

A THEORY OF PITCH ACCENT IN ENGLISH*

DWIGHT L. BOLINGER

1. THE FUNCTION OF PITCH

Like most other languages, English has contrasts of pitch. Unlike the pitch contrasts of certain languages, the English contrasts are not "phonemic" in the sense that raising or lowering the pitch of one syllable in a word such as *lower* will change its meaning in the way that replacing /t/ with /d/ will change it, causing it to point to something completely different in the world beyond language.

This much pitch in English does not do. No other phenomenon in language, however, has more firmly resisted efforts to find out what it does do. In the last two decades, linguists in the United States have attempted to take the garrison by main force. They have based their operations on a set of propositions which may be regarded either as assumptions or as conclusions from the evidence, depending on how far the evidence is credited. Among these propositions are the following:

1. That pitch functions in the same way as the segmental phonemes—the vowels and consonants—and that a sequence of different pitches will produce something potentially meaningful in the same way that a sequence of vowels and consonants may produce a word. In order to manage this, it is necessary to decide when one pitch is different from another. The segmental

* Research at Haskins Laboratories, 1956-57, supported by a grant from the Carnegie Corporation of New York. The author gratefully acknowledges the cooperation of the Laboratory staff, particularly the hand-drawing of spectrograms by Prof. Pierre Delattre and the technical and editorial help of Dr. Louis J. Gerstman.

phonemes can be distinguished with relative ease because they are complex: each contains certain "distinctive features," whose presence or absence rather sharply separates one from another. Pitch has only one ingredient, the fundamental frequency of the voice. As a speaker moves up or down the normal range, there is no point at which the ear can detect that one thing ceases to occur and another begins—there is only up and down. So it is impossible to isolate a pitch as one can isolate a segmental phoneme: an untrained listener can identify a prolonged English /s/ without benefit of context; no amount of training will enable him to distinguish one linguistically significant pitch from another, however much prolonged, unless it occurs in context. Relying on contexts chosen for the purpose, analysts have discovered, or invented, "contrastive levels," each level representing a phoneme. Four is the usual number of levels.¹

2. That the meaningful something constituted by pitch phonemes is an intonation morph. The morph 231 (numbers refer to the "contrastive levels" of pitch) differs from the morph 221 as *lower* differs from *bower*. For those who follow Smith and Trager² the morphs are more complex, including, besides the pitch levels, "terminal junctures," which are ways of going from phrase to phrase or from phrase to silence, and which involve both pitch (direction, not level) and tempo. Such an intonation morph would be 231 ∇ , which can be described phonetically as "mid-low level followed by mid-high level followed by low level followed by

¹ Zellig S. Harris, in *Language*, XX (1944), 189, marked seven. Rulon Wells, in *Language*, XXI (1945), 27-39, marked four, as did Kenneth L. Pike in his *Intonation of American English* (Ann Arbor, 1945). George L. Trager and Henry Lee Smith Jr., in their *Outline of English Structure* (Norman, Okla., 1951), have followed Pike, and recent textbooks based on the *Outline* have propagated the number four. (For advanced classes: Charles F. Hockett, *A Course in Modern Linguistics* [New York, 1958]; A. A. Hill, *Introduction to Linguistic Structures* [New York, 1958]; Henry A. Gleason, *An Introduction to Descriptive Linguistics* [New York, 1955]; Velma Pickett, *An Introduction to the Study of Grammatical Structures* [Glendale, Calif., 1956]. For beginning college classes: Donald J. Lloyd and Harry R. Warfel, *American English in Its Cultural Setting* [New York, 1956]; W. Nelson Francis, *The Structure of American English* [New York, 1958]; Paul Roberts, *Understanding English* [New York, 1958]. At least one for high school: Paul Roberts, *Patterns of English* [New York, 1956].) There has been no serious attempt to collect a real corpus of examples to test the theory, nor has it been proved experimentally. For intonational contrasts that the Trager-Smith *Outline* fails to account for, see my "Intonation: Levels versus Configurations," *Word*, VII (1951), 199-210.

² *Outline*, especially p. 46.

commonplace. For something as inherently unlikely as a bombing,

A A
bomb wrecked
A had it

with two A accents, or

A
bomb
A had wrecked it

with one A accent, on *bomb* itself, is more probable.

Accent C is anti-assertive. It may be merely lackadaisical, as in a descending series like

It's C
never too C
late to C
mend

or it may be clearly and even strenuously restrained, as in

D'you think I'm C
crazy?

Accent B means something like 'connectedness' and 'incompleteness.' In the question

B
better?
Were they

it is terminal, and 'incomplete'. In the first clause of

B A
better they'd be more accep^{table}
Were they

it is non-terminal and 'connected' to the following Accent A.

Test 14 relates this connectedness to compound words. A live-voice recording of *attorney-at-law* was made. The natural pitch pattern was removed, and the following artificial ones substituted:

- (1) (90) al(110)tor(90)ney-at-(110-90)law
- (2) (90) al(110)lorney-at-(110-90)law
- (3) (90) al(110)lorney-(90)al-(110-90)law

The stimuli were mixed with others involving *mother-in-law*, and

phasis on the utterance as a whole. For this purpose, an item that is ordinarily incapable of receiving it is selected—an auxiliary verb, a preposition, or something of the sort. In the following example there is no contrast between 'in' and 'out of', nor is there any special attention of any other kind bestowed on the word *in*; it merely carries a sentence accent which, if it were to fall anywhere else, might be mistaken as contrastive: the speaker refers to someone who had been denied re-entry to the United States because of having worked for the Communist government of Rumania, and says

B ll.. the De fense maintained that everybody in Rumania worked for the government.

The 'separateness' and 'newness' of A may be illustrated by what happens in a narrative. If one is telling a story in which a fog has been introduced, at a later point one may say

B A
The fog had deepened

But a B accent for something new and unexpected is unnatural. If we encountered

B A
bomb had wrecked it

at all, it would probably be in a time of violence when bombs⁵ are

total number of syllables and mechanical placement (in Arnold's terms, p. 440, "TONIC STRONG stress is undoubtedly free and not tied to any given position within the word"), but also other syllables may be raised to the status of potential-carrier when it is desired to lay on more than one accent. This is especially true when one word occupies the dimensions of a whole utterance. Note the successions⁶ of AAA and CCA in the following:

A A A
ab lute fu
He so ly re used

C C A
He so ly re fu
ab lute used

It also operates to shift the potential toward the end of a word when the word occupies the climactic position in the sentence, as when Thomas Mitchell said, in a TV program,

you ti
What need is a little adver sing

terminal fall and fade to silence, with a slowing down toward the end."³

3. That pitch and stress are phonemically independent. It is recognized that changes in stress may affect any level of intonation that happens to be running at the moment, but not, for example, in such a way as to raise a Level 2 pitch to a Level 3 pitch; the changes are phonetically slight and phonemically non-distinctive.

My purpose in this article is to deny the third assumption, and to reverse the roles of stress and pitch. I shall offer evidence that far from being a non-distinctive by-product or a completely independent variable, pitch is our main cue to stress.

2. EARLIER NOTIONS OF PITCH AS A CUE TO STRESS

The idea that stress may depend on pitch is not new. The experiment of John Muyskens in 1931, using kymographic records, purported to show that the familiar noun-verb pairs like *pérmit*—*permít* are distinguished by higher pitches on their stressed syllables. Kenneth L. Pike⁴ and Daniel Jones⁵ demolish this argument by pointing out, in Jones' words, that "it often happens in a language that strong stresses are found on low-pitched syllables and weak stresses on high-pitched syllables." To demonstrate this, all we need to do is turn the *permít* example into a question: *pérmit?*

The refutation, however, is based on a persistent fallacy: that in order to serve as a cue to stress, pitch must rise. Recent discussions and descriptions continue to look for this kind of relationship,⁶ and, failing to find it, enter a verdict against pitch in general.

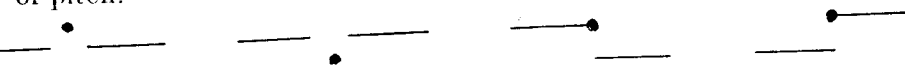
³ The analysis of Spanish juncture and intonation made by Stockwell, Bowen, and Silva-Fuenzalida, *Language*, XXXII (1956), 641-665, which follows Smith and Trager, throws some doubt on the morphemic status of intonation patterns: "A sequence of pitches up to and including a terminal juncture will be referred to as an INTONATION PATTERN. Whether or not such a sequence of suprasegmental elements is a morph remains to be demonstrated on the morphological level of analysis" (p. 661). It is hard to see on what basis the levels and junctures are contrastive units, if sequences of them are not morphs.

⁴ *Intonation*, pp. 16, 83.

⁵ *An Outline of English Phonetics* (New York, 1956), § 912 footnote.

⁶ Compare G. F. Arnold, "Stress in English Words," *Lingua*, VI (1957), 226, and K. L. Pike and W. Kindberg, "A Problem in Multiple Stresses," *Word*, XII (1956), 421.

