieder. Bei Anwendung der Hervorhebung durch die Sprechstärke ergibt sich so eine dynamische Abstufung der Glieder . . . " (p. 118).[1]

The confusion is multiplied when formal definitions of ' stress ' are considered. At one extreme we have Jones (1956) : " Stress may be described as the degree of force with which a sound or syllable is uttered. It is essentially a subjective action " (p. 245). This statement is a minimal linkage between the linguistic concept of stress and the speaker. Jones and his purest followers would not consider it necessary for there to be any neuro-muscular or acoustic result of this ' force '.

Kingdon (1958) defining stress as " the force employed in uttering a syllable, giving it a certain degree of prominence " (p. ix) makes essentially the same statement but includes a perceptual element. This addition conflicts with Jones' definition since although in the absence of auditory cues one might consider that stress could still be perceived by ' empathy ', i.e. from contextual cues, such a percept would be due to the language experience of the listener, not the ' force '. (In one sense this perception would be illusory, but phonemically and subjectively it would be real.) Pike's definition, already quoted, seems identical with Jones' (with " initiator pressure " for " force "), but in his ' Phonemics ', Pike (1947) muddies the term ; " Stress : a degree of intensity upon some syllable which makes it more prominent or louder than unstressed syllables " (p. 250). The equation of ' stress ' with ' louder ' is one which Jones would reject, nor is the definition appropriate for any more general use of the term, since a ' stressed syllable ' in this sense may differ from the unstressed syllables in duration or pitch.

Jones in his use of ' force ', which, while appropriate as a subjective description, has objective overtones of ' intensity ' and so ' loudness ', has inadvertantly contributed to the confusion. In ' The Phoneme ' (1950), he also states " Prominence of a syllable may be due to strong stress, but it may also be due to other features of pronunciation, and particularly to the inherent quality of sounds, to length of syllables . . . or to intonation " (para. 435). According to this usage ' stress ' is amenable to study neither through the speaker nor the listener, who perceives ' prominence '.

However, while other writers have produced definitions of ' stress ' similar to those cited, there is a growing literature which presents the results of experimental work based on measurements of physiological activity of the speaker, measurements of various aspects of the speech wave and the perception of natural or synthetic utterances. One common feature of this work is its regarding ' stress ' as one of the " observable

[1] " *In spoken words consisting of several prosodic units, one group of sounds usually stands out against the others by virtue of some kind of emphasis, and this prevails against them as the controlling member, the others being the controlled members. The application of emphasis through force of utterance results in a dynamic graduation of the members.*"

linguistic phenomena " (Ladefoged *et al.*, 1958, p. 14). This is strictly incompatible with Jones' use of the word, but this is an incompatibility of usage, not a dispute as to what some essential ' stress ' *is*.

While it is apparent that there is need for general agreement as to the precise use of the various terms used in discussing the general topic (' stress ', ' prominence ', ' accent ', etc.) we will follow other experimentalists in using ' stress ' as a general term to describe a structurally significant phonetic entity, identifiable and definable at any level of speech communication : the psychological level of the speaker, his neural and articulatory level, the level of the sound wave in the transmitting medium, as well as the perceptive and higher analytic level of the hearer. Speech as a communicative activity represents a chain of code transformations, many of them conditional, in which any missing link is equivalent to an open position of a switch in a circuit. Some of the levels may be at the moment less accessible by direct experiment than others so that it may be convenient or even necessary to operate at one particular level, but this need not lead to an assumption that a linguistic-phonetic item intrinsically belongs to any one of the levels rather than any other. It is quite justifiable to investigate or to describe only one aspect of speech, say the articulatory, but then, a full structural statement has to define *all* elements at that level and must not relegate one of them to another, say the neural or psycho-acoustic. Having defined one's units at one particular level one may refer to correlates of all these units or any one of them on some of the other levels, but one must remember that it is theoretically possible to start from any one of the different levels.

Thus, when Lehiste and Peterson (1959) say, " It is our conclusion that stress is actually physiological stress " (p. 429) and Ladefoged *et al.* say " statements about stress are best regarded as statements about the speaker's muscular behaviour (or about the action of the listener's muscles which would have to be made in order to produce similar sounds) " (p. 9), we feel they are being over-restrictive. We would prefer another remark of Ladefoged *et al.* " differences in stress . . . can be *ordered more simply* in terms of the human behaviour producing them than in terms of the accompanying acoustic phenomena " (p. 9, our italics). While this is by no means demonstrated rigorously, it agrees with the conclusions of the Haskins workers, among others, concerning the production of consonants (e.g. Liberman *et al.*, 1962). In the latter case, it is observable that the acoustic nature of a particular phoneme is dependent upon its context ; but we doubt whether anyone would claim that /d/, for example, *is* a psychological phenomenon ; it has (in current usage) existence at many levels, the acoustic definition being merely more complex. It may also be remarked that there is a considerable consistency in the *listener's* ordering of stress (as with the identification of /d/), as the present paper will show.

A good review of the experimental literature until about 1956 is contained in Fónagy (1958). We would now like to review briefly some of the more recent publications dealing with the acoustical (physical) properties of stress. While all authors seem to agree that more than one property of the speech wave are involved, there is no agreement as to the significance of each of them. Linguistically, there are two possibilities : either there is one particular acoustic feature which is *always* present in a stressed syllable (others being incidental), and then stress can simply be defined at the acoustic level in terms of that single parameter ; or more parameters are involved significantly in such a way that there is complementary (or possibly free) variation in the distributional sense. If there is complementary variation, then the variants must be positionally defined.