The experiments that I shall report suggest that it is not pitch RISE, but rather pitch PROMINENCE, that is essential to what we react to as stress. By prominence I mean a rapid and relatively wide departure from a smooth or undulating contour. A rise is only one kind of pitch prominence, though it is certainly the commonest kind. In the following simplified diagrams, the "corners," or "sharp points," represented by dots, indicate some of the ways in which a syllable can be made to stand out by means of pitch:



3. STRESS AND INTENSITY

According to Bloomfield, "stress—that is, intensity or loudness—consists in greater amplitude of sound waves."⁷ Jones gives more attention to gestural accompaniments, but he too insists on "the objective impression of loudness."⁸ Nearly all linguists

The latter deals with a complex relationship of stress and pitch in Campa; the wording on the page cited indicates that the investigators gave up associating pitch with stress when it became apparent that HIGH pitch was not systematically related to it. If applied to English—I cannot judge its application to Witoto—the following statement would illustrate the current fear of mixing levels: "There is a complex suprasegmental phoneme of stress accompanied by high pitch" (Eugene A. Minor, "Witoto Vowel Clusters," *International Journal of American Linguistics*, XXII [1956], 137)—rather like saying that the dawn breaks accompanied by sunshine. Uriel Weinreich draws my attention to the following from N. I. Žinkin, "The Perception of Stress in Russian Words" (in Russian), *Izvestija Akademii pedagogičeskix nauk R. S. F. S. R.*, LIV (1954), 7-82, page 9: "The claim that under stress the fundamental pitch of the vowel is raised, is subject to doubt.... This raising should be attributed to sentence intonation, not to word stress. A pronunciation is possible without such raising. Furthermore, whispered pronunciation is possible, where the fundamental pitch ... is excluded altogether." Again pitch is dismissed as a cue to stress as a result of failure to find a pitch rise. It is reasonable, of course, to assign to the role of sentence intonation, instead of to stress, a particular manifestation of pitch change, e. g., rise rather than fall. If, however, we find that either a rise or a fall, when certain conditions are met, is responded to as "stress," then we have pitch playing a dual role; I have preferred to keep the two functions intact, labeling them *accent*. (As for the absence of intonational contrasts in whispered speech, see my speculation on a change of vowel quality to substitute for it, *Studies in Linguistics* V [1947], 77, confirmed experimentally by Werner Meyer-Eppeler, "Realization of Prosodic Features in Whispered Speech," *Journal of the Acoustical Society of America*, XXVIII [1956], 760.) These comments, of course, are beside the point if Russian stress and intonation are essentially different from English.

⁷ Leonard Bloomfield, *Language* (New York, 1933), § 7.3.

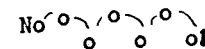
⁸ *Outline*, § 909.

8. THE ACCENTS AS MORPHEMES

The procedure that I have followed in grouping the accents about certain norms has been first to look for similarities and differences in meaning, and then to try to match them with similarities and differences in form. This reverses the approved order of business, but had to be adopted because pitch contours are if anything more fluid than meanings.

To take meaning as the starting point can be justified only if the accents are meaningful units. It assumes that they are morphemes, not phonemes. This assumption holds even in the most restricted sense of accent as mere fact-of-prominence. It opposes the current American theory, which is that stresses are meaningless sub-units that build themselves into morphemes ("superfixes") which give, in effect, a phonemic shape to syntax. I regard this view as untenable.³⁸ The older theory, that accent signals emphasis, has been re-stated by Weinreich for Yiddish: phonemic stress within the morpheme (i.e., word stress) is "the place at which relative loudness [sc. pitch change] occurs if the morpheme is emphasized."³⁹ This is to say that stress is phonemic only in the sense that a given syllable and not some other syllable within the morpheme carries the potential for pitch accent. When the accent occurs, it signals emphasis, i.e., is meaningful.

Accent A is assertive. It is used with items that are separately important, contrastive, and/or new to the discourse. It usually singles out the morph on which it falls, but this function of separation ("contour separation" in Pike's terminology) may be absent. I have heard a single syllable given as many as three successive A accents:⁴⁰

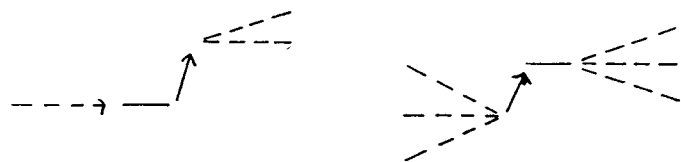


Also, instead of singling out one item, the A accent may put em-

³⁸ See my "Stress and Information," *American Speech*, XXXIII (1958), 5-20, and "Intonation and Grammar," *Language Learning*, VIII (1957-58), 31-38.

³⁹ Weinreich, "Stress and Word Structure," p. 2. The relevant part of this statement is that accent is not phonemic, whereas potential for pitch accent is. Whether Weinreich is right in not admitting any other kind of word stress than the potential is another matter. It has been disputed by Stockwell, *Language*, XXXII (1956), 374-383.

⁴⁰ This freedom to "spread" syllables in order to accommodate more accents reflects a certain degree of freedom in the location of the potential for pitch accent in English. Not only is the normal locus of the potential free, i. e. unpredictable in terms of the



Examples of the accents:

B

Do you really hate your brother?

B

Do you really hate your brother?

B

Do you really hate your brother?

B

Do you really hate your brother?

A

Do you really hate your brother?

A

Do you really hate your brother?

A

Do you really hate your brother?

A

Do you really hate your brother?

C

Do you really hate your brother?

C

Do you really hate your brother?

B

Do you really hate your brother?

B

Do you really hate your brother?

B

Do you really hate your brother?

Minimal contrast between A and B

(As if impatiently repeated)

(An echo question)

have followed this lead.⁹ The unanimity of their opinion is a roadblock that must be got out of the way before an explanation based on pitch can gain acceptance. The removal will occupy a large part of this article.

I shall refer to *intensity*, the physical term for amplitude of sound waves, rather than *loudness*, the psychological impression that varies directly with amplitude, because the experiments are based partly on measurements of intensity. On the other hand, I shall refer to *pitch*, the psychological term, rather than *fundamental frequency*, the physical term, for the sake of brevity.

DEFINITION. By *stressed syllable* I mean one that occurs IN AN UTTERANCE with the kind of prominence that listeners identify as "stress." The same distinction can be made for English that Charles A. Ferguson makes for Persian: "The syllable on which the stress falls when a given word is uttered in isolation is said to have 'inherent' or 'potential' stress, or simply the WORD STRESS. It must be noted that this concept of word stress is essentially morphological. A statement of the type 'The word X in Persian has (word) stress on the third syllable' means in effect that the word X has two alternants, one with stress on the third syllable, one with no stress at all."¹⁰ Stress as I use it does not refer to potential, or word, stress, but to stress that is actually there, imposed within an utterance. And for the most part the stress that I shall use as an example is the most prominent one in each of the utterances tested.

EXPERIMENTS BY OTHER INVESTIGATORS. Though a number of competent investigators have expressed doubts about intensity,¹¹ the only recent published experimental study to my

⁹ Pike (pp. 83, 96) gives first place to intensity, but allows that duration may be a factor. Trager-Smith (*Outline*, § 1.61) correlate stress with loudness, and Edith C. Trager (in *General Linguistics*, II [1956], 2) says unequivocally, "there is only one component that matters—loudness."

¹⁰ "Word Stress in Persian," *Language*, XXXIII (1957), 124-125. See also Uriel Weinreich, "Stress and Word Structure in Yiddish," in *The Field of Yiddish: Studies in Language, Folklore, and Literature* (New York, 1954), especially § 3.1.

¹¹ Arnold says (pp. 440-441), "articulatory force is frequently a difficult and, sometimes, an impossible yardstick for the recognition of linguistic stress in English." He is not concerned with the syllable that carries the main stress, but, given that syllable, with predicting where other stresses will fall, which he does on rhythmic principles. His statement therefore is a denial of the importance of intensity not on the main stress, but on the stresses that Smith and Trager would mark with /' and /˘/. They, of course, insist on loudness there also.

Using an approach similar to Arnold's, Wiktor Jassem, in "Stress in Modern English,"

knowledge that calls it into question is "The Linguistic Relevance of Intensity in Stress," by H. Mol and E. M. Uhlenbeck.¹² The authors adduce arguments from Franz Saran (1907), who pointed to the ear's well-known sensitivity to pitch and its relative unresponsiveness to intensity, but who failed to dissuade other phoneticians from their adherence to dynamic stress. Mol and

Bulletin de la Société Polonaise de Linguistique, Fascicule XI, pp. 23-49, is even more positive: "A definition of stress in terms of absolute force (or loudness) is an obvious impossibility. That the absolute breath-force (or loudness) of what are generally termed 'stressed' syllables may, under certain circumstances, be much less than the absolute force (or loudness) of 'unstressed' syllables hardly requires elucidation" (p. 29), and he cites Jespersen's *Lehrbuch*, p. 116, on this point. For him "the 'tonal accent' [is] wrongly called 'stress'" (p. 38).

Both Arnold and Jassem agree with the position that I adopt, which is that there is a special kind of syllabic prominence due to pitch, which should not be confused with other phases of stress.

H. E. Palmer "regards stressed syllables as those which take 'nucleus tones' in contrast to all other syllables which are pronounced with equal or nearly equal force" (quoted by Newman, p. 178; see below). Palmer's nuclei (see his *English Intonation* [Cambridge, 1922]) are the nearest to the pitch accents that I describe in this article. The principal difference is that the nuclei are sentence stresses (p. 7), and while other syllabic prominences due to pitch are scored in his notation, he lumps them under "heads." For example,

What a ^{re}mark^{ably} pretty little house.

contains, for him, a nuclear tone on *house*, and what precedes is merely a "broken scendent head" (p. 46). As set forth in § 7 of this article, I would put pitch accents on *what*, *-mark-*, and *pret-*, those on *-mark-* and *pret-* being the same as the "nuclear" in Palmer's (p. 39)

Could you spare me one?

on the word *spare*.