According to Gimson (1956) stress is acoustically a " complex . . . of pitch, quantity and quality variations " (p. 149). He points out that " prominence " (which in this connection is substituted for stress as an objective manifestation thereof) is mostly produced by pitch changes in English, though other factors may occasionally prevail. No general rules are given as to when pitch is replaced by other factors. Gimson does not include amplitude as one of the " objective ", i.e. physical, parameters directly signalling stress (or prominence). For Lehiste and Peterson (1959) " . . . stress . . . in English . . . is reflected in at least four acoustical parameters : speech power, fundamental voice frequency, phonetic quality and duration ". They do not commit themselves definitely as to which parameter they consider predominant (or linguistically relevant), but they show by experiment that in the perception of stress " correction factors " are applied for the amplitudes according to vowel quality. In a recent article Lieberman (1960) investigates " the relevance of changes in the fundamental frequency, envelope amplitude and duration to mechanical recognition of the stressed syllable . . ." While this ostensibly is a study of the acoustic properties of spoken stress, in fact it relates to perceived stress since some of the original utterances (of pairs such as " ˈexport " and " exˈport ") were rejected so that " The stressed syllables of all the utterances accepted for processing could be clearly perceived " (p. 451). (Some implications of this will be taken up in the discussion.)

Both because of the procedure used in the investigation (large-scale measurements of several parameters with statistical treatment of the relative contribution of each) and the results obtained this is a very important contribution towards a clarification of the physical properties of stress. According to Lieberman pitch is the predominant factor with amplitude and duration following in that order. Yet Lieberman's conclusions must be taken with some essential reservations. His criterion of stress is simply quantitative, i.e., a stressed syllable has higher pitch, less frequently higher amplitude and still less frequently greater duration than the unstressed syllable. But it is well known that whether pitch is considered predominant or not, it is not its height that ' signals ' stress but the *change of its direction* (cf. Gimson, 1956). That special configurations of pitch rather than its peaks are related to stress was demonstrated by Bolinger (1958) who also showed that what is, or what at least was at that time, believed to be a stressed syllable by virtue of being relatively loud is in point of fact a syllable with specific pitch characteristics. Bolinger therefore introduced the term *pitch accent*. That out of the three parameters—intensity, duration and fundamental frequency—the last is mainly responsible for the judgement of stress by English-speaking listeners was also demonstrated by Fry (1955, 1958). One would be tempted to infer from the results of the experiments performed by Bolinger, Fry and Lieberman that the relevant or distinctive feature of stress (accent) in English is a set of specific pitch configurations. Yet there are syllables in longer English utterances which are described as " stressed " in some sense of the term though they have apparently none of the features that may be ascribed to pitch as a " signal " of stress. Numerous examples may be found, for instance, in O'Connor and Arnold (1960) (e.g. " By ˈall ˌmeans ask him "—the low pitch on the last three syllables, cf. pairs like " ˈbookˌworm " *vs.* " ˈbookworm " in Kingdon, 1958 p. 195 ; cf. also " It ˈisn't a ˌriding ˌschoolgirl " *vs.* " It ˈisn't a ˌriding-school ˌgirl ", etc.). O'Connor and Arnold distinguish accent and stress in English, though they are somewhat non-committal on the matter of the definitions, except that accent is apparently pitch configurations.

Using synthetic speech produced by PAT, Rigault (1962) proved pitch to be considerably more effective in producing the impression of stress in French listeners than intensity or duration, and Jassem (1959, 1962) concluded from his measurements of the same three acoustic parameters that Polish stress, which had quite unanimously been described as " dynamic ", i.e. due to relative force of articulation, or amplitude, or loudness, is in fact melodic, i.e., that it is signalled by, or that it simply *is*, physically, a set of fundamental frequency patterns.

The fact that recent experimental work has shown pitch to be the predominant acoustic parameter related to stress (or ' signalling ' stress or ' physically defining stress ') in three rather distantly related languages (English, French, Polish) makes it sensible to pose the question whether it would not be worth while to compare the nature of stress in many different languages at one level of description, by some kind of unified or even standard method, so that results obtained for such different languages could be directly comparable. The present article suggests one such possible method. It is hoped that even in their application to one specific language such tests as those here presented may contribute toward solving some of the basic problems of the phonetic nature of stress. In particular, the tests are intended to show to what extent each of the three parameters (fundamental frequency, duration, amplitude) is, or may be, responsible for an impression identified by listeners as linguistic stress. A restricted application of stress is being assumed, *viz.* one that refers to the minimum sequential comparison situation involving two items only, i.e., in our case, two consecutive syllables constituting an utterance. If, therefore, a language has more than one relevant stress contrast (more ' degrees ' of stress) our experiment will do nothing to reveal this.

The materials consist of utterances which may be transcribed as [sisi], [sɔsɔ], [sasa], synthesized on PAT with specified variations of the appropriate parameters. The choice of the ' phones ' (or ' phonetic segments ') was based on the following considerations :

(1) Phonetically the same vowel should be used in both syllables. This restriction removes the complication that vowels of different quality may have ' intrinsically ' different intensities (cf. Lehiste and Peterson, 1959).

(2) The experiment should be tested on at least three different vowels to see whether results depend on the spectral properties of the vowels.

(3) " Extreme " vowels (in the sense of both articulatory features and formant frequencies in real speech) should be used as far as possible, but naturalness is an essential desideratum. With PAT working, as it did at the time of this synthesis, with parallel connection of the formant circuits [u], which we would like to have had, was not entirely satisfactory, and was consequently replaced by [ɔ].

(4) The stimuli should be phonetically as simple as possible to avoid distraction from the main issue (see also Discussion).

(5) The utterances should contain sounds and sound sequences which are likely to be permissible in different languages or at least not sound very un-speech-like to listeners with different linguistic backgrounds.[2]

The synthesis of the original stimuli was based on parameters extracted from actual utterances in which every effort was made by the speaker (a phonetician) to keep any stress out, i.e. to pronounce the two syllables of each nonsense-word so that neither syllable should in any way be more prominent than the other. It is however not crucial for the experiments whether the original utterance or the subsequent synthesis thereof were actually quite successful in this respect because we were mainly interested in the differential results. After the three basic utterances were produced on PAT, the individual parameters were varied in a pre-determined manner, and each new version, as well as the basic version, was tape-recorded. The values selected for the variations of the parameters were somewhat arbitrary, but we were trying to choose them in such a way that under favourable conditions of comparison (e.g. with a relatively short interval of silence) no two versions of the final set of stimuli should be indistinguishable. The variations introduced in the overall level parameter (amplitude) were −3, −6 and −9 db. *re* the original level. Duration was varied by 40, 20, −20 and −40 per cent of the vowel length.[3] (Larynx) frequency was varied so that (a) the original frequency of 120 c.p.s. was replaced by steady tones of 76, 96, 110, 130, 151 and 190 c.p.s. and (b) a slope up or down, to or from 120 c.p.s. was substituted for a steady tone of 120 c.p.s. with the frequencies just quoted as the maximum. Of each " word "[4] 48 versions were thus synthesized, in which at least one syllable was as in the basic version and the variation, if any, affected no more than one parameter at a time. 32 more versions were produced in which two parameters were varied. In 16 of these, both variations were introduced in the same syllable

[2] *Experiments similar to those here described have been conducted with the same material at the Department of Phonetics, Poznan University, Poland, with Polish listeners, and the results, it is hoped, will soon be ready for publication.*

[3] *The durations of the vowels in the standard stimuli were 210 msec. for the first vowel and 260 msec. for the second vowel in* [sɔsɔ] *and 280 msec. for* [sisi] *and* [sasa].

[4] *The inverted commas will subsequently be omitted.*

Morton, John, Applied Psychology Research Unit, Cambridge, *Acoustic correlates of stress* ,
Language and Speech, 8:3 (1965:July/Sept.) p.159

(either the first or the second) and in the remaining 16 one parameter was altered in the first, and another in the second syllable. The set of 80 versions of each [sisi], [sasa], and [sɔsɔ] and the order of playing are shown in Table 3.

The final set of stimuli consisted of three tapes for each of the three words, with the 80 versions randomized. The first five stimuli on each tape were regarded as practice items.

Procedure

The subjects were Royal Navy personnel with normal hearing. Their ages were between 17 and 23. In all cases the subjects listened to the tapes in groups of up to six in a normally damped room. The stimuli were reproduced by a good quality tape recorder and loudspeaker, at a level of about 70 db. *re* 0.0002 dynes/cm.^{2}

Instructions

On the blackboard in the experimental room were written words designed to illustrate the concept of stress. These were usually the pairs ˈpermit, perˈmit ; ˈtransport, transˈport ; ˈforearm, foreˈarm. Occasionally one of the pairs ˈinsult, inˈsult, and ˈrefuse, reˈfuse were substituted.[5] Also on the board were written the test words ' Soso ', ' Sisi ', ' Sasa '. At no time were any of these words spoken by the experimenter.

The following instructions were given to the subjects : " This is an experiment in the perception of word stress. By ' stress ' we mean the way words which have the same spelling—like those on the blackboard—differ in the way they are said. They differ usually because one is the noun form and the other is the verb form. (The words were then defined in pairs.) If you will say these words over quickly to yourselves you should hear the difference. Please do that for a few moments so that you can all understand what I mean . . . Are you all sure that you know what I am talking about ?

" In this test you will hear samples of these three two-syllable words (the test words were indicated on the blackboard). What I want you to do is to say which of the two syllables you think is stressed.

" You are given three choices. If you think the first syllable is stressed, then underline it ; if you think the second syllable is stressed, underline it ; if you think neither syllable is stressed more than the other, or if you find it impossible to make a judgement, then underline the whole word.

" I want to point out that there is no right or wrong as far as we are concerned at the moment. We want your opinion as to which syllable is stressed. This is, in fact, an opinion poll more than a recognition test. You are not being given spoken

[5] *The difference between /s/ and /z/ in* ˈrefuse *and* reˈfuse *was overlooked by the experimenter (JM). It did not confuse any of the subjects however.*

TABLE 1

Experimental Design for Part I

ORDER OF PLAYING THE TAPES

	Day 1	Day 2
Group 1 (n = 6)	Soso (2-1), Sasa	Soso (2-2), Sisi, Soso (5)
Group 2 (n = 5)	Soso (2-1), Soso (2-2)	Sisi, Soso (5), Sasa
Group 3 (n = 5)	Soso (2-1), Sisi	Soso (5), Sasa, Soso (2-2)
Group 4 (n = 6)	Soso (2-1), Soso (5)	Sasa, Soso (2-2), Sisi

examples because spoken examples might possibly influence your judgement.

" I shall now play you a few examples of the test items so that you can hear what they sound like. Any questions ? "

All but two or three of the 60 subjects had an immediate apprehension of the concept of stress. That is they realised that whatever it was that distinguished between the words in one pair, also distinguished between the words in the other pairs. When questioned at the end of the experiment, they sometimes called this property ' stress ', more often ' emphasis ' and rarely ' accent '.

Two examples of the stimuli were played before the experiment started.

The design of the first experiment is shown in Table 1 ; it had two objects. First, we wished to discover any effects of practice upon the subjects' responses. Accordingly all groups heard the ' Soso ' tape initially—' Soso (2-1) ', and a second time in a serial position depending upon the group—' Soso (2-2) '. A comparison of the pooled results on these tapes would reveal any systematic changes in behaviour during the experiment.

The second object was to investigate the effects of the interval between stimuli. To this effect a tape of the ' Soso ' stimuli was produced with a five second interval between the stimuli, ' Soso (5) '. The results for this tape can be compared with those for ' Soso (2-2) '.

Results

For each stimulus the responses were summed for the five tapes separately. This yields three figures—termed *1*, *x* and 2—which are the number of subjects marking respectively the first syllable, the whole word and the second syllable. The significance of the deviation of these three figures from random may be tested by the formula :

$$\chi^2 = 2(r_1 \log_e r_1 + r_x \log_e r_x + r_2 \log_e r_2 + n \log_e 3 - n \log_e n)$$

which reduces to :

$$\chi^2 = 1.3862 \sum_{i = 1, x, 2} (r_i \log_2 r_i - n \log_2 (n/3)), \text{ where } r_1, r_x, r_2$$

Morton, John, Applied Psychology Research Unit, Cambridge, ***Acoustic correlates of stress*** ,
Language and Speech, 8:3 (1965:July/Sept.) p.159

are the values of *1*, *x* and *2* for any item, and *n* is the total number of judgements made on that item.

130 triplets out of the 400 were significantly non-random with $p < 0.001$ and a further 65 were significant at $p < 0.01$. This indicated that the group was behaving coherently.

It should be noted that this test says nothing about *what* is significant. It merely indicates the probability of the three numbers occurring by chance. Thus while the highly significant result (18 4 0) is clearly in favour of *1*, the result (11 11 0), also highly significant, might be better described as ' not *2* '.