Stanley S. Newman, "On the Stress System of English," *Word*, II (1946), 171-178, holds to the intensity theory but with reservations: "Altho force of articulation is the primary medium thru which the stress phonemes are externalized, this phonetic feature is not the exclusive medium of stress" (p. 171). He distinguishes between "expressive accents," in which articulatory force is secondary to or at least equalled by pitch and quantity (p. 173), and "stress accents," where intensity predominates. My view differs in that a large part of what Newman puts under the stress accents I would put under the expressive accents.

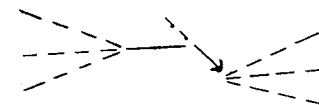
E. A. Glikina, in "An Attempt at an Experimental Study of the Elements of Dynamic Stress (With Reference to English)" (in Russian), *Voprosy jazykoznanija*, 1958, no. 5, 18-85, showing no awareness of recent laboratory investigations in the United States, and employing no synthetic devices, found that in 250 pronunciations of compounds by native Englishmen, accentual prominence was generally achieved by a combination of intensity, duration and pitch, although any one factor could render the distinction when the others were equal.

¹² *Lingua*, V (1956), 205-213. The same authors adduce other theoretical arguments against intensity in a later article, *Lingua*, VI (1957), 346.

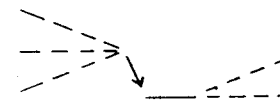
least as common (the affective nuances need to be correlated as much with dominant social attitudes as with "dialect"). One rather sharp deviation, which for semantic reasons I would class as a sub-type of A, puts the accentable syllable at a lower pitch than the one immediately following, but requires that only that one weak syllable remain high—the syllable after it must come down rapidly. An example from a radio dramatization,³⁷ with accent on *out*:

I am seldom out of my castle

The least common denominator in all A's is the abrupt fall rarely more than two syllables after the accentable syllable. In the following diagram the arrow represents a skip or skip-like motion, and solid lines denote essential movements while dotted lines indicate optional ones:



Accent C: This is a kind of anti-accent A, both in form and in meaning. The accentable syllable is approached from above, and skipped down to. What follows may level off or rise, but a further fall seems to be avoided:



Accent B: The characteristic of this accent is upmotion. It is neither skipped down to nor skipped down from. It may be approached from below and skipped up to, with the following motion continuing level, or rising (the usual thing), or falling slightly (an abrupt drop would create an A). Or it may be approached from a relative level and skipped up from, after which the movement usually continues upward slightly or levels off. This makes two diagrams necessary:

³⁷ Charles Dickens, *The Poor Relation*, broadcast from Station WABC, New York, 7:55 P. M., December 24, 1956.

Pike's '4-3-4 contour (pp. 56-57) is an example of this type of A accent compressed on a single syllable, *yes* and *no*, prolonged for the purpose.

Because each distinctive accent is thus capable of maintaining its identity despite broader intonational changes, I believe that it is practical to recognize more than the mere fact of prominence, of obtrusion per se, i.e., to describe and identify the different ways of achieving prominence. These are the pitch accents.

Test 13 undertakes to show identifications among similar manifestations of accent in dissimilar intonational settings, and discriminations between dissimilar accents in similar intonational settings. The experimenter arranged the phrase *I did* in the following four ways:

- (1) I (2) I (3) I di (4) I di
 did di_d d d

These were presented in 12 separately recorded groups of 3 as an ABX test, i.e., the listeners were asked to match X with whichever of the two preceding stimuli it resembled more.³⁶ Pattern (1) was present in each set of 3. Six listeners participated. A majority in every instance treated (1) as the orphan. This supports the theory that the four utterances have only two patterns of accent: in (1), the accent is formed by skipping down to the accentable syllable. In (2), (3), and (4) it is formed by moving rapidly down FROM it.

There seem to be sufficient resemblances and sharp enough differences among accents to group them about three kinds of obtrusion, which I describe and label as follows:

Accent A: A relative leveling off of the accentable syllable followed by a relatively abrupt drop, either within the accentable syllable (which is prolonged for the purpose) or in the immediately following syllable. In very rapid speech the drop may be postponed to the second following syllable, but rarely beyond this. There are affective (intonational) differences between a drop within the accentable syllable and a drop later.

The commonest type of approach ("head" in Palmer's terminology, "precontour" in Pike's) to the A accent in American English is from a lower pitch, with a skip up to the accentable syllable. In British Received Pronunciation, if descriptions are accurate, an approach from a higher pitch would seem to be at

Uhlenbeck experimented with the *pérnil*—*permil* pair, altering intensities in such a way that the stresses could not be signaled by them, and finding that nevertheless the stresses were clearly heard. They are perhaps too categorical about the absolute irrelevance of intensity, but their conclusions are probably true as regards its relative importance: "It is obvious that the decoding-system of the ear does not use any amplitude information. The system is concerned with the recognition of shapes or forms." The prominences to which I have referred are, of course, shapes—configurations of pitches.

At about the same time, Dennis Fry was conducting experiments at Haskins Laboratories using similar sets of minimal pairs, to determine whether pitch or duration was the better cue to stress. He had previously used the sets in pitting intensity against duration, and found that duration on the whole is a better cue.¹³ The later experiments showed pitch to be superior to duration,

¹³ "Duration and Intensity as Physical Correlates of Linguistic Stress," *Journal of the Acoustical Society of America*, XXVII (1955), 765ff.; "Experiments in the Perception of Stress," *Language and Speech* I (1958), 126-152.

In a sense the separation of duration and intensity may be a false dichotomy, if what we are seeking is determinations of "loudness." It is possible that increasing only the duration of a syllable, and not its intensity, may cause at least some hearers to report it as "louder," owing to the integration of intensities over a period of time: where two otherwise identical syllables have the same intensity level but one lasts twice as long as the other, the longer syllable could then be said to have twice as much acoustic energy. Another way of putting it is that there is a point below which the failure to discriminate between changes of intensity and changes of duration is determined biologically (sensory discrimination) rather than culturally (function of naming). Apparently duration does operate in some such way, at threshold—a sound with a given intensity which would be altogether inaudible becomes audible when prolonged (see W. R. Garner and G. A. Miller, "The Masked Threshold of Pure Tones as a Function of Duration," *Journal of Experimental Psychology*, XXXVII [1917], 293-303). More recent evidence indicates, however, that above threshold, and at durations significant for judgments of syllabic stress, this may no longer hold. The experiments of Irwin Pollack ("Loudness of Periodically Interrupted White Noise," *Journal of the Acoustical Society of America*, XXX [1958], 181-185) suggest that for noises lasting longer than 1/10 second there is little change in loudness when duration is increased (provided the intervening "silences" are less than 1/3 second—a condition which is handily met in the normal stream of speech); and he refers to another study which sets the critical duration at around 1/20 second. A glance at Fig. 4 will show that durations of syllables normally regarded as stressable are well beyond the 1/20 to 1/10 second minimum, so that a further increase in duration would not create an impression of greater loudness, if these experiments have any bearing. I mention the point only because of the incidental observations in this article about duration and intensity, not because it affects the priority of pitch over intensity, or duration, or duration-intensity, however one may choose to look at the latter complex.

³⁶For an example of ABX procedure, see Liberman, Harris, Hoffman, and Griffith, "The Discrimination of Speech Sounds Within and Across Phoneme Boundaries," *Journal of Experimental Psychology*, LIV (1957), 358-368.

from which it is reasonable to infer that pitch is superior to intensity.

A more direct attack was made by J M. Cowan, in an experiment whose results have not yet been published but which he has been kind enough to explain in a personal letter. It consists, essentially, in producing speech from which all intensity-control has been cut out. A harmonic-rich tone is fed from an artificial larynx through a tube to the back wall of the pharynx, where the operator, articulating in a whisper, re-emits it. The pitch of the tone is continuously variable; besides the tone-control there is an on-off switch to contrast voicing and non-voicing, but there is no control for intensity. Listeners report very good "stress" contrasts.

TESTS 1, 2, AND 3: INTENSITY AND PITCH IN NATURAL SPEECH.

Test 1. The sentence *Wouldn't it be easier to wait?* was recorded by the experimenter in two patterns of pitch, as seen in Fig. 1. In one (solid line), the syllable *wait* was "obtruded" from the pitch line by giving it a rapid fall. In the other (dotted line), the syllable *ea-* was obtruded by putting the rapid fall immediately after it, while the syllable *wait* was "embedded" in the following gradual curve. (This pitch contrast is the same one reported by James Sledd for the sentence *It's utterly ridiculous*, opposing the syllables *ut-* and *-dic-*.¹⁴) Intensities (peak value on vowel of syllable) were uniform in the four syllables: *ea* = *wait* = *ea* = *wait*.

Eight listeners¹⁵ were asked to say whether *easier* or *wait* contained the major stress. They had no difficulty, despite the lack of contrast in intensity: the vote in each instance was in favor of the similarly obtruded syllable.

Test 2. Smith and Trager¹⁶ utilize an example that has been widely quoted: *The Pennsylvania Railroad is the main Pennsylvania railroad.* Without giving further details, it is enough to say (1) that *Pennsylvania Railroad* and *Pennsylvania railroad* are supposed to be distinguished by the arrangement of dynamic stresses (loudnesses), and (2) that in both instances in normal speech the principal stress on *railroad* is supposed to be louder than that on *Pennsylvania*.

¹⁴ In *Litara*, III (1956), 38.

¹⁵ Except as otherwise noted, listeners in all experiments were researchers and technicians at Haskins Laboratories.