At this stage we are primarily interested in the systematic differences between the scores on Soso (2-1) and Soso (2-2) to ascertain any effects due to the serial order of presentation, and between Soso (5) and Soso (2-2) to check on effects due to the interval between successive stimuli. Comparisons were made by obtaining the signed difference between the scores for each stimulus on a pair of tapes for each of *1*, *x* and 2. Wilcoxon tests (Siegel, 1956) were applied to each distribution to determine whether or not there were any significant overall differences in the judgements. These tests were only significant in comparing Soso (2-1) with the other tapes. The meaning of this was that the first tape played had fewer judgments of *2* and more of *x* and *1*, compared with the other tapes. When the stimuli are divided into groups, depending on the nature of the difference between the two syllables (i.e. fundamental frequency, intensity, etc.) these systematic differences in judgment are found in all the groups, and in the sub-groups (e.g. frequency rise as opposed to frequency fall) to the same extent. In as much as the object of the experiment was to compare the effects of different kinds of acoustic variation, we can ignore these differences.

There were no such systematic differences found between the Soso (5) and Soso (2-2) tapes.

As a consequence of this analysis it was decided to omit the Soso (5) tape from the subsequent testing. In addition, since the effect of the position of the tape applied to all the kinds of stimuli used, the second part of the experiment consisted of only three tapes.

Experiment 2

Design

In the second part of the experiment only tapes Soso (2), Sisi and Sasa were played, in a single session of half-an-hour. There were six groups of six subjects each, and the order of presentation of the tapes varied according to two 3 × 3 Latin Square designs.

Results

A total of 60 subjects were finally used ; 36 from Experiment 2, 22 from Experiment 1, and two other subjects were tested singly to make up Groups 2 and 3 in Experiment

TABLE 2

Tests of the stimulus groups for differences between the *D*-scores of the 3 words

Section	*N*	χ^2 (significance)	Sign. Test & Level
Neutral	5	6.7 ($p < 0.04$)	Soso > Sasa ($p < 0.031$)
Intensity	6	9.0 ($p < 0.008$)	Soso > Sisi ($p < 0.016$) Soso > Sasa ($p < 0.016$)
Duration	8	6.1 ($p < 0.05$)	Soso > Sasa ($p < 0.035$)
Frequency	24	7.85 ($p < 0.02$)	Soso > Sasa ($p < 0.02$)
Intensity-Duration	32	3.15 n.s.	
Total	80	18.1 ($p < 0.001$)	Sisi > Sasa ($p < 0.008$) Soso > Sasa ($p < 0.005$)

NOTE: Soso > Sasa means that the *D*-scores for Soso were significantly more positive than those for Sasa at the *p*-level indicated.

1. The results for the 60 subjects were pooled for Sisi, Soso and Sasa separately. For Soso, the results from Soso (2-2) were used, since that tape was balanced for serial order effects with Sisi and Sasa. For each stimulus we then calculated the *D*-scores. This measure is the result of substracting the number of subjects marking the second syllable from the number marking the first syllable : i.e. $D = (1 - 2)$. It expresses the majority in favour of *1* rather than *2*. If *D* is negative, then the majority favoured 2.

The scores for the three words were then compared using the Friedman two-way analysis of variance and the Sign Test (Siegel, 1956). Taking all the 80 stimuli, a difference in the *D*-scores was found between the words with $\chi^2 = 18.1$ ($p < 0.001$). The *D*-scores for Sasa were biased towards the second syllable (i.e. more negative) compared with the other two words. Using the Sign Test it was found that

$$D\,(\text{Sisi}) > D\,(\text{Sasa}) \quad (p = 0.008)\ (\text{two-tailed})$$
$$D\,(\text{Soso}) > D\,(\text{Sasa}) \quad (p = 0.006)\ (\text{two-tailed})$$

Each of the different stimulus groups were tested separately using these tests with the results shown in Table 2. With the exception of the stimuli involving changes in both intensity and duration all groups show a significant variance in the *D*-scores contributed by the differences between the words. Sasa produces most responses in favour of stress on the second syllable, and Soso most in favour of the first syllable. Since the ' neutral ' stimuli have the same bias as the other group, we can conclude that this is not an interaction effect between acoustic change and word, but is a consequence of differences between the balance of the word as originally spoken.

Morton, John, Applied Psychology Research Unit, Cambridge, *Acoustic correlates of stress* ,
Language and Speech, 8:3 (1965:July/Sept.) p.159

Soso differs acoustically from the other words in that the second syllable is slightly shorter (260 msec. compared with 280 msec.) which would result in judgments on that word being relatively biased towards the first syllable. However, there is no acoustic difference between Sasa and Sisi which could reasonably be expected to produce the differences between their *D*-scores. We are left then with the hypothesis that these differences result from some factors intrinsic in the perception of the vowel sounds of the words.

The differences in the *D*-scores here discussed are small, and it is not surprising that they were not detected by the informal listening initially used to check the stimuli. Indeed, they were not detectable to any degree of confidence from the results of Experiment 1 either by the techniques described in the appropriate section or by a Friedman analysis of variance subsequently carried out. It is of course questionable whether effects which only become statistically significant after use of large groups of subjects have any vital phonetic or psychological significance.

The results were also tested to find systematic differences in the *x* scores. No such differences could be found for any stimulus group except the intensity-duration group. In this group, Sasa had significantly fewer *x* responses than either Sisi ($p < 0.03$) or Soso ($p < 0.003$) on a Sign Test. This tendency is more marked in the ambiguous stimuli—where intensity and duration effects are working in opposite directions (items 44 - 59)—than in the other stimuli. This is not a result of favouring either intensity or duration but more of a general tendency to make some positive stress marking.

To summarize the differences between the tapes we can say that compared with the other words there is an overall tendency with Sasa to place the stress on the second syllable. In addition, where there is a possibility of confusion between two acoustic variables, subjects are more likely to make a positive decision with Sasa.

Overall results

The results for the three words were then pooled, giving a total of 180 responses for each stimulus item. These pooled results are given in Table 3. Of the 80 items, 73 gave distributions which deviated significantly from random at $p < 0.001$ using the test mentioned previously. Of the other items, six involved a duration change of -40% and the remaining one a conflict between maximum changes of both intensity and duration. The results will be discussed by groups of items.

Neutral Stimuli

Items 4 - 8 were the original versions of the stimuli. The average score for the five was (19 142 19). The range of the *D*-scores was from 4 to -3. These 'basic' items can thus be described as neutral with respect to stress placement. Thus any differences we may find between the effects of changes in the two syllables are a result of the interaction of acoustic change and position, and not of response bias.

TABLE 3

Results pooled for three words × 60 subjects = 180 responses/item

Item No.	*Position No.*	*Stimulus Description*	*1*	*x*	*2*	*D*
1	47	−9 N	18	107	55	−37
2	57	−6 N	12	141	27	−15
3	52	−3 N	24	128	28	−4
4	36	N N	17	143	20	−3
5	67	N N	23	134	23	0
6	71	N N	11	155	14	−3
7	77	N N	26	132	22	4
8	7	N N	19	143	18	1
9	23	N −3	16	149	15	1
10	18	N −6	69	89	22	47
11	12	N −9	86	70	24	62
12	72	40 N	70	83	27	43
13	20	20 N	24	144	12	12
14	15	−20 N	43	94	43	0
15	60	−40 N	61	66	53	8
16	14	N −40	57	75	48	9
17	61	N −20	21	140	19	2
18	53	N 20	23	140	17	6
19	76	N 40	15	65	100	−85
20	80	190 - 120 St	139	27	14	125
21	56	151 - 120 St	155	8	17	138
22	39	130 - 120 St	100	56	24	76
23	33	110 - 120 St	101	40	39	62
24	49	96 - 120 St	108	43	29	79
25	27	76 - 120 St	130	20	30	100
26	48	120 - 190 St	7	6	167	−160
27	19	120 - 151 St	12	11	157	−145
28	38	120 - 130 St	22	12	146	−124
29	44	120 - 110 St	37	67	76	−39
30	70	120 - 96 St	25	20	135	−110
31	42	120 - 76 St	30	19	131	−101
32	79	190 - 120 Sl	155	10	15	140
33	22	151 - 120 Sl	154	13	13	141
34	45	130 - 120 Sl	57	112	11	46
35	40	110 - 120 Sl	62	96	22	40
36	58	96 - 120 Sl	134	26	20	114
37	65	76 - 120 Sl	131	24	27	104
38	43	120 - 190 Sl	6	3	171	−165
39	50	120 - 151 Sl	10	14	156	−146
40	31	120 - 130 Sl	10	17	153	−143

Morton, John, Applied Psychology Research Unit, Cambridge, *Acoustic correlates of stress* ,
Language and Speech, 8:3 (1965:July/Sept.) p.159

TABLE 3 (CONTINUED)

Item No.	*Position No.*	*Stimulus Description*	*1*	*x*	*2*	*D*
41	35	120 - 110 St	29	30	121	−92
42	41	120 - 96 St	18	14	148	−130
43	29	120 - 76 St	29	10	141	−112
44	17	N −20 −6 N	35	87	58	−23
45	32	N −20 −9 N	24	93	63	−39
46	10	N −40 −6 N	62	63	55	7
47	74	N −40 −9 N	52	54	74	−22
48	24	20 N −6 N	26	91	63	−37
49	55	20 N −9 N	22	76	82	−60
50	63	40 N −6 N	74	70	36	38
51	9	40 N −9 N	82	43	55	27
52	66	−20 N N −6	31	129	20	11
53	62	−20 N N −9	70	81	29	41
54	28	−40 N N −6	31	106	43	−12
55	69	−40 N N −9	99	51	30	69
56	73	N 20 N −6	34	118	28	6
57	59	N 20 N −9	58	100	22	36
58	30	N 40 N −6	30	121	29	1
59	13	N 40 N −9	59	52	69	−10
60	6	−20 N −6 N	13	124	43	−30
61	64	−20 N −9 N	30	76	74	−44
62	25	−40 N −6 N	66	50	64	2
63	78	−40 N −9 N	73	49	58	15
64	11	N 20 −6 N	10	81	89	−79
65	68	N 20 −9 N	20	20	80	−60
66	8	N 40 −6 N	10	57	113	−103
67	46	N 40 −9 N	13	70	97	−84
68	75	N −20 N −6	36	128	16	20
69	16	N −20 N −9	80	80	20	60
70	26	N −40 N −6	73	78	29	44
71	34	N −40 N −9	69	91	20	49
72	54	20 N N −6	66	101	13	53
73	37	20 N N −6	83	76	21	62
74	51	40 N N −6	66	102	12	54
75	21	40 N N −9	109	52	19	90
76	1	N 40	14	91	75	−61
77	2	N N	33	106	41	−8
78	3	−40 N −9 N	54	39	87	−33
79	4	190 - 120 Sl	158	6	16	142
80	5	190 - 120 St	135	24	21	144

NOTES: (1) −3, −6, −9 refer to changes in intensity; ± 20, ± 40 refer to changes in duration; N means normal. Thus item No. 66, labelled N 40 −6 N had the first syllable 6 db. lower than normal and the second syllable 40% longer than normal.
(2) Items 26 - 43 involved changes in fundamental frequency as indicated with a Step (St) or Slope (Sl).
(3) Items 4 - 8 were the original ones.

TABLE 4

D-scores for items involving changes in fundamental frequency

		F.F. Raised			F.F. Lowered		
Percentage change		58%	25%	9%	−9%	−25%	−58%
Frequency (c.p.s.)		190	151	130	110	96	76
STEP	FN	125	138	76	62	79	100
	NF	−160	−145	−124	−39	−110	−101
SLOPE	FN	140	141	46	40	114	104
	NF	−165	−146	−143	−92	−130	−112

The stimuli are classified by the percentage change up or down to the stated frequency of either the first syllable (FN) or the second syllable (NF) with either a step or a slope transition.

From the results on these items it is also apparent that we can expect scores of up to at least 20 on either syllable resulting from purely random behaviour.