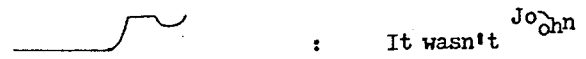
¹⁶ § 4.3.

left-hand one and at the trough in the right-hand one, we get entirely different meanings:

(1) It *wasn't* Jo^hhn

(2) It *wasn't* John

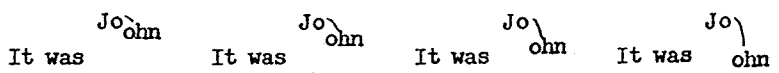
For a contrastive context in which *John* comes as a potentially new datum, (1) is appropriate: "I wish I knew who did a thing like that." — "It wasn't John (that much I'm sure of, though as to who else it might have been I'm not saying)." But (2) is uncomplicated denial; *John* is not a new datum: "I think it was John." — "It wasn't John—what are you talking about?" This contrast is pointed out by Maria Schubiger.³⁵ On the other hand, if we take a superficially disparate shape:



and put *John* at the peak, we again have implications of new-datum and contrast. The over-all similar curves carry something more dissimilar in meaning than the over-all dissimilar curves.

The phonetically similar fact in the two shapes with similar meanings is the way in which the accent is formed: it levels off from whatever kind of motion precedes it, and is followed by an abrupt drop. This differs from the other shape in that there the accent is formed by a drop to the accentable syllable.

Now compare this all-or-none contrast (which is like the one on *intrigue* in the preceding section) with the gradient differences between



which can be laid on a smooth scale of 'degree of finality'. The pitch differences between the first and last in the series are great, but the accents are the same. There is no point along this scale, as there was in the utterance of Test 1 and Test 10 in which the listeners had to make up their minds whether they were or were not hearing a given kind of accent on the word *easier*, where it is necessary to say "yes" or "no" rather than "more" or "less." The all-or-none is the domain of the pitch accents; the gradient is the domain of intonation.

³⁵ "Again: Fall-Rise Intonations in English," p. 2, reprinted from *English Studies*, XXXVII (1956). Her example is *I hadn't expected to see him there*. She credits "colleagues of University College, London," with having called the contrast to her attention.

the accent. So in *who's*
I don't know *coming*
either *know* or *who's* can be accented, depending on which is reduced
in this fashion. Similarly

In
trigue
can be distinguished as to whether we have a speaker of one
dialect saying *intrigue* with one kind of accent (Accent A; see next
section) or a speaker of another dialect saying *intrigue* with another
kind of accent (Accent C).

7. PITCH ACCENTS A, B, AND C

It would be possible to view pitch accent as a matter of obtrusion
per se, relegating manners of approach, sustention, and take-off
to another level of analysis, i.e., to intonation. This might
seem the more advisable now that we know accent to be cued by
the same kind of phonetic stimulus that we have always known
operated in the sphere of intonation, namely, pitch change, and in
view of the obvious difficulty of separating the different manners
of achieving accent—apart from the fact of accent itself—from the
domain of intonation.

That pitch can function in two equally complex ways at the
same time, however, is well demonstrated by the tone-and-into-
nation languages. Chinese and Japanese use pitch phonemically,
and despite this fact have intonations in many ways similar to
those of English.³⁴ What happens in English with certain abrupt
changes in a wider curve of pitch suggests that the same kind
of line can be drawn between the pitch accents and intonation:



Two shapes like these are more alike, superficially, than they are
different. Yet if we put an accentable syllable at the kink in the

³⁴ See Y. R. Chao, "A Preliminary Study of English Intonation (with American
Variants) and its Chinese Equivalents," reprinted from *The Ts'ai Yüan P'ei Anniversary
Volume* (Supplementary Vol. I of the *Bulletin of the Institute of History and Philology
of the Academia Sinica*) (Peiping, 1932). See also Isamu Abe, "Intonational Patterns
of English and Japanese," *Word*, XI (1955), 386-398. Weinreich, "Stress and Word
Structure," treats stress in similar overlapping functions (phonemic, constructive,
contrastive, expressive).

Fig. 1

ACCENT A IN TWO DIFFERENT LOCATIONS FOR THE SAME SENTENCE

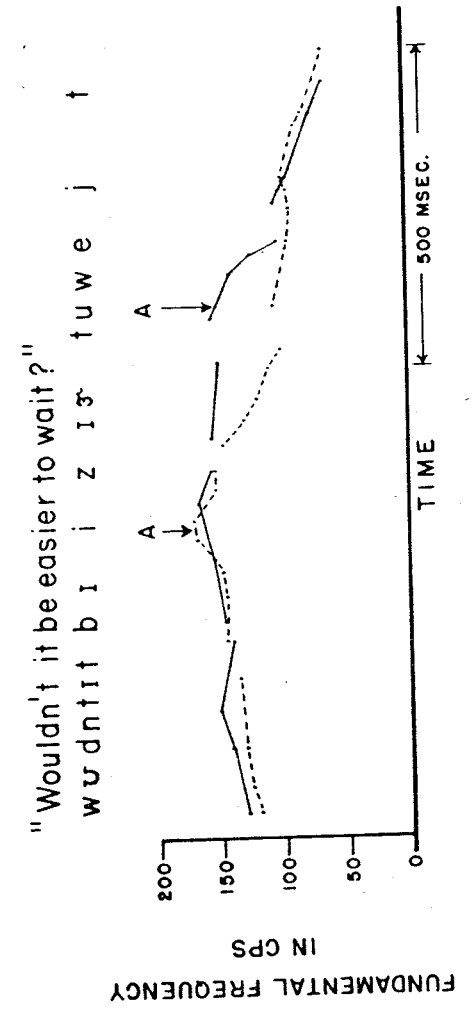
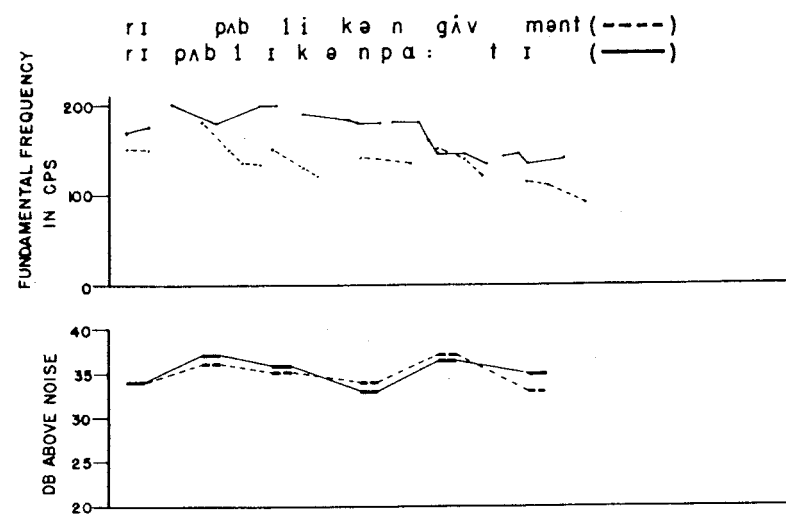


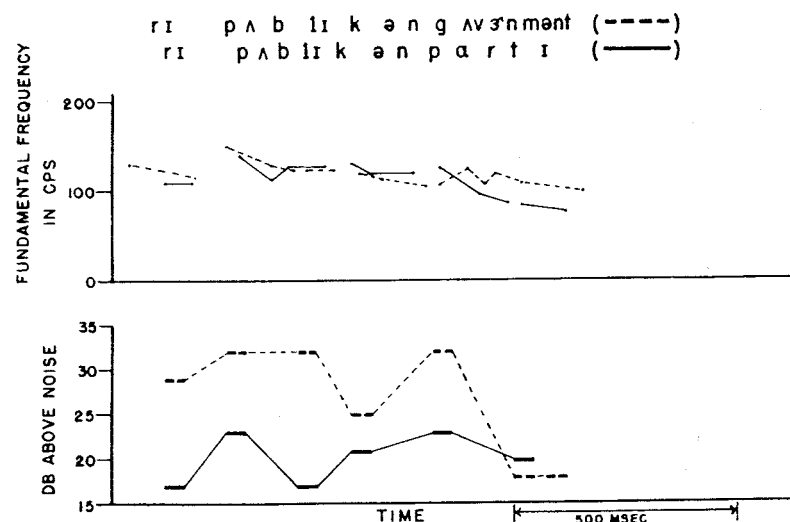
Fig. 2

"Republican Party" vs "Republican government"

SPEAKER A



SPEAKER B



hiccup (Outline § 920, footnote)—other than the one carrying the potential for pitch accent tends to be unstable. This is demonstrated by the history of loanwords in which the potential for pitch accent shifts its position. The nouns *hurricane*, *program*, *crayon*, *address*, *rupee*, *caravan*, and *mascol* were imported with the potential on the final syllable. They accommodated themselves to the normal accentual pattern of English nouns by shifting the potential forward. The syllable originally having the potential, however, did not lose its length immediately; for some speakers it is still long; for others it is short, with correspondingly centralized vowel (the ACD recognizes both stages for *mascol*, the earlier uncentralized [ɒ] and the later centralized shwa); for others it wavers. *Turquoise* has shifted without losing length, perhaps because of its syllable structure. *Intrigue* and *Portuguese* are shifting for some speakers, and have not had time to lose their length (which is furthermore supported rhythmically in *Portuguese*). This "medial stress" in many other words is uncertain: *cantlōn*, *cōncērt*, *infidēl*, *rēcōrd*, *cōnvēnt*, *lābōo*, *āncēstor*.

3. A short syllable to which the potential is shifted becomes long. I do not have measurements of this, but the evidence of vowel-decentralization seems clear enough: *pecan*, from [pi-] in *pecān* to [pi:-] in *pēcān*, in the dialects that have the latter; verb *survėy* [sɜːrˈveɪ] giving noun *sŭrvėy* [ˈsɜːrveɪ], etc. Verb-based nouns such as *cōme-ōn*, *cōmbine*, *impōrt*, *pėrvėrl*, *dīscārd*, and *trānsfėr* (or *trānsfer*) are typical. A favorite among linguists is the verb to *segmēt* (or *sėgmēt*), from the noun *sėgmēt*, with shwa.

4. A long syllable other than the one carrying the potential is often the syllable that carries the potential in a cognate or popularly associated word—the pitch accent here is an indirect source of length. Newman lists³³ many examples of this "underlying theme" relationship: *nāturalization* ~ *nāturalize*; *matērialistic* ~ *matērial*. In citation forms, and sometimes elsewhere, the syllable in question actually develops a secondary pitch accent.

If my assumption is correct, the speaker has the gradient co-variable of duration at his command in any case of ambiguity: by reducing the syllable (and also centralizing the vowel somewhat), he throws the balance toward another syllable as the receiver of

³³ Pp. 184-185. I would add most if not all of his examples of "sonorous weak" (pp. 186-187): *lēgality* ~ *lēgal*, *vībration* ~ *vībrate*, *īncūlcation* ~ *īncūlcate*.

the accents are laid. The English speaker's knowledge of the individual morphs and of morph and syllable types of his language gives permanent cues that serve as a foil to the phonetic nonce cues of pitch.³²

Looking a little farther, however, we find that pitch accent depends also on certain other phonetic cues, nonce like itself, of which one is especially important: duration. From the synchronic standpoint, duration is to a large extent a co-variable with accent. From the diachronic standpoint, duration accounts in part for how the unaccentables got that way.

A pitch obtrusion requires time for its execution. When the pitch accent is embraced completely by a single syllable, the syllable is lengthened to accommodate the necessary range of pitches; so *go* is probably longer in the first than in the second of the following two examples:

When did you *go*? When did you *go* there?

And the second is in turn longer than the *go* of

the
When did you *go* re?

in which *go* is embedded in what precedes the accent.

The experiments have made it clear that in the duration-pitch complex it is pitch that primarily signals accent. I therefore assume that duration is ancillary. Figuratively speaking, it is there IN ORDER to make room for the accent (though I would not say that duration has no other function). This is supported by the following considerations:

1. Accented syllables are normally longer than unaccented ones in comparable positions within the utterance (i.e., in reference to junctures). (Cf. Jones, *Outline*, § 870.)

2. A long syllable—and here I follow Jones in assigning to length and concomitant vowel quality rather than to intensity the prominence of the “medial stress” in words like *asphalt*, *leapold*,

³² The fact that the speaker must know the morphs disposes of the suggestion, often made, that nonsense syllables be used to test these patterns. The suggestion is a plausible one, for the worst obstacle to agreement about pitch stimuli is that there are always syntactic, morphological, and other cues present, and untrained listeners are apt to go off on a tangent, while trained listeners are apt to interpret them in terms of their training. Tests in which everything is stripped off except pitch would seem to be an answer; but, as I have tried to show, they will not work with the pitch accents.

As a less complicated example, since *Pennsylvania* contains what is usually marked as a secondary stress, I substituted the word *republican* in two contexts, which I infer to be analogous to those of *Pennsylvania*: (1) *He's a Methodist and belongs to the Republican Party* and (2) *It's too bad that Spain doesn't have a more republican government*. This test is reported more fully elsewhere,¹⁷ but I reproduce the tracings in Fig. 2 to show a visual display of pitches in natural speech.

Six speakers recorded the two sentences. Only two could be judged by ear to have made a distinction between *Republican* and *republican*; the others apparently inferred a contrast with *Democratic*. Spectrograms were made of the phrases *Republican Party* and *republican government* as uttered by these two speakers. The pitch and intensity information appears in Fig. 2.

As the curves show, there are two kinds of pitch prominence. In *Republican Party* (solid line), the syllable *-pub-* stands out by reason of being skipped up to; after that the line remains fairly uniform until the syllable *Par-*, where there is a sharp drop. In *republican government* (dotted line), however, *-pub-* is obtruded both by an upskip and by following rapid downmotion, and *gov-* is obtruded in the same way.

The information from the intensity profile is irrelevant, and in one instance actually runs counter to the Smith-Trager markings: Speaker A gives slightly more intensity to *-pub-* than to *Par-*. His recording gave no impression of a contrastive stress.

Test 3. In a suggestive article comparing German and English intonation patterns,¹⁸ Hugo Mueller observes that English numeral-plus-noun combinations tend to place the numeral on a higher pitch than the noun. He follows the Smith-Trager pitch-and-stress markings, giving an example such as ²*That look* ³*twenty minutes*¹#, and interprets as follows: “In English, the number tends to have the highest pitch in the phrase, although it does not bear the strongest stress.”

I judge this description to be correct as regards pitch. I would generalize it, however, to refer to all QUANTIFYING MODIFIERS, i.e., modifiers that show an amount or degree of something rather than a characteristic that distinguishes it from other things. I would expect the word *single* in *I couldn't contribute because I*

¹⁷ In my article “On Certain Functions of Accents A and B,” *Litara*, IV (1957), pp. 80-89.

¹⁸ “Some German Intonation Patterns and Their Relation to Stress,” *Modern Language Journal*, XL (1956), 28-30.

accept the downward syllabic prominence, but not always to interpret it as something "stressed."

What emerges from this is that while prominence as such gives the accent, accents are of more than one kind. It is best then to speak of PITCH ACCENTS, in the plural, and to look for kinships and differences. The search will also illumine that other part of the accentual complex, the unaccentable syllable.

6. FAVORED ACCENTS AND THE LEARNING OF ACCENTUAL STRUCTURE

Of the two arrangements of *It's the man who broke it* (p. 18), one gives *who* the same kind of pitch prominence that *man* has in the other. If accent were a matter of pitch only, *who* would have the same kind of accent as *man*. But I have used *who* as a means of warding off an accent, not of itself acquiring one, and referred to it as "unaccentable."³⁰ There is an apparent contradiction.

The explanation probably lies in the manner in which the lexicon is learned. A child confronted with the two patterns

daɪ dʒɛstəbəl daɪdʒɛs təbəl

would have no cue, from pitch, to tell whether the accent is on

to have "a protesting or exclamatory character" (*English Intonation*, p. 80). Jassem regards

You've de sɜ:rtəd hɪm?

as "stupefaction, surprise combined with either incredulity or disapproval" (*Intonation of Conversational English* [Wrocław, 1952], p. 29). The verbal context does suggest this, of course, but the speakers are probably restraining themselves. The words and pitches pulling against each other are reminiscent of familiar paradoxes of literature, "an icy smile," "a cheerless laugh." For Maria Schubiger, *English Intonation: Its Form and Function* (Tübingen, 1958), the lowered pitches are "regardful."

³⁰ Actually this is not quite true. *Who* may be accented as a means of affirmation. Normally the configuration would differ slightly from the way it was set forth in § 5 above: specifically, the *who* would rise in pitch above what precedes:

It's the man who broke it

'He is the very person who did it,' analogous to the rise in pitch on other function words for the same purpose of affirmation, commonest in the auxiliary verbs, e. g.

The man did break it

but also found on other words:

There's no chance to do it There's no way for doing it

Note the example in *Rumania*, text § 8. But these instances are too infrequent to matter in the stigmatizing of most function words as unaccentables.

Test 4. The synthetic sentence *Break both apart* was given various treatments of intensity and pitch, designed to throw stress on one or another element. Seven listeners were asked simply to indicate the syllable or syllables that they heard as stressed. Here are some of the stimuli and their results:

1. Syllables *break* and *-part* given an inflected pitch of 120 cps dropping to 100 (about 3 semitones), other syllables held at 100. Syllable *both* 10.5 db more intense than *break* (at least "twice as loud"), and 7.5 db more intense than *-part*. Majority hear stress on *-part*.

2. Syllable *break* inflected, 120-100 cps, other syllables at 100 cps monotone. *Both* 11 db more intense than *break*, 9 db more intense than *-part*. Majority hear stress on *both*. A massive increase in intensity overcomes a comparatively small inflection of pitch, in this particular pitch pattern.

3. Syllable *break* at 120 cps, rest at 100 cps. *Both* 2 db more intense than *break* and 1 db more intense than *-part*. A large majority hear stress on *break*. The smaller difference of intensity is overwhelmed by the difference of pitch.

4. All syllables at 80 cps monotone. *Break* 10.5 db more intense than *both* and *-part*. A small majority hear stress on *break*. The massive increase of intensity here does not do as well as the comparatively small rise of pitch in (3), in putting the stress on *break*.

5. Syllable *both* given an inflected pitch of 90-80 cps, the rest at 80 cps monotone. Syllable *both* 6 db less intense than *break* and *-part*, but heard as stressed by the majority.

Other patterns and treatments could be listed, but these are typical, and indicate the ease with which changes in pitch register as stress and the difficulty that changes of intensity have in competing with them.

Test 5 is designed to show the relative power of certain changes of pitch and intensity. The synthetic sentence *Many are taught to breathe through the nose* was played at a monotone, and without manipulations of its intensities, to six listeners, to find how they considered it to be stressed. In 7 out of 12 judgments they favored *breathe* (3 *many*, and 1 each *nose* and *taught*). Knowing that the listeners already favored *breathe*, the experimenter now set about increasing the preference by adding intensity to this word, making

it 3 to 7 db more intense than *laught* or *nose*.²¹ The object was to see how it would fare when changes of pitch were applied to these two other words.