Effects of changes in Fundamental Frequency

The normal fundamental frequency for the stimuli was 120 c.p.s. Items involving fundamental frequency change are defined by the frequency to which one of the two syllables was raised or lowered. This is not to be confused with a ' rise ' or ' fall ' of frequency, which is a statement of the relationship of the frequency of the second syllable to that of the first. The *D*-scores for this group, pooled over the three words, are given in Table 4. For convenience we will distinguish between fundamental frequency variations on the first and second syllable by *FN*, and *NF* respectively, where *N* stands for normal. Similarly later, for intensity, *I*, and duration, *D*.

The first, and overwhelming effect is that any syllable with a fundamental different from the normal 120 c.p.s. has the stress marking. This applies for a rise or fall in fundamental, to or from the normal frequency with either a step or a slope transition. In Table 4 this is reflected in the fact that all the *FN* items have a positive *D*-score, and all the *NF* items a negative *D*-score.

The other effects were tested by means of a 5-way analysis of variance performed on the *D*-scores. In the analysis all the *D*-scores were treated as positive in sign, i.e. ignoring the main effect, to normalise the data. A summary of the analysis is given in Table 5. None of the fourth-order interactions were significant when tested by the original error term (the fifth order term), and the final *F*-ratios were obtained by using a pooled-sum-of-squares from all these sources. The following conclusions may be drawn from Table 5.

TABLE 5

Summary of analysis of variance for stimuli involving changes in fundamental frequency.

SOURCE	DF	SS	MS	F	Sig. Level
Between Words (W)	2	2200.1	100.1	3.40	$p < 0.05$
„ Transitions (T)	1	180.5	180.5	6.14	$p < 0.025$
„ Syllables (S)	1	1266.7	1266.7	43.07	$p < 0.001$
„ Directions (D)	1	3016.1	3016.1	102.55	$p < 0.001$
„ Proportions (P)	2	4075.0	2037.5	69.28	$p < 0.001$
Interactions W.T.	2	25.7	12.9	<1	—
„ W.S.	2	196.7	98.4	3.34	$p < 0.05$
„ W.D.	2	114.6	57.3	1.95	—
„ W.P.	4	105.4	21.4	<1	—
„ T.S.	1	150.3	150.3	5.11	$p < 0.05$
„ T.D.	1	107.5	107.5	3.66	$p < 0.05$
„ T.P.	2	32.3	16.2	<1	—
„ S.D.	1	242.0	242.0	8.23	$p < 0.01$
„ S.P.	2	338.3	169.2	5.72	$p < 0.01$
„ D.P.	2	27.0	13.5	<1	—
„ T.S.P.	2	502.7	251.4	8.55	$p < 0.001$
„ D.S.P.	2	494.0	247.0	8.40	$p < 0.001$
Remainder	41	1204.8	29.39		
TOTAL	71	12259.7			

(1) The effects due to differences between Words have a significance level of about 5%. The effects of fundamental frequency change are overall greatest with Sisi and least with Soso. The difference between these words is however marginal and would have to be confirmed before any other conclusions could be drawn.

(2) There is a significant effect due to the two kinds of transition used, slope generally being more effective than step. The significant Transition × Direction interaction indicates that the superiority of slope is greater when the fundamental is lowered rather than raised.

(3) The effects of changing the fundamental of the second syllable are very much greater than those resulting from changes in the first syllable. The significant Syllable × Transition interaction indicates that the difference is reliably greater when a slope as opposed to a step transition is used.

(4) The largest single effect is the comparison between raising and lowering the fundamental frequency, a rise producing the larger effects.

(5) The proportion by which the fundamental frequency is varied is also highly significant, the 9% variation being less effective than the other two but the 25% variation being as effective as the 58% variation.

(6) Of the interaction effects of Words with the other variables, only that with Syllables reaches significance level, and that only just, being caused by a greater difference between the two syllables with Sasa. It was pointed out above that Sasa produced more 2 responses generally than the other words, so that this interaction would be expected.

(7) Syllable × Direction interaction, which is highly significant, is due to the increasing superiority of the second syllable changes when this syllable is raised in fundamental frequency. This might be taken as illustrating the basic nature of the upwards intrusion (cf. Bolinger, 1958, p. 149). Syllable also interacts highly with Proportion, the difference between the syllables being greatest with the 9% variations (see Table 4).

(8) Two of the third order interactions are highly significant, both of them involving Syllable and Proportion. The effect of these two with Transition may be observed in Table 4, where both of the 9% changes on the first syllable reverse the normal superiority of slope over step. That interaction involving Direction is largely due to the high D-score with a 9% change where the second syllable is raised, compared with the other item involving a 9% change in pitch.

Effects of Intensity Change

The *D*-scores for the items involving an intensity change are given in Table 6. It will be seen that a 3 db. reduction has no effect and that the effect of a −9 db. change is greater than that of −6 db. The stress is given to the more intense syllable, and the effects are greatest when the second syllable is altered, i.e. when the first syllable receives the stress marking. The *D*-scores are small compared with those caused by pitch changes.

Effects of Duration Changes

From previous work (e.g. Fry, 1955) we would expect the longer of the two syllables to be marked with the stress. Table 7 shows this to be true for a 40% increase in duration, the effect being stronger on the second syllable than on the first. Changes of ± 20% have no effect and neither, apparently, has a −40% change. However, the items involving a −40% change (nos. 15 & 16) produce non-significant distributions and this was unusual enough to warrant further investigation. Accordingly the scores of the items involving a duration change were examined in detail. These

Morton, John, Applied Psychology Research Unit, Cambridge, *Acoustic correlates of stress* ,
Language and Speech, 8:3 (1965:July/Sept.) p.159

TABLE 6

Effects due to intensity alone – *D*-scores

	–9	–6	–3
IN	–37	–17	–4
NI	62	47	1

TABLE 7

Effects due to duration alone – *D*-scores

	40%	20%	–20%	–40%
DN	43	12	0	8
ND	–85	6	2	9

TABLE 8

Scores for duration changes

	DN			ND		
	1	*x*	*2*	*1*	*x*	*2*
40%	70	83	27	10	65	100
20%	24	144	12	23	140	17
–20%	43	94	43	21	140	19
–40%	61	66	53	57	75	48

scores are retabulated in Table 8. It will be seen that the *D*-score conceals the fact that the scores on both *1* and *2* are higher for the –40% items than the –20% items. In other words, more people placed the stress on the longer syllable when one was shortened further, but in addition there was a relatively high proportion of times when the shorter syllable was marked as stressed. This point will be taken up below.