The word *laught* was given 12 simple rises of pitch in 5 cps steps from 95 to 150 cps, the rest of the sentence being kept at a 90 cps monotone. The same treatment was given to *nose*, and in addition, the entire series was done again on *nose* with an inflected rise instead of a simple rise, the first half of the syllable being at the higher pitch and the second half at the reference pitch. Responses of the six listeners to the stimuli are shown in Fig. 3.²²

A pitch rise of from 5 cps (less than a diatonic semitone) to 15 cps (less than a minor third) is sufficient to overcome both the initial bias in favor of *breathe* and its added intensity.

Test 6. The Smith-Trager system allows for rises in pitch associated with stress, but considers them "allophones of the pitch phonemes."²³ One to whom this point of view has become second nature is likely to say, "Of course if pitch rises when intensity rises, the hearer who is given a stimulus in which there is a change in pitch automatically interprets it as a change in intensity; the pitch is only an indirect cue to the stress, but since it is the only cue present, it becomes effective." That is to say, a feedback is assumed from pitch, through intensity, to stress. While this argument hardly refutes the experiments where both pitch cues and intensity cues are present, it needs to be tested for whatever other plausibility it may have. This can be accomplished by asking the question in reverse: "When there are changes in intensity, can they be heard as changes in pitch?" If the answer is yes, this means that the gun shoots both ways.

The synthetic sentences *Break both apart* and *But would many return?* were given various treatments of pitch and intensity, including some in which the utterance was held at a monotone but one syllable was given an increase of intensity, and others in which the same utterance had both pitch and intensity changes.

²¹ The 3-7 fluctuation was not intended, but resulted from an inherent characteristic of the Voback.

²² It should not be necessary to point out that normal precautions were taken in administering this and other tests. The stimuli were randomized and four other stimuli, two with pitch changes on *many*, were introduced to divert attention from the systematic changes lest the hearers assume that *breathe*, *laught*, and *nose* were the only words that could contain the major stress.

²³ P. 43.

Test 12 shows this graphically. From the hand-drawn patterns of *undertaking* in Test 8, the one was chosen which had shown the clearest bias in favor of *un-* (the last pattern in Fig. 4). This pattern was then given pitch treatments similar to those in Test 5, with successive 5 cps rises on the syllable *-la-*. In addition to the upward obtrusion, downward obtrusions also in 5 cps steps were put on the same syllable. The resulting stimuli, randomized and mixed with other stimuli, were played to 46 naive listeners at the University of Connecticut. The data are schematized in Fig. 7. The top diagram, representing the upward obtrusions, is similar to those of Fig. 3. The bottom diagram, representing the downward obtrusions, leaves little doubt that a prominence in that direction is effective, but also makes it clear that upward obtrusions are superior to downward ones.

The hypothesis that I advance to account for at least part of the difference in effectiveness between the two directions is that when we ask listeners to point to a "stress," and then give them a downward obtrusion, we miscue them. Downward obtrusions give perfectly normal syllabic prominences, as the judgments of Test 11 prove. But "stress" popularly means 'being emphatic about something.' The downward obtrusion seems to contradict this. Its use in situations where gentleness or restraint is called for is widely recognized.²⁹ Listeners are therefore willing to

²⁹ Pike (*Intonation*, § 4.3.1) gives 'encouragement' as the meaning of

0
kaʔ

I have pointed out the frequency of the low pitch in questions addressed to a stranger, in "The Intonation of Accosting Questions," *English Studies*, XXIX (1948), 109-144. In Test 7 (text, above), listeners were apparently more reluctant to approve of added intensity on the bottom pitch than on higher pitches, suggesting a preference for "softness" there. Jones notes (§ 1069) that *All right* is usually

all
right

and that if the directions are reversed one "may have the effect of a threat." Fred W. Householder, Jr., calls my attention to an example by Martin Joos involving a full answer to a yes-no question: in response to *Is this an ashtray?* it would probably be

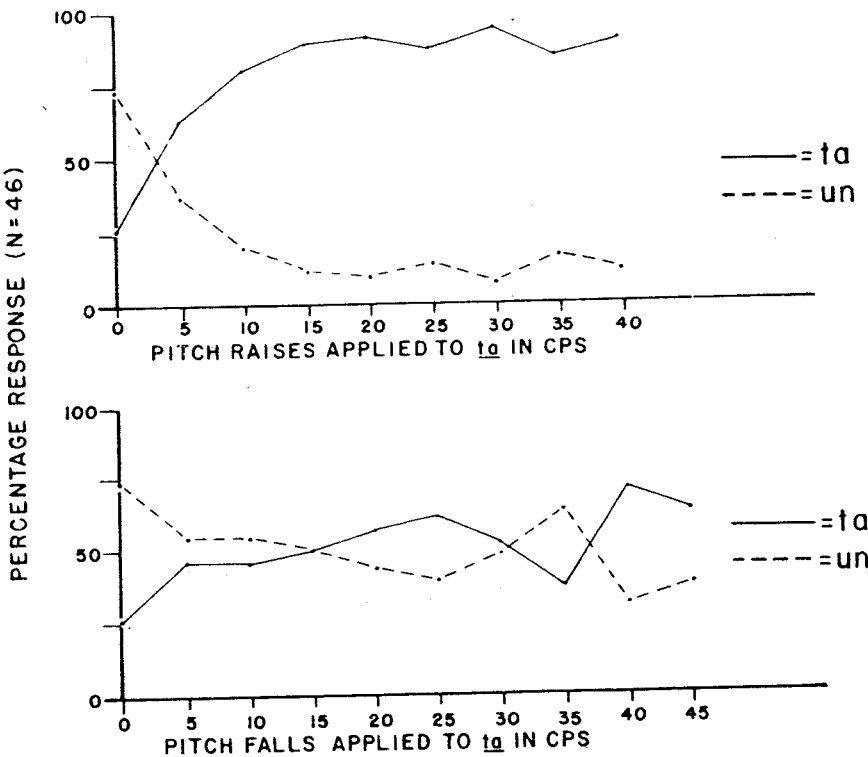
an
Is^o it's ashtraʔ

He suggests the meaning 'de-emphasis or restraint.' Palmer and Jassem, however, take their metaphors literally. Palmer considers

Be
patient

Fig. 7

VARIABLE PITCH VS CONSTANT INTENSITY - DURATION
AS A CUE TO LOCATION OF SYLLABLE STRESS

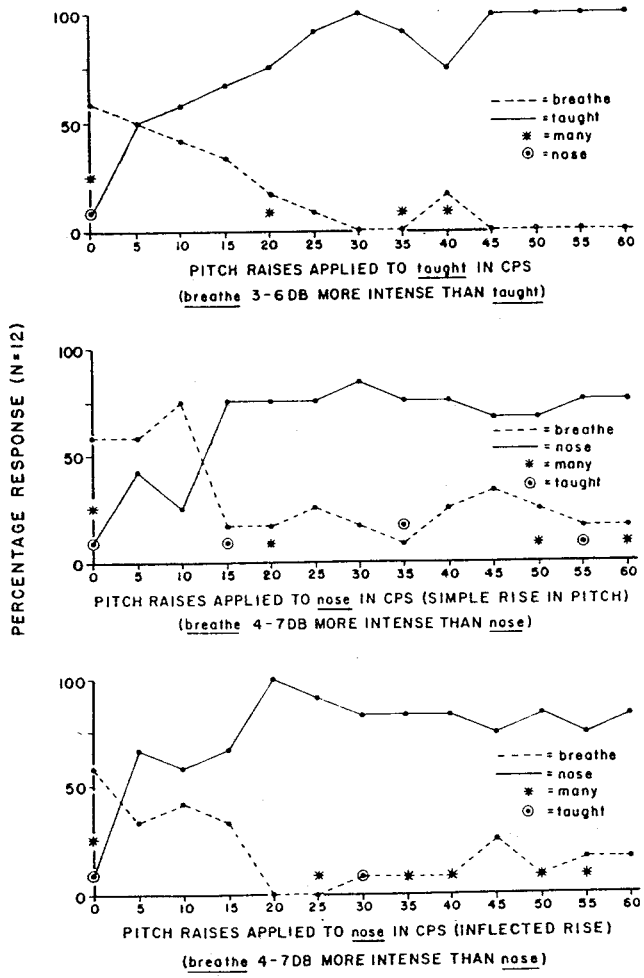


un	der	ta	king
180	70 45	170	250

DURATION IN MSEC.
(CONSTANT FOR ALL TEST STIMULI)

Fig. 3

VARIABLE PITCH VS CONSTANT INTENSITY - DURATION
AS A CUE TO LOCATION OF SENTENCE STRESS



ma	ny	are	taught	to	breathe	through	the	nose
182	238	58 168	182 126 112 70	476	196	182	84	518

DURATION IN MSEC.
(CONSTANT FOR ALL TEST STIMULI)

The listeners were asked to indicate the syllables that went higher or lower in pitch than the reference level.

Responses showed a tendency to hear a marked rise in intensity as a rise in pitch, when the stimulus was actually a monotone. This tells us that it is just as possible to mistake a change in intensity for a change in pitch as it is to mistake a change in pitch for a change in intensity. The feedback works both ways.

On the other hand, the responses showed no tendency to hear even a large drop in intensity as a drop in pitch. Instead, the subjective impression was that of some external interference, as if the speaker had turned his face away or the wind had suddenly snatched part of a word. Intensity seems to be significant only when it rises—as we would expect if it were a voice qualifier rather than something linguistic in the narrower sense. Configurations of pitch work in either direction—up or down; in fact, some of the listeners responded to changes in pitch without knowing which way they went. Pitch is again the more reliable cue, in respect of this greater flexibility.