Effects of Intensity and Duration Interaction

Items 44-75 involved all possible combinations of –6 and –9 db. change in intensity and ± 20% and ± 40% change in duration. The *D*-scores for these items

TABLE 9

D-scores for stimuli involving intensity and duration

DURATION CHANGES	INTENSITY CHANGES −9 N	−6 N	O	N −6	N −9
N 40	Q_1 −84	−103	−90	Q_2 1	−10
N 20	−60	−79	6	6	36
O	−37	−17	0	47	62
N −20	Q_3 −39	−23	2	Q_4 20	60
N −40	−22	7	9	44	49

DURATION CHANGES	INTENSITY CHANGES −9 N	−6 N	O	N −6	N −9
−40 N	Q_5 15	2	8	Q_6 −12	69
−20 N	−44	−30	0	11	41
O	−37	−17	0	47	62
20 N	Q_7 −60	−37	12	Q_8 53	62
40 N	27	38	43	54	90

and those involving no change in one or other of the variables are given in Table 9. For convenience of reference the eight quadrants of the Table are numbered Q_1 - Q_8. Stimuli involving a −40% change in duration are again abnormal. Apart from these, the following observations may be made.

(1) Where the two factors operate in the same direction they reinforce each other (Quadrants, 1, 5, 8, though not for N—20 in Q_4). There appears to be no simple relationship between the *D*-scores for the two components separately and the score for the composite.

(2) Where the factors operate in opposite directions :
(a) +40% change in duration is equal to (Q_2) or outweighs (Q_7) a −6 db. or a −9 db. change in intensity.
(b) In general, changes of ± 20% in duration are outweighed by both −9 db. and −6 db. changes in intensity.

These results could have been predicted from the size of the *D*-scores for the components separately, apart from the fact that the ± 20% changes in duration exert more influence in combination with intensity changes than they do by themselves.

An analysis of variance was carried out on the data, taking the three words as a further variable. The results are summarized in Table 10. As would be expected the effects of duration and intensity are both highly significant. There is a significant

TABLE 10

Summary of analysis of variance on the *D*-scores for stimuli involving intensity and duration changes

SOURCE	*DF*	*SS*	*MS*	*F*	*Sig. Level*
Between words (W)	2	31.83	215.42	4.48	$p < 0.025$
Between durations (D)	8	11205.12	1400.64	29.11	$p < 0.001$
Between intensities (I)	4	15046.23	3761.55	78.19	$p < 0.001$
Interactions W × D	16	1457.24	91.08	1.89	—
,, D × I	32	3472.95	108.52	2.26	$p < 0.01$
,, W × I	8	316.48	39.56	< 1	—
Residue: W × D × I	64	3079.10	48.11		
TOTAL	134	35008.95			

TABLE 11

Stress marking given to the syllables for all ten stimuli involving the various duration changes

	Stress marking given to:		
Duration Change	*Short Syllable*	*x response*	*Long Syllable*
−40%	556	679	565
−20%	363	1032	405
20%	336	987	447
40%	273	704	818

difference between words, Soso being different from Sasa with the latter more negative (cf. Table 2). There is however no significant *interaction* of the words with either of the other variables. That is, changes in both duration and intensity have a similar effect on all words. The significant intensity × duration interaction merely indicates departures from additivity in Table 9. These are fairly complex, and, apart from the ± 20% duration change phenomena mentioned in the previous paragraph, do not seem to be phonetically or psychologically significant.

Stimuli Involving Changes of −40% in Duration

It has been pointed out that the stimuli involving changes of −40% in duration are unusual in that a high proportion of responses placed the stress on the shorter syllable. For the 10 items with a −40% change, the total figures were 556 for the shorter syllable, 679 *x* responses and 565 for the longer syllable. The three corresponding figures for the items involving other duration changes may be seen in Table 11. There are no significant differences between the three words in this respect.

The hypothesis was suggested that individual subjects would have a tendency to favour either the short or the long syllable, an even split of these subjects leading to the even distribution of votes in the total. Accordingly, the 30 responses to items containing a −40% change (10 responses to each of three words) were summed for each of the 60 subjects. Two figures were obtained for each subject; the number of times a stress was marked on the short syllable and the number of marks on the long syllable. 17 of the 60 scores were significantly different from random at the 5% level, 8 in favour of the long and 9 in favour of the short syllable. This proportion of deviations is highly significant ($p < 0.0001$). Thus we can say that with a duration change of −40%, individual subjects tended to favour either the longer syllable or the shorter syllable. This contrasts with all other conditions where all subjects tended to favour the longer syllable or show no preference. The vowel duration of the stimuli under discussion were approximately 210 msec. and 110 msec. respectively for the two syllables when the second syllable was shortened, and 90 msec. and 270 msec. respectively when the first syllable was shortened.

Summary of results

(1) The responses to the neutral stimuli showed that there was no generalized response bias in the situation.

(2) On the first test tape, compared with subsequent tapes, there is a general tendency away from the second syllable for all types of acoustic variation.

(3) There were no significant differences between a tape with 2-sec. intervals between successive stimuli and one with 5-sec. intervals.

(4) In general the word Sasa produced more judgements of stress on the second syllable than either Sisi or Soso.

(5) Within the ranges of variation used, fundamental frequency changes were by far the most effective in producing universally accepted stress-marking.

(6) All changes in fundamental frequency from the standard or 'context' frequency result in a stress marking on that syllable whose fundamental has been changed.

(7) A raised fundamental frequency is more effective than a lowered one, and changes of 58% in the fundamental are no more effective than changes of 25% in producing consistent stress marking.

(8) Changes from the normal on the second syllable produced more stress markings than the equivalent changes on the first syllable in all cases except one. For all

kinds and degrees of fundamental frequency variation and for the positive duration changes this superiority led to a greater marking of stress on the second syllable. For intensity reduction, the greater effect when the second syllable was changed meant that the first syllable was marked as having the stress more often than was the second syllable in the equivalent stimuli.

(9) The exception to the above rule, a -40% change in duration, produced in addition the interesting result that for some people the shorter of the two syllables was consistently marked as being stressed.

(10) When intensity and duration were varied together, the effects, by the measure used (*D*-score), were mainly additive, but there were significant deviations from this.

Discussion

The results here reported are strictly valid only for the situation employed involving two consecutive phonetically identical symbols in an utterance. Before applying these results to some general theory or application (such as automatic analysis or synthesis of speech) several other factors must be taken into consideration. The first of these is the context of the word in question. We have already drawn attention to the ambiguity in interpretation inherent in the experiment of Lieberman (1960). The lack of complete correspondence between spoken and perceived stress which he reports probably arises from his method of obtaining the utterances. He presented subjects with sentences of the form "We had a *contract*" and "*Rebel* from your state of misery". Subjects read each sentence silently except for the italicized word which was read aloud. This procedure is similar in certain respects to having subjects read the whole sentence and subsequently gating out the relevant words. The latter method was used by Fry (1955) who found "There was considerable variation in the behaviour of speakers with respect to the placing of the accent in different words" (p. 765). In both cases the critical word is part of a longer utterance and is defined by the utterance. Thus the acoustic cues for stress are redundant. It seems highly probable that this procedure would produce different results from those obtained from presenting for reading a formal definition such as "*Rebel;* to oppose authority". We would expect the acoustic cues for stress in the latter case to be more marked and more consistent.

The situation is analogous to a further experiment by Lieberman (1963). He asked subjects to read sentences of the form "Neither a borrower nor a lender be" and "Never listen to a man who wants to be a borrower", gated out the versions of "borrower" and presented them in an intelligibility test. The version from the first sentence, in which the word was bilaterally highly redundant, was relatively unintelligible in isolation compared with the other, non-redundant version. Thus the speaker produces utterances which require a knowledge of the whole for the correct identification of some of the parts.

The complementary phenomenon is that a context strongly influences the perception of a word regardless of its physical characteristics (see Morton, 1964). Thus we would expect that certain values of duration, intensity and pitch of the vowel sounds of " con'tract ", which lead to a perception of stress on the second syllable when the word is presented in isolation, would not seem abnormal when the word is presented in the sentence " We had a contract ". Likewise, although Lieberman's (1960) subjects had " practised to their satisfaction " we would not expect them always to produce utterances which were unambiguous in isolation since the speakers always had the context as part of the feed-back.

The third point is that we were using nonsense syllables with which there was no overall preference for identifying one of the syllables rather than the other. With pairs such as " 'transport—trans'port ", there can be no such guarantee since the different frequencies of occurrence of the alternatives would affect the response bias (see Morton, 1964).

Finally we would note that the vowels in the two syllables were identical in the present experiment. Where vowels differ, allowance would have to be made following the results of Lehiste and Peterson (1958).

In this paper we have presented basic data concerning the relative effects of intensity, duration and pitch upon the perception of stress. We hope that upon this data, more complex relationships between the acoustic and perceptual levels of description may be formulated allowing for factors such as linguistic context, intonation patterns, vowel quality and the like.

We are deeply grateful to James Anthony of the Department of Phonetics in Edinburgh for his participation in the synthesis of the material employed, and to Professor David Abercrombie for the opportunity he gave Wiktor Jassem to work under contract in his department. We are also grateful to Mrs. Elizabeth Uldall and John Trim for useful discussions and to Peter Freeman for statistical advice.

References

Bloomfield, L. (1957). Language (London).

Bolinger, D. L. (1958). A theory of pitch accent in English. *Word*, **14**, 109.

Essen, O. von (1953). Allgemeine und angewandte Phonetik (Berlin).

Fónagy, I. (1958). Elektrophysiologische Beiträge zur Aksentfrage. *Phonetica*, **2**, 12.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *J. acoust. Soc. Amer.*, **27**, 765.

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

Gimson, A. C. (1956). The linguistic relevance of stress in English. *Zeitschr. f. Phonetik u. allg. Sprachwiss.*, **9**, 143.

Jassem, W. (1959). The phonology of Polish stress. *Word*, **15**, 252.

Jassem, W. (1962). Akcent jezyka polskiego. Prace Jezykoznawcze 31 (Wroclaw).

Jones, D. (1950). The Phoneme : its Nature and Use (Cambridge).

Jones, D. (1956). An Outline of English Phonetics (Cambridge).

Kingdon, R. (1958). The Groundwork of English Stress (London).

LADEFOGED, P., DRAPER, M. H. and WHITTERIDGE, D. (1958). Syllables and stress. *Miscellanea Phonetica* III, 1.

LEHISTE, I. and PETERSON, G. E. (1959). Vowel amplitudes and phonemic stress in American English. *J. acoust. Soc. Amer.*, 31, 428.

LIBERMAN, A. M. (1957). Some results of research on speech perception. *J. acoust. Soc. Amer.*, 29, 117.

LIBERMAN, A. M., COOPER, F. S., HARRIS, K. S. and MACNEILAGE, P. F. (1963). A motor theory of speech perception. In *Proc. Stockholm Speech Communication Seminar* (Stockholm), Vol. II.

LIEBERMAN, P. (1960). Some acoustic correlates of word stress in American English. *J. acoust. Soc. Amer.*, 32, 451.

LIEBERMAN, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172.

MORTON, J. (1964). The effect of context on the perception of words. *Brit. J. Psychol.*, 55, 165.

O'CONNOR, J. D. and ARNOLD, G. F. (1961). Intonation of Colloquial English (London).

PIKE, K. L. (1943). Phonetics (Ann Arbor).

PIKE, K. L. (1947). Phonemics (Ann Arbor).

RIGAULT, A. (1962). Rôle de la fréquence, de l'intensité et de la durée vocalique dans la perception de l'accent en français. *Proc. IV Int. Congress Phon. Sci.* ('s-Gravenhage), 735.

SIEGEL, S. (1956). Nonparametric Statistics (New York).

WISE, C. M. (1957). Applied Phonetics (Englewood Cliffs, N.J.).