Test 7. Assuming now what seems to have been amply demonstrated—that intensity is at best UNNECESSARY as a cue to stress and that pitch alone will serve so long as an utterance is kept reasonably close to the normal range of intensity and duration—one may ask whether intensity perhaps at least contributes to the QUALITY of a stress, to making it “sound right.” To inform ourselves on this point, we devised a test the results of which are published in another journal²⁴ and will merely be summarized here.

The synthetic sentence *Alexander's an intelligent conversation-alist* was given pitch-marked stresses at points where the intensity was BELOW the maximum for the utterance. Boosts of intensity were then applied at those points, and the resulting stimuli, with and without added intensity, were judged for quality by 62 listeners.²⁵ Minimum intensities turned out to be slightly preferred to small additions of intensity; small additions were somewhat preferred over large additions of intensity; and the greatest preference was shown for no addition of intensity as against large additions.

²⁴ Bolinger, “On Intensity as a Qualitative Improvement of Pitch Accent,” *Lingua*, VII (1958), 175-182.

²⁵ A pre-test by Laboratory personnel plus a large-scale test by naive listeners at the University of Connecticut.

Copies were made of the recording. In one copy, the second *go on* was cut out, and the remainder constituted Stimulus II. In the other copy, everything was eliminated except *Go on, go on*. A second voice now recorded a long passage which was interrupted in the middle, at which point the *Go on, go on* was inserted, whereupon the second voice resumed. The effect was that of someone reading, stopping before reaching the end, being urged by another person to continue, and finally completing what was begun. This interchange constituted Stimulus I.

The listeners were asked to indicate which word carried the accent (instructions were explicit, using the terms *stress*, *accent*, and *emphasis*, to avoid a possible tendency to judge phonetically without regard to context), and also to say whether what they heard sounded like normal English. Three groups of varying degrees of sophistication participated: A, Laboratory staff and researchers; B, a class in Spanish phonology; C, a sophomore-level college class. Responses are shown in Table 1.

TABLE 1. JUDGMENTS OF ACCENT AND ACCEPTABILITY IN TWO CONTEXTS OF *go on*. (In Group B the same two marked “not normal” as marked *ón*; in Group C, however, none of those who marked “not normal” were among those who were judging ‘continue’ as *go ón*. There is therefore no correlation here.)

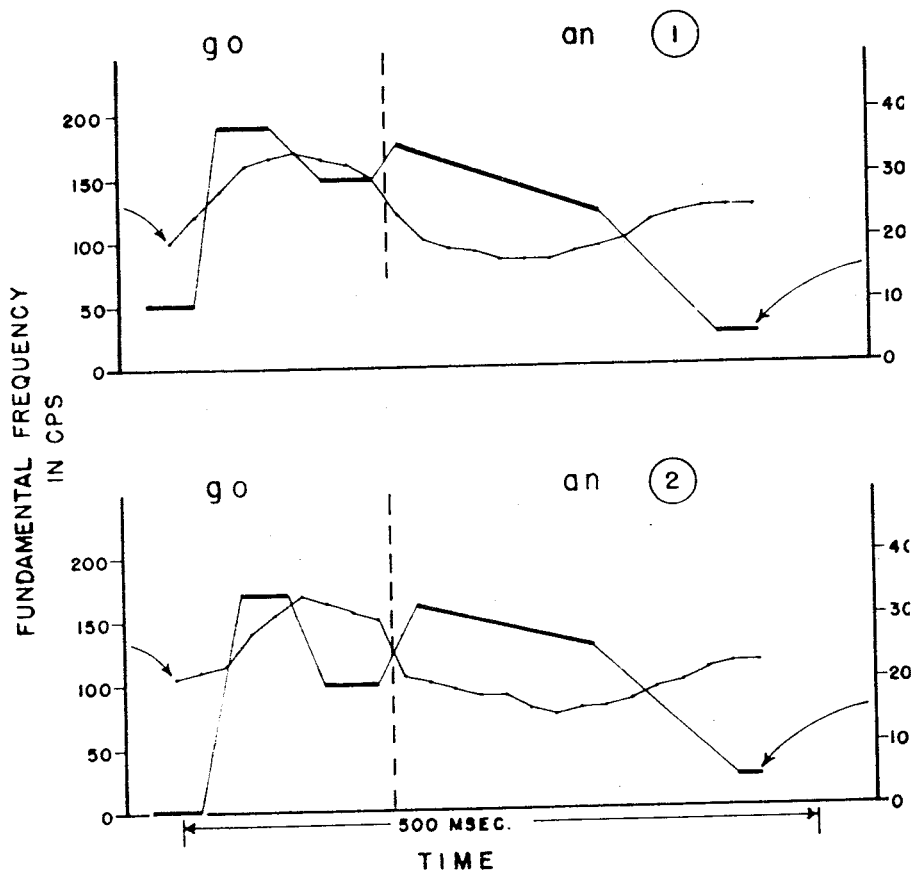
Group	Stimulus I ‘continue’				Stimulus II ‘rely on’			
	Accent on		Normal?		Accent on		Normal?	
	<i>go</i>	<i>on</i>	yes	no	<i>go</i>	<i>on</i>	yes	no
A	5	4	9	0	8	1	9	0
B	4	2	9	2	6	0	6	0
C	19	5	16	8	20	4	18	6
Totals	28	11	29	10	34	5	33	6

The experimenter's judgment that both instances are normal was confirmed. On the other hand, the greater tendency to locate the accent as *go ón* for ‘continue’ and *gó on* for ‘rely on’, while probably significant, is certainly not impressive. In all three groups the majority marked *gó* in both instances. Yet the analyst “knows” that ‘continue’ has to be *go ón*.

The reluctance of the listeners to hear the bottom pitch as an intentional prominence confirms the difficulty that we experienced in other tests, and suggests a reason why speakers—including phoneticians—instinctively look for correlations of HIGH pitch and stress. In the other tests, a downward obtrusion nearly always gave data which while reliable were not as clean as those obtained with an upward obtrusion.

Fig. 6

TWO PRODUCTIONS OF "Go on."



TESTS 8 AND 9: INTENSITY OPPOSED TO OTHER CUES THAN PITCH.

Listeners might well feel that in sequences like *Pennsylvania Railroad, republican government*, etc., the stress at the end gives the subjective impression of greater prominence than the earlier stress or stresses. With this I would agree. As we have seen, however, measurements of intensity do not confirm it. Neither do the measurements of pitch in *republican government*, where the same kind of prominence (upskip followed by rapid downmotion) is present in both, but the second, instead of being more marked than the first, is less marked.

The only remaining measurable factor that is usually recognized to play a role in stress is duration. But a fourth possibility suggests itself: position. It is conceivable that stress is climactic, and that we attribute extra intensity to the position at the end, even when it lacks it phonetically. This is suggested by our tendency to shift the sentence stress toward the end at the price, sometimes, of distorting the word stress as a result.²⁶ If position overrides pitch, which in turn overrides intensity, we have one explanation of why the end stresses are so consistently marked as "louder."

Test 8 pits the cues of pitch and position against each other. Five patterns of the word *undertaking* were hand-drawn with varying syllable lengths, as shown in Fig. 4. Various combinations of pitch were superimposed, and the resulting stimuli were judged by several listeners, who were told to indicate whether they heard the word *undertaking*, 'what a mortician does,' or *undertaking*, 'enterprise.' There were 16 pitch patterns comprising a total of 615 individual responses. In all but 3 patterns the majority of listeners reacted as the experimenter had predicted on the basis of pitch, and in only one of the 3 could the discrepancy be correlated with duration (i.e., the fault lay with a wrong interpretation of the pitches, not with the influence of duration). This confirms Fry's experiments where pitch overrides duration.

Fig. 4 tabulates the responses that bear on the problem of pitch and position. The first pitch pattern, a monotone, shows a bias in favor of *un-*. When *un-* and *-la-* are given equal pitch rises, however, the preference shifts radically to *-lá-*, and even when *un-*

²⁶ For examples, see Bolinger, "Intersections of Stress and Intonation," *Word*, XI (1955), 199-201, and "English Stress: the Interpenetration of Strata," § 6, in *Study of Sounds* (Tokyo, 1957), pp. 295-315.

STRESS JUDGMENTS FOR SEVERAL VERSIONS OF THE WORD "UNDERTAKING"

Fig. 4

FUNDAMENTAL FREQUENCY PATTERN IN CPS (UN-DER-TA-KING)																																							
80-80-80-80					90-80-90-80					100-80-90-80																													
STRESS HEARD ON:																																							
UN					TA																																		
170					65					170					65					155					70					320					KING				
210					40					95					60					180					30					310									
230					30					80					55					200					50					260									
145					80					85					215					50					250														
180					30					70					45					170					25					250									
RESPONSE TOTALS										22 18										6 29										16 24									
NUMBER OF SUBJECTS										8										7										8									

2. That when one of the items is involved in some movement of pitch, but not enough to make it stand out from its environment, the other, if more clearly marked, is heard as accented. For example:

a) Movement down to the accented syllable, other syllable embedded in gradual upmotion:

80 *Wouldn't* 90 *it* 95 *be* 100 *easy* 110 *er* 115
 to 80 wait? 6-3, *wait*
 80 *Wouldn't* 90 *it* 95 *be* 80 *easier to wait?* 8-1, *easier*

b) Movement down from the accented syllable, other syllable embedded in more gradual downmotion:

130 *Wouldn't it be* *easy* 110-100 *sier* 98-92 *to 90-70*
 wait ? 7-2, *easier*
 130 *Wouldn't it be* *easy* 130-120 *sier* 118-108 *to 100-70*
 wait ? 7-2, *wait*

These last two patterns reproduce the crux of Test 1—note the accent-forming 30 cps drop across *wait* in the second pattern as against the non-accent-forming 20 cps drop in the first.


c) Movement down from the accented syllable, with an undulation on the other syllable:

130 *Wouldn't it be* *easy* 110-100 *sier* 96-86 *to 94-70 wait?* 9-0, *easier*

3. That when two similar accents are obtruded with equal clarity, the second is more prominent (confirming the remark above about position). Example of two accents both skipped down to:

130 *Wouldn't it be* 115 *easier to* 100 *wait?* 8-1, *wait*

5. PREDICTABLE AMBIGUITY.

If prominences of the kind that I have described are the true basis of the accents, it should be possible to predict when a given utterance may be taken in more than one way. For example, in a configuration of the shape  there are two "corners," either of which may be the significant one. If the top corner, we describe it as an accent that is "skipped down from." If the bottom corner, it is an accent that is "skipped down to." The fact that the linguistic signal is mixed does not refute it, but

is made twice as prominent as *-la-*, *-lá-* is still favored by the majority. This confirms the tendency to hear the latter of two pitch-marked stresses as "louder."

Test 9. In the Smith-Trager system, constructs such as *light-house-keeper* and *light housekeeper* ('housekeeper who is light rather than heavy') are supposed to be distinguished by different loudnesses at different points. As the naive speaker is likely to think that the difference depends more on degrees of separation than on loudness, an experiment was performed to oppose the two cues of disjuncture (separation of syllable centers) and intensity. Here again it developed that the information from intensity was irrelevant, and that the difference was in fact signaled by separation.²⁷

With all the phonetic evidence pointing away from intensity as the thing that matters most in utterance stress, how can we account for its hold on the imagination? The answer lies, I think, in our folklore, and reflects not the acoustic signal but its linguistic function. The parts of our utterance that we stress most noticeably are the ones about which we want to be most forceful, and the kinetics of that force is felt and seen in a number of ways. A writer underlining the important parts of a message does it energetically. A speaker bobs his head and swings his arms in time with his stresses.²⁸ With this pugilistic obligato to the linguistic tune, it is hardly surprising that the tune became identified with the exercise that went with it; hence force, intensity, loudness. Finally, the pitch direction most fundamental to stress is up, and rising pitch bears a "pre-linguistic," physiological relation to physical tension.

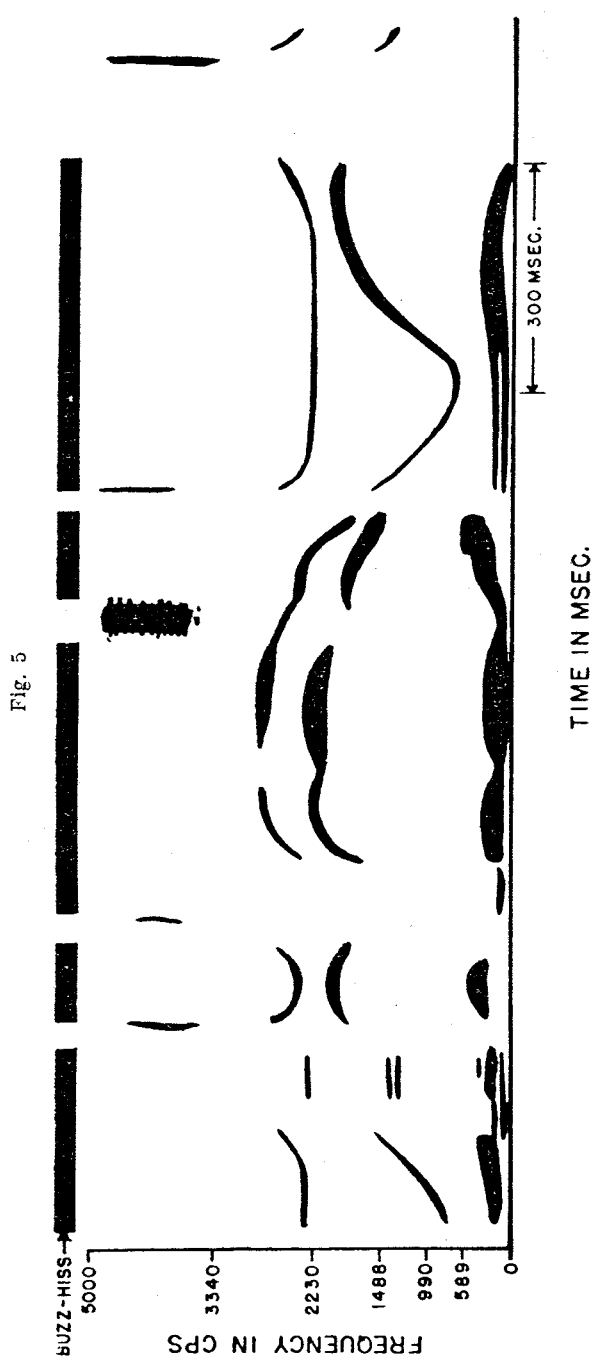
Having given up the more usual definition of stress, I think it is wise, because of associations, to give up the term also. From this point on I shall therefore refer not to stress but to PITCH ACCENT, or simply ACCENT, meaning prominence due to the configuration of pitches.

4. PITCH PROMINENCE IN GENERAL

The experiments relegating intensity to a secondary position have shown that it is unnecessary to take great precautions about

²⁷ See Bolinger and Louis J. Gerstman, "Disjuncture as a Cue to Constructs," *Word*, XIII (1957), 246-255.

²⁸ A. J. Vanvik, observing the orators at Hyde Park, noted that how they stressed a word such as *salvation* could be determined by their gestures without actually hearing the syllabic contrasts. See *Maître Phonétique*, No. 103, Jan.-June 1955, p. 8.



small random differences in intensity between one syllable and another in a synthetic utterance. Since the Voback does not deliver perfectly uniform intensities throughout the pitch range (nor does the human voice, though it is somewhat more stable in this regard), without these proofs it might have seemed necessary to specify intensities at every point. As it is, we may attend simply to pitch markings and let the minor fluctuations of intensity fall where they may.

Test 10 uses the same sentence, *Wouldn't it be easier to wail?*, as Test 1, but synthetic. The hand-drawn spectrogram is shown in Fig. 5. (This is done with white paint on transparent film. The pitch information, none of which appears here, is painted in a solid design at the top, with its upper edge tracing the rises and falls. The broken block lines directly beneath are for buzz (voice) and hiss (whisper). The rest of the painting controls the segmental phonemes—formants, transitions, bursts, releases, and frictions of the vowels and consonants.) Some 38 patterns of pitch were overlaid and judged by 9 listeners, who were told to indicate whether the principal accent ("stress" was the term used in the instructions) fell on *easier* or on *wail*. The responses indicated:

1. That when only one item is given pitch prominence, it is heard as accented. The pitch movement may be UP TO, DOWN TO, or DOWN FROM the accented syllable. Examples (the frequencies hold for all following syllables up to the next indicated frequency):

PATTERN	DIRECTION	VOTE FAVORING
100 <i>Wouldn't it be easier to</i> 120 <i>wail?</i>	up to	9-0, <i>wail</i>
100 <i>Wouldn't it be</i> 120 <i>easier to</i> <i>wail?</i>	— —	9-0, <i>easier</i>
100 <i>Wouldn't it be easier</i> 120 <i>to</i> 100 <i>wail?</i>	down to	9-0, <i>wail</i>
100 <i>Wouldn't it</i> 120 <i>be</i> 100 <i>easier</i> <i>to wail?</i>	— —	8-1, <i>easier</i>
100 <i>Wouldn't it</i> 120 <i>be ea</i> 100 <i>sier</i> <i>to wail?</i>	down from	9-0, <i>easier</i>
130 <i>Wouldn't it be ea</i> 110-100 <i>sier</i> 100 <i>to wail?</i>	— —	9-0, <i>easier</i>

Or the movement may be within the syllable:

100 <i>Wouldn't it be easier to</i> 100- 130 <i>wail?</i>	up in	8-1, <i>wail</i>
--	-------	------------------

Furthermore, the problem of terminal junctures needs to be re-studied in the light of pitch accent. Something like the endings that Smith and Trager describe is probably correct, but with refinements. For example, it seems to make little difference after a B accent whether there is a slight fall, a level, or a rise; the differences here are gradient. But after an A accent there seems to be an all-or-none difference between a level and a rise, but a gradient difference between a level and a fall. In answer to *George asked for his shovel back* one might get

A

I thought he ^{gave} it to you

with *to you* at a fairly high tonal level, as if the speaker were wondering and half-questioning; but if there is a rise in pitch on *you*, the utterance either takes on an admonitory tone or, though these words scarcely lend themselves, becomes a perplexed iterative question. On the other hand, if *to you* remains level, it may stand at any height, with gradient differences of suspension and conclusiveness; or it may fall for still greater conclusiveness, but without the sharp difference created by the rise.

10. CONCLUSIONS

1. Tests with both natural and artificial speech have shown that the primary cue of what is usually termed STRESS in the utterance is pitch prominence.

2. Intensity is found to be negligible both as a determinative and as a qualitative factor in stress.

3. To avoid unwanted associations, it is better to speak of PITCH ACCENT and to leave the term STRESS to the domain of word stress. In the latter domain, one possible kind of phonemic stress is POTENTIAL FOR PITCH ACCENT.

4. While the upward obtrusion is basic, pitch prominence need not be merely upward, as commonly supposed, but may take other directions. The differences of form respond to differences of meaning, giving the accents a morphemic status. Differentiating them often calls for a repertorial cue (the user's knowledge of the morphs of his language, and what syllables have the potential for pitch accent) or a gradient phonetic cue (length of syllable and grade of vowel